

Marktmodellanpassung bezüglich stilisierter Fakten mittels Bootstrapping

Annalena Schröppel
(1313974)

Masterarbeit
am Lehrstuhl für Angewandte Mathematik
bei Prof. Dr. L. Grüne

am Mathematischen Institut
der Fakultät für Mathematik, Physik und Informatik
der Universität Bayreuth

27. November 2018

Inhaltsverzeichnis

0.1	Motivation	5
1	Einleitung	9
1.1	Grundlagen stochastischer Prozesse	10
1.2	Grundlagen empirischer Momente	12
2	Überblick über stilisierte Fakten	15
3	Messung stilisierter Fakten	25
3.1	Autokorrelationsfunktion der Erträge	29
3.2	Autokorrelationsfunktion der absoluten oder quadrierten Erträge . .	31
3.3	Hurst Exponent	33
3.3.1	R/S Analyse	34
3.3.2	Korrigierte R/S Analyse von Anis und Lloyd	35
3.3.3	Durchführung der R/S Analyse mit R	38
3.3.4	Empirische Konfidenzintervalle von Weron	42
3.4	Kurtosis	43
3.5	Hill Schätzer für den Tail Index	44
4	Verfügbarkeit geeigneter Datensätze	51
4.1	Datensatz von Robert Shiller	52
4.2	Anwendung des Datensatzes	57
5	Ermittlung des Fundamentalwertes	61
5.1	Definitionsansätze für den Fundamentalwert	61
5.1.1	Residual Income Valuation Model	62
5.1.2	Dividend Discount Model	63
5.2	Berechnungsmethode von Robert Shiller	64
5.2.1	Trendbereinigung realer Preise und Dividenden	64
5.2.2	Berechnung geschätzter Fundamentalwerte	75

6	Bootstrappingansatz	81
6.1	Bootstrappingmethode von Bradley Efron	81
6.2	Methode und Anwendung von Block Bootstrapping	82
7	Marktmodellanpassung	91
7.1	Aufstellung eines allgemeinen Marktmodells	91
7.2	Method of Simulated Moments	95
7.3	Schätzung der Modellparameter	100
8	Fazit	111
	Quellcodes	113

0.1 Motivation

Stellen wir uns vor, wir könnten Kursverläufe an Finanzmärkten perfekt realitätsnah simulieren: Wofür könnten wir ein solches Modell benutzen? Wir könnten das Resultat von externen Schocks wie zum Beispiel plötzliche Zinsanstiege oder -abfälle simulieren und untersuchen, ohne dass diese tatsächlich eintreten müssen. Außerdem könnten wir den mittleren Gewinn oder Verlust einer beliebigen Handelsstrategie ermitteln, oder auch mehrere Kursverläufe simulieren und dadurch herausfinden, wie stark die Preisentwicklung von zufälligen Einflussfaktoren abhängt. Die Kenntnis eines perfekten Simulationsmodells würde uns in einigen Fragestellungen weiterhelfen. Doch welche Eigenschaften muss ein solches Modell erfüllen, um realitätsnah zu simulieren? Wie können wir zwei verschiedene Modelle bewerten und beurteilen, welches von beiden besser ist? Wie können wir, wie in Abbildung 1, eine Simulationszeitreihe mit einer Zeitreihe aus der Realität vergleichen und eine Aussage darüber treffen, ob die Simulation gut oder schlecht ist? Mit diesen Fragen werden wir uns in dieser Arbeit beschäftigen.

Nachdem wir in Kapitel 1 einige grundlegende Begrifflichkeiten definiert haben, werden wir in den anschließenden Kapiteln 2 und 3 untersuchen, welche statistischen Eigenschaften reale Finanzmarktzeitreihen aufweisen und wie wir diese messen können. Die statistischen Eigenschaften sind unter dem Namen *stilisierte Fakten* bekannt. Ihre Messung hilft uns später bei der Beurteilung von Marktmodellen. Wir werden diejenigen Modelle als *gut* bewerten, deren Simulationszeitreihen ähnliche Messwerte ergeben wie reale Zeitreihen. Als Eigenleistung sind hier vor allem die Strukturierung der vielfältigen Literatur zum Thema stilisierte Fakten sowie die Erzeugung von Abbildungen zu deren Veranschaulichung zu bewerten. Im Allgemeinen habe ich alle Abbildungen in dieser Arbeit selbst erstellt. Außerdem gebe ich für einige stilisierte Fakten eigene Messideen an. In Abschnitt 3.3 erläutere ich ausführlich den klassischen Hurst Exponent, den wir zur Messung verwenden wollen, sowie die Korrektur von Anis und Lloyd und hinterfrage selbstständig die Bedeutung der verschiedenen, mit R berechneten Werte. Die Extremwerttheorie ist notwendig, um den Hill Schätzer zu verstehen, den wir ebenfalls als Maßzahl verwenden wollen. Aus diesem Grund erläutere ich die Grundlagen der Theorie in Abschnitt 3.5 in eigenen Worten und führe einen Beweis.

Zum Vergleich benötigen wir stets reale Finanzmarktzeitreihen. Deshalb beschäftigen wir uns in Kapitel 4 mit der Frage, wo wir geeignete Zeitreihen finden können. Hier sind die genaue Recherche zur Bedeutung der Daten von Shiller (2018) sowie die angegebenen R-Quellcodes als Eigenleistung zu sehen. Allgemein wurden sämtliche, in dieser Arbeit angegebenen Quellcodes von mir erstellt.

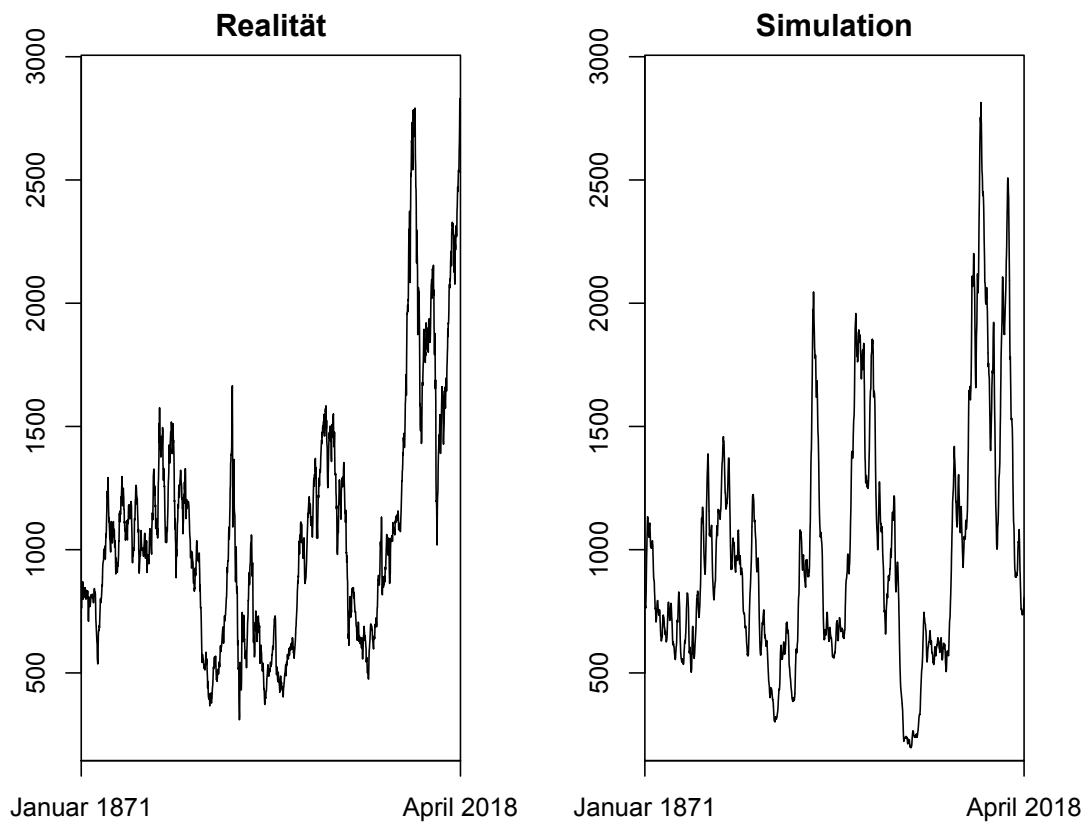


Abbildung 1: Links ist der reale und trendbereinigte Kursverlauf des Aktienindex S&P 500 zwischen Januar 1871 und April 2018 abgebildet. Rechts ist eine model-
 lergezeugte Simulationszeitreihe für den selben Zeitraum zu sehen. Wir stellen uns die Frage, wie gut das Modell die Realität nachahmt

Einige stilisierte Fakten sowie Handelsstrategien setzen die Kenntnis des sogenannten *Fundamentalwertes* voraus. Dies ist eine Art theoretisch hergeleiteter Wert des Finanzproduktes. In Kapitel 5 gehen wir genauer auf diesen Wert ein und beschreiben eine Berechnungsmethode. Als Vorarbeit müssen die realen Preise und Dividenden zunächst um eine gemeinsame Wachstumsrate bereinigt werden. Ich habe hierfür eine Idee entwickelt, wie wir die gemeinsame Trendbereinigung in einem geeigneten Sinne optimal durchführen können. Der Beweis der Optimalität, die Erzeugung von Abbildungen und der angegebene Quellcode sind Eigenleistung. Außerdem stelle ich in Theorem 5.2.2 eine geschlossene Formel der geschätzten Fundamentalwerte von Shiller (1981) auf und beweise diese.

Bei der Modellbewertung werden wir uns fragen, wie stark die einzelnen Maßzahlen ausgewertet an der realen Zeitreihe und an der Simulationszeitreihe voneinander entfernt sein dürfen, um das Modell noch als gutes Modell bewerten zu können. Als Hilfsmittel verwenden wir dabei das sogenannte *Block-Bootstrapping-Verfahren*, das wir in Kapitel 6 genauer erläutern werden. Mit diesem Ansatz können wir analysieren, wie stark die betrachtete Maßzahl in der Realität schwanken würde, wenn wir mehrere, reale Zeitreihen ziehen könnten. Bei einer hohen Schwankung in der Realität sind wir auch bei der Modellbewertung toleranter, d. h., wir akzeptieren eine große Abweichung zwischen der Maßzahl ausgewertet an der realen Zeitreihe und an der Simulationszeitreihe, und umgekehrt. Mittels Block Bootstrapping bestimmen wir also die Gewichtung der einzelnen Maßzahlen in unserer Modellbewertung. Eine wichtige Stellschraube des Block-Bootstrapping-Verfahrens ist die Wahl einer geeigneten Blocklänge. Deshalb gebe ich in Abschnitt 6.2 eine eigene Idee an, wie wir diese für unseren, konkreten Anwendungsfall definieren können.

Das erarbeitete Wissen können wir schließlich in Kapitel 7 nutzen, um eine Modellbewertung herzuleiten. Dort definieren wir zunächst ein allgemeines Marktmodell, das einige, frei wählbare Parameter enthält. Die Wahl der Parameterwerte beeinflusst die erzeugten Simulationszeitreihen und bestimmt, ob diese einen stabilen oder einen unregelmäßigen Verlauf mit vielen Schwankungen aufweisen. Gesucht sind Parameterwerte, sodass die Modellbewertung ein bestmögliches Ergebnis liefert, d. h., im Sinne der Messung stilisierter Fakten und deren Gewichtung sollen sich Simulationszeitreihen und reale Zeitreihen ähneln. Als Eigenleistung ist hier der gesamte Abschnitt 7.3 zu sehen. Dort führe ich unter anderem eine Rastersuche durch, um eine erste, gute Parameterkombination zu finden. Außerdem beschäftige ich mich mit der Rolle der verschiedenen Modellparameter, vergleiche Simulationszeitreihen und stelle mir die Frage, wie stark unsere Modellbewertung vom Zufall abhängt. Zuletzt hinterfrage ich das Ergebnis der Modellbewertung kritisch und gebe Verbesserungs- und Erweiterungsmöglichkeiten an.

Kapitel 1

Einleitung

Wir werden in dieser Arbeit Preise von Finanzprodukten, wie zum Beispiel Aktien, im Zeitverlauf betrachten. Die Zeiteinheiten werden dabei diskret angegeben, sodass wir von *Perioden* $t \in \mathbb{N}$ sprechen werden. Die Preise unterliegen in gewisser Weise dem Zufall und werden vom Bekanntwerden preisbeeinflussender Informationen gesteuert. Es wird deshalb sinnvoll sein, die Preise als stochastischen Prozess $(P_t)_{t \in \mathbb{N}}$ zu definieren.

Ein für uns wichtiger Begriff ist außerdem der *Ertrag* im Zeitverlauf. Darunter verstehen wir die relative Preisänderung innerhalb einer Periode $t \in \mathbb{N}$, d. h., die Erträge lassen sich direkt aus den Preisen berechnen und stellen somit selbst wiederum einen stochastischen Prozess $(R_t)_{t \in \mathbb{N} \setminus \{1\}}$ dar:

$$R_t := \frac{P_t - P_{t-1}}{P_{t-1}}, \quad t \in \mathbb{N} \setminus \{1\} \quad (1.1)$$

Der Ertrag R_t entspricht somit der Rendite, die beim Kauf des Finanzproduktes zum Zeitpunkt $t - 1$ und Verkauf zum Zeitpunkt t realisiert wird. Zur vereinfachten Notation werden wir auch für die Erträge die Indexmenge \mathbb{N} anstatt $\mathbb{N} \setminus \{1\}$ verwenden, d. h., wir nehmen an, dass sich der stochastische Prozess der Preise durch eine Zufallsvariable P_0 erweitern lässt, sodass durch (1.1) auch R_1 definiert ist.

Da stochastische Prozesse in dieser Arbeit eine zentrale Rolle spielen, werden wir im folgenden Abschnitt 1.1 einige Grundlagen behandeln und Begrifflichkeiten definieren. Außerdem wollen wir später Stichproben stochastischer Prozesse betrachten und diese statistisch auswerten. Deshalb werden wir in Abschnitt 1.2 auf notwendige, empirische Momente eingehen.

1.1 Grundlagen stochastischer Prozesse

Wir werden nun einige, grundlegende Definitionen bezüglich stochastischer Prozesse formulieren. Die Definitionen sind dem Lehrbuch von Schlittgen und Streitberg (1999, S. 90 ff.) entnommen.

Definition 1.1.1. *Ein zeitdiskreter, stochastischer Prozess ist eine Folge $(X_t)_{t \in T}$ von Zufallsvariablen*

$$X_t : \Omega \rightarrow \mathbb{R}, \omega \mapsto X_t(\omega),$$

wobei der (als Zeitparameter bezeichnete) Index t Element einer höchstens abzählbaren Indexmenge T ist. Dabei bezeichne $(\Omega, \mathcal{F}, \mathbb{P})$ den zugrundeliegenden Wahrscheinlichkeitsraum.

Wir werden uns in dieser Arbeit stets auf diskrete Zeiträume konzentrieren. Deshalb werden wir das Wort *zeitdiskret* in der Regel weglassen. Außerdem wählen wir üblicherweise die Menge der natürlichen Zahlen \mathbb{N} als Indexmenge T . Wir werden die Konvention befolgen, Zufallsvariablen groß zu schreiben und ihre Realisationen

$$x_t := X_t(\omega)$$

durch Kleinschreibung zu kennzeichnen.

Da X_t für jeden festen Zeitindex eine Zufallsvariable darstellt, kann ihr Erwartungswert betrachtet werden, sofern dieser existiert. Die Mittelwertsfunktion ergibt sich als entsprechende Funktion im Zeitindex $t \in \mathbb{N}$. Analoges gilt für die Varianzfunktion.

Definition 1.1.2. *Sofern für jedes $t \in T$ der Erwartungswert bzw. die Varianz der Zufallsvariable X_t existiert, ist die Mittelwertsfunktion $\mu_X(t)$ des stochastischen Prozesses $(X_t)_{t \in T}$ gegeben durch*

$$\mu_X(t) := E[X_t]$$

und die Varianzfunktion $\sigma_X^2(t)$ von $(X_t)_{t \in T}$ durch

$$\sigma_X^2(t) := \text{Var}[X_t] := E[(X_t - E[X_t])^2].$$

In der Finanzmathematik sagt man zur Standardabweichung in den Erträgen manchmal auch *Volatilität*. Wir werden uns in den folgenden Abschnitten dafür interessieren, inwiefern Preise bzw. Erträge zu unterschiedlichen Zeitpunkten voneinander abhängen, wobei wir die Preise und Erträge jeweils als stochastische Prozesse interpretieren. Wir benötigen also ein geeignetes Maß, um das Abhängigkeitsverhältnis eines stochastischen Prozesses zu untersuchen. Deshalb definieren wir nun die Kovarianzfunktion und die Korrelationsfunktion.

Definition 1.1.3. Sei $(X_t)_{t \in T}$ ein stochastischer Prozess, dessen Varianzfunktion existiert. Die Kovarianzfunktion $\gamma(s, t)$ ordnet jedem Paar von Zeitpunkten $s, t \in T$ die Kovarianz der Zufallsvariablen X_s und X_t zu:

$$\gamma(s, t) := \text{Cov}[X_s, X_t] := E[(X_s - \mu_X(s))(X_t - \mu_X(t))]$$

Die Korrelationsfunktion $\varrho(s, t)$ des stochastischen Prozesses ordnet jedem Paar von Zeitpunkten $s, t \in T$ die Korrelation

$$\varrho(s, t) := \text{Corr}[X_s, X_t] := \frac{\text{Cov}[X_s, X_t]}{\sqrt{\sigma_X^2(s) \cdot \sigma_X^2(t)}}$$

der entsprechenden Zufallsvariablen zu.

Es lässt sich zeigen, dass die Kovarianzfunktion eine positiv semidefinite, symmetrische Bilinearform im Raum der quadratisch integrierbaren Zufallsvariablen ist. Also können wir die Cauchy-Schwarz'sche Ungleichung anwenden und erhalten $\gamma(s, t)^2 \leq \sigma_X^2(s) \cdot \sigma_X^2(t)$. Somit folgt aus der Existenz der Varianzfunktion die Wohldefiniertheit der Kovarianzfunktion. Gilt zusätzlich $\sigma_X^2(t) > 0 \forall t \in T$, dann ist auch die Korrelationsfunktion wohldefiniert.

Häufig sagt man auch *Autokovarianzfunktion* und *Autokorrelationsfunktion* zu den definierten Funktionen. Verwendet man die Definition der Varianz, so erhält man einen Zusammenhang zwischen der Autokovarianzfunktion und der Varianzfunktion:

$$\gamma(t, t) = E[(X_t - \mu_X(t))^2] = \text{Var}[X_t] = \sigma_X^2(t) \quad \forall t \in T \quad (1.2)$$

Eine weitere, wichtige Begrifflichkeit ist die sogenannte Stationarität. Hierbei wird charakterisiert, welche Rolle der Zeitpunkt $t \in T$ in der Mittelwerts-, Varianz- und Kovarianzfunktion spielt.

Definition 1.1.4. Ein stochastischer Prozess $(X_t)_{t \in T}$ heißt

- mittelwertstationär, wenn $\mu_X(t)$ konstant ist:

$$\mu_X(t) =: \mu \quad \forall t \in T$$

- varianzstationär, wenn $\sigma_X^2(t)$ konstant ist:

$$\sigma_X^2(t) =: \sigma^2 \quad \forall t \in T$$

- kovarianzstationär, wenn die Kovarianzfunktion $\gamma(s, t)$ des Prozesses nur von der Entfernung $s - t$ abhängt. Deshalb definieren wir für kovarianzstationäre Prozesse $\gamma(s - t)$ als

$$\gamma(s - t) := \gamma(s, t) \quad \forall s, t \in T.$$

- schwach stationär, wenn er mittelwert- und kovarianzstationär ist.

Schwache Stationarität erlaubt somit die Vereinfachung, den Zeitparameter beim Mittelwert wegzulassen und die Kovarianzfunktion nur in der zeitlichen Differenz zu betrachten. Diese zeitlichen Differenzen $\tau := s - t$, $s, t \in T$ werden häufig als *Lags* bezeichnet, sodass man für schwach stationäre Prozesse die Kovarianzfunktion

$$\gamma(\tau) = \gamma(s - t)$$

als Funktion im Lag schreiben kann. Schwach stationäre Prozesse sind außerdem stets varianzstationär, denn es gilt

$$\sigma_X^2(t) \stackrel{(1.2)}{=} \gamma(t, t) = \gamma(0) = \gamma(s, s) \stackrel{(1.2)}{=} \sigma_X^2(s) \quad \forall s, t \in T.$$

Damit ergibt sich zusammen mit der Definition der Korrelationsfunktion sowie dem Zusammenhang (1.2) für schwach stationäre Prozesse, dass

$$\varrho(\tau) = \frac{\gamma(\tau)}{\sigma^2} = \frac{\gamma(\tau)}{\gamma(0)}$$

gilt.

1.2 Grundlagen empirischer Momente

Wie bereits erwähnt, werden wir die Preise $(P_t)_{t \in \mathbb{N}}$ bzw. die Erträge $(R_t)_{t \in \mathbb{N}}$ als stochastische Prozesse interpretieren. Eines unserer Ziele ist es, Eigenschaften dieser stochastischen Prozesse zu formulieren. Dabei können wir lediglich reale Datensätze der Finanzmärkte betrachten und versuchen, Rückschlüsse über den zugrundeliegenden stochastischen Prozess zu ziehen. Wir stellen uns vor, uns sei ein $(N + 1)$ -elementiger Datensatz $(p_t)_{t=0, \dots, N}$ gegeben. Mit den Preisdaten können wir die Erträge gemäß (1.1) bestimmen und nach Entfernung des Elements p_0 erhalten wir zwei N -elementige Zeitreihen $(p_t)_{t=1, \dots, N}$ und $(r_t)_{t=1, \dots, N}$. Um die Zeitreihe analysieren und schließlich Aussagen über den stochastischen Prozess treffen zu können, benötigen wir nun geeignete, statistische Kenngrößen. Hierbei spielen die Momente erster und zweiter Ordnung eine zentrale Rolle. Die im Folgenden angegebenen Definitionen orientieren sich an Schlittgen und Streitberg (1999, S. 3 ff.).

Im Folgenden bezeichne allgemein $(X_t)_{t \in \mathbb{N}}$ den zugrundeliegenden, stochastischen Prozess. Wir müssen nun annehmen, dass es sich um einen schwach stationären stochastischen Prozess handelt, damit die hier definierten Größen sinnvoll sind. Diese Annahme werden wir in Kapitel 3 für unseren Kontext rechtfertigen.

Definition 1.2.1. Das arithmetische Mittel oder auch der empirische Mittelwert einer Zeitreihe $(x_t)_{t=1,\dots,N}$ ist gegeben durch

$$\bar{x} := \frac{1}{N} \sum_{t=1}^N x_t.$$

Die empirische Varianz der Zeitreihe wird durch

$$s^2 := \frac{1}{N-1} \sum_{t=1}^N (x_t - \bar{x})^2$$

berechnet. Die Wurzel der Varianz

$$s = \sqrt{s^2}$$

heißt empirische Standardabweichung.

Die beiden Kenngrößen \bar{x} und s^2 bzw. s dienen als Maße für die zentrale Lage und die Stärke der Schwankungen in den Werten der betrachteten Zeitreihe. Die Division durch $N-1$ anstatt durch N in der Definition der empirischen Varianz sorgt für eine bessere Güte des Schätzers für die wahre Varianz. Genauer gesagt erhält man mit s^2 einen erwartungstreuen Schätzer für die Varianz. Nutzt man stattdessen den Nenner N , dann erhält man den nicht erwartungstreuen, d. h. verzerrten Maximum-Likelihood Schätzer für σ^2 (vergleiche Hartung et al., 1993, S.157). Um die beiden Varianten klar voneinander abzugrenzen, werden wir sie mit *empirischer* und *verzerrter* Varianz bzw. Standardabweichung bezeichnen. Die verzerrte Varianz wird manchmal auch *Stichprobenvarianz* genannt.

Zudem sind wir an der Abhängigkeit zwischen verschiedenen Zeitpunkten interessiert. Um auch hierfür ein geeignetes Maß zu erhalten, leiten wir das empirische Analogon zur Autokovarianz- bzw. Autokorrelationsfunktion her. Wir beginnen dafür zunächst beim Begriff der klassischen, empirischen Kovarianz, mit der ein linearer Zusammenhang zweier Zeitreihen spezifiziert werden kann.

Definition 1.2.2. Für zwei N -elementige Zeitreihen $x = (x_t)_{t=1,\dots,N}$ und $y = (y_t)_{t=1,\dots,N}$ sei die empirische Kovarianz gegeben durch

$$\text{Cov}_e(x, y) := \frac{1}{N} \sum_{t=1}^N (x_t - \bar{x})(y_t - \bar{y}).$$

Der Korrelationskoeffizient von Bravais-Pearson entsteht daraus durch Normierung mit dem Produkt der einzelnen, verzerrten Standardabweichungen

$$\text{Corr}_e(x, y) := \frac{\text{Cov}_e(x, y)}{\sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - \bar{x})^2} \cdot \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})^2}}.$$

Diese klassische Definition können wir leicht modifizieren, um nicht den Zusammenhang zwischen *zwei* Zeitreihen sondern zwischen Werten unterschiedlicher Zeitpunkte mit festem Abstand innerhalb *einer* Zeitreihe zu messen. Gesucht ist also eine statistische Kennzahl, die beispielsweise angibt, ob der jeweilige Kurs mit dem Kurs von vor einem Jahr zusammenhängt oder nicht.

Sei $0 \leq \tau \leq N - 1$ die Zeitdifferenz, bezüglich der wir den linearen Zusammenhang untersuchen wollen. Im Beispiel ist also $\tau = 12$, wenn eine Periode einem Monat entspricht. Nun können wir aus der Zeitreihe $(x_t)_{t=1, \dots, N}$ zwei $(N - \tau)$ -elementige Zeitreihen $(x_t^{(1)})_{t=1, \dots, N-\tau}$ und $(x_t^{(2)})_{t=\tau+1, \dots, N}$ bilden und deren Kovarianz

$$Cov_e(x^{(1)}, x^{(2)}) = \frac{1}{N - \tau} \sum_{t=1}^{N-\tau} (x_t - \overline{x^{(1)}}) (x_{t+\tau} - \overline{x^{(2)}}) \quad (1.3)$$

betrachten. Aufgrund statistischer Vorteile werden wir nun durch N anstatt durch $N - \tau$ dividieren (vergleiche Schlittgen und Streitberg, 1999). Da wir außerdem annehmen, dass der zugrundeliegende stochastische Prozess schwach stationär ist, sollten sich die beiden Mittelwerte in (1.3) nicht stark unterscheiden. Deshalb ersetzen wir sie durch den Mittelwert \bar{x} der gesamten Zeitreihe. Zudem sollen auch negative Werte für τ zulässig sein. Für $-(N - 1) \leq \tau < 0$ wendet man einfach das beschriebene Vorgehen auf $-\tau > 0$ an. Damit ist die Modifikation abgeschlossen und wir haben ein passendes Maß zur Untersuchung des Zusammenhangs der Werte unterschiedlicher Zeitpunkte gefunden:

Definition 1.2.3. Die Autokovarianzfunktion $\gamma_e(\tau)$ einer Zeitreihe $(x_t)_{t=1, \dots, N}$ ist die durch

$$\gamma_e(\tau) := \frac{1}{N} \sum_t (x_t - \bar{x})(x_{t+\tau} - \bar{x})$$

definierte Funktion des Zeitabstands oder Lags $\tau = -(N - 1), \dots, -1, 0, 1, \dots, (N - 1)$. Die Autokorrelationsfunktion $\rho_e(\tau)$ ist durch

$$\rho_e(\tau) := \frac{\gamma_e(\tau)}{\gamma_e(0)}$$

gegeben.

Nun haben wir die für diese Arbeit grundlegenden Begrifflichkeiten definiert. Im folgenden Kapitel 2 gehen wir auf die sogenannten stilisierten Fakten, d. h., typische, statistische Eigenschaften realer Finanzzeitreihen ein.

Kapitel 2

Überblick über stilisierte Fakten

In den vergangenen Jahrzehnten wurden Daten der Finanzmärkte in zahlreichen Studien untersucht, um typische Eigenschaften zu identifizieren. Dabei ist eine Reihe statistischer Charakteristiken entstanden, die bei Finanzzeitreihen üblicherweise beobachtbar sind. Diese sogenannten *stilisierten Fakten* sollen in diesem Kapitel detailliert erklärt werden. Zur Veranschaulichung werden dabei Daten des Aktienindex S&P 500 verwendet, auf die wir in Kapitel 4 genauer eingehen werden. Der entsprechende Datensatz der monatlichen Werte seit Januar 1871 wird auf der Homepage von Shiller (2018) bereitgestellt. Da wir die zugehörigen Erträge erst ab Februar 1871 definieren können, werden wir grundsätzlich die Daten ab Februar 1871 zur Veranschaulichung verwenden.

Die Menge der stilisierten Fakten ist in der Literatur nicht einheitlich festgelegt. Welche stilisierten Fakten aufgezählt werden und welche nicht, hängt von der Einschätzung der Autoren ab. Somit weichen die Aufzählungen in bekannten Arbeiten zu diesem Thema teilweise stark voneinander ab (vergleiche z. B. Cont, 2001; Lux und Ausloos, 2002; Hommes, 2005, 2006; Lux, 2008; Franke und Westerhoff, 2011). Auffällig ist allerdings, dass in allen Werken vier spezielle stilisierte Fakten vorkommen und zum Teil besonders hervorgehoben werden: keine Autokorrelation in den Erträgen, Volatilitätsclustering, Langzeitgedächtnis (*long memory*) und *Fat Tails* in der Ertragsverteilung. Aus diesem Grund werden wir auf die genannten Charakteristiken besonders eingehen. Die übrigen, hier aufgezählten, stilisierten Fakten orientieren sich vor allem an den Arbeiten von Hommes (2005, 2006). Es wird sinnvoll sein, als erstes den stilisierten Fakt Unvorhersehbare Vermögenspreise und -erträge zu betrachten.

Unvorhersehbare Vermögenspreise und -erträge

Betrachtet man die Preisentwicklung von Finanzprodukten, so gleicht diese auf den ersten Blick einer Zufallsbewegung. Preisänderungen werden durch

neue Informationen beispielsweise über unternehmensinterne Entscheidungen, politische Beschlüsse oder auch das Wetter beeinflusst, wobei die Einflussfaktoren vom Finanzprodukt abhängig sind. Gibt zum Beispiel eine Aktiengesellschaft eine neue strategische Ausrichtung bekannt, so beeinflusst diese Information die Kursentwicklung. Rohstoffpreise sind dagegen unter anderem von Witterungsbedingungen abhängig, da diese den Rohstoffabbau beeinträchtigen können. Offenbar können wir zum heutigen Zeitpunkt nicht vorhersagen, welche preisbeeinflussenden Informationen in der Zukunft bekannt werden. Dies ist die Ursache für unseren ersten, hier betrachteten, stilisierten Fakt: Die Unvorhersehbarkeit von Vermögenspreisen und -erträgen.

Um dies mathematisch zu konkretisieren, verwendet Lux (2008) die sogenannte *Martingaleigenschaft*, welche durch

$$E[P_{t+1}|\mathcal{F}_t] = P_t$$

beschrieben wird. Dabei bezeichnet \mathcal{F}_t die zum Zeitpunkt t zur Verfügung stehenden Informationen über die Einflussfaktoren und P_t den Preis zum Zeitpunkt t . Die Familie $(\mathcal{F}_t)_{t \in \mathbb{N}}$ bildet eine sogenannte Filtration auf dem zugrundeliegenden Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$, d. h., es gilt $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}$, denn im Zeitverlauf wächst die Menge der zur Verfügung stehenden Informationen stets an. Der tatsächliche Preis P_{t+1} in der nächsten Periode $t + 1$ bildet sich aus dem vorherigen Preis P_t und den neuen Informationen, die seit dem Zeitpunkt t bekannt wurden. Nun hat die Entstehung neuer Informationen nach obiger Beschreibungen einen zufälligen Charakter. Die Zufälligkeit überträgt sich auf die Preisänderung, sodass wir zur Verdeutlichung

$$\delta_t := P_{t+1} - P_t = P_{t+1} - E[P_{t+1}|\mathcal{F}_t],$$

schreiben, wobei δ_t eine Zufallsvariable mit $E[\delta_t] = 0$ ist. Die Preisänderung δ_t hängt von den zum Zeitpunkt t unbekannt Informationen ab, die bis zum Zeitpunkt $t + 1$ bekannt werden und verursachen, dass P_{t+1} von P_t abweichen kann. Aus der Unvorhersehbarkeit dieser Informationen folgt demnach die Unvorhersehbarkeit der zukünftigen Vermögenspreise. Da der Vermögensertrag der relativen Preisänderung

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

entspricht, folgt sofort, dass auch Erträge unvorhersehbar sind.

An dieser Stelle können wir uns die Frage stellen, was genau bei der Bekanntgabe von preisbeeinflussenden Informationen passiert. Durch die Veröffentlichung der Information entsteht kurzfristig eine *Arbitragemöglichkeit*, d. h. eine Möglichkeit, risikolos Gewinn zu erzielen. Wirkt sich die Information beispielsweise positiv auf den Wert einer Aktie aus, so ist sie zum Zeitpunkt der Bekanntgabe kurzfristig unterbewertet. Marktteilnehmer versuchen schließlich, dies durch den Kauf von Aktien zur eigenen Gewinnmaximierung auszunutzen. Dadurch steigt die Nachfrage und simultan sinkt das Angebot. Da Marktpreise stets die Nachfrage und das Angebot ins Gleichgewicht bringen, kommt es zur Preisanpassung und im betrachteten Beispiel steigt der Preis solange an, bis die Arbitragemöglichkeit behoben ist. Aufgrund des technischen Fortschritts in den vergangenen Jahrzehnten werden Finanztransaktionen heute sehr schnell abgewickelt und die beschriebenen Preisanpassungen finden innerhalb kürzester Zeit statt. Der geringe zeitliche Abstand zwischen der Informationsbekanntgabe und der entsprechenden Preisanpassung erlaubt es somit kaum, die Arbitragemöglichkeit auszunutzen. Man spricht deshalb im Allgemeinen auch von der *Arbitragefreiheit* an Finanzmärkten.

Keine Autokorrelation in den Erträgen

Eng verbunden mit dem ersten, beschriebenen stilisierten Fakt ist die Beobachtung, dass die Erträge keine Autokorrelation aufweisen. Mit der Definition der Autokorrelationsfunktion ausgedrückt soll also gelten:

$$\varrho(s, t) = \text{Corr}[R_s, R_t] = \frac{\text{Cov}[R_s, R_t]}{\sqrt{\sigma_R^2(s) \cdot \sigma_R^2(t)}} = 0 \quad \forall s, t \in \mathbb{N} \quad (2.1)$$

Die Gleichung besagt, dass kein statistischer Zusammenhang der Erträge zu unterschiedlichen Zeitpunkten besteht. Insbesondere kann man anhand der Erträge vergangener Perioden nicht auf zukünftige Erträge schließen. Eine einfache Extrapolation vergangener Preisentwicklungen ist somit kein sinnvolles Mittel zur Prognose zukünftiger Preisentwicklungen. Hier liegt also die Verknüpfung mit der oben beschriebenen Unvorhersehbarkeit.

Zu beachten ist, dass wir hierbei die *relative* Preisänderung und nicht die Preise selbst betrachten. Denn in der reinen Preisentwicklung $(P_t)_{t \in \mathbb{N}}$ besteht natürlich eine Korrelation zwischen den Werten verschiedener Zeitpunkte. Der Preis P_{t+1} wird mit hoher Wahrscheinlichkeit eine ähnliche Größenordnung haben wie der Preis P_t , da extreme Preisänderungen mit nur sehr geringer Wahrscheinlichkeit eintreten. Die Gleichheit (2.1) gilt also nicht, wenn $(P_t)_{t \in \mathbb{N}}$ als stochastischer Prozess eingesetzt wird.

Große und anhaltende Handelsvolumen

An den Finanzmärkten lässt sich beobachten, dass es stets zum Handel kommt. Täglich wird eine Vielzahl an Transaktionen an den Finanzmärkten abgeschlossen und das Handelsvolumen wird nicht Null.

Später werden wir mathematische Modelle für Finanzmärkte bewerten, indem wir prüfen, inwiefern sie die betrachteten stilisierten Fakten erfüllen. Für die hier beschriebene Charakteristik muss also sichergestellt werden, dass das Modell dauerhaft große Handelsvolumen generiert. Hommes (2006) weist in diesem Zusammenhang auf Probleme von Modellen hin, die ausschließlich rationale, risikoscheue Händler mit asymmetrischen Informationen betrachten. In solchen Modellen stehen unterschiedlichen Händlern am Finanzmarkt verschiedene Informationen zur Verfügung. Außerdem gehen alle Händler keinerlei Risiko ein und fällen ihre Entscheidungen ausschließlich rational, d. h. faktenbasiert. Besitzt in solchen Modellen ein Händler eine preisbeeinflussende Information, die andere Händler nicht besitzen, dann kann er dieses Wissen nicht ausnutzen. Der Grund liegt darin, dass kein Handelspartner gefunden wird, da die anderen Marktteilnehmer aufgrund Ihrer Rationalität wissen, dass der Händler zusätzliche Informationen haben muss und deshalb keinen Handel mit ihm eingehen. In den beschriebenen Modellen kommt es also niemals zu einem Handel, was der Beobachtung an tatsächlichen Märkten widerspricht. Modelle mit ausschließlich rationalen, risikoscheuen Händlern mit asymmetrischen Informationen können also reale Finanzmärkte in diesem Sinne nicht gut abbilden.

Übermäßige Volatilität und lang anhaltende Abweichung vom Fundamentalwert

Eine zentrale Frage, die sich Akteure an Finanzmärkten stellen, ist, ob ein Finanzprodukt aktuell über- oder unterbewertet ist. Im Falle einer Unterbewertung wird ein zukünftiger Preisanstieg erwartet. Der heutige Kauf des Finanzproduktes kann also zu einem profitablen, zukünftigen Verkauf führen. Ist das Finanzprodukt dagegen überbewertet, geht man umgekehrt von einem zukünftigen Preisabfall aus. Akteure können versuchen, diesen durch Leerverkäufe zur eigenen Profitgenerierung auszunutzen. Um eine Aussage über die aktuelle Bewertung eines Finanzproduktes zu treffen, betrachtet man den sogenannten *Fundamentalwert*. Im Gegensatz zum Marktpreis wird der Fundamentalwert nicht auf Grundlage von aktuellem Angebot und Nachfrage ermittelt. Stattdessen bestimmt man den Wert durch einen theoretisch hergeleiteten Ansatz. Dabei sind jedoch unterschiedliche Überlegungen möglich, sodass es keine einheitliche Definition des Fundamentalwertes gibt. Auf

Details über den Fundamentalwert werden wir in Kapitel 5 eingehen. An dieser Stelle ist zunächst nur wichtig, dass der Fundamentalwert ein alternatives Bewertungskonzept zum Preis des Finanzproduktes darstellt.

Der hier betrachtete stilisierte Fakt besagt nun, dass die Preise im Zeitverlauf eine hohe Variabilität aufweisen, d. h., die Volatilität ist sehr hoch. Sind die Schwankungen in der Preisentwicklung größer als in der Fundamentalwertentwicklung, dann folgt sofort, dass die Preise vom Fundamentalwert abweichen können. Die Abweichungen können über lange Zeiträume hinweg anhalten.

Volatilitätsclustering

Das Volatilitätsclustering wird als einer der wichtigsten stilisierten Fakten betrachtet. Es beschreibt den unregelmäßigen Wechsel zwischen ruhigen Phasen mit geringen Schwankungen in der Preisentwicklung und turbulenteren Phasen mit großen Preisschwankungen.

Wir haben bereits den stilisierten Fakt der fehlenden Autokorrelation in den Erträgen betrachtet. Infolgedessen ging man lange Zeit davon aus, dass die Preisentwicklung einem *Random-Walk*-Prozess gleicht.

Definition 2.0.1. (*Schlittgen und Streitberg, 1999, S. 92*) Als White-Noise-Prozess oder reinen Zufallsprozess bezeichnet man eine Folge $(\epsilon_t)_{t \in T}$ von identisch verteilten und unabhängigen Zufallsvariablen ϵ_t .

Es sei $(\epsilon_t)_{t \in \mathbb{N}}$ ein solcher White-Noise-Prozess über der Menge der natürlichen Zahlen. Ein Random-Walk-Prozess $(X_t)_{t \in \mathbb{N}}$ ist definiert durch die Rekursion

$$X_t = \begin{cases} \epsilon_1 & \text{für } t = 1, \\ X_{t-1} + \epsilon_t & \text{für } t = 2, 3, \dots \end{cases}$$

Auf dem ersten Blick sehen Finanzmarktzeitreihen aus, als wären die absoluten Preisänderungen stochastisch unabhängig und identisch verteilt. Tatsächlich ist es jedoch wahrscheinlicher, dass es während einer turbulenteren Phase auch weiterhin große Preisschwankungen gibt - und umgekehrt. Die absoluten Preisinkremente, und damit auch die Absoluterträge, sind also nicht stochastisch unabhängig.

Durch die erstmalige Beobachtung von Volatilitätsclustering durch Mandelbrot (1963) musste somit die Ansichtswiese des Random-Walk-Prozesses verworfen werden. Engle (1982) entwickelte deshalb ein Modell, bei dem die

Erträge aus einer Normalverteilung stammen, deren Varianz einem autoregressiven Prozess folgt. Dieses Modell ist unter dem Namen *generalized autoregressive conditional heteroscedasticity (GARCH)* bekannt. Das GARCH-Modell hat jedoch Defizite in der Nachbildung der wichtigen stilisierten Fakten Langzeitgedächtnis und Fat Tails in der Ertragsverteilung, zu denen wir noch kommen werden. Deshalb werden wir uns in Kapitel 7 auf eine andere Modellklasse konzentrieren.

Langzeitgedächtnis

Die Langzeitgedächtnis- oder auch long memory-Eigenschaft besagt, dass eine langfristige Abhängigkeit in den absoluten sowie quadrierten Erträgen existiert. Erstmals beobachtet wurde das Phänomen durch Hurst (1951) bei der Analyse von Wasserständen des Nils. Anschließend stellte man fest, dass auch Zeitreihen von Finanzprodukten diese Eigenschaft besitzen.

Zur Beobachtung der Langzeitgedächtniseigenschaft müssen die Absoluterträge $(|R_t|)_{t \in \mathbb{N}}$ bzw. die quadrierten Erträge $(R_t^2)_{t \in \mathbb{N}}$ als stochastische Prozesse interpretiert werden. In beiden Fällen wird das Vorzeichen und damit die Richtung der Preisänderung ignoriert, sodass es sich um geeignete Größen zur Darstellung der Volatilität handelt. Sind die genannten Prozesse schwach stationär, dann besagt die Langzeitgedächtniseigenschaft, dass die jeweilige Autokorrelationsfunktion $\rho(\tau)$ im Lag τ fällt, der Abfall jedoch nur langsam stattfindet. Demnach sind auch für große zeitliche Differenzen $\tau \in \mathbb{N}$ die Werte $|R_t|$ und $|R_{t+\tau}|$ bzw. R_t^2 und $R_{t+\tau}^2$ für alle $t \in \mathbb{N}$ korreliert, die Stärke der Korrelation nimmt aber mit wachsendem Wert für τ langsam ab.

Temporäre Blasen und Trendfolgestrategie sowie plötzliche Abstürze und *Mean Reversion*

Auf Finanzmärkten lässt sich immer wieder das Phänomen beobachten, dass Kurse über lange Zeit ansteigen und anschließend plötzlich abstürzen. Eine wesentliche Rolle spielen hierbei Anhänger der sogenannten Trendfolgehändlerstrategien. Wie der Name schon sagt, orientieren sie sich am aktuellen Trend im Preisverlauf. Befindet sich der Preis in einem Aufwärtstrend, dann entscheiden sich die Anhänger zum Kauf. In einer Phase, in der der Preis dagegen fällt, wird das Finanzprodukt verkauft. Es lässt sich an Finanzmärkten beobachten, dass Anhänger von Trendfolgehändlerstrategien im Mittel positive Gewinne erzielen. Das Phänomen der temporären Blasen, plötzlichen Abstürze und Mean Reversion lässt sich nun an dem folgenden, typischen Szenario beschreiben.

Befindet sich der Preisverlauf in einem Aufwärtstrend, so wollen Anhänger von Trendfolgehandelstrategien diesen Trend ausnutzen und entscheiden sich deshalb für den Kauf des Finanzproduktes. Dadurch kommt es zur erhöhten Nachfrage und die Preise steigen weiter an. Es entsteht eine temporäre Blase, bei der das Finanzprodukt stark überbewertet ist, d. h., der Preis liegt währenddessen weit über dem Fundamentalwert. Der Begriff Mean Reversion beschreibt die Überzeugung von Akteuren an Finanzmärkten, dass Preise immer wieder zu ihrem langfristigen Mittelwert zurückkehren. Je stärker die Überbewertung des Finanzproduktes also ausfällt, desto stärker ist auch die Angst vor einem zukünftigen Preisabfall. Dadurch kommt es zu einem unbestimmten Zeitpunkt zur Trendwende und die Marktteilnehmer entscheiden sich zunehmend zum Verkauf des Finanzproduktes. Sobald die Trendwende vom Kursanstieg zum -abfall eintritt, reagieren Händler oft sehr schnell, um von der Überbewertung noch profitieren zu können. Das Angebot steigt also innerhalb kürzester Zeit stark an, sodass der Kursabfall sehr viel schneller stattfindet als der vorherige Kursanstieg.

In Abbildung 2.1 ist der Kursverlauf des Aktienindex S&P 500 zwischen April 2010 und April 2018 basierend auf dem Datensatz von Shiller (2018) dargestellt. Temporäre Blasen und plötzliche Abstürze sind beispielsweise in der Jahresmitte 2011 sowie zum Jahresende 2015 erkennbar.

Fat Tails in der Ertragsverteilung

Betrachtet man die Dichtefunktion der Ertragsverteilung, dann lässt sich feststellen, dass diese typischerweise glockenförmig und symmetrisch ist. Der Vergleich zur Normalverteilung ist somit naheliegend. Fama (1963) und Mandelbrot (1963) haben jedoch festgestellt, dass die Verteilung der Erträge nicht normal ist, da eine höhere Wahrscheinlichkeit im Zentrum sowie in den Tails, d. h., in den Randbereichen der Verteilung liegt. Man spricht deshalb von Fat Tails der Ertragsverteilung.

In Abbildung 2.2 stellt die durchgezogene Linie die empirische Dichtefunktion der Erträge des Aktienindex S&P 500 im Zeitraum Februar 1871 bis April 2018 dar. Zum Vergleich ist gestrichelt die Dichtefunktion der Normalverteilung mit den empirischen Verteilungsparametern

$$\bar{r} := \frac{1}{N} \sum_{t=1}^N r_t, \quad s := \sqrt{\frac{1}{N-1} \sum_{t=1}^N (r_t - \bar{r})^2}$$

eingezeichnet. Die Abweichung zur Normalverteilung ist deutlich erkennbar.

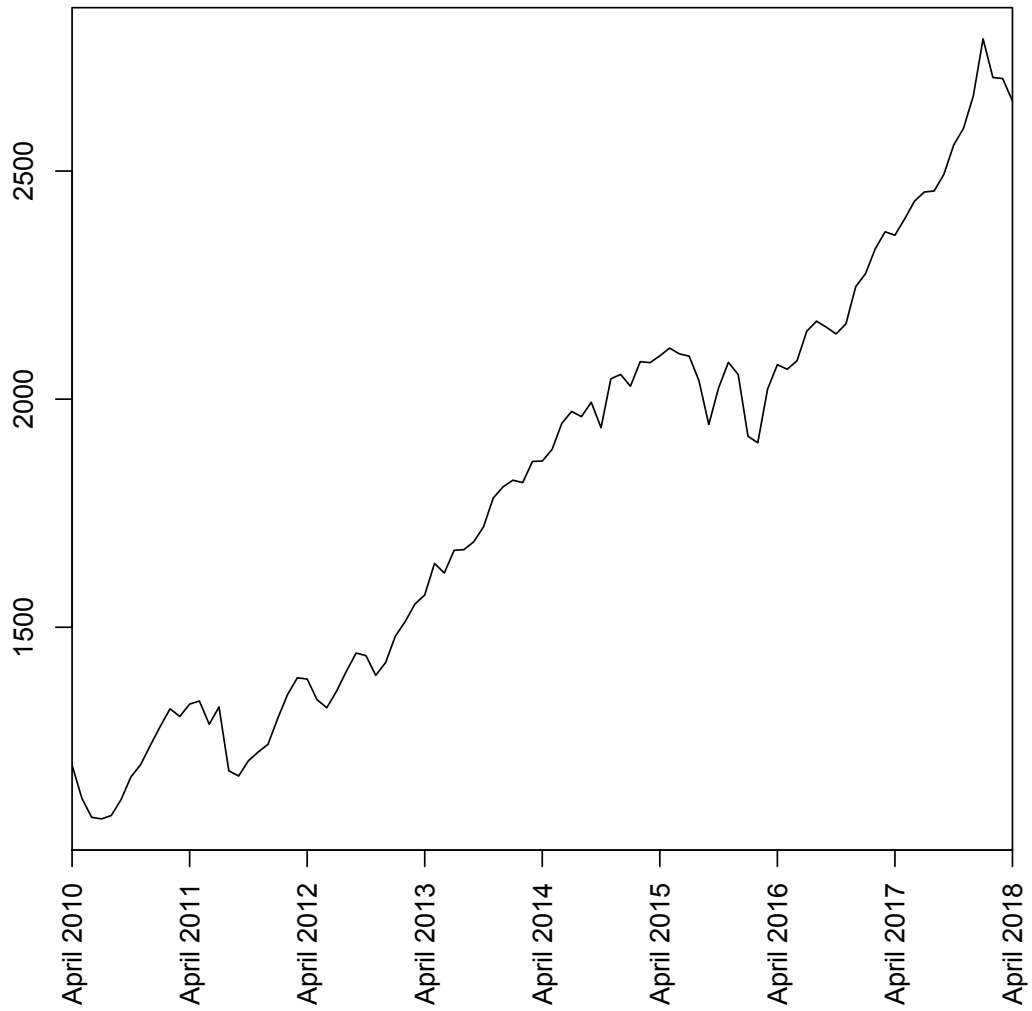


Abbildung 2.1: S&P 500-Kursverlauf im Zeitraum April 2010 bis April 2018 mit erkennbaren, plötzlichen Abstürzen in der Jahresmitte 2011 und zum Jahresende 2015

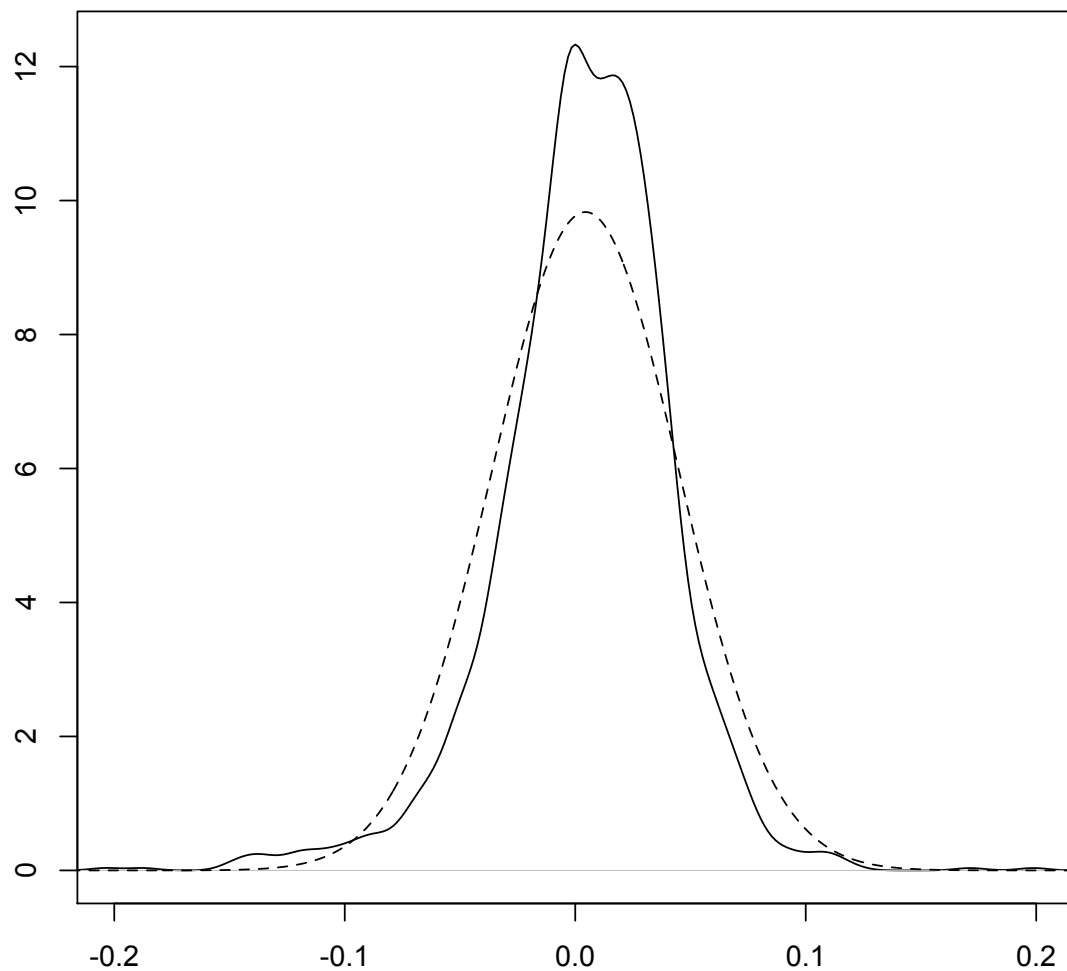


Abbildung 2.2: Empirische Dichtefunktion der S&P 500 Aktienindexerträge im Zeitraum Februar 1871 bis April 2018 (durchgezogene Linie) und Dichte der Normalverteilung mit dem empirischen Mittelwert und der empirischen Varianz als Verteilungsparameter (gestrichelte Linie)

Kapitel 3

Messung stilisierter Fakten

Nachdem wir in Kapitel 2 die beobachtbaren, stilisierten Fakten aufgezählt haben, stellt sich die Frage, wie wir diese messen können. Schließlich können wir nur mit geeigneten Maßzahlen später Simulationsergebnisse spezieller Modelle bewerten. Ein Modell werden wir dann als gut empfinden, wenn die gemessenen Werte der Simulationsergebnisse eine ähnliche Größenordnung haben wie die der vorher ausgewählten real-world Stichprobe. Ist dies der Fall, so können aus der Perspektive der Maßzahlen die Simulationsergebnisse nicht unterschieden werden von *real-world data generation process (DGP)*-generierten Daten. Somit scheint das Modell den DGP gut nachzuahmen. Die Notwendigkeit geeigneter Kennzahlen liegt demnach auf der Hand.

Die unterschiedlichen Kennzahlen werden an verschiedenen Zeitreihen ausgewertet, die jeweils aus realen Beobachtungen oder aus Simulationen stammen können. Zum besseren Überblick gibt die folgende Aufzählung alle für die Kennzahlen relevanten Zeitreihen sowie ihre Abkürzungen an. Auf die letzten zwei Zeitreihen kann erst in späteren Abschnitten genauer eingegangen werden:

- Zeitreihe der Preise $(p_t)_{t=1,\dots,N}$
- Zeitreihe der Erträge $(r_t)_{t=1,\dots,N}$ mit $r_t := \frac{p_t - p_{t-1}}{p_{t-1}}$
- Zeitreihe der Absoluterträge $(|r_t|)_{t=1,\dots,N}$
- Zeitreihe der quadrierten Erträge $(r_t^2)_{t=1,\dots,N}$
- Zeitreihe der Handelsvolumen $(h_t)_{t=1,\dots,N}$
- Zeitreihe der inflations- und trendbereinigten Preise $(p_t^{\text{real,detrended}})_{t=1,\dots,N}$, welche in den Abschnitten 4.1 und 5.2.1 genauer behandelt werden

- Zeitreihe der geschätzten Fundamentalwerte $(\hat{p}_t^*)_{t=1,\dots,N}$, die sich auf die inflations- und trendbereinigten Preise beziehen und deren Berechnung in Kapitel 5 erläutert wird

Zur Auswertung und Veranschaulichung der beschriebenen Kennzahlen werden wir auch in diesem Kapitel auf den S&P 500-Datensatz für den Zeitraum Februar 1871 bis April 2018 von Shiller (2018) zurückgreifen. Außerdem wird die Softwareumgebung R bei der Auswertung statistischer Funktionen sowie der Erstellung geeigneter Grafiken verwendet.

Wir werden in den folgenden Abschnitten gezielt für die vier wichtigsten, stilisierten Fakten im Detail geeignete Kennzahlen herleiten. Dazu zählen: Keine Autokorrelation in den Erträgen, Volatilitätsclustering, Langzeitgedächtnis und Fat Tails in der Ertragsverteilung. Der Vollständigkeit halber werden nun zumindest oberflächlich auch für die übrigen, in Kapitel 2 aufgezählten, stilisierten Fakten Maßzahlen angeben. Dabei werden wir jeweils kurz auf eine Idee zur Messung eingehen, ohne Einzelheiten, wie zum Beispiel die genaue Programmierung, zu behandeln.

Unvorhersehbare Vermögenspreise und -erträge

Wir haben in Kapitel 2 bereits erwähnt, dass die Unvorhersehbarkeit sehr eng mit der fehlenden Autokorrelation in den Erträgen verknüpft ist. Da wir für diese in Abschnitt 3.1 eine geeignete Messung angeben werden, können wir auf eine Maßzahl für die Unvorhersehbarkeit verzichten.

Große und anhaltende Handelsvolumen

Es sei $(H_t)_{t \in \mathbb{N}}$ der stochastische Prozess des Handelsvolumens. Wir nehmen an, dass die Zufallsvariablen paarweise stochastisch unabhängig und identisch verteilt mit Erwartungswert μ_H und Varianz σ_H^2 sind. Gilt der stilisierte Fakt, so sollte μ_H signifikant von Null verschieden sein. Um zu beurteilen, ob eine Stichprobe $(h_t)_{t=1,\dots,N}$ den stilisierten Fakt erfüllt, können wir einen zweiseitigen Einstichproben-t-Test für die Nullhypothese $H_0: \mu_H = 0$ durchführen. Dazu ist es wichtig, dass die Stichprobengröße N hinreichend groß ist, sodass mithilfe des zentralen Grenzwertsatzes näherungsweise die Normalverteilung des empirischen Mittelwertes \bar{h} folgt, welche für den Test notwendig ist. Die Teststatistik lautet

$$T(h) := \sqrt{N} \cdot \frac{\bar{h} - 0}{s_h^2},$$

wobei s_h^2 die empirische Standardabweichung der Stichprobe bezeichnet. Testen wir bezüglich des Signifikanzniveaus α , dann lehnen wir die Nullhypothese ab, falls

$$|T(h)| \geq t_{1-\frac{\alpha}{2}, N-1}$$

gilt. Dabei steht $t_{1-\frac{\alpha}{2}, N-1}$ für das $(1 - \frac{\alpha}{2})$ -Quantil der t-Verteilung mit $N - 1$ Freiheitsgraden (auch Studentsche t-Verteilung; vergleiche Christmann, 2016, S. 52 ff.). Eine geeignete Maßzahl für den stilisierten Fakt ist somit eine boolesche Variable, die den Ausgang des Tests, beispielsweise bezüglich des Niveaus $\alpha = 5\%$, repräsentiert.

Übermäßige Volatilität und lang anhaltende Abweichung vom Fundamentalwert

Zur Messung können wir die Summe der absoluten Abweichungen der Preise vom geschätzten Fundamentalwert betrachten. Da sich der geschätzte Fundamentalwert $(\hat{p}_t^*)_{t=1, \dots, N}$ auf die inflations- und trendbereinigten Preise $(p_t^{\text{real, detrended}})_{t=1, \dots, N}$ bezieht, müssen wir diese auch in unserer Messung verwenden. Die Kennzahl lautet somit

$$\sum_{t=1}^N \left| p_t^{\text{real, detrended}} - \hat{p}_t^* \right|.$$

Temporäre Blasen

Der stilisierte Fakt besagt, dass wir immer wieder Zeiträume beobachten können, in denen der Preis weit über dem Fundamentalwert liegt. Zur Quantifizierung können wir beispielsweise zählen, wie viele zusammenhängende Zeiträume es gibt, in denen $p_t^{\text{real, detrended}}$ durchgehend größer als $1,25 \cdot \hat{p}_t^*$ ist, d.h., die inflations- und trendbereinigten Preise übersteigen innerhalb dieser Zeiträume den geschätzten Fundamentalwert ununterbrochen um mehr als 25%. An dieser Stelle ist natürlich jeder beliebige Prozentsatz wählbar. Zusätzlich kann es sinnvoll sein, eine Mindestlänge der Zeiträume zu fordern, um kurzfristige Abweichungen nicht als temporäre Blasen zu interpretieren. Zum Beispiel können wir drei Monate als minimale Phasenlänge festlegen.

Trendfolgestrategie

Hierfür können wir messen, welchen Gewinn bzw. Verlust ein Anhänger einer konkreten Trendfolgestrategie bei der betrachteten Stichprobe des Preises $(p_t)_{t=1, \dots, N}$ erzielt hätte. Dazu stellen wir uns vor, der Händler hat zum Zeitpunkt $t = 1$ das Finanzprodukt nicht im Besitz und sein Kontostand liegt bei 0€ . Nun durchlaufen wir die Zeitpunkte $t = 2, \dots, N - 1$ und prüfen jeweils, ob der Preis in der vergangenen Periode gestiegen ist oder nicht. Bei einer Preissteigerung $p_t > p_{t-1}$ und falls der Händler zum Zeitpunkt t das Finanzprodukt nicht besitzt, kauft er das Produkt und der Kontostand verringert sich um den Betrag $p_t\text{€}$. Fällt dagegen der Preis $p_t < p_{t-1}$ und besitzt der Händler das Finanzprodukt, so wird es zum Preis $p_t\text{€}$ verkauft und der Kontostand erhöht sich um diesen Betrag. Zum Zeitpunkt N wird

das Finanzprodukt gegebenenfalls verkauft und dem Konto $p_N \text{€}$ gutgeschrieben. Den anschließend bestehenden Kontostand verwenden wir als Maßzahl für die Trendfolgestrategie. Ist er positiv, so wurde mit der Strategie ein Gewinn erzielt, andernfalls ein Verlust. Natürlich kann an dieser Stelle ebenso das Ergebnis einer anderen Trendfolgestrategie betrachtet werden. Die Strategien unterscheiden sich im wesentlichen im Zeithorizont, der bei der Kaufentscheidung berücksichtigt wird (vergleiche z. B. Ivanova et al., 2014, S. 4 ff.).

Plötzliche Abstürze

Während eines plötzlichen Absturzes fällt der Preis in kurzer Zeit relativ stark. Als Kennzahl können wir Zeiträume zählen, in denen der Preis durchgehend je Periode um beispielsweise mindestens 1% fällt. Der Prozentsatz sollte dabei an die gewählte Periodenlänge angepasst werden. Entspricht eine Periode zum Beispiel einem Monat, dann wird es sinnvoll sein, einen geringeren Prozentsatz zu wählen, als bei jährlichen Perioden, da eine starke Preisänderung innerhalb eines Monats unwahrscheinlicher ist als innerhalb eines ganzen Jahres. Außerdem kann es sinnvoll sein, zu fordern, dass die so definierten Zeiträume einen angemessenen Mindestabstand voneinander haben. Andernfalls würden wir einen plötzlichen Absturz, der durch eine kurzfristige Abflachung unterbrochen wird, als zwei Abstürze zählen. Darüber hinaus können wir, ähnlich wie bei der Messung temporärer Blasen, eine Mindestlänge der Zeiträume von beispielsweise zwei Monaten fordern, um Kursverluste, die sich auf einen Monat beschränken, nicht als plötzlichen Absturz zu deuten.

Mean Reversion

Hier benötigen wir eine Maßzahl, die angibt, wie oft sich der Preis *in der Nähe* des Fundamentalwertes befindet. Dazu müssen wir natürlich definieren, wie groß der Abstand dabei höchstens sein darf. Beispielsweise können wir je Periode t betrachten, ob der inflations- und trendbereinigte Preis $p_t^{\text{real,detrended}}$ innerhalb des Intervalls $[0,95 \cdot \hat{p}_t^*; 1,05 \cdot \hat{p}_t^*]$ liegt. Als Kennzahl können wir die Anzahl der Perioden definieren, in denen dies erfüllt ist. Alternativ können wir ebenso zählen, wie oft der inflations- und trendbereinigte Preis in dieses Intervall zurückkehrt, nachdem er im vorherigen Zeitraum mit einer bestimmten Mindestlänge durchgehend ober- bzw. unterhalb der entsprechenden Intervallgrenze lag.

Um in den folgenden Abschnitten Maßzahlen für die vier wichtigsten, stilisierten Fakten herzuleiten, werden wir empirische Momente der reinen, absoluten und quadrierten Erträge benötigen. Wie bereits in Kapitel 1.2 angedeutet, müssen wir

annehmen, dass es sich bei den stochastischen Prozessen der Erträge $(R_t)_{t \in \mathbb{N}}$, der Absoluterträge $(|R_t|)_{t \in \mathbb{N}}$ und der quadrierten Erträge $(R_t^2)_{t \in \mathbb{N}}$ um schwach stationäre Prozesse handelt, sodass die Momente sinnvoll definiert sind. Insbesondere werden wir annehmen, dass die Erträge im Zeitverlauf keinerlei Trend aufweisen, da der stochastische Prozess ansonsten nicht mittelwertstationär wäre. Die Annahme der schwachen Stationarität lässt sich anhand in der Realität beobachteter Zeitreihen durchaus rechtfertigen. Besteht dennoch Zweifel an der Korrektheit der Annahme, so müssen die betrachteten Zeitreihen zunächst bereinigt werden, indem Trends und zyklische Schwankungen, wie beispielsweise Saisoneffekte, mit geeigneten Transformationen entfernt werden (vergleiche Schlittgen und Streitberg, 1999, Kapitel 1).

3.1 Autokorrelationsfunktion der Erträge

Um die fehlende Autokorrelation in den Erträgen zu messen, ist es naheliegend, die in Kapitel 1.2 definierte Autokorrelationsfunktion für die Erträge der betrachteten Zeitreihe zu bilden. In Abbildung 3.1 sind die berechneten Werte für die S&P 500-Zeitreihe mit Lags bis zu 120 Monaten, d. h. 10 Jahren, abgebildet. Außerdem ist dort das 95%-Konfidenzintervall eingezeichnet, für den Fall, dass die Werte aus einem normalverteilten White-Noise-Prozess stammen. Es lässt sich zeigen, dass die empirischen Autokorrelationen $\varrho_e(\tau)$ dann asymptotisch normalverteilte, unabhängige Zufallsvariablen mit Erwartungswert $(-1/N)$ und Varianz $(1/N)$ sind. Um das 95%-Konfidenzintervall aufzustellen, werden die 2,5%- und 97,5%-Quantile der entsprechenden Normalverteilung $\mathcal{N}(-1/N, 1/N)$ berechnet (vergleiche Schlittgen und Streitberg, 1999, S. 243 ff.). Die R-Funktion `acf` aus dem Paket `stats` berechnet dies über das 97,5%-Quantil $z_{0,975}$ der Standardnormalverteilung $\mathcal{N}(0, 1)$ und einer anschließenden Transformation mit dem angegebenen Erwartungswert und der Varianz. Für große Stichprobengrößen gilt $-1/N \approx 0$ und deshalb lauten die Intervallgrenzen näherungsweise

$$0 \pm \frac{z_{0,975}}{\sqrt{N}}.$$

Damit können nun die berechneten Autokorrelationsfunktionswerte mit dem Intervall verglichen werden. Liegen sie außerhalb es Intervalls, dann sind sie bezüglich des 5%-Signifikanzniveaus signifikant von Null verschieden. Jedoch ist bei der Interpretation darauf zu achten, dass selbst normalverteilte White-Noise-Werte das Intervall mit der Wahrscheinlichkeit 5% verlassen.

Für $\tau = 0$ ergibt sich per Definition stets der Wert $\varrho_e(0) = \gamma_e(0)/\gamma_e(0) = 1$. Dieser Funktionswert hat somit keine Aussagekraft. Für $\tau = 1$ ist die Autokorrelationsfunktion im Beispiel signifikant von Null verschieden. Für Lags $\tau \geq 2$ liegen die

Werte jedoch größtenteils innerhalb des 95%-Konfidenzintervalls und sind somit selten signifikant von Null verschieden. Die Abbildung unterstützt somit die These *keine Autokorrelation in den Erträgen*.

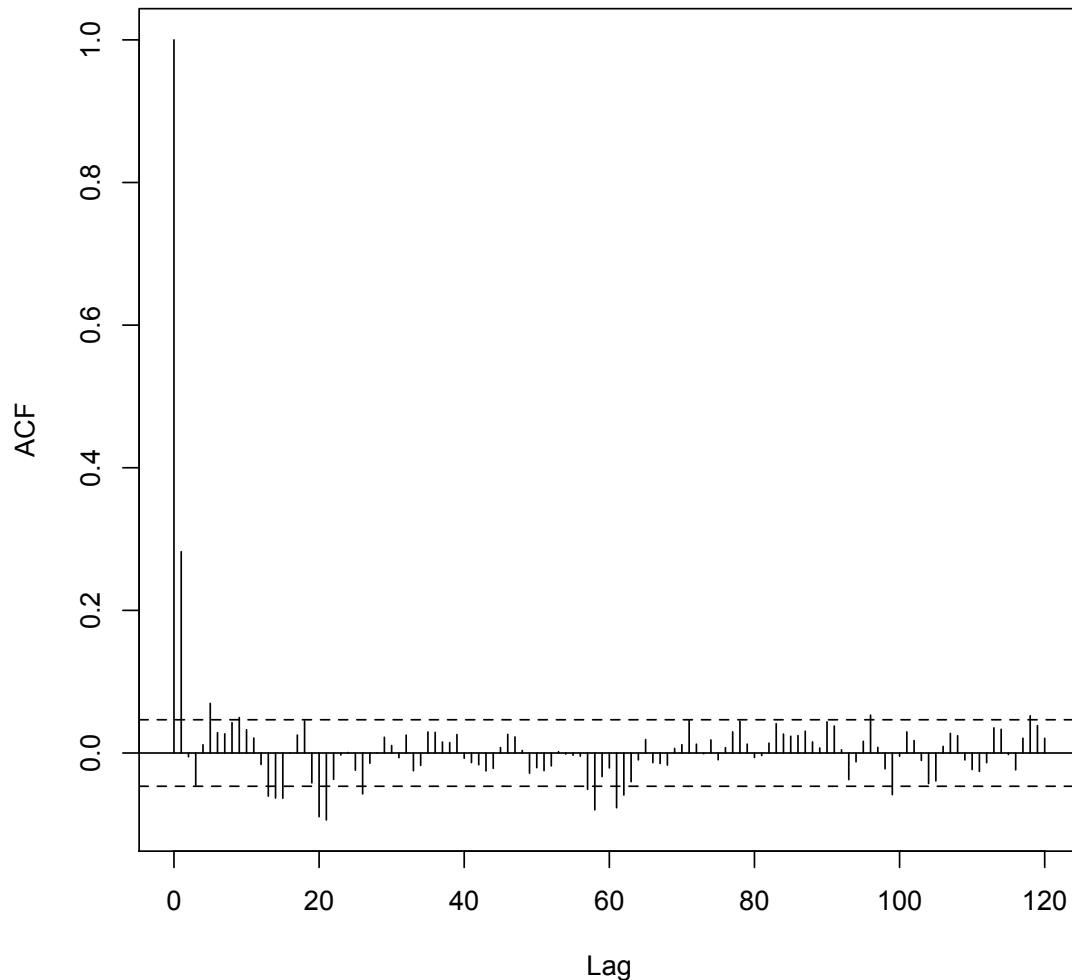


Abbildung 3.1: Autokorrelationsfunktion der monatlichen S&P 500-Erträge von Februar 1871 bis April 2018 mit Lags bis zu 120 Monaten sowie dem 95%-Konfidenzintervall für den normalverteilten White-Noise-Fall

Wie zu Beginn des Kapitels beschrieben, benötigen wir jedoch eine konkrete Maßzahl und keine Funktion, mit der wir den stilisierten Fakt messen. An dieser Stelle ist es naheliegend, den Wert der Autokorrelationsfunktion bezüglich eines bestimmten Lags zu wählen. Franke und Westerhoff (2011) schlagen hierfür $\rho_e(1)$, d. h. den Funktionswert für $\tau = 1$ vor. Man sagt dazu auch *Autokorrelationskoeffizient erster Ordnung*. Obwohl dies gerade das Lag ist, bei dem wir eine signifikante

Abweichung von Null in Abbildung 3.1 feststellen, werden wir den Vorschlag von Franke und Westerhoff (2011) übernehmen. Alternativ sind ebenso andere Lags denkbar. In unserer S&P 500-Zeitreihe beträgt die Maßzahl $\varrho_e(1) = 0,2823089$. Ein Modell soll somit bezüglich des ersten stilisierten Fakt es als gut empfunden werden, wenn die Simulationsstichprobe einen Autokorrelationskoeffizienten erster Ordnung derselben Größenordnung ergibt. Allerdings könnte man bei der Modellbewertung zusätzlich auch die gesamte Autokorrelationsfunktion der Simulationsstichprobe betrachten und überprüfen, ob diese ein ähnliches Bild wie in Abbildung 3.1 ergibt. Mit zunehmendem Lag sollten die Autokorrelationsfunktionswerte also nicht signifikant von Null verschieden sein.

3.2 Autokorrelationsfunktion der absoluten oder quadrierten Erträge

Will man die Volatilität in der Preisentwicklung betrachten, so dienen im Allgemeinen die Absoluterträge

$$|R_t| = \left| \frac{P_t - P_{t-1}}{P_{t-1}} \right| = \frac{|P_t - P_{t-1}|}{P_{t-1}}$$

sowie die quadrierten Erträge

$$R_t^2 = \frac{(P_t - P_{t-1})^2}{P_{t-1}^2}$$

als sinnvolle Größen, da sie die Preisänderung messen, ohne deren Vorzeichen zu berücksichtigen. Sind nun beispielsweise $|R_{t-1}|$ und $|R_t|$ positiv korreliert, dann besteht ein Zusammenhang zwischen den Preisänderungen von aufeinanderfolgenden Perioden. Eine turbulente Phase mit starken Preisschwankungen setzt sich dann tendenziell fort. Gleiches gilt für ruhigere Phasen. Dies entspricht gerade der Definition von Volatilitätsclustering. Der stilisierte Fakt sollte sich somit durch signifikant von Null verschiedene, positive Autokorrelationsfunktionswerte der Absoluterträge bzw. der quadrierten Erträge bemerkbar machen. Aber auch das beschriebene Langzeitgedächtnis sollten wir bei der Betrachtung der Autokorrelationsfunktion der absoluten bzw. quadrierten Erträge erkennen. Der stilisierte Fakt besagt, dass die Funktion mit wachsendem Lag τ nur langsam fällt.

In Abbildung 3.2 lassen sich die erwarteten Effekte gut erkennen. Dort sind die Autokorrelationsfunktionswerte für Lags bis zu 120 Monaten der absoluten und quadrierten Erträge der S&P 500-Daten für den Zeitraum Februar 1871 bis April 2018 abgebildet. Besonders bei der Autokorrelationsfunktion der Absoluterträge

fällt auf, dass die Funktionswerte auch für große Lags noch signifikant von Null verschieden und positiv sind und mit wachsenden Werten für τ nur langsam abfallen. Diese Beobachtung ist bei den Absoluterträgen deutlich besser erkennbar als bei den quadrierten Erträgen, sodass wir uns im Folgenden auf diese Bezugsgröße konzentrieren werden.

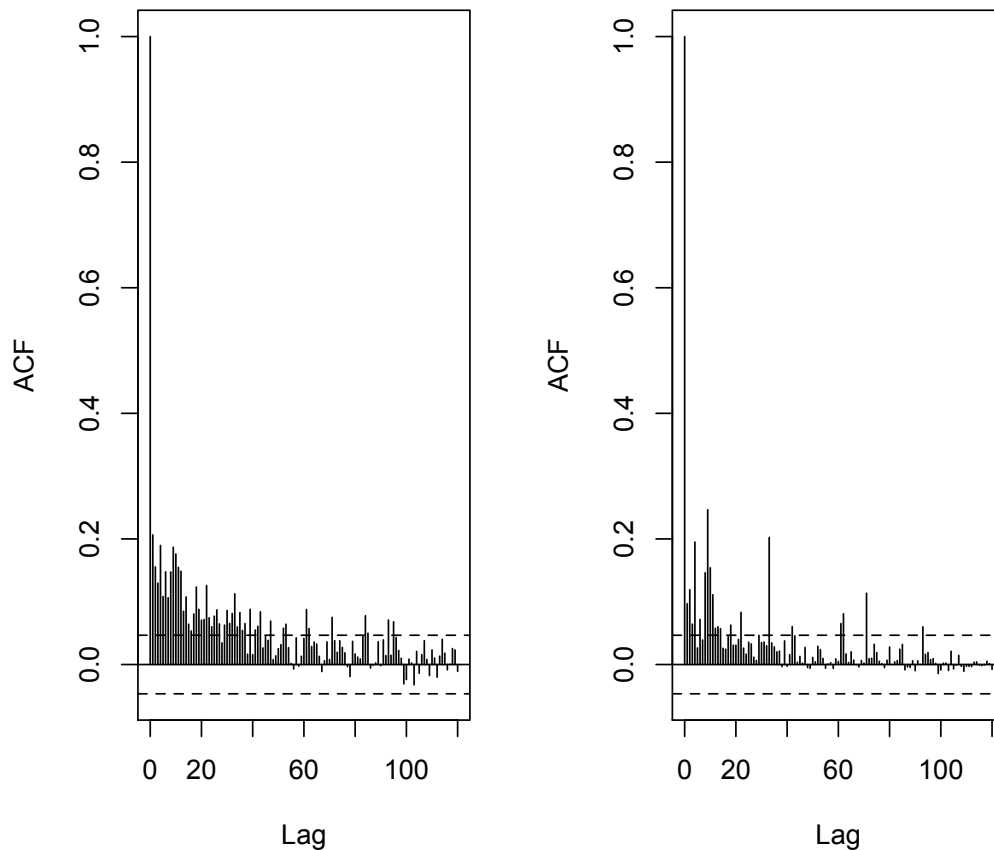


Abbildung 3.2: Autokorrelationsfunktion der monatlichen S&P 500-Absoluterträge (links) bzw. quadrierten Erträge (rechts) von Februar 1871 bis April 2018 mit Lags bis zu 120 Monaten sowie dem 95%-Konfidenzintervall für den normalverteilten White-Noise-Fall

Wie bereits erwähnt, benötigen wir konkrete Maßzahlen, mit denen wir die stilisierten Fakten messen. Für das Volatilitätsclustering und das Langzeitgedächtnis bietet es sich an, die Autokorrelationsfunktionswerte für mehrere, feste Funktionsargumente τ zu verwenden. Die Wahl der festen Argumente hängt natürlich von der gewählten Periodenlänge ab. So macht es einen großen Unterschied, ob die Stichprobenwerte beispielsweise im täglichen oder monatlichen Abstand gemessen

wurden. Da in unserem Beispiel monatliche Kurse bzw. Erträge des Index S&P 500 betrachtet werden, sollten kleinere Werte für τ gewählt werden, als bei einem täglichen Datensatz wie beispielsweise bei Franke und Westerhoff (2011). Wir werden hier die Funktionsargumente $\tau = 1, 3, 6, 12$ und 24 , d. h. Zeitabstände bis zu zwei Jahren, wählen. Die Wahl kann je nach Bedarf angepasst werden. In Tabelle 3.1 sind die genannten Kennzahlen der hier betrachteten Zeitreihe angegeben. Ein Modell bildet das Volatilitätsclustering und das Langzeitgedächtnis demnach gut nach, wenn die Autokorrelationsfunktion der Absoluterträge an den genannten Lags für Simulationsstichproben ähnliche Werte liefert.

Tabelle 3.1: Autokorrelationsfunktionswerte der S&P 500-Absoluterträge von Februar 1871 bis April 2018

τ	$\varrho_e(\tau)$
1	0,20649391
3	0,12958505
6	0,14763076
12	0,14847419
24	0,06014666

3.3 Hurst Exponent

Eine weitere Möglichkeit, die langfristige Abhängigkeit der Erträge zu untersuchen, bildet der Hurst Exponent. Dieser wurde erstmals durch Hurst (1951) während der Zeitreihenanalyse von Wasserständen des Nils eingeführt und ist heute ein beliebtes Mittel zur Messung der long memory-Eigenschaft in unterschiedlichen Anwendungsgebieten. Während seiner Analysen stellte Hurst (1951) die Behauptung auf, dass die sogenannten *adjusted rescaled range*-Werte $(R/S)_n$ den Zusammenhang

$$\left(\frac{R}{S}\right)_n \sim cn^H \quad (3.1)$$

genügen. Unter der *adjusted rescaled range* versteht er die Spannweite der kumulierten Abweichungen vom Mittelwert normiert mit der Standardabweichung der betrachteten Zeitreihe. Im Wesentlichen handelt es sich also um ein bestimmtes, statistisches Maß für die Variabilität. Der Zusammenhang (3.1) besagt nun, dass die *adjusted rescaled range* polynomial in der betrachteten Zeitreihenlänge n wächst und das genaue Wachstum durch den sogenannten Hurst Exponenten H bestimmt wird. Hurst (1951) stellte fest, dass Zeitreihen eines White-Noise-Prozesses den theoretischen Hurst Exponent $H = 0,5$ ergeben. Liegt dagegen eine langfristige, positive Autokorrelation vor, so gilt $H > 0,5$ und bei einer langfristigen,

negativen Autokorrelation erhält man $H < 0,5$. In unserem Kontext interessieren wir uns also für den theoretischen Hurst Exponent der absoluten oder quadrierten Erträge. Wir erwarten einen Ergebniswert $H > 0,5$.

Nun stellt sich natürlich die Frage, wie man den Hurst Exponenten schätzen kann. Wir werden uns hier auf die bekannteste Schätzmethode *R/S Analyse* sowie eine korrigierte Version davon konzentrieren. Die Beschreibung der Methoden orientiert sich an den Arbeiten von Weron (2002) und Hamed (2007). An dieser Stelle ist darauf hinzuweisen, dass in der Literatur Differenzen bei der Wahl der empirischen oder verzerrten Standardabweichung festzustellen sind. Aus diesem Grund werden wir im Folgenden stets betonen, welche Variante gemeint ist.

3.3.1 R/S Analyse

Basierend auf der Theorie von Hurst (1951) entwickelten Mandelbrot und Wallis (1968) die R/S-Analyse zur Schätzung des Parameters H . Die Idee liegt dabei in einer möglichst vielfältigen Teilfolgenzerlegung der Gesamtzeitreihe. Für jede Teilfolge berechnen Mandelbrot und Wallis (1968) den adjusted rescaled range-Wert. Somit erhalten sie für verschiedene Folgenlängen n jeweils Werte $(R/S)_n$. Die berechneten Werte setzen sie schließlich in (3.1) ein und schätzen den Hurst Exponent H so, dass der Zusammenhang möglichst gut erfüllt ist. Genauer lässt sich das Vorgehen wie folgt beschreiben.

Gegeben sei eine Zeitreihe x_1, \dots, x_N der Gesamtlänge N . Diese wird gleichmäßig in d nicht überlappende Teilfolgen der Länge n aufgeteilt. Insbesondere muss dabei n ein Teiler von N sein. Für jede der Teilfolgen $(x_{i,m})_{1 \leq i \leq n}$, $m = 1, \dots, d$ sind nun schrittweise die folgenden Werte zu berechnen:

1. Berechnung des empirischen Mittelwertes \bar{x}_m sowie der verzerrten Standardabweichung s_m
2. Berechnung der Abweichungen vom Mittelwert

$$y_{i,m} := x_{i,m} - \bar{x}_m, \quad i = 1, \dots, n$$

3. Berechnung der kumulierten Abweichungen vom Mittelwert

$$z_{i,m} := \sum_{j=1}^i y_{j,m}, \quad i = 1, \dots, n$$

4. Ermittlung der Spannweite der kumulierten Abweichungen vom Mittelwert

$$range_m = \max_{1 \leq i \leq n} z_{i,m} - \min_{1 \leq i \leq n} z_{i,m}$$

5. Skalierung der Spannweite mit der verzerrten Standardabweichung

$$\frac{\text{range}_m}{s_m}$$

Anschließend kann der Mittelwert der d skalierten Spannweiten der einzelnen Teilfolgen mittels

$$\left(\frac{R}{S}\right)_n := \frac{1}{d} \sum_{m=1}^d \frac{\text{range}_m}{s_m}$$

berechnet werden. Gilt die Behauptung (3.1), dann stehen offenbar $\log((R/S)_n)$ und $\log(n)$ in einem linearen Zusammenhang, d. h., es existieren Konstanten c und H , sodass

$$\log\left(\left(\frac{R}{S}\right)_n\right) \sim \log(c) + H \cdot \log(n) \quad (3.2)$$

gilt. Wiederholt man die oben beschriebenen Schritte für verschiedene Werte für $n \mid N$ und veranschaulicht die Ergebnisse in einer Graphik, dann sollte also nach Hurst (1951) ein linearer Zusammenhang erkennbar sein. Man erwartet also ein ähnliches Bild wie in Abbildung 3.3. Es liegt nahe, die Steigung der Näherungsgerade zu schätzen und dies als Schätzer für den Hurst Exponent zu verwenden. Hierfür setzen Mandelbrot und Wallis (1968) eine einfache lineare Regression ein und der geschätzte Steigungsparameter ist das Ergebnis der R/S Analyse.

3.3.2 Korrigierte R/S Analyse von Anis und Lloyd

Anis und Lloyd (1976) stellten fest, dass die Steigung der Regressionsgerade manchmal stark vom theoretisch korrekten Wert 0,5 abweicht, obwohl es sich um White-Noise-Daten handelt. Dieses Phänomen ist vor allem dann beobachtbar, wenn die Zeitreihe x_1, \dots, x_N zu wenig Daten enthält, d. h., wenn N zu klein ist. Um dies genauer zu untersuchen, betrachteten sie den Erwartungswert der adjusted rescaled range für normalverteilte White-Noise-Daten analytisch und kamen zum folgenden Ergebnis.

Theorem 3.3.1. *(Anis und Lloyd, 1976) Seien X_1, \dots, X_n stochastisch unabhängige, identisch normalverteilte Zufallsvariablen. Es bezeichne range die Spannweite der kumulierten Abweichungen vom Mittelwert, s die verzerrte Standardabweichung und Γ die Euler'sche Gammafunktion. Dann gilt*

$$E\left[\frac{\text{range}}{s}\right] = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{n}{2}\right)} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}}. \quad (3.3)$$

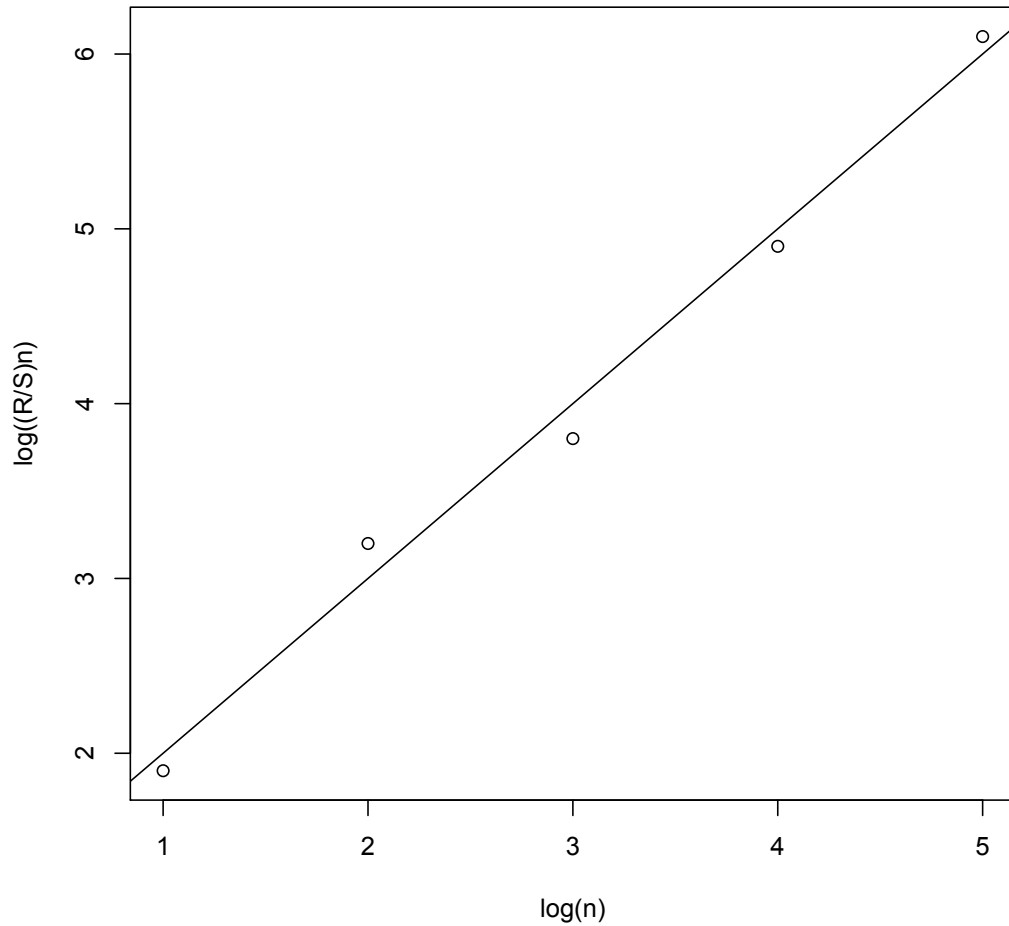


Abbildung 3.3: Gilt die Behauptung (3.1) und berechnet man für verschiedene Werte für $n \mid N$ den Wert $(R/S)_n$, dann erwartet man einen linearen Zusammenhang zwischen den logarithmierten Größen, welcher hier durch die lineare Regressionsgerade angedeutet ist. Die konkreten, eingezeichneten Werte sollen einen linearen Zusammenhang veranschaulichen und haben keine weitere Bedeutung

Mit unserer Notation aus dem Kapitel 3.3.1 und der Linearität des Erwartungswertes erhalten wir

$$E \left[\left(\frac{R}{S} \right)_n \right] = \frac{1}{d} \sum_{m=1}^d E \left[\frac{\text{range}_m}{s_m} \right].$$

Die Teilfolgen $(x_{i,m})_{1 \leq i \leq n}$, $m = 1, \dots, d$ haben alle dieselbe Länge n . Falls es sich also um normalverteilte White-Noise-Daten handelt, dann können wir für jeden Summanden die Formel (3.3) einsetzen und erhalten schließlich

$$E \left[\left(\frac{R}{S} \right)_n \right] = \begin{cases} \frac{n-\frac{1}{2}}{n} \frac{\Gamma(\frac{n-1}{2})}{\sqrt{\pi} \Gamma(\frac{n}{2})} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} & , \text{ für } n \leq 340 \\ \frac{n-\frac{1}{2}}{n} \frac{1}{\sqrt{\frac{n\pi}{2}}} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} & , \text{ für } n > 340, \end{cases} \quad (3.4)$$

wobei die Fallunterscheidung sowie der Vorfaktor $(n - 1/2)/n$ durch Peters und Peters (1994) ergänzt wurden, um eine bessere Performance für kleine Werte für n zu erreichen.

Damit ist uns eine exakte Formel für die erwartete adjusted rescaled range von normalverteilten White-Noise-Daten bekannt. Unter der Behauptung (3.1) sollte die Gleichung (3.4) logarithmiert eine Gerade mit Steigung 0,5 in $\log(n)$ ergeben. Im asymptotischen Sinne, d. h. für sehr große Werte für n , ist dies tatsächlich erfüllt. Um dies zu zeigen, verwenden wir die folgenden Abschätzungen (Hamed, 2007).

- Nach Abramowitz und Stegun (1972, Formel 6.1.46, S. 257) gilt

$$\lim_{x \rightarrow \infty} (x)^{b-a} \frac{\Gamma(x+a)}{\Gamma(x+b)} = 1.$$

Wir setzen $x = n/2$, $a = -1/2$, $b = 0$ ein und erhalten eine erste asymptotische Abschätzung

$$\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})} \approx \sqrt{\frac{2}{n}}.$$

- Solari und Anis (1957) untersuchen die Summe in (3.4) asymptotisch. Diese lässt sich für große Werte für n durch

$$\sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} \approx \frac{n\pi}{2}$$

abschätzen.

- Für große Werte für n gilt $(1/2)/n \approx 0$. Deshalb gilt asymptotisch für den Vorfaktor

$$\frac{n - \frac{1}{2}}{n} \approx 1.$$

Insgesamt erhalten wir die asymptotische Abschätzung

$$E \left[\left(\frac{R}{S} \right)_n \right] \approx \left\{ \begin{array}{l} 1 \cdot \frac{1}{\sqrt{\pi}} \cdot \sqrt{\frac{2}{n}} \cdot \frac{n\pi}{2} \quad , \text{ für } n \leq 340 \\ 1 \cdot \frac{1}{\sqrt{\frac{n\pi}{2}}} \cdot \frac{n\pi}{2} \quad , \text{ für } n > 340 \end{array} \right\} = \sqrt{\frac{n\pi}{2}}. \quad (3.5)$$

Logarithmiert ergibt sich also tatsächlich die Gerade

$$\log \left(E \left[\left(\frac{R}{S} \right)_n \right] \right) \approx \frac{1}{2} \log \left(\frac{\pi}{2} \right) + \frac{1}{2} \log(n)$$

mit Steigung 0,5. Jedoch ist hierbei zu betonen, dass die hergeleitete Abschätzung nur im *asymptotischen* Sinne gilt. Für endliche Stichproben liegt $E[(R/S)_n]$ dagegen stets unter dem Wert $\sqrt{n\pi/2}$. Schätzt man also für normalverteilte White-Noise-Daten mithilfe einer linearen Regression den linearen Zusammenhang zwischen $\log(E[(R/S)_n])$ und $\log(n)$, dann übertrifft die geschätzte Steigung den theoretisch korrekten Wert 0,5. Dies ist in Abbildung 3.4 veranschaulicht. Der Graph der exakten, logarithmierten Formel verläuft offenbar steiler als die asymptotische Gerade, sodass die R/S Analyse in Erwartung den Hurst Exponenten überschätzt. Dies ist der Grund für die Abweichung, die Anis und Lloyd (1976) feststellten.

Um diese theoretisch begründete Verzerrung zu kompensieren, kann anstelle von (3.2) die Korrektur

$$\log \left(\left(\frac{R}{S} \right)_n - \left(E \left[\left(\frac{R}{S} \right)_n \right] - \sqrt{\frac{n\pi}{2}} \right) \right) \sim \log(c) + H \cdot \log(n) \quad (3.6)$$

betrachtet werden. Schätzt man ihre Steigung H durch eine lineare Regression, dann erhält man den sogenannten *korrigierten empirischen Hurst Exponent*.

3.3.3 Durchführung der R/S Analyse mit R

Mit der Softwareumgebung R lassen sich verschiedene Varianten der R/S-Analyse leicht durchführen. Dafür kann die Funktion `hurstexp` aus dem Paket `pracma` verwendet werden. Als Funktionsargumente wird der Datensatz sowie optional die gewünschte, minimale Länge n_{min} der Teilfolgen übergeben. Ohne spezielle Angabe wird diese minimale Länge mit $n_{min} = 50$ initialisiert. An dieser Stelle ist darauf hinzuweisen, dass `hurstexp` stets die empirische Standardabweichung bei der Durchführung der verschiedenen R/S-Analysen verwendet.

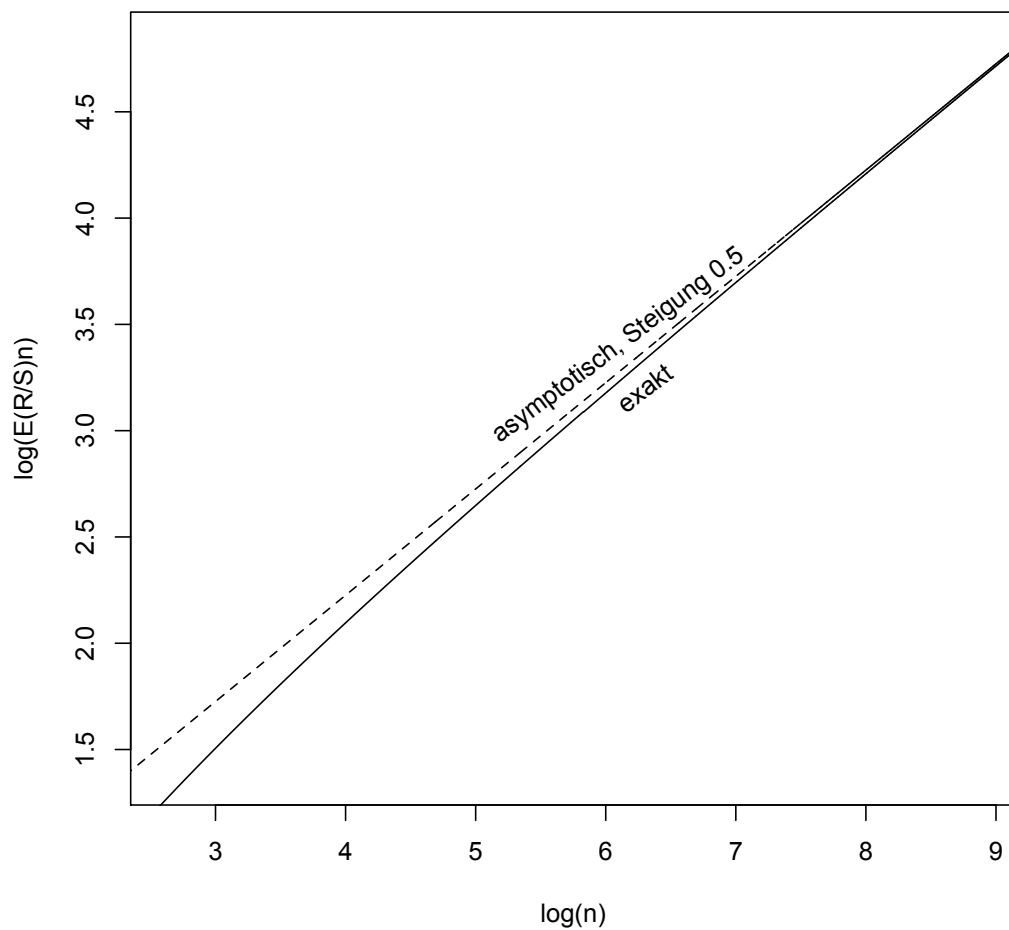


Abbildung 3.4: Vergleich der exakten Formel (3.4) für $E[(R/S)_n]$ und der asymptotischen Abschätzung (3.5) im log-log-Diagramm

Wir rufen die Funktion mit unseren S&P 500-Absoluterträgen von Februar 1871 bis April 2018 ($N=1767$) auf. Dabei kommt R zum folgenden Ergebnis:

```

hurstexp(absolutreturns)
Simple R/S Hurst estimation: 0.6667151
Corrected R over S Hurst exponent: 0.7482771
Empirical Hurst exponent: 0.7491803
Corrected empirical Hurst exponent: 0.7171259
Theoretical Hurst exponent: 0.5327066

```

Die fünf angegebenen Werte haben unterschiedliche Bedeutung und sollen nun etwas genauer erläutert werden. Zunächst sucht die Funktion `hurstexp` nach einer optimalen, verkürzten Stichprobenlänge N' , sodass sich möglichst viele Divisoren $n \mid N'$ ergeben. Dies ist wichtig für eine möglichst vielfältige Teilfolgenzerlegung. Dabei wird sichergestellt, dass die ursprüngliche Stichprobe um höchstens ein Prozent verkürzt wird, damit nicht zu viele Daten verloren gehen. In unserem Beispiel wird auf die Stichprobenlänge $N' = 1764$ verkürzt, sodass elf verschiedene Divisoren der minimalen Größe 50 möglich sind. Der Einfachheit halber wird die verkürzte Stichprobe im Folgenden wieder mit $(x_t)_{t=1,\dots,N}$ bezeichnet.

Bei der *Simple R/S Hurst estimation* handelt es sich um die einfachste R/S Analyse, bei der vollständig auf die Teilfolgenzerlegung verzichtet wird. Deshalb wird allein das Verhältnis $(R/S)_N$ über die gesamte, gegebenenfalls verkürzte, Stichprobe berechnet und der Schätzer für den Hurst Exponenten ergibt sich dann als

$$\hat{H}_{simple} := \frac{\log\left(\left(\frac{R}{S}\right)_N\right)}{\log(N)}.$$

Beim *Empirical Hurst exponent* findet dagegen die vorab ermittelte Menge der Divisoren ihren Einsatz. Dort wird für jeden Divisor $n \mid N$ der empirische adjusted rescaled range-Wert $(R/S)_n$ berechnet und der Schätzer ergibt sich, wie bereits in Abschnitt 3.3.1 erwähnt, als Steigung der linearen Regressionsgerade. In Abbildung 3.5 sind unsere elf berechneten $(R/S)_n$ -Werte im log-log-Diagramm abgebildet. Wir können leicht einen linearen Zusammenhang zwischen $\log(n)$ und $\log((R/S)_n)$ erkennen. Die geschätzte, lineare Regressionsgerade ist in die Abbildung eingezeichnet. Ihre Steigung hat den berechneten Wert 0,7491803.

Nach dem gleichen Vorgehen wird auch der *Corrected R over S Hurst exponent* ermittelt. Dabei entstehen die betrachteten Teilfolgen jedoch durch schrittweise Halbierung der vorher betrachteten Teilfolgen. Dies wird solange wiederholt, bis eine Teilfolge weniger als acht Elemente besitzt. Die festgelegte Minimallänge n_{min}

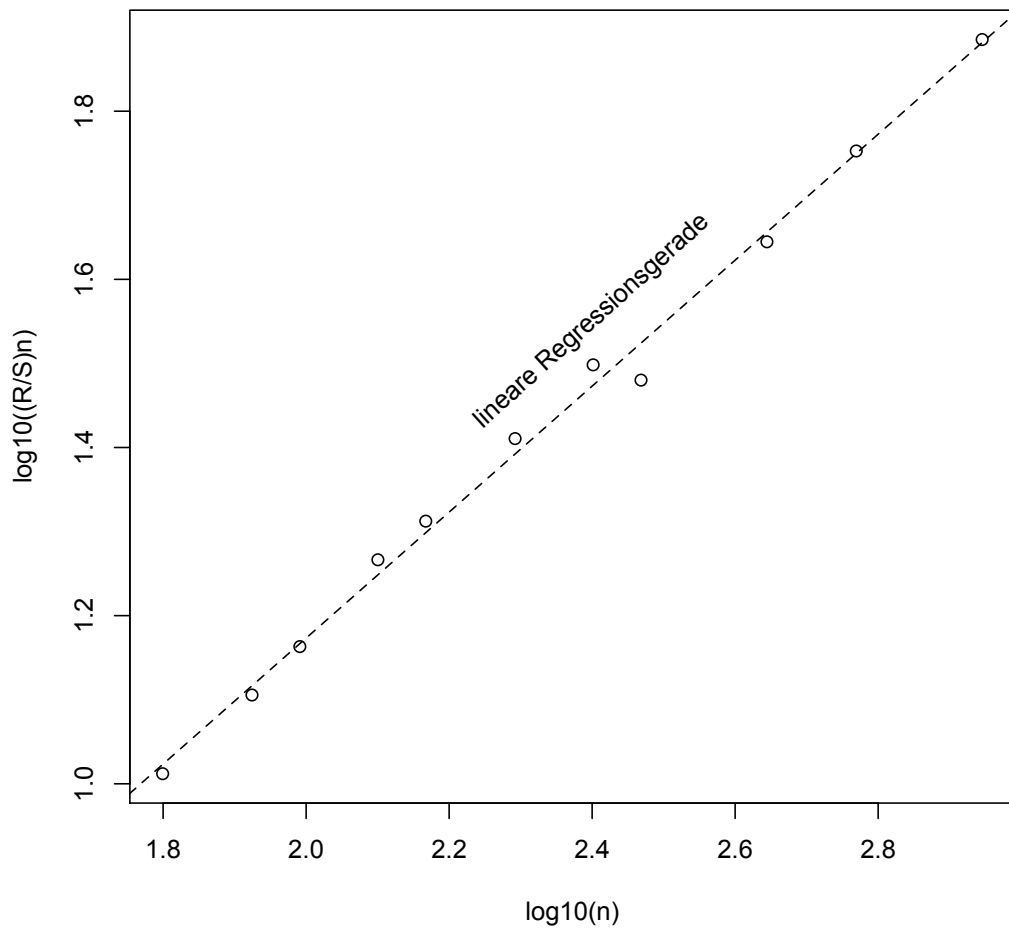


Abbildung 3.5: Log-log-Diagramm mit den elf berechneten adjusted rescaled range-Werten $(R/S)_n$ und der linearen Regressionsgerade

wird bei diesem Verfahren ignoriert. Je Iteration werden auch hier die mittleren adjusted rescaled range-Werte berechnet und schließlich wird eine lineare Regression durchgeführt. Der einzige Unterschied zur vorherigen Variante sind also die unterschiedlichen, betrachteten Teilfolgen.

Der *Theoretical Hurst exponent* wird durch die Schätzung der Steigung von

$$\log \left(E \left[\left(\frac{R}{S} \right)_n \right] \right) \sim \log(c) + H \cdot \log(n)$$

berechnet, wobei für n alle ermittelten Divisoren eingesetzt werden und $E[(R/S)_n]$ gemäß (3.4) berechnet wird. Wären unsere Daten aus einem normalverteilten White-Noise-Prozess, so würden man diesen Wert für den Empirical Hurst exponent erwarten. In unserem Fall liegen die Werte Empirical Hurst exponent und Theoretical Hurst exponent weit voneinander entfernt, sodass nicht von Daten aus einem normalverteilten White-Noise-Prozess auszugehen ist.

Die verbleibende Variante *Corrected empirical Hurst exponent* entsteht durch die Schätzung der Steigung des korrigierten Zusammenhangs (3.6), wobei auch hier für n alle gefundenen Divisoren eingesetzt werden. Ausschließlich bei diesem Schätzer für den Hurst Exponenten ist also die theoretisch hergeleitete Verzerrung berücksichtigt. Aus diesem Grund ist der Corrected empirical Hurst exponent wohl der aussagekräftigste, berechnete Schätzer. Deshalb werden wir diesen Wert als Maßzahl der long memory-Eigenschaft verwenden.

3.3.4 Empirische Konfidenzintervalle von Weron

Nachdem die Schätzwerte bestimmt und erläutert wurden, stellt sich die Frage, wie die Ergebnisse interpretiert werden können. Der typische, statistische Ansatz wäre, die Nullhypothese $H_0: H = 0,5$ bezüglich eines Signifikanzniveaus α zu testen. Um einen entsprechenden Test festzulegen, ist jedoch die asymptotische Verteilung der Teststatistik notwendig. Leider sind für die R/S Analyse keine asymptotischen Eigenschaften bekannt.

Einen Ausweg bieten die von Weron (2002) ermittelten, empirischen Konfidenzintervalle. Weron (2002) erzeugte hierzu für verschiedene Zeitreihenlängen viele Zeitreihen aus einer Normalverteilung und berechnete dann für $\alpha \in \{10\%, 5\%, 1\%\}$ die $(\alpha/2)$ - und $(1 - \alpha/2)$ -Quantile. Mit den Ergebnissen schätzte er anschließend die 90%-, 95%- und 99%-Konfidenzintervalle. In Tabelle 3.2 sind die empirischen Konfidenzintervalle für den korrigierten empirischen Hurst Exponenten mit Zeitreihenlängen der Form $N = 2^K$ und $n_{min} = 50$ angegeben.

Tabelle 3.2: Empirische Konfidenzintervalle von Weron (2002) für den korrigierten empirischen Hurst Exponenten mit $N = 2^K$ und $n_{min} = 50$

α	untere Grenze	obere Grenze
10%	$0,5 - \exp(-7,35 \cdot \log(\log K) + 4,06)$	$\exp(-7,07 \cdot \log(\log K) + 3,75) + 0,5$
5%	$0,5 - \exp(-7,33 \cdot \log(\log K) + 4,21)$	$\exp(-7,20 \cdot \log(\log K) + 4,04) + 0,5$
1%	$0,5 - \exp(-7,19 \cdot \log(\log K) + 4,34)$	$\exp(-7,51 \cdot \log(\log K) + 4,58) + 0,5$

Um die angegebenen Intervallgrenzen auf unser Beispiel anwenden zu können, bestimmen wir $K = \log_2(1764) \approx 10,78463$ und erhalten die in Tabelle 3.3 angegebenen Konfidenzintervalle für unsere S&P 500-Absoluterträge. Damit liegt unser geschätzter Wert 0,7171259 über den oberen Grenzen aller drei Konfidenzintervalle von Weron (2002). Dies deutet somit auf eine langfristige Abhängigkeit der Absoluterträge hin.

Tabelle 3.3: Empirische Konfidenzintervalle von Weron (2002), ausgewertet am verkürzten Datensatz der S&P 500-Absoluterträge im Zeitraum Februar 1871 bis Januar 2018 (N=1764)

α	untere Grenze	obere Grenze
10%	0,4004791	0,5930312
5%	0,3823524	0,611087
1%	0,3487435	0,6457303

Der stilisierte Fakt Langzeitgedächtnis bestätigt sich somit in der durchgeführten R/S Analyse. Der korrigierte empirische Hurst Exponent stellt eine geeignete Maßzahl für die Charakteristik dar.

3.4 Kurtosis

Auch die Fat Tails in der Ertragsverteilung wollen wir mithilfe von geeigneten Maßzahlen messbar machen. Eine naheliegende Kennzahl ist hierbei die *empirische Kurtosis* oder auch *Wölbung*. Sie ist ein typisches Maß für das Verhalten einer Verteilung an den Tails und entspricht für eine Zeitreihe $(x_t)_{t=1,\dots,N}$ dem standardisierten, empirischen, zentralen Moment vierter Ordnung:

$$\kappa := \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{\left[\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right]^2} \quad (3.7)$$

Ein eng verwandter Begriff ist der *Exzess* $\nu := \kappa - 3$ (vergleiche Hartung et al., 1993, S. 49). Der Wert $\nu = 0$ bzw. $\kappa = 3$ charakterisiert dabei die *mesokurtische* Normalverteilung. Ergibt sich $\nu > 0$, dann handelt es sich um eine *leptokurtische* oder auch *steilgipflige* Verteilung. Bei $\nu < 0$ spricht man von einem *platykurtischen* oder *flachgipfligen* Tail-Verhalten der Verteilung. In Abbildung 3.6 sind die Dichtefunktionen der steilgipfligen *t-Verteilung mit drei Freiheitsgraden* sowie der flachgipfligen *stetigen Gleichverteilung auf dem Intervall $[-3, 3]$* jeweils im Vergleich mit der Normalverteilung $\mathcal{N}(0, 3)$ dargestellt. Die genannten Verteilungen haben einheitlich den Mittelwert Null und die Varianz drei, sodass die Vergleichbarkeit sichergestellt ist. Es ist erkennbar, dass bei steilgipfligen Verteilungen mehr Wahrscheinlichkeit im Zentrum sowie in den Tails liegt. Da dies gerade der Charakteristik Fat Tails in der Ertragsverteilung entspricht, erwarten wir einen positiven Exzess für die Zeitreihe der Erträge.

Mit R können wir die Kurtosis mit der Funktion `kurtosis` aus dem Paket `moments` ausrechnen. Da es sich nicht um den Exzess handelt, müssen wir bei der Interpretation des Ergebnisses mit dem Wert drei anstelle von null vergleichen. Ausgewertet an der S&P 500-Zeitreihe der Erträge von Februar 1871 bis April 2018 kommt der Funktionsaufruf zum Ergebnis $\kappa = 20,19594 > 3$. Die empirische Kurtosis deutet somit tatsächlich auf eine steilgipflige Verteilung der Erträge hin. Damit dient die empirische Kurtosis der Ertragszeitreihe $(r_t)_{t=1, \dots, N}$ als geeignetes Maß für den stilisierten Fakt Fat Tails in der Ertragsverteilung.

3.5 Hill Schätzer für den Tail Index

Alternativ zur Kurtosis kann mit dem Ansatz der Extremwerttheorie eine weitere Kennzahl für die Eigenschaft Fat Tails in der Ertragsverteilung hergeleitet werden. Der zentrale Begriff ist hierbei der sogenannte *Tail Index* α . Dieser Index beschreibt das Tailverhalten einer Verteilung. Je größer der Tail Index ist, desto dünner sind die Tails der Verteilung. Gauß'sche, exponentielle oder noch flachgipfligere Verteilungen haben Tail Index $+\infty$. Bei einer *Power-Law* Verteilung, d. h., die Dichte ist für große Werte für x dominiert durch einen Term der Form $x^{-\alpha}$, entspricht der Tail Index gerade dem Parameter α . Eine Messung des Tail Index der Ertragsverteilung würde uns demnach eine geeignete Kennzahl für den betrachteten stilisierten Fakt liefern.

Für eine solche Messung müssen wir uns zunächst mit Erkenntnissen der Extremwerttheorie vertraut machen. Die Theorie beschäftigt sich mit der Wahrscheinlichkeit extremer Ereignisse und damit auch mit den Tails von Verteilungen. Ein grundlegender Satz ist dabei das sogenannte *Extremal Types Theorem* oder auch

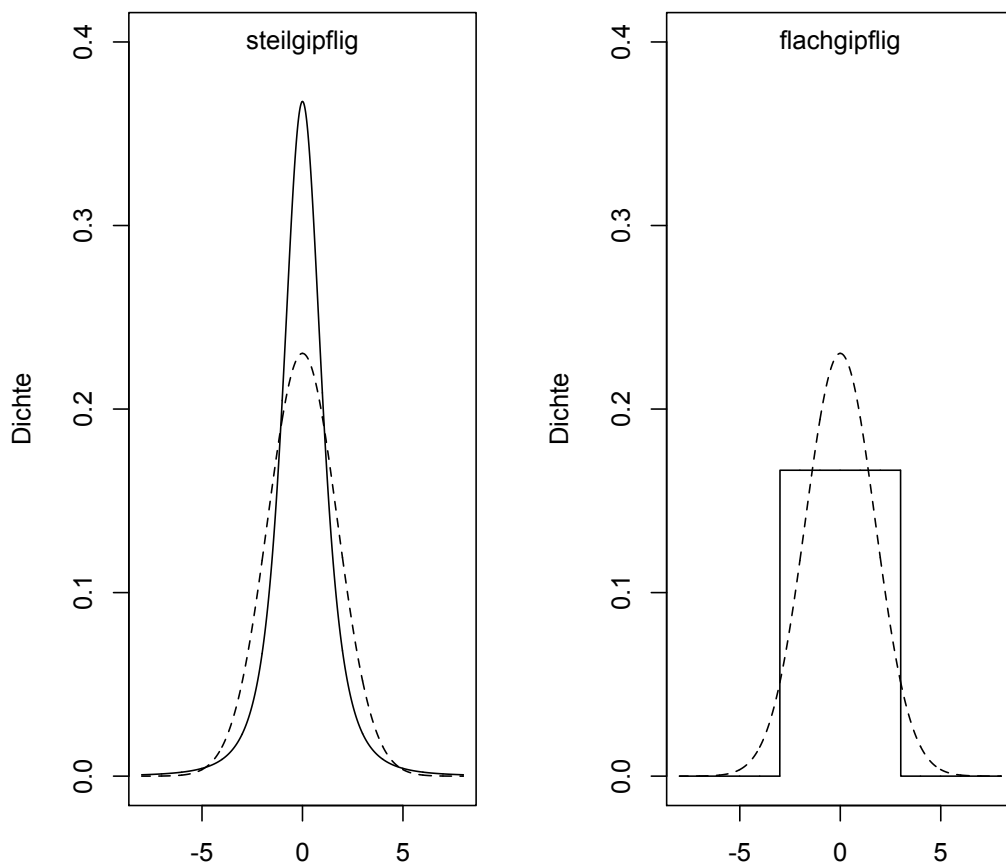


Abbildung 3.6: Dichtefunktion der t-Verteilung mit drei Freiheitsgraden (durchgezogene Linie links) bzw. Dichtefunktion der stetigen Gleichverteilung auf $[-3, 3]$ (durchgezogene Linie rechts) jeweils verglichen mit der Dichtefunktion der Normalverteilung (gestrichelte Linie), wobei alle drei Verteilungen den Mittelwert Null und die Standardabweichung $\sqrt{3}$ haben

Fisher-Tippett-Gnedenko-Theorem, welches auf die Arbeiten von Fisher und Tippett (1928) und Gnedenko (1943) zurückgeht. Das Theorem charakterisiert die Grenzverteilungsfunktion des Maximums mehrerer, stochastisch unabhängiger und identisch verteilter Zufallsvariablen und besagt, dass es im Wesentlichen nur drei Typen für diese Grenzverteilungsfunktion gibt.

Theorem 3.5.1. (Fisher-Tippett-Gnedenko-Theorem) (vergl. Lux, 2008; Cont, 2001) Seien X_1, \dots, X_n stochastisch unabhängig und identisch verteilt mit gemeinsamer Verteilungsfunktion F und sei $M_n = \max\{X_1, \dots, X_n\}$. Falls es normalisierende Konstanten (a_n, b_n) , $a_n > 0$ und eine nichtausgeartete Grenzverteilungsfunktion G für das normalisierte Maximum gibt, d. h.

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \xrightarrow{n \rightarrow \infty} G(x),$$

dann muss G von einem der drei Typen sein:

$$\text{Typ I („Gumbel“): } G_1(x) = \exp(-\exp(-x)), \quad x \in \mathbb{R};$$

$$\text{Typ II („Fréchet“): } G_{2,\alpha}(x) = \begin{cases} 0 & , x \leq 0 \\ \exp(-x^{-\alpha}) & , x > 0 \end{cases}, \text{ für ein } \alpha > 0;$$

$$\text{Typ III („Weibull“): } G_{3,\alpha}(x) = \begin{cases} \exp(-(-x)^\alpha) & , x \leq 0 \\ 1 & , x > 0 \end{cases}, \text{ für ein } \alpha > 0$$

Die drei Verteilungen unterscheiden sich stark im Tailverhalten. Die Gumbelverteilung hat einen exponentiellen Abfall, d. h. sehr dünne Tails. In diese Klasse fallen beispielsweise die Normalverteilung und die Exponentialverteilung. Die Fréchetverteilungen dagegen zeichnen sich durch einen hyperbelförmigen Abfall aus, d. h., es handelt sich um Verteilungen mit Fat Tails. Weibullverteilungen haben endliche Endpunkte. In Abbildung 3.7 sind die drei Verteilungsfunktionen sowie die zugehörigen Dichtefunktionen abgebildet, wobei für die Fréchet- und Weibullverteilung der Parameter $\alpha = 1$ gewählt wurde. Das unterschiedliche Tailverhalten ist anhand der Dichtefunktionen gut erkennbar.

Anstelle des Tail Index α liest man in der Literatur häufig den Begriff *shape parameter* ξ . Die Begriffe dürfen nicht als Synonyme verwendet werden, sind aber eng miteinander verknüpft. Für den Zusammenhang sorgt die sogenannte *Cramer-von Mises-Parametrisierung*. Dabei handelt es sich wiederum um eine parametrisierte Verteilungsfunktion, die definiert ist als

$$C_\xi(u) := \begin{cases} \exp\left(-\left(1 + \xi u\right)^{-\frac{1}{\xi}}\right) & , \xi \neq 0 \\ \exp(-\exp(-u)) & , \xi = 0, \end{cases} \quad (3.8)$$

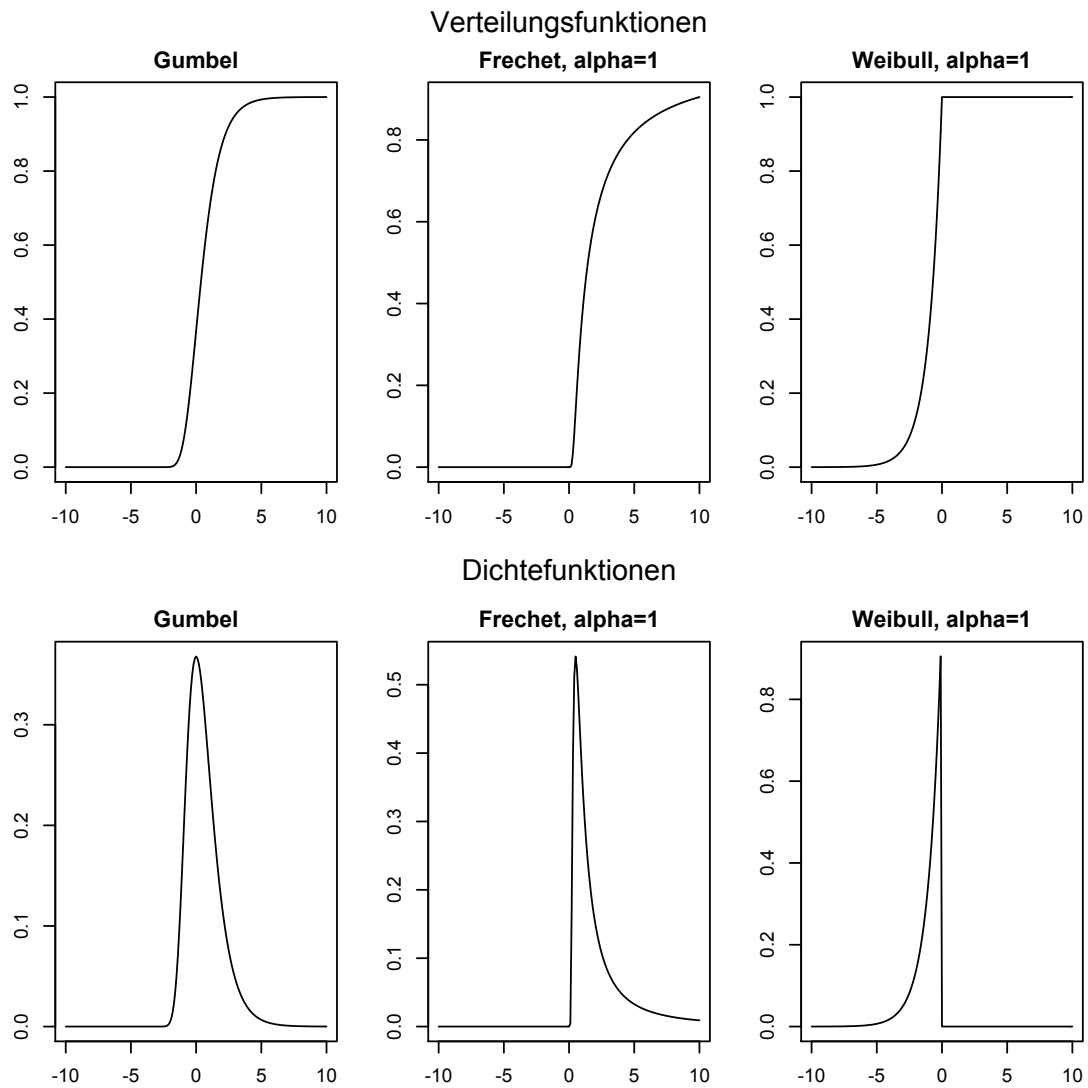


Abbildung 3.7: Verteilungs- (erste Zeile) und Dichtefunktionen (zweite Zeile) der Gumbelverteilung, der Fréchetverteilung mit $\alpha = 1$ und der Weibullverteilung mit $\alpha = 1$

wobei an Definitionslücken so fortgesetzt wird, dass sich eine Verteilungsfunktion ergibt. Insbesondere müssen die Bedingungen

$$\lim_{u \rightarrow -\infty} C_\xi(u) = 0 \quad (3.9)$$

und

$$\lim_{u \rightarrow +\infty} C_\xi(u) = 1 \quad (3.10)$$

erfüllt sein.

Theorem 3.5.2. (vergl. Lux, 2008; Cont, 2001) *Die Cramer-von Mises-Parametrisierung umfasst alle drei Klassen der Gumbel-, Fréchet- und Weibullverteilungen, d. h., für alle $\xi, u \in \mathbb{R}$ existieren ein $x \in \mathbb{R}$ und gegebenenfalls ein $\alpha > 0$, sodass eine der drei Gleichungen*

$$\begin{aligned} C_\xi(u) &= G_1(x) \\ C_\xi(u) &= G_{2,\alpha}(x) \\ C_\xi(u) &= G_{3,\alpha}(x) \end{aligned}$$

erfüllt ist.

Beweis. Wir zeigen die Aussage durch Fallunterscheidung für ξ . Sei also $u \in \mathbb{R}$ beliebig, aber fest.

$\xi = 0$: In diesem Fall erhält man mit $x := u$ sofort die Gumbelverteilung, denn es gilt

$$C_0(u) = \exp(-\exp(-u)) = G_1(u) = G_1(x).$$

$\xi > 0$: Im zweiten Fall ergibt sich für $1 + \xi u \leq 0$ eine Definitionslücke in der Definition (3.8), da dann die Basis der Potenz nichtpositiv und der Exponent reell und negativ ist. Aufgrund der Bedingung (3.9) setzt man die Verteilungsfunktion folgendermaßen fort:

$$C_\xi(u) := \begin{cases} 0 & , 1 + \xi u \leq 0 \\ \exp\left(- (1 + \xi u)^{-\frac{1}{\xi}}\right) & , 1 + \xi u > 0 \end{cases}$$

Mit den Definitionen $x := 1 + \xi u$ und $\alpha := 1/\xi$ gilt $\alpha > 0$ und wir erhalten gerade die Fréchetverteilung:

$$C_\xi(u) = \left\{ \begin{array}{ll} 0 & , x \leq 0 \\ \exp(-x^{-\alpha}) & , x > 0 \end{array} \right\} = G_{2,\alpha}(x)$$

$\xi < 0$: Im Gegensatz zum vorherigen Fall ist hier der Exponent positiv, sodass die Null als Basis keine Definitionslücke darstellt. Die Definitionslücke beschränkt sich somit auf den Fall $1 + \xi u < 0$. Gemäß der Bedingung (3.10) setzt man hier mittels

$$C_\xi(u) := \begin{cases} \exp\left(- (1 + \xi u)^{-\frac{1}{\xi}}\right) & , 1 + \xi u \geq 0 \\ 1 & , 1 + \xi u < 0 \end{cases}$$

fort, sodass sich eine Verteilungsfunktion ergibt. Wir definieren $x := -(1 + \xi u)$ und $\alpha := -1/\xi > 0$ und erhalten die Weibullverteilung:

$$C_\xi(u) = \begin{cases} \exp(-(-x)^\alpha) & , x \leq 0 \\ 1 & , x > 0 \end{cases} = G_{3,\alpha}(x)$$

□

Wir haben also gesehen, dass der Tail Index α und der shape parameter ξ direkt miteinander verknüpft sind und damit beide charakterisierende Parameter für das Tailverhalten der Grenzverteilungsfunktion darstellen. Wir wollen die beschriebene Theorie nun auf die Erträge $(R_t)_{t \in \mathbb{N}}$ anwenden, wobei wir die Annahme treffen müssen, dass diese stochastisch unabhängig und identisch verteilt sind. Dann liefert uns das Fisher-Tippett-Gnedenko-Theorem die Aussage, dass extreme, d. h. maximale, Erträge nach einer Normierung im Grenzwert einer Gumbel-, Fréchet- oder Weibullverteilungsfunktion folgen. In Abbildung 2.2 sowie bei der Auswertung der Kurtosis haben wir bereits festgestellt, dass die Ertragsverteilung offenbar Fat Tails hat. Wir werden uns deshalb auf den Typ II, also auf die Fréchetverteilungen, konzentrieren. Mit der Cramer-von Mises-Parametrisierung befinden wir uns somit im Fall $\xi > 0$ mit dem Zusammenhang $\alpha = 1/\xi$. Ziel ist es, den für das Tailverhalten entscheidenden Tail Index α bzw. den shape parameter ξ anhand einer gegebenen Zeitreihe $(r_t)_{t=1,\dots,N}$ zu schätzen und damit eine Kennzahl für die tatsächliche Größe der Tails der Ertragsverteilung zu erhalten.

Der in unserem Kontext üblicherweise verwendete Schätzer heißt *Hill-Schätzer* (Hill, 1975). Er schätzt den Tail Index $\alpha = 1/\xi$ für den Fréchet-Fall durch einen *bedingten Maximum Likelihood Ansatz*. Im ersten Schritt wird dazu die Zeitreihe $(r_t)_{t=1,\dots,N}$ absteigend geordnet, sodass

$$r_{(N)} \geq r_{(N-1)} \geq \dots \geq r_{(1)}$$

gilt. Anschließend muss durch die Wahl eines Wertes $1 \leq k \leq N - 1$ festgelegt werden, bezüglich welcher Tailgröße der Hill-Schätzer berechnet werden soll. Beispielsweise kann man k gerade so wählen, dass nur die 2, 5%, 5% oder 10% größten,

gemessenen Erträge der Zeitreihe in die Schätzung eingehen. Der Hill-Schätzer ist nun definiert als

$$\hat{\alpha}_k^{(H)} := \frac{1}{\hat{\xi}_k^{(H)}} := \frac{1}{\frac{1}{k} \sum_{i=1}^k [\log r_{(N-i+1)} - \log r_{(N-k)}]} \quad (3.11)$$

(vergleiche Lux und Ausloos, 2002).

Mit R kann der beschriebene Hill-Schätzer für den Tail Index mit der Funktion `thetaHill` aus dem Paket `laeken` berechnet werden. Der Funktion wird die betrachtete Zeitreihe sowie der gewünschte Wert k übergeben.

Wir werten die Funktion `thetaHill` an der Zeitreihe der S&P 500-Erträge von Februar 1871 bis April 2018 aus und wählen k dabei so, dass nur die 2,5%, 5% bzw. 10% größten, gemessenen Erträge in die Schätzung einfließen. Dies liefert das Ergebnis

$$\begin{aligned} \hat{\alpha}_{2,5\%}^{(H)} &= 3,32094; \\ \hat{\alpha}_{5\%}^{(H)} &= 3,630764; \\ \hat{\alpha}_{10\%}^{(H)} &= 2,77868. \end{aligned}$$

Die Werte lassen sich mit den Ergebnissen zahlreicher Studien für Finanzzeitreihen vereinbaren, die ebenfalls Hill-Schätzer zwischen 2,5 und 4 ergaben (vergleiche Lux und Ausloos, 2002). Mit dem Hill-Schätzer für den Tail Index haben wir nun eine weitere, geeignete Maßzahl für den stilisierten Fakt Fat Tails in der Ertragsverteilung gefunden.

An dieser Stelle ich nochmals zu betonen, dass wir für die Verwendung des Hill-Schätzers annehmen müssen, dass die Werte der Zeitreihe aus einer Fréchetverteilung stammen, da die theoretische Herleitung des Schätzers diese Annahme voraussetzt. Insbesondere kann anhand des Hill-Schätzers nicht zwischen den drei Grenzverteilungstypen unterschieden werden. Der Hill-Schätzer liefert vielmehr einen Schätzer für die tatsächliche Dicke der Tails im Fréchet-Fall. Um uns von der Richtigkeit der Annahme zu überzeugen, sollten wir vor der Verwendung des Hill-Schätzers also stets die Kurtosis betrachten und darauf achten, dass diese auf Fat Tails der Verteilung und damit auf den Fréchet-Fall hindeutet.

Kapitel 4

Verfügbarkeit geeigneter Datensätze

In den Kapiteln 2 und 3 haben wir gesehen, welche Charakteristiken in realen Finanzzeitreihen üblicherweise beobachtbar sind und mit welchen statistischen Mitteln man diese messbar machen kann. Dieses Wissen werden wir in Kapitel 7 nutzen, um mathematische Modelle zu bewerten. Dazu werden wir Simulationszeitreihen der Modelle erzeugen und die in Kapitel 3 beschriebenen Kennzahlen an ihnen auswerten. Zur Beurteilung der Güte des Modells benötigen wir nun natürlich einen Vergleich mit den entsprechenden Werten einer realen Zeitreihe. In diesem Kapitel stellen wir uns die Frage, woher wir geeignete Datensätze der realen Finanzmärkte nehmen und wie wir mit diesen arbeiten können.

Ein beliebter, kostenlos verfügbarer Datensatz ist der S&P 500-Datensatz, der auf der Homepage des Ökonomen Robert J. Shiller bereitgestellt wird (Shiller, 2018). Der Aktienindex S&P 500, ausgeschrieben *Standard and Poor's 500*, gilt als einer der wichtigsten, US-amerikanischen Indizes und umfasst die 500 größten, börsennotierten Unternehmen der USA. Die Werte werden dabei nach der Marktkapitalisierung gewichtet. Darunter versteht man den aktuellen Börsenwert eines börsennotierten Unternehmens, d. h., die Marktkapitalisierung ergibt sich aus dem aktuellen Aktienkurs multipliziert mit der Gesamtzahl der herausgegebenen Anteilsscheine (vergleiche Gruppe Deutsche Boerse, 2018b). Shiller (2018) präsentiert in seinem Datensatz monatliche Werte wie Preise, Dividenden und Gewinne beginnend mit Januar 1871, wobei aktuelle Daten kontinuierlich ergänzt werden. Da der Aktienindex erst 1957 eingeführt wurde, handelt es sich bei vorherigen Daten um rückgerechnete Werte. Natürlich gibt es neben dem Datensatz von Shiller (2018) zahlreiche weitere, verwendbare Zeitreihen. Beispielsweise können bei Yahoo (2018) historische Kurswerte verschiedener Aktien und Indizes kostenlos heruntergeladen werden. Für den S&P 500-Aktienindex sind dort tägliche Kurswerte ab 1950 verfügbar. Außerdem gibt es zahlreiche Anbieter kostenpflichtiger Datensätze.

Unabhängig von der Herkunft des gewählten Datensatzes sollte man sich vor der Verwendung zunächst mit der genauen Bedeutung der Werte befassen. Da wir uns für den Datensatz von Shiller (2018) entscheiden, werden wir in Abschnitt 4.1 die dort bereitgestellten Werte genauer erläutern. Anschließend werden wir in Abschnitt 4.2 auf das Einlesen und die Anwendung des Datensatzes mit R eingehen.

4.1 Datensatz von Robert Shiller

Für die in Kapitel 3 erläuterten Maßzahlen benötigen wir vor allem eine Zeitreihe $(p_t)_{t=0,\dots,N}$ der Preise, aus denen wir die reinen, absoluten und quadrierten Erträge berechnen können. Aber auch Zeitreihen realer, d. h. inflationsbereinigter Preise und Dividenden sowie Zinssätze sind für die Berechnung der geschätzten Fundamentalwerte in Kapitel 5 und damit auch für einige Kennzahlen aus Kapitel 3 notwendig. Der S&P 500-Datensatz von Shiller (2018) enthält alle diese Werte. Im Folgenden soll auf jede bereitgestellte Zeitreihe genauer eingegangen werden.

Preise

Shiller (2018) gibt monatliche Preise ab Januar 1871 in seinem Datensatz an. Nun stellt sich natürlich die Frage, wie sich die monatlichen Werte ergeben. Schließlich werden Preise kontinuierlich angepasst, sodass innerhalb eines Monats unterschiedliche Preise angenommen werden. Shiller (2018) berechnet seine monatlichen Preise, indem er den Mittelwert der täglichen Schlusskurse des jeweiligen Monats bildet. Unter dem Schlusskurs versteht man den letzten Kurs eines Handelstages, zu dem noch Umsätze generiert wurden. Diese täglichen Werte werden vom *Standard and Poor's Statistical Service* aus dem *Security Price Index Record* bereitgestellt und von Shiller (2018) einfach übernommen (vergleiche Shiller, 1989, S. 439).

Dividenden

Ein Teil des Bilanzgewinns eines börsennotierten Unternehmens wird in der Regel in Form von Dividendenzahlungen an die Aktionäre ausgeschüttet (vergleiche Gruppe Deutsche Boerse, 2018a). In Deutschland findet die Ausschüttung üblicherweise jährlich statt, in den USA dagegen meist vierteljährlich. Shiller (2018) gibt in seinem Datensatz unter *Dividenden* die Summe der gewichteten Jahresdividenden der 500 enthaltenen Unternehmen an, wobei jeweils die aktuellen S&P 500-Gewichte verwendet werden. Mit Jahresdividende ist die Summe der vergangenen vier Quartalsdividenden gemeint.

Die dafür notwendigen historischen Werte bezieht Shiller (2018) aus zwei verschiedenen Quellen. Ab 1926 sind die Quartalsdividenden aus der Tabelle *Earnings, Dividends and Price-Earnings Ratios - Quarterly* des Standard and Poor's Statistical Service (Security Price Index Record) entnommen. Für frühere Zeitpunkte werden dort keine Werte angegeben. Shiller (2018) übernimmt deshalb für den Zeitraum 1871-1925 die Dividendenzeitreihe von Cowles (1939), die er mit einer Konstante multipliziert, um den Übergangswert anzupassen. Die beiden angegebenen Quellen liefern jedoch nur jährliche bzw. vierteljährliche Werte. Um monatliche Dividenden auszurechnen, ist es naheliegend, die gegebenen Werte zu interpolieren. Sind beispielsweise die Dividenden $d_{\text{Sep},2017}$ und $d_{\text{Dez},2017}$ für September und Dezember 2017 gegeben, so bildet Shiller (2018) die verbindende Gerade und wählt als Oktober- und Novemberdividenden $d_{\text{Okt},2017}$, $d_{\text{Nov},2017}$ die entsprechenden Werte auf dieser Gerade. In Abbildung 4.1 ist das Prinzip der linearen Interpolation am beschriebenen Beispiel graphisch veranschaulicht. Anhand der September- und Dezemberwerte 2017 aus dem Datensatz berechnen wir

$$\begin{aligned} d_{\text{Sep},2017} = 48,17 & & d_{\text{Okt},2017} = \frac{2}{3} \cdot d_{\text{Sep},2017} + \frac{1}{3} \cdot d_{\text{Dez},2017} \approx 48,42 \\ \Rightarrow & & \\ d_{\text{Dez},2017} = 48,93 & & d_{\text{Nov},2017} = \frac{1}{3} \cdot d_{\text{Sep},2017} + \frac{2}{3} \cdot d_{\text{Dez},2017} \approx 48,68 \end{aligned}$$

und diese stimmen mit den Oktober- und Novemberwerten 2017 von Shiller (2018) überein.

An dieser Stelle ist nochmals zu betonen, dass die genannten Dividendenwerte einen jährlichen Bezug haben. Sie geben also die theoretischen Dividendenzahlungen des Aktienindex im vergangenen Jahr an. Sobald wir mit einer monatlichen Zeitskala rechnen, müssen wir die Werte von Shiller (2018) durch zwölf teilen, um einen monatlichen Bezug zu erhalten.

Gewinne

Der betrachtete Datensatz enthält außerdem die Zeitreihe der Gewinne. Diese sind aus den gleichen Quellen wie die Dividenden entnommen und auch hier wird lineare Interpolation verwendet, um monatliche Werte aus den jährlichen bzw. vierteljährlichen Werten zu erhalten. Die angegebenen S&P 500-Gewinne entsprechen, analog zu den Dividenden, den gewichteten Unternehmensgewinnen der jeweils beteiligten 500 Unternehmen. Dabei ist es natürlich entscheidend, dass die Unternehmensgewinne einheitlich berechnet wurden, da ansonsten die Vergleichbarkeit nicht gegeben und die Gewichtung nicht sinnvoll ist. Zu diesem Zweck legt das Rechnungslegungsgremium *Financial Accounting Standards Board* buchhalterische Berechnungsstandards fest, welche unter dem Namen *generally accepted accounting principles*

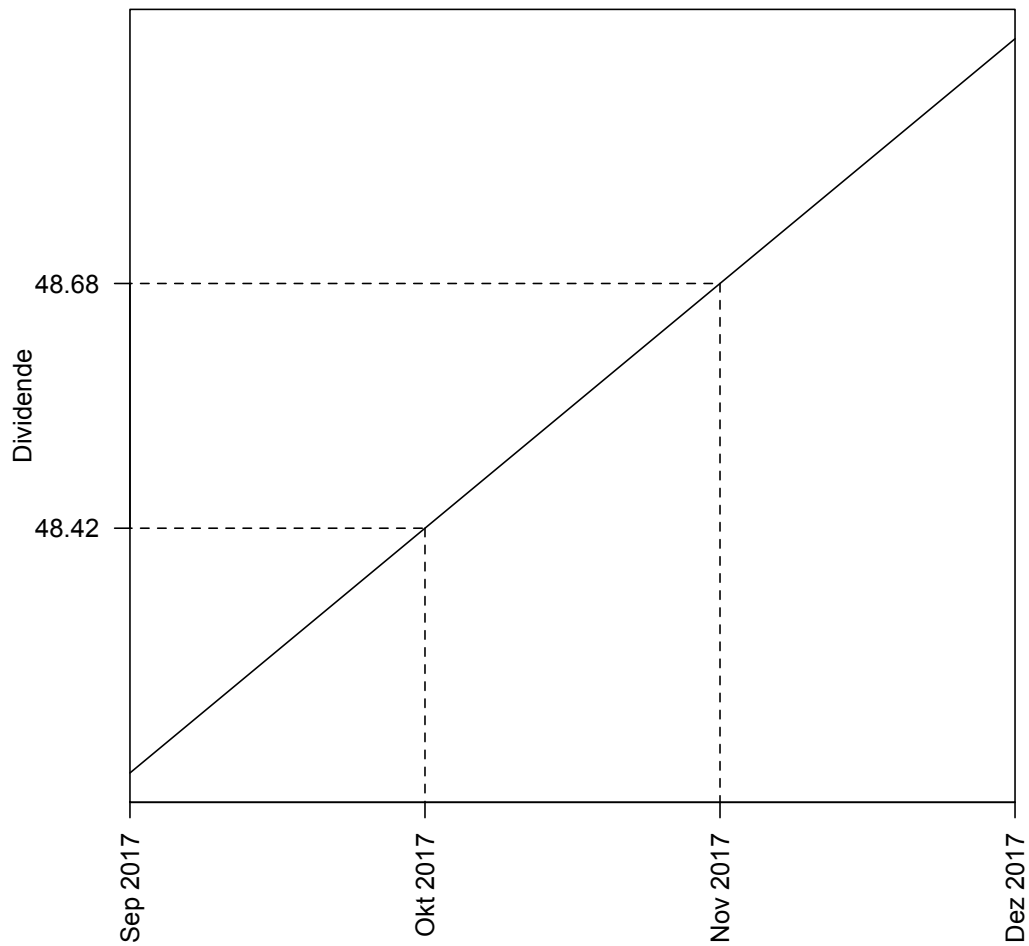


Abbildung 4.1: Mittels linearer Interpolation erhält Shiller (2018) aus den gegebenen Dividenden für September und Dezember 2017 die Werte für Oktober und November 2017

(GAAP) bekannt sind (vergleiche Siegel, 2016). Da die von Standard and Poor's dokumentierten Gewinne auf Grundlage von GAAP bestimmt wurden, ist die einheitliche Berechnungsweise sichergestellt und die gewichteten S&P 500-Gewinne ergeben eine sinnvolle Größe.

Für Zeitpunkte vor 1926 greift Shiller (2018), wie auch bei den Dividenden, auf Werte von Cowles (1939) zurück. Dieser musste sich jedoch mit dem Problem konfrontieren, dass nicht alle notwendigen Gewinndaten der beteiligten Unternehmen vorhanden sind. Cowles (1939) berechnet deshalb für die verfügbaren Unternehmensgewinne das jeweilige Gewinn-Preis-Verhältnis und nutzt diese Verhältnisse, um durch Multiplikation mit dem jährlichen, durchschnittlichen Kurs einen jährlichen S&P 500-Gewinn auszurechnen. Die Hochrechnung dient als Hilfsmittel, die fehlenden Daten zu kompensieren, führt aber auch zu einer Ungenauigkeit, die wir beachten sollten.

Consumer Price Index

Mit dem *Consumer Price Index (CPI)* gibt Shiller (2018) eine Größe an, die das Preisniveau zum jeweiligen Zeitpunkt widerspiegelt. Es handelt sich genauer um die Größe *CPI for all Urban Consumers (CPI-U)*, die monatlich vom *Bureau of Labor Statistics (BLS)* bereitgestellt wird. Dabei wird ein Warenkorb alltäglicher Ausgaben gemäß dem durchschnittlichen Kaufverhalten von Bewohnern städtischer Räume und Metropolregionen in den USA aufgestellt und sein aktueller Preis bestimmt. Die Auswahl der Güter und Dienstleistungen, die Festlegung ihrer Gewichtung und ihre Preisbestimmung beruhen auf umfangreichen Studien und komplexen Messungen, sodass sich ein möglichst präziser Wert des Warenkorbs ergibt. Damit spiegelt der CPI das Preisniveau des alltäglichen Lebens gut wieder und dient als übliche Größe zur Deflationierung, d. h. zur Beseitigung des Inflationseinflusses von Zeitreihen (Bureau of Labor Statistics, 2018). Bei der inflationsbehafteter Zeitreihe spricht man von *nominalen Größen* und durch Deflationierung mithilfe des CPI erhält man die entsprechenden *realen Größen*. Vergleichbar mit dem amerikanischen CPI ist in Deutschland der *Verbraucherpreisindex* des *Statistischen Bundesamts* (Statistisches Bundesamt, 2018).

Ab dem Jahr 1913 übernimmt Shiller (2018) die CPI-Werte des BLS. Zu vorherigen Zeitpunkten liefert das BLS keine Angaben. Deshalb bezieht sich Shiller (2018) für die Jahre 1871-1912 auf den monatlichen Preisindex von Warren und Pearson (1935, Tabelle 1), der mit dem CPI vergleichbar ist. Um die Werte von Warren und Pearson (1935) an den Übergangswert im Januar 1913 anzupassen, multipliziert Shiller (2018) sie mit dem Verhältnis

der beiden Januarwerte 1913:

$$\text{CPI}_{\text{Shiller}}(t) := \begin{cases} \text{Preisindex}_{\text{W\&P}}(t) \cdot \frac{\text{CPI}_{\text{BLS}}(1913)}{\text{Preisindex}_{\text{W\&P}}(1913)} & , t < \text{Januar 1913} \\ \text{CPI}_{\text{BLS}}(t) & , t \geq \text{Januar 1913} \end{cases}$$

Date Fraction

Diesen Wert gibt Shiller (2018) lediglich an, um Diagramme wie zum Beispiel die Preisentwicklung im Zeitverlauf erstellen zu können. Die Zeitreihe hat keine weitere Bedeutung.

Long Interest Rate

Die nächste, angegebene Zeitreihe gibt die Zinsentwicklung von US-Staatsanleihen im Zeitverlauf an. Genauer gesagt handelt es sich um die *10-Year Treasury Constant Maturity Rate (GS10)*, d. h., es wird der jeweilige Zinssatz einer US-Staatsanleihe mit Laufzeitende in exakt zehn Jahren angegeben. Nun ist es natürlich möglich, dass zum gewünschten Berechnungstichtag keine US-Staatsanleihe offen ist, die *exakt* in zehn Jahren endet. Deshalb bildet man zunächst die Zinsstrukturkurve. Diese bildet den Zusammenhang zwischen dem Zinssatz und der Restlaufzeit einer Anlage ab. Mittels Interpolation erhält man aus den Werten der tatsächlich gehandelten Anleihen eine kontinuierliche Kurve, an welcher der Wert für die Restlaufzeit von exakt zehn Jahren abgelesen werden kann (vergleiche Board of Governors of the Federal Reserve System, 2018).

Die von Shiller (2018) angegebenen, monatlichen Werte ab 1953 werden durch die Datenbank Federal Reserve Economic Data (FRED) (2018) bereitgestellt. Für den Zeitraum 1871-1952 übernimmt Shiller (2018) die jährlichen Januarwerte aus Homer und Sylla (1996) und verwendet erneut lineare Interpolation, um die restlichen Monatswerte zu bestimmen.

Reale Preise, Dividenden und Gewinne

Den bereits beschriebenen CPI können wir zur Deflationierung der Preise, Dividenden und Gewinne nutzen und erhalten so die jeweiligen Zeitreihen mit realen Größen. Dazu rechnen wir alle Werte durch Multiplikation mit dem entsprechenden CPI-Verhältnis auf das gegenwärtige Preisniveau um, wobei mit Gegenwart der letzte, betrachtete Zeitpunkt im Datensatz gemeint ist. Am Beispiel der Preise berechnet Shiller (2018) den realen Preis p_t^{real} zum Zeitpunkt t durch Multiplikation der nominalen Größe p_t mit dem CPI-Verhältnis zwischen der Gegenwart und dem Zeitpunkt t :

$$p_t^{\text{real}} := p_t \cdot \frac{\text{CPI}(\text{Gegenwart})}{\text{CPI}(t)} \quad (4.1)$$

Cyclically Adjusted Price Earnings Ratio

Die letzte, von Shiller (2018) angegebene Zeitreihe gibt das von ihm entwickelte *Cyclically Adjusted Price Earnings Ratio (CAPE)* an. Die Kennzahl definiert er als das Verhältnis des realen Preises zum mittleren, realen Gewinn der vergangenen zehn Jahre:

$$\text{CAPE}_t := \frac{p_t^{\text{real}}}{e_t^{\text{real}}}$$
$$\text{mit } \overline{e_t^{\text{real}}} := \frac{1}{10 \cdot 12} \sum_{j=t-10 \cdot 12}^{t-1} e_j^{\text{real}}$$

und e_t^{real} = realer Gewinn zum Zeitpunkt t

Das Bilden des Mittelwertes dient dabei der Glättung von Konjunkturzyklen, sodass die Kennzahl nicht von üblichen Schwankungen beeinflusst wird. Das CAPE erfährt große Beliebtheit bei der Beurteilung, ob ein Finanzprodukt über- oder unterbewertet ist. Ein hohes CAPE deutet definitionsgemäß auf eine Überbewertung hin und umgekehrt. Natürlich gibt es auch Kritiker der Kennzahl. Siegel (2016) kritisiert, dass die GAAP-Richtlinien in den vergangenen Jahren stark verändert wurden. Somit wurden die Gewinndaten, die in die Berechnung von CAPE einfließen, nicht konsistent bestimmt, sodass das CAPE tendenziell zu einer zu pessimistischen Interpretation führt.

4.2 Anwendung des Datensatzes

Nachdem wir uns für einen Datensatz entschieden haben und auf die genaue Bedeutung seiner enthaltenen Daten eingegangen sind, wollen wir die in Kapitel 3 beschriebenen Maßzahlen an ihm auswerten. Dazu verwenden wir wieder die Softwareumgebung R. Um mit den Daten arbeiten zu können, muss zunächst der Datensatz eingelesen werden, wofür wir die Funktion `read.csv` aus dem Paket `utils` verwenden wollen. Dafür ist es notwendig, dass der Datensatz im einfachen `csv`-Format vorliegt, bei dem die Daten durch einheitliche Zeichen getrennt sind. Sollte dies, wie zum Beispiel beim hier betrachteten `xls`-Datensatz von Shiller (2018), nicht der Fall sein, so muss zunächst das Format geändert werden. Da wir uns stets für den Zeitraum Februar 1871 bis April 2018 interessieren, ist darauf zu achten, dass die Deflationierung bezüglich April 2018 durchgeführt wurde. Andernfalls können die realen Größen leicht gemäß der Formel (4.1) korrigiert werden. Vor der Speicherung des Datensatzes im gewünschten `csv`-Format können wir außerdem in einer weiteren Spalte die Zeitreihe der Erträge nach ihrer Definition (1.1) berechnen.

Um die `csv`-Datei mit R korrekt einlesen zu können, ist es wichtig, die verwendeten Zeichen zur Werte- und Dezimaltrennung zu kennen, um sie als Argumente `sep` und `dec` der Funktion `read.csv` übergeben zu können. In unserem Fall werden hierfür der Strichpunkt und das Komma verwendet. Dies kann kontrolliert werden, indem die gespeicherte `csv`-Datei mit einem einfachen Texteditor geöffnet wird.

Der folgende Befehl in R sorgt nun für das Einlesen der Datei. Dabei muss der Platzhalter `[Dateipfad.csv]` durch den vollständigen Dateipfad inklusive der `csv`-Endung ersetzt werden:

```
data_Shiller <- read.csv('[Dateipfad.csv]', sep = ';', dec = ',',
  , header = FALSE)
```

Anschließend kann für jede notwendige Zeitreihe jeweils ein separater Vektor angelegt werden. Da die Erträge erst ab der zweiten Zeile definiert sind, entfernen wir grundsätzlich das erste Element, sodass alle Zeitreihen dieselbe Länge haben. Außerdem müssen wir die Long Interest Rates durch hundert teilen, weil die Werte von Shiller (2018) in Prozent angegeben sind:

```
prices <- data_Shiller[-1,1]
long_interest_rates <- data_Shiller[-1,6]/100
prices_real <- data_Shiller[-1,7]
dividends_real <- data_Shiller[-1,8]
returns <- data_Shiller[-1,11]
```

Die folgende R-Funktion wertet diejenigen, in Kapitel 3 beschriebenen Maßzahlen aus, die sich auf die vier wichtigsten, stilisierten Fakten beziehen. Dazu zählen keine Autokorrelation in den Erträgen, Volatilitätsclustering, Langzeitgedächtnis und Fat Tails in der Ertragsverteilung:

```
#folgende Funktion wertet die beschriebenen Maßzahlen für eine
Ertragszeitreihe aus
evaluate_moments <- function(data_returns) {
  evaluation_vector <- numeric(length=11)

  #Autokorrelationsfunktion der Erträge für Lag = 1 (
nötiges Paket: stats)
  evaluation_vector[1] <- acf(data_returns, lag=1, type='
  correlation', plot=FALSE)[[1]][2]

  #Autokorrelationsfunktion der Absoluterträge für Lags =
1,3,6,12,24 (nötiges Paket: stats)
  acf_absolut_returns <- acf(abs(data_returns), lag=24,
  type='correlation', plot=FALSE)[[1]]
  evaluation_vector[2] <- acf_absolut_returns[2]
```

```

evaluation_vector[3] <- acf_absolut_returns[4]
evaluation_vector[4] <- acf_absolut_returns[7]
evaluation_vector[5] <- acf_absolut_returns[13]
evaluation_vector[6] <- acf_absolut_returns[25]

#Korrigierter empirischer Hurst Exponent (nötiges Paket:
  pracma)
evaluation_vector[7] <- Inf
evaluation_vector[7] <- try(hurstexp(abs(data_returns),
  display=FALSE)[[4]], silent=TRUE)

#Kurtosis (nötiges Paket: moments)
evaluation_vector[8] <- kurtosis(data_returns)

#Hill-Schätzer bezüglich der 2,5%, 5% und 10% größten,
  gemessenen Erträge (nötiges Paket: laeken)
evaluation_vector[9] <- thetaHill(data_returns, ceiling(
  length(data_returns)*0.025))
evaluation_vector[10] <- thetaHill(data_returns, ceiling(
  length(data_returns)*0.05))
evaluation_vector[11] <- thetaHill(data_returns, ceiling(
  length(data_returns)*0.1))

  evaluation_vector
}

empirical_moments <- evaluate_moments(returns)
#Ergebnis:
#empirical_moments[1] = 0,2823089
#empirical_moments[2] = 0,2064939
#empirical_moments[3] = 0,1295851
#empirical_moments[4] = 0,1476308
#empirical_moments[5] = 0,1484742
#empirical_moments[6] = 0,06014666
#empirical_moments[7] = 0,7171259
#empirical_moments[8] = 20,19594
#empirical_moments[9] = 3,32094
#empirical_moments[10] = 3,630764
#empirical_moments[11] = 2,77868

```

Die in Kommentaren angegebenen Ergebniswerte entsprechen natürlich gerade den genannten Maßzahlenwerten in Kapitel 3, da wir uns auch dort auf den Datensatz von Shiller (2018) im Zeitraum Februar 1871 bis April 2018 bezogen haben. Da

die Funktion `hurstexp` zu Fehlermeldungen führen kann, verwenden wir die R-Funktion `try`. Falls `hurstexp` fehlschlägt, hat die entsprechende Maßzahl den Wert unendlich. Wir haben nun gesehen, wie wir R verwenden können, um Datensätze einzulesen und Maßzahlen an ihnen auszuwerten. In den folgenden Kapiteln werden wir weitere Quellcodes kennenlernen. Diese sind stets als Fortsetzung des hier betrachteten Codes zu verstehen. Insbesondere werden in den folgenden Quellcodes die eingelesenen Zeitreihen sowie der Vektor `empirical_moments` verwendet.

Kapitel 5

Ermittlung des Fundamentalwertes

Wir haben den Begriff des Fundamentalwertes bereits verwendet, ohne ihn genauer zu definieren. Dabei haben wir ihn oberflächlich als einen theoretisch hergeleiteten Wert eines Vermögens beschrieben und die lang anhaltende Abweichung des Preises vom Fundamentalwert als stilisierten Fakt in Kapitel 2 aufgezählt. Wir wollen in diesem Kapitel genauer auf den Fundamentalwert eingehen und eine Möglichkeit betrachten, wie man diesen berechnen kann. Die hier hergeleitete Berechnungsmethode orientiert sich an der Arbeit von Shiller (1981).

5.1 Definitionsansätze für den Fundamentalwert

Der Fundamentalwert oder auch *innere Wert* eines Vermögens ergibt sich, im Gegensatz zu seinem Marktwert, nicht aus Angebot und Nachfrage, sondern aus theoretisch begründeten Überlegungen, wobei verschiedene Ansätze denkbar sind und sich somit auch unterschiedliche Werte ergeben können. Es gibt keine einheitliche Definition oder Berechnungsformel für den Fundamentalwert. Betrachtet man den inneren Wert P_t^* einer Aktie zum Zeitpunkt t , dann lautet die grundlegende Überlegung

$$P_t^* = \frac{\text{Unternehmenswert zum Zeitpunkt } t}{\text{Anzahl ausgegebener Aktien}},$$

aber es ist vollkommen offen, wie genau der Unternehmenswert definiert ist. Denkbar ist beispielsweise die Verwendung von Bilanzwerten. Allerdings besteht hierbei häufig das Problem, dass diese nicht immer realistische Werte darstellen und tendenziell zu einer Unterbewertung des Unternehmenswertes führen. Alternativ könnte man zukünftig erwartete Dividendenzahlungen zur Schätzung des Unternehmenswertes heranziehen, da Dividenden eng mit dem Unternehmensgewinn verknüpft sind. Dabei stellt sich natürlich die Frage, wie Erwartungswerte gebildet werden können. Wir werden in den folgenden zwei Abschnitten auf zwei konkrete,

beliebte Ansätze zur Fundamentalwertdefinition eingehen. Anschließend werden wir einen Ansatz konkretisieren und eine Berechnungsmethode herleiten.

5.1.1 Residual Income Valuation Model

Ein möglicher Ansatz ist die Überlegung, dass sich der innere Wert zum einen aus dem bereits investierten Kapital und zum anderen aus den zukünftig erwarteten Einnahmen abzüglich zukünftig erwarteter Kosten zusammensetzt. Dies lässt sich an dem folgenden Beispiel veranschaulichen.

Wir stellen uns vor, ein Aktienunternehmen hat in der Vergangenheit eine Maschine zu Produktionszwecken gekauft. Die Maschine steht seit dem Kauf als Anlagevermögen in der Bilanz, wobei die Abnutzung durch Abschreibungen berücksichtigt wird, sodass der Bilanzwert einen realistischen, aktuellen Maschinenwert darstellen soll. Der Bilanzwert zum Zeitpunkt t der Maschine sollte also in den Unternehmenswert und damit auch in den inneren Wert der Aktie einfließen. Zum anderen trägt die Maschine innerhalb ihrer Nutzungsdauer als Teil des Produktionsprozesses zur Generierung des Unternehmensgewinns bei. Im inneren Wert sollten also außerdem zukünftig erwartete, durch die Investition in die Maschine verursachte Gewinne berücksichtigt werden. Wir können an dem Beispiel erkennen, dass Investitionen sowohl einen vergangenheitsbezogenen als auch einen zukunftsbezogenen Anteil am inneren Wert verursachen können.

Den beschriebenen Ansatz verfolgt das sogenannte *Residual Income Valuation Model* (vergleiche Lee et al., 1999, S. 1699 ff.). In dem Modell setzt sich der Fundamentalwert aus dem Buchwert je Aktie B_t des Unternehmens zum Zeitpunkt t und dem zukünftig erwarteten Residualeinkommen, d. h. Nettoeinkommen abzüglich Eigenkapitalkosten, zusammen. Der Buchwert basiert auf Werten der Bilanzbuchhaltung. Dort wird je Vermögensgegenstand ein möglichst aktueller Wert ermittelt, indem die Anschaffungs- und Herstellkosten um planmäßige, aber auch außerplanmäßige Abschreibungen verringert und um Zuschreibungen erhöht werden. Mit dem Buchwert und dem Residualeinkommen haben wir also vergangenheits- sowie zukunftsbezogene Summanden in der Berechnung des Fundamentalwertes. Die zukunftsbezogenen Summanden müssen wir natürlich auf den betrachteten Zeitpunkt t abzinsen, um Zinseffekte zu berücksichtigen. Das zukünftig erwartete Residualeinkommen der Periode $t+k$ können wir dazu mit einem entsprechenden Diskontierungsfaktor γ^k multiplizieren, wobei sich γ aus dem Zinssatz i durch

$$\gamma := \frac{1}{1+i}$$

ergibt. Man kann also gemäß dem Residual Income Valuation Model den Funda-

mentalwert einer Aktie zum Zeitpunkt t definieren als Addition des Buchwertes mit dem Gegenwartswert der zukünftig erwarteten Residualeinkommen:

$$P_t^* := B_t + \sum_{k=1}^{\infty} \gamma^k E [\text{Residualeinkommen}_{t+k} | \mathcal{F}_t]$$

Das Modell bezieht sich durch den Buchwert auf Bilanzwerte. Es kann kritisiert werden, dass die dort angegebenen Werte nicht realistisch sind. Aufgrund des *Vorsichtsprinzips* im Rechnungswesen kann es zur Unterbewertung von Aktiva und Überbewertung von Passiva kommen. Damit spiegelt der Buchwert je Aktie nicht unbedingt ein geeignetes Maß für den Wert des bis zum Zeitpunkt t investierten Kapitals wider und deshalb wollen wir dieses Berechnungsverfahren nicht weiter verfolgen.

5.1.2 Dividend Discount Model

Einen anderen Ansatz verfolgt das *Dividend Discount Model*. Hierbei versetzen wir uns in die Lage eines Aktionärs, der die Aktie zum Zeitpunkt t kauft und sie unendlich lange hält. Wir können uns die Frage stellen, welchen Preis der Aktionär für die Aktie zahlen würde. Dieser erlangt durch den Aktienkauf das Recht auf den Erhalt sämtlicher Dividendenausschüttungen ab dem Zeitpunkt t . Es ist also leicht nachvollziehbar, dass der Aktionär bereit ist, den addierten Gegenwartswert sämtlicher, zukünftig erwarteter, realer Dividendenzahlungen zu zahlen, wobei bei der Erwartungsbildung die zum Zeitpunkt t verfügbaren Informationen optimal berücksichtigt werden müssen. Durch die Bildung des Gegenwartswertes werden auch hier Zinseffekte berücksichtigt und durch die Verwendung von realen anstatt nominalen Dividenden werden inflationäre Effekte entkräftet. In Abbildung 5.1 ist das Prinzip der Diskontierung veranschaulicht. Wir nehmen dabei an, dass die Dividenden stets zum Periodenende ausgeschüttet werden. Mit dem Zeitpunkt t ist aber der Anfang der Periode t gemeint. Also müssen wir zur Diskontierung der Dividende aus der Periode $t+k$ mit dem Faktor γ^{k+1} multiplizieren. Das Dividend Discount Model definiert nun den Fundamentalwert gemäß

$$P_t^* := \sum_{k=0}^{\infty} \gamma^{k+1} E [D_{t+k}^{\text{real}} | \mathcal{F}_t], \quad (5.1)$$

wobei $(D_t^{\text{real}})_{t \in \mathbb{N}}$ den stochastischen Prozess der realen Dividenden bezeichnet. Das Modell wurde erstmals von Williams (1938) aufgestellt. Zu beachten ist, dass das Dividend Discount Model offenbar nur für Aktien anwendbar sind, die eine Dividendenausschüttung vorsehen.

In (5.1) ist der Faktor γ konstant, d. h., unabhängig von der betrachteten Periode $t + k$ wird der gleiche Zinssatz i zur Diskontierung verwendet. Alternativ können wir natürlich auch zeitvariable Zinssätze i_t zulassen. Da zukünftige Zinssätze unbekannt sind, bilden wir den Erwartungswert. Mit zeitvariablen Zinssätzen kommen wir auf die folgende Verallgemeinerung der Formel (5.1)

$$P_t^* := \sum_{k=0}^{\infty} E_t [D_{t+k}^{\text{real}}] \prod_{j=0}^k \frac{1}{1 + E_t [i_{t+j}]}, \quad (5.2)$$

wobei wir zur vereinfachten Notation $E_t[\cdot]$ anstatt $E[\cdot|\mathcal{F}_t]$ schreiben.

5.2 Berechnungsmethode von Robert Shiller

Shiller (1981) verwendet in seiner Berechnungsmethode den Ansatz des Dividend Discount Models. Für zukünftige Zeitpunkte sind die Werte $E_t[D_{t+k}^{\text{real}}]$ und $E_t[i_{t+j}]$ natürlich unbekannt, sodass wir den Fundamentalwert nicht exakt ausrechnen können. Stattdessen versuchen wir, eine möglichst gute Approximation zu definieren. Die grundlegende Idee von Shiller (1981) liegt dabei in der *nachträglichen Berechnung* der Fundamentalwerte. Wir stellen uns vor, uns seien Zeitreihen realer Dividenden $(d_t^{\text{real}})_{t=1,\dots,N}$ und Zinssätze $(i_t)_{t=1,\dots,N}$ gegeben. In unserem Fall werden wir wieder den S&P 500-Datensatz für Februar 1871 bis April 2018 von Shiller (2018) betrachten und die dort angegebenen realen Dividenden und Long Interest Rates verwenden. Mithilfe der real-world Zeitreihen werden wir im Nachhinein eine Zeitreihe $(\hat{p}_t^*)_{t=1,\dots,N}$ der geschätzten Fundamentalwerte aufzustellen. Dabei ersetzen wir in der Formel (5.2) die unbekannteren Erwartungswerte $E_t[D_{t+k}^{\text{real}}]$ und $E_t[i_{t+j}]$ durch die tatsächlich beobachteten Dividenden d_{t+k}^{real} und Long Interest Rates i_{t+j} . Außerdem verkürzen wir die unendliche Summe durch eine geeignete Schätzung, sodass alle für die Auswertung benötigten Dividenden- und Zinssatzwerte auch in den Zeitreihen enthalten sind. Damit diese Schätzung Sinn macht, werden wir *trendbereinigte* reale Preise und Dividenden benötigen. Aus diesem Grund werden wir im folgenden Abschnitt zunächst eine Trendbereinigung dieser Zeitreihen durchführen.

5.2.1 Trendbereinigung realer Preise und Dividenden

Vor der Berechnung der geschätzten Fundamentalwertzeitreihe $(\hat{p}_t^*)_{t=1,\dots,N}$ führt Shiller (1981) eine Trendbereinigung der realen Preise und Dividenden durch. Die trendbereinigten Zeitreihen werden wir in Abschnitt 5.2.2 benötigen, um die unendliche Summe geeignet zu verkürzen.

In Abbildung 5.2 bilden die durchgezogenen Linien die realen Preise $(p_t^{\text{real}})_{t=1,\dots,N}$ und die realen Dividenden $(d_t^{\text{real}})_{t=1,\dots,N}$ aus dem S&P 500-Datensatz für Februar 1871 bis April 2018 von Shiller (2018) ab. Wir können jeweils einen durchgehenden, positiven Trend erkennen. Um diesen zu eliminieren, müssen wir natürlich zunächst den Trendtyp ermitteln. Besonders bei der Betrachtung der realen Preise, aber auch bei den realen Dividenden können wir ein exponentielles Wachstum vermuten. Wir stellen also die exponentiellen Zusammenhänge

$$\begin{aligned} P_t^{\text{real}} &\sim P_0 \cdot (1 + g_P)^t, \\ D_t^{\text{real}} &\sim D_0 \cdot (1 + g_D)^t \end{aligned} \tag{5.3}$$

auf, wobei g_P und g_D als *Wachstumsraten* bezeichnet werden. Durch Logarithmieren ergeben sich die linearen Zusammenhänge

$$\begin{aligned} \log(P_t^{\text{real}}) &\sim \log(P_0) + t \cdot \log(1 + g_P), \\ \log(D_t^{\text{real}}) &\sim \log(D_0) + t \cdot \log(1 + g_D). \end{aligned} \tag{5.4}$$

Dabei müssen wir natürlich annehmen, dass P_t^{real} , D_t^{real} , P_0 und D_0 positiv und $g_P, g_D > -1$ sind, um die Wohldefiniertheit sicherzustellen. Diese Annahmen lassen sich leicht rechtfertigen, da Preise und Dividenden stets positiv sein sollten und wir positive Wachstumsraten vermuten. Die Zusammenhänge (5.4) können wir nun mit linearer Regression schätzen. In Abbildung 5.2 entspricht die gestrichelte Linie dem so geschätzten, exponentiellen Zusammenhang. Zum Vergleich ist dort außerdem als gepunktete Linie das Ergebnis einer einfachen, direkten, linearen Regression der realen Preise bzw. Dividenden eingezeichnet. Wir können erkennen, dass die exponentielle Regressionskurve eine bessere Näherung als die lineare Regressionsgerade darstellt. Der Trend scheint also tatsächlich eher exponentieller anstatt linearer Form zu sein. Um dies zu quantifizieren, können wir das Bestimmtheitsmaß R^2 der jeweiligen Regressionen ausrechnen. Als Verhältnis zwischen der erklärten Streuung und der gesamten Streuung dient das Bestimmtheitsmaß als beliebte Kennzahl zur Beurteilung der Güte einer linearen Regression. In Tabelle 5.1 sind die Bestimmtheitsmaße der linearen Regressionen für den linearen sowie für den exponentiellen Zusammenhang der realen Preise und Dividenden angegeben. Wir können sehen, dass der exponentielle Ansatz in beiden Fällen zu einem besseren, d. h. höheren, Bestimmtheitsmaß führt als der lineare Ansatz. Das Ergebnis bestätigt somit unsere Annahme des exponentiellen Wachstums.

Wir haben uns nun davon überzeugt, dass P_t^{real} und D_t^{real} einem exponentiellen Wachstum gemäß (5.3) unterliegen. Für das weitere Vorgehen ist es wichtig, dass wir eine *gemeinsame* Wachstumsrate g für die realen Preise und Dividenden ermitteln und damit die Trendbereinigung durchführen. Den genauen Grund für die

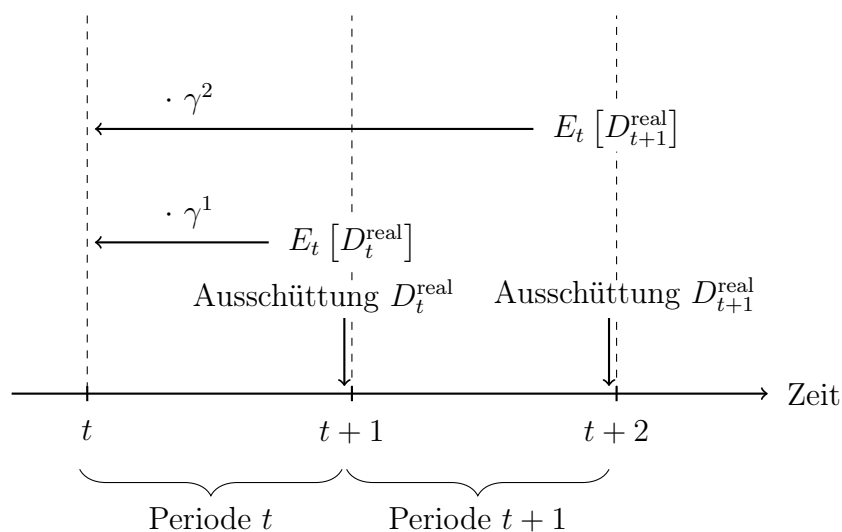


Abbildung 5.1: Zur Diskontierung der zukünftig erwarteten Dividende in Periode $t + k$ muss der Diskontierungsfaktor γ^{k+1} verwendet werden, weil die Dividende D_{t+k}^{real} stets zum Ende der Periode $t + k$ ausgeschüttet wird und mit dem Zeitpunkt t der Anfang der Periode t gemeint ist

Tabelle 5.1: Für die realen Preise und Dividenden aus dem S&P 500-Datensatz von Shiller (2018) zwischen Februar 1871 bis April 2018 wird sowohl der direkte lineare Zusammenhang als auch der exponentielle Zusammenhang aus (5.3) durch eine lineare Regression geschätzt. Für die linearen Regressionen können jeweils die Bestimmtheitsmaße R^2 ausgerechnet werden

Zu schätzender linearer Zusammenhang	R^2
$P_t^{\text{real}} \sim a_P + b_P t$	0,5922143
$\log(P_t^{\text{real}}) \sim \log(P_0) + t \cdot \log(1 + g_P)$	0,7822638
$D_t^{\text{real}} \sim a_D + b_D t$	0,739724
$\log(D_t^{\text{real}}) \sim \log(D_0) + t \cdot \log(1 + g_D)$	0,8518733

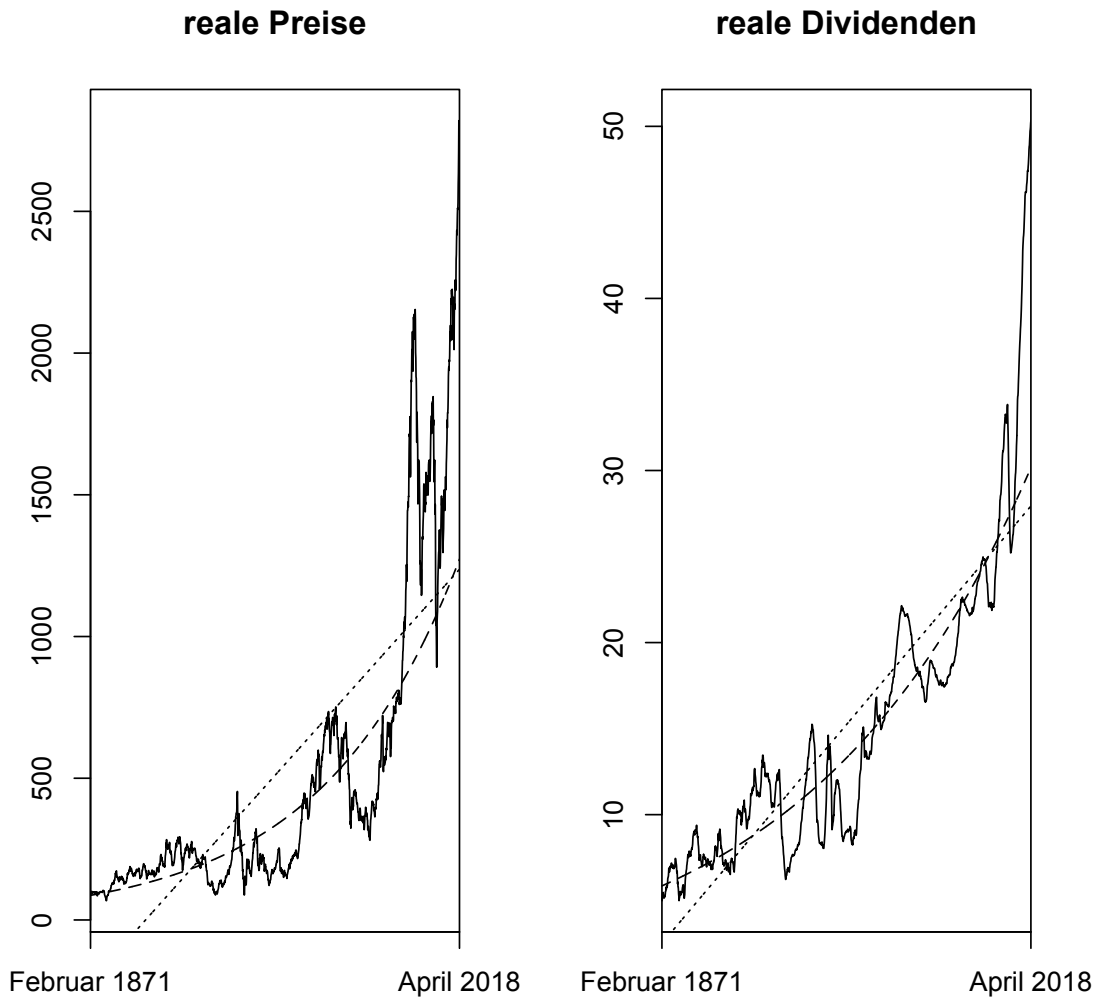


Abbildung 5.2: Reale Preise und Dividenden zwischen Februar 1871 und April 2018 (durchgezogene Linie) und Trendbestimmung durch die lineare Regressionsgerade (gepunktete Linie) und die exponentielle Regressionskurve (gestrichelte Linie), wobei die exponentielle Regressionskurve im Vergleich zur lineare Regressionsgerade den Trend besser abzubilden scheint

Notwendigkeit der gemeinsamen Wachstumsrate werden wir in Abschnitt 5.2.2 erläutern. Shiller (1981) geht nicht genauer darauf ein, wie er diese gemeinsame Wachstumsrate definiert. Deshalb möchte ich an dieser Stelle einen eigenen Ansatz zur möglichst optimalen Definition von g , P_0 und D_0 herleiten.

Wir betrachten die exponentiellen Zusammenhänge aus (5.3), wobei wir die beiden Wachstumsraten g_P und g_D durch eine gemeinsame Wachstumsrate g ersetzen:

$$\begin{aligned} P_t^{\text{real}} &\sim P_0 \cdot (1 + g)^t, \\ D_t^{\text{real}} &\sim D_0 \cdot (1 + g)^t \end{aligned} \tag{5.5}$$

Erneut erhalten wir durch Logarithmieren die linearen Darstellungen

$$\begin{aligned} \log(P_t^{\text{real}}) &\sim \log(P_0) + t \cdot \log(1 + g), \\ \log(D_t^{\text{real}}) &\sim \log(D_0) + t \cdot \log(1 + g). \end{aligned} \tag{5.6}$$

Gesucht sind nun Werte für g , P_0 und D_0 , sodass die Zusammenhänge in einem bestimmten Sinne für N -elementige Zeitreihen möglichst gut erfüllt sind. Ähnlich wie bei einer einfachen linearen Regression schlage ich vor, die Summe der quadrierten Residuen zu minimieren.

Theorem 5.2.1. *Man betrachte die Zusammenhänge (5.6) an N -elementigen Zeitreihen mit $N > 1$. Im Sinne der minimalen Summe der quadrierten Residuen bildet der Punkt (g^*, P_0^*, D_0^*) mit den Werten*

$$\begin{aligned} g^* &:= \exp \left[\frac{\sum_{t=1}^N (-t) \cdot \left(\log(P_t^{\text{real}}) - \overline{\log(P^{\text{real}})} + \log(D_t^{\text{real}}) - \overline{\log(D^{\text{real}})} \right)}{2 \cdot \sum_{t=1}^N t \cdot \left(\frac{N+1}{2} - t \right)} \right] - 1 \\ P_0^* &:= \exp \left[\overline{\log(P^{\text{real}})} - \log(1 + g^*) \cdot \frac{N+1}{2} \right] \\ D_0^* &:= \exp \left[\overline{\log(D^{\text{real}})} - \log(1 + g^*) \cdot \frac{N+1}{2} \right] \end{aligned}$$

einen kritischen Punkt.

Beweis. Wir betrachten die Summe der quadrierten Residuen, d. h. die quadrierten Abweichungen der linken und rechten Seite, für (5.6) und erhalten so das folgende Minimierungsproblem:

$$\begin{aligned} S(g, P_0, D_0) &:= \sum_{t=1}^N \left(\log(P_t^{\text{real}}) - \log(P_0) - t \cdot \log(1 + g) \right)^2 \\ &\quad + \left(\log(D_t^{\text{real}}) - \log(D_0) - t \cdot \log(1 + g) \right)^2 \rightarrow \min_{g, P_0, D_0} \end{aligned}$$

Ein kritischer Punkt ist ein Punkt, der die Bedingungen erster Ordnung erfüllt. Wir stellen fest, dass die Funktion S partiell differenzierbar ist und stellen die Bedingungen erster Ordnung auf:

$$\begin{aligned}
\text{(I)} \quad \frac{\delta S}{\delta g}(g, P_0, D_0) &= \sum_{t=1}^N 2 \cdot (\log(P_t^{\text{real}}) - \log(P_0) - t \cdot \log(1+g)) \cdot \left(\frac{-t}{1+g}\right) \\
&\quad + 2 \cdot (\log(D_t^{\text{real}}) - \log(D_0) - t \cdot \log(1+g)) \cdot \left(\frac{-t}{1+g}\right) \stackrel{!}{=} 0 \\
\text{(II)} \quad \frac{\delta S}{\delta P_0}(g, P_0, D_0) &= \sum_{t=1}^N 2 \cdot (\log(P_t^{\text{real}}) - \log(P_0) - t \cdot \log(1+g)) \cdot \left(-\frac{1}{P_0}\right) \stackrel{!}{=} 0 \\
\text{(III)} \quad \frac{\delta S}{\delta D_0}(g, P_0, D_0) &= \sum_{t=1}^N 2 \cdot (\log(D_t^{\text{real}}) - \log(D_0) - t \cdot \log(1+g)) \cdot \left(-\frac{1}{D_0}\right) \stackrel{!}{=} 0
\end{aligned}$$

Die Bedingung (II) können wir auflösen, sodass wir eine Formel für P_0 in Abhängigkeit von g erhalten

$$\begin{aligned}
\text{(II)} \Leftrightarrow \sum_{t=1}^N (\log(P_t^{\text{real}}) - \log(P_0) - t \cdot \log(1+g)) &= 0 \\
\Leftrightarrow \sum_{t=1}^N \log(P_t^{\text{real}}) - N \cdot \log(P_0) - \log(1+g) \cdot \sum_{t=1}^N t &= 0 \\
\Leftrightarrow \log(P_0) = \overline{\log(P^{\text{real}})} - \log(1+g) \cdot \frac{N+1}{2} &\quad (5.7)
\end{aligned}$$

$$\Leftrightarrow P_0 = \exp \left[\overline{\log(P^{\text{real}})} - \log(1+g) \cdot \frac{N+1}{2} \right], \quad (5.8)$$

wobei wir bei der vorletzten Umformung die Gauß'sche Summenformel verwenden. Vollkommen analog erhalten wir eine Formel für D_0 aus der Bedingung (III):

$$\text{(I)} \Leftrightarrow \log(D_0) = \overline{\log(D^{\text{real}})} - \log(1+g) \cdot \frac{N+1}{2} \quad (5.9)$$

$$\Leftrightarrow D_0 = \exp \left[\overline{\log(D^{\text{real}})} - \log(1+g) \cdot \frac{N+1}{2} \right] \quad (5.10)$$

Nun können wir die Bedingung (I) betrachten und die hergeleiteten Formeln (5.7)

und (5.9) einsetzen:

$$\begin{aligned}
(I) &\Leftrightarrow \sum_{t=1}^N t \cdot \left(\log(P_t^{\text{real}}) - \log(P_0) - t \cdot \log(1+g) \right. \\
&\quad \left. + \log(D_t^{\text{real}}) - \log(D_0) - t \cdot \log(1+g) \right) = 0 \\
&\Leftrightarrow \sum_{t=1}^N t \cdot \left(\log(P_t^{\text{real}}) - \left(\overline{\log(P^{\text{real}})} - \log(1+g) \cdot \frac{N+1}{2} \right) - t \cdot \log(1+g) \right. \\
&\quad \left. + \log(D_t^{\text{real}}) - \left(\overline{\log(D^{\text{real}})} - \log(1+g) \cdot \frac{N+1}{2} \right) - t \cdot \log(1+g) \right) = 0 \\
&\Leftrightarrow \sum_{t=1}^N t \cdot \left(\log(P_t^{\text{real}}) - \overline{\log(P^{\text{real}})} + \log(D_t^{\text{real}}) - \overline{\log(D^{\text{real}})} \right) \\
&\quad + 2 \cdot \log(1+g) \cdot \sum_{t=1}^N t \cdot \left(\frac{N+1}{2} - t \right) = 0 \\
\stackrel{N \geq 1}{\Leftrightarrow} \log(1+g) &= \frac{\sum_{t=1}^N (-t) \cdot \left(\log(P_t^{\text{real}}) - \overline{\log(P^{\text{real}})} + \log(D_t^{\text{real}}) - \overline{\log(D^{\text{real}})} \right)}{2 \cdot \sum_{t=1}^N t \cdot \left(\frac{N+1}{2} - t \right)} \\
\Leftrightarrow g &= \exp \left[\frac{\sum_{t=1}^N (-t) \cdot \left(\log(P_t^{\text{real}}) - \overline{\log(P^{\text{real}})} + \log(D_t^{\text{real}}) - \overline{\log(D^{\text{real}})} \right)}{2 \cdot \sum_{t=1}^N t \cdot \left(\frac{N+1}{2} - t \right)} \right] - 1 \quad (5.11)
\end{aligned}$$

Dabei müssen wir bei der vorletzten Äquivalenz annehmen, dass N ungleich und damit echt größer als eins ist, um die Wohldefiniertheit sicherzustellen:

$$\begin{aligned}
&2 \cdot \sum_{t=1}^N t \cdot \left(\frac{N+1}{2} - t \right) \stackrel{?}{=} 0 \\
&\Leftrightarrow \frac{N+1}{2} \sum_{t=1}^N t - \sum_{t=1}^N t^2 = 0 \\
&\Leftrightarrow \frac{N+1}{2} \cdot \frac{N \cdot (N+1)}{2} - \frac{N \cdot (N+1) \cdot (2N+1)}{6} = 0 \\
&\Leftrightarrow \frac{3 \cdot N \cdot (N+1)^2 - 2 \cdot N \cdot (N+1) \cdot (2N+1)}{12} = 0 \\
&\Leftrightarrow N \cdot (N+1) \cdot [3 \cdot (N+1) - 2 \cdot (2N+1)] = 0 \\
&\stackrel{N \geq 0}{\Leftrightarrow} 3 \cdot (N+1) - 2 \cdot (2N+1) = 0 \\
&\Leftrightarrow -N+1 = 0 \\
&\Leftrightarrow N = 1
\end{aligned}$$

Bei der zweiten Äquivalenz verwenden wir wieder die Gauß'sche Summenformel sowie die verallgemeinerte Summenformel für quadrierte Summanden (vergleiche

Abramowitz und Stegun, 1972, S. 804, Formeln 23.1.2 - 23.1.4). Da N die Länge der real-world Zeitreihen $(p_t^{\text{real}})_{t=1,\dots,N}$ und $(d_t^{\text{real}})_{t=1,\dots,N}$ bezeichnet und diese natürlich nicht nur aus einem Zeitpunkt bestehen sollten, können wir die Annahme $N > 1$ problemlos treffen.

Damit lösen also die im Theorem 5.2.1 angegebenen Formeln für g^* , P_0^* und D_0^* die Bedingungen erster Ordnung unseres Minimierungsproblems. Es handelt sich somit um einen kritischen Punkt. \square

Wir können nun das Theorem 5.2.1 auf den S&P 500-Datensatz von Shiller (2018) von Februar 1871 bis April 2018 anwenden. Dies liefert den kritischen Punkt:

$$\begin{aligned} g^* &= 0,001224568 \\ P_0^* &= 112,7469 \\ D_0^* &= 4,506226 \end{aligned} \tag{5.12}$$

Beim gefundenen, kritischen Punkt handelt es sich tatsächlich um ein strikt lokales Minimum der Funktion $S(g, P_0, D_0)$. Dies lässt sich zeigen, indem die Hessematrix an (g^*, P_0^*, D_0^*) aufgestellt und ihre Definitheit untersucht wird. Nach dem Satz von Schwarz ist die Hessematrix symmetrisch, da S offenbar zweimal stetig differenzierbar ist. Damit ist das Hauptminorenkriterium anwendbar und wir können die Definitheit untersuchen, indem die drei Hauptminoren berechnet werden. Sind sie alle positiv, dann folgt, dass die Hessematrix am betrachteten, kritischen Punkt positiv definit ist und es sich somit um ein strikt lokales Minimum handelt. Die Hauptminoren lauten in unserem Fall

1. Hauptminor = 7.344.360.288 > 0
2. Hauptminor = 1.275.902.488 > 0
3. Hauptminor = 88.761.196.421 > 0

und damit hat unsere Zielfunktion S am Punkt (g^*, P_0^*, D_0^*) ein strikt lokales Minimum. Die Hauptminoren wurden mit dem am Ende dieses Abschnitts angegebenen R-Quellcode berechnet.

Abbildung 5.3 zeigt als durchgezogene Linien die realen Preise und Dividenden zwischen Februar 1871 und April 2018 sowie als gestrichelte Linien die exponentielle Regressionskurven aus (5.5) mit den berechneten Werten aus (5.12) für g , P_0 und D_0 . Wir können feststellen, dass die Regressionskurven den jeweiligen Trend gut abzubilden scheinen.

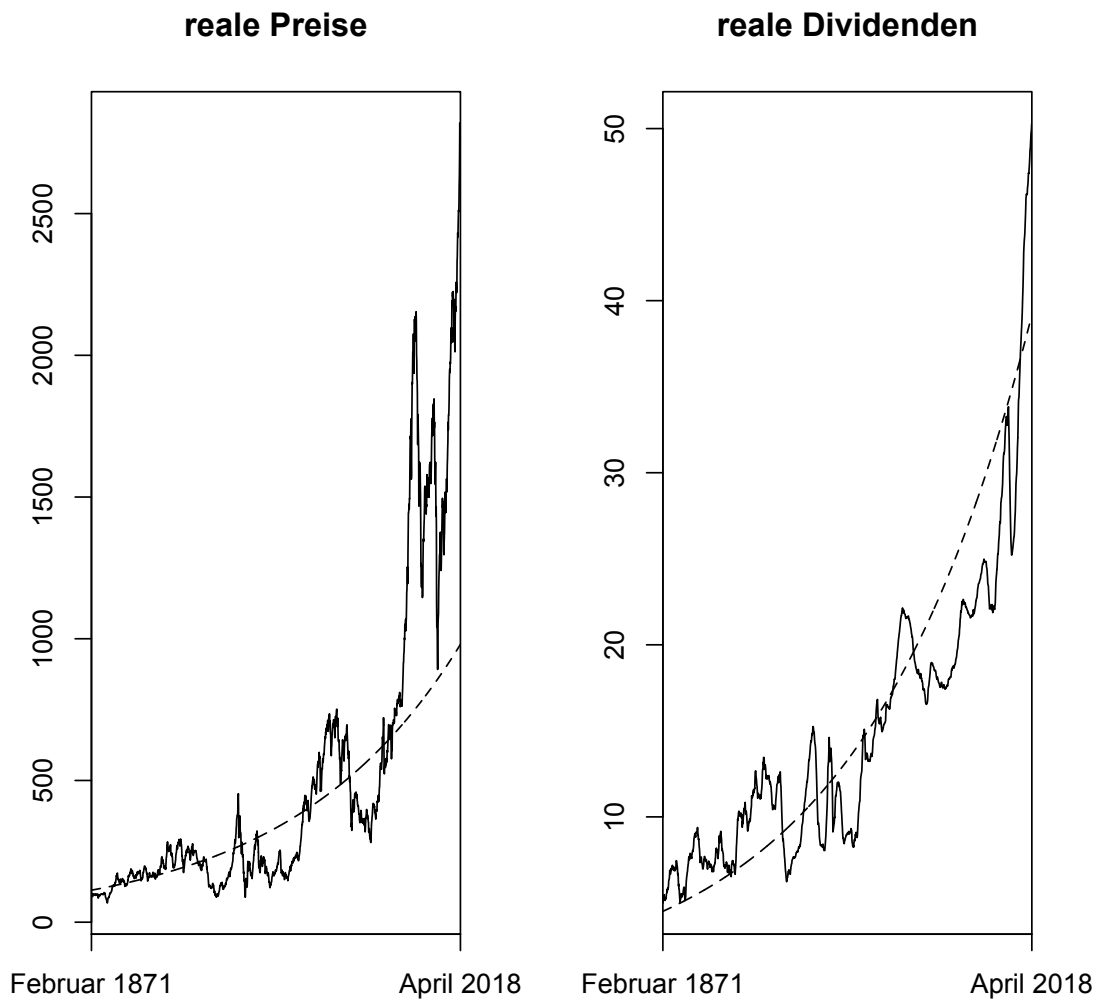


Abbildung 5.3: Reale Preise und Dividenden zwischen Februar 1871 bis April 2018 (durchgezogene Linie) und Trendbestimmung durch die exponentielle Regressionskurve mit der berechneten, optimalen, gemeinsamen Wachstumsrate g^* und den berechneten, optimalen Werten P_0^* und D_0^* aus (5.12) (gestrichelte Linie)

Mithilfe der gemeinsamen Wachstumsrate g können wir nun die Trendbereinigung durchführen. Die neuen Zeitreihen $(p_t^{\text{r,d}})_{t=1,\dots,N}$ und $(d_t^{\text{r,d}})_{t=1,\dots,N}$ definiert Shiller (1981) durch

$$p_t^{\text{r,d}} := p_t^{\text{real,detrended}} := \frac{p_t^{\text{real}}}{(1+g)^{t-N}} = p_t^{\text{real}} \cdot (1+g)^{N-t},$$

$$d_t^{\text{r,d}} := d_t^{\text{real,detrended}} := \frac{d_t^{\text{real}}}{(1+g)^{t+1-N}} = d_t^{\text{real}} \cdot (1+g)^{N-t-1}, \quad t = 1, \dots, N.$$

Durch die Multiplikation mit $(1+g)^{N-t}$ berücksichtigen wir das trendbedingte, exponentielle Wachstum zwischen den Zeitpunkten t und N . Damit beziehen sich die trendbereinigten, realen Preise einheitlich auf das Niveau zum Zeitpunkt N , sodass Trendeffekte in der Zeitreihe $(p_t^{\text{r,d}})_{t=1,\dots,N}$ eliminiert sind. Bei den realen Dividenden gehen wir analog vor, wobei sich die Dividende d_t^{real} auf das *Ende* der Periode t bezieht und wir deshalb den Exponenten anpassen müssen.

Der folgende R-Quellcode berechnet den kritischen Punkt (g^*, P_0^*, D_0^*) gemäß Theorem 5.2.1, erzeugt Abbildung 5.3, berechnet die Hessematrix und ihre drei Hauptminoren und führt die Trendbereinigung mit der berechneten, gemeinsamen Wachstumsrate durch. Dabei bezeichnen `prices_real` und `dividends_real` die eingelesenen Zeitreihen der realen Preise und Dividenden aus dem S&P 500-Datensatz im Zeitraum Februar 1871 bis April 2018 von Shiller (2018).

```
N <- length(prices_real)
x <- c(1:N)

#Berechnung von g
g_numerator <- 0
g_denominator <- 0
mean_log_prices_real <- mean(log(prices_real))
mean_log_dividends_real <- mean(log(dividends_real))
mean_x <- (N+1)/2
for(t in 1:N) {
  g_numerator <- g_numerator + (-t) * (log(prices_real[t])
    - mean_log_prices_real + log(dividends_real[t]) -
    mean_log_dividends_real)
  g_denominator <- g_denominator + (2*t) * (mean_x-t)
}
g <- exp(g_numerator/g_denominator) - 1

#Berechnung von P0 und D0
P0 <- exp(mean_log_prices_real - log(1+g) * mean_x)
```

```

D0 <- exp(mean_log_dividends_real - log(1+g) * mean_x)

#Erstellung der Abbildung der realen Preise und Dividenden und
  der jeweiligen exponentiellen Regressionskurve mit
  gemeinsamer Wachstumsrate g
par(mfrow = c(1,2))

plot(x, prices_real, type = 'l', main = 'reale_Preise', ylab = '
', xlab = '', xaxt = 'n', xaxs = 'i')
x_text <- c('Februar_1871', 'April_2018')
axis(1, at = c(1,N), labels = x_text, col.axis = 'black')
lines(x, P0 * (1+g)^x, type = 'l', lty = 'dashed')

plot(x, dividends_real, type = 'l', main = 'reale_Dividenden',
      ylab = '', xlab = 'l', xaxt = 'n', xaxs = 'i')
axis(1, at = c(1,N), labels = x_text, col.axis = 'black')
lines(x, D0 * (1+g)^x, type = 'l', lty = 'dashed')

#Berechnung der Hesse-Matrix
sum_log_prices <- sum(log(prices_real))
sum_log_dividends <- sum(log(dividends_real))
sum_t <- N * (N+1)/2
sum_t_square <- N * (N+1) * (2*N+1)/6
sum_t_times_log_prices <- sum(x * log(prices_real))
sum_t_times_log_dividends <- sum(x * log(dividends_real))
A <- (2/((1+g)^2)) * (sum_t_times_log_prices - log(P0) * sum_t -
  2 * log(1+g) * sum_t_square + sum_t_times_log_dividends -
  log(D0) * sum_t) + (4/((1+g)^2)) * sum_t_square
B <- (2/(P0 * (1+g))) * sum_t
C <- (2/(D0 * (1+g))) * sum_t
D <- B
E <- (2/(P0^2)) * (sum_log_prices - N * log(P0) - log(1+g) * sum
  _t) + ((2*N)/(P0^2))
F <- 0
G <- C
H <- F
I <- (2/(D0^2)) * (sum_log_dividends - N * log(D0) - log(1+g) *
  sum_t) + ((2*N)/(D0^2))

#Die Hesse-Matrix ist symmetrisch
Hesse_matrix <- matrix(c(A,B,C,D,E,F,G,H,I), byrow = TRUE, nrow
  = 3)

```

```

#erster Hauptminor
A
#zweiter Hauptminor
A * E - D * B
#dritter Hauptminor (Regel von Sarrus)
A * E * I + B * F * G + C * D * H - G * E * C - H * F * A - I *
  D * B

#Trendbereinigung
prices_real_detrended <- vector(length = N)
dividends_real_detrended <- vector(length = N)
lambda <- 1 + g
for(t in 1:N) {
  prices_real_detrended[t] <- prices_real[t] * lambda^(N-t)
  dividends_real_detrended[t] <- dividends_real[t] *
    lambda^(N-t-1)
}

```

Auch hier gilt wieder, dass alle folgenden Quellcodes als Fortsetzung des hier angegebenen Codes zu sehen sind. Mit `prices_real_detrended` und `dividends_real_detrended` sind also stets die hier berechneten Zeitreihen gemeint und `N` steht für die Länge unserer eingelesener Zeitreihen.

5.2.2 Berechnung geschätzter Fundamentalwerte

Nach der Trendbereinigung können wir nun die Zeitreihe $(\hat{p}_t^*)_{t=1,\dots,N}$ der geschätzten Fundamentalwerte aufstellen. Dazu ersetzen wir in der Formel (5.2) des Dividend Discount Models mit zeitvariablen Zinssätzen die Erwartungswerte $E_t[D_{t+k}^{\text{real}}]$ und $E_t[i_{t+j}]$ durch die Realisationen $d_{t+k}^{\text{r,d}}$ und i_{t+j} . Dabei machen wir natürlich einen Fehler, denn die Erwartungswerte in (5.2) bezieht sich ausschließlich auf die zum Zeitpunkt t verfügbaren Informationen. Die tatsächlichen Realisationen werden dagegen zur jeweiligen Periode festgelegt und berücksichtigen somit auch Informationen, die erst nach dem Zeitpunkt t bekannt wurden. Bei der folgenden Umformung soll das Rundungszeichen \approx den beschriebenen Fehler verdeutlichen und die Kleinschreibung von p_t^* steht wie gewohnt für den Übergang auf Realisationen:

$$\begin{aligned}
p_t^* &\approx \sum_{k=0}^{\infty} d_{t+k}^{r,d} \prod_{j=0}^k \frac{1}{1+i_{t+j}} \\
&= \frac{1}{1+i_t} \cdot \left(\sum_{k=0}^{\infty} d_{t+k}^{r,d} \prod_{j=1}^k \frac{1}{1+i_{t+j}} \right) \\
&= \frac{1}{1+i_t} \cdot \left(d_t^{r,d} + \sum_{k=1}^{\infty} d_{t+k}^{r,d} \prod_{j=1}^k \frac{1}{1+i_{t+j}} \right) \\
&= \frac{1}{1+i_t} \cdot \left(d_t^{r,d} + \sum_{k=0}^{\infty} d_{t+1+k}^{r,d} \prod_{j=0}^k \frac{1}{1+i_{t+1+j}} \right) \\
&\approx \frac{1}{1+i_t} \cdot \left(d_t^{r,d} + p_{t+1}^* \right)
\end{aligned}$$

Die Umformung liefert uns die folgende Idee. Wenn wir uns eine geeignete Schätzung \hat{p}_N^* für p_N^* definieren, können wir schrittweise für $t = N-1, N-2, \dots, 1$ geschätzte Fundamentalwerte \hat{p}_t^* berechnen, indem wir die rekursive Gleichung

$$\hat{p}_t^* := \frac{1}{1+i_t} \cdot \left(d_t^{r,d} + \hat{p}_{t+1}^* \right) \quad (5.13)$$

auswerten. Shiller (1981) schlägt die Schätzung

$$\hat{p}_N^* := \frac{1}{N} \sum_{t=1}^N p_t^{r,d} \quad (5.14)$$

vor. An dieser Stelle lässt sich erklären, warum wir im vorherigen Abschnitt 5.2.1 die gemeinsame Trendbereinigung der realen Preise und Dividenden durchgeführt haben. Die Bildung des Mittelwertes in (5.14) ist natürlich nur sinnvoll, wenn sich die realen Preise alle auf dasselbe Niveau beziehen, was wir durch die vorherige Trendbereinigung erreichen. Dies soll in Abbildung 5.4 veranschaulicht werden. Dort sind links die realen S&P 500-Preise zwischen Februar 1871 und April 2018 ohne Trendbereinigung und rechts mit Trendbereinigung als durchgezogene Linie eingezeichnet. Die gestrichelte Linie stellt jeweils den Mittelwert der Zeitreihe dar. Es ist offensichtlich, dass der Mittelwert ohne vorheriger Trendbereinigung keine gute Schätzung für p_N^* ergibt. Die Trendbereinigung der realen Preise ist also notwendig, damit die Formel (5.14) sinnvoll ist. Es bleibt noch zu klären, warum wir eine *gemeinsame* Trendbereinigung der realen Preise und Dividenden benötigen. Das liegt an der rekursiven Vorgehensweise in Formel (5.13). Diese führt dazu, dass

\hat{p}_N^* in die restlichen, berechneten Werte \hat{p}_t^* , $t = N - 1, \dots, 1$ eingeht. Da in Formel (5.13) die realen und trendbereinigten Dividenden vorkommen, sind die Werte \hat{p}_t^* , $t = N - 1, \dots, 1$ also sowohl von den realen, trendbereinigten Preisen als auch Dividenden abhängig. Die Verrechnung der beiden trendbereinigten Zeitreihen ist natürlich nur dann sinnvoll, wenn die Trendbereinigung mit derselben Wachstumsrate durchgeführt wurde.

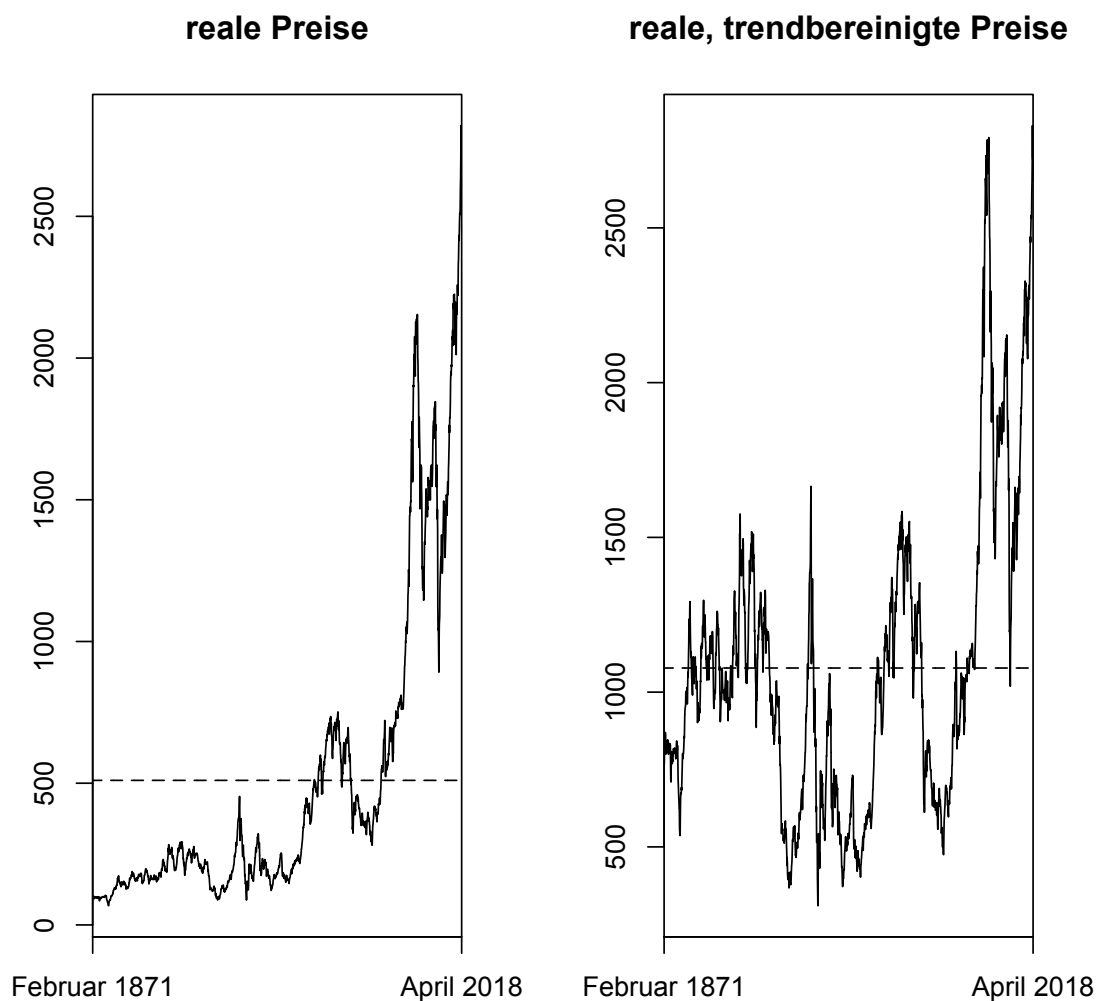


Abbildung 5.4: Die durchgezogenen Linien stellen die realen S&P 500-Preise zwischen Februar 1871 und April 2018 ohne (links) bzw. mit (rechts) Trendbereinigung dar. Der jeweilige Mittelwert ist durch die gestrichelte Linie eingezeichnet. Offenbar dient der Mittelwert nur dann als sinnvolle Schätzung für p_N^* , wenn er sich auf die trendbereinigten Preise bezieht

Theorem 5.2.2. *Aus der rekursiven Gleichung (5.13) folgt die Formel*

$$\hat{p}_t^* = \sum_{k=0}^{N-t-1} \left(d_{t+k}^{r,d} \prod_{j=0}^k \frac{1}{1+i_{t+j}} \right) + \left(\prod_{j=0}^{N-t-1} \frac{1}{1+i_{t+j}} \right) \cdot \hat{p}_N^* \quad \forall t = 1, \dots, N. \quad (5.15)$$

Beweis. Die Formel (5.15) lässt sich per Induktion über $t = N, N-1, \dots, 1$ beweisen. Wir kürzen die rechte Seite der Formel durch RS ab.

Induktionsanfang $t = N$: Es ergeben sich leere Summen und Produkte und wir erhalten

$$RS = 0 + 1 \cdot \hat{p}_N^* = \hat{p}_N^*.$$

Induktionsvoraussetzung (IV): Die Formel (5.15) gelte für $t > 1$.

Induktionsschritt $t \rightarrow t-1$: Nach geschickter Ausklammerung und Verschiebung der Summen- und Produktgrenzen können wir die Induktionsvoraussetzung und die rekursive Gleichung (5.13) verwenden und erhalten das gewünschte Ergebnis:

$$\begin{aligned} RS &= \sum_{k=0}^{N-(t-1)-1} \left(d_{(t-1)+k}^{r,d} \prod_{j=0}^k \frac{1}{1+i_{(t-1)+j}} \right) + \left(\prod_{j=0}^{N-(t-1)-1} \frac{1}{1+i_{(t-1)+j}} \right) \cdot \hat{p}_N^* \\ &= \frac{1}{1+i_{t-1}} \cdot \left[d_{t-1}^{r,d} + \sum_{k=1}^{N-t} \left(d_{t-1+k}^{r,d} \prod_{j=1}^k \frac{1}{1+i_{t-1+j}} \right) + \left(\prod_{j=1}^{N-t} \frac{1}{1+i_{t-1+j}} \right) \cdot \hat{p}_N^* \right] \\ &= \frac{1}{1+i_{t-1}} \cdot \left[d_{t-1}^{r,d} + \sum_{k=0}^{N-t-1} \left(d_{t+k}^{r,d} \prod_{j=0}^k \frac{1}{1+i_{t+j}} \right) + \left(\prod_{j=0}^{N-t-1} \frac{1}{1+i_{t+j}} \right) \cdot \hat{p}_N^* \right] \\ &\stackrel{(IV)}{=} \frac{1}{1+i_{t-1}} \cdot \left[d_{t-1}^{r,d} + \hat{p}_t^* \right] \\ &\stackrel{(5.13)}{=} \hat{p}_{t-1}^* \end{aligned}$$

□

An der nun bewiesenen Formel (5.15) können wir sofort erkennen, dass mit wachsender Länge N der real-world Zeitreihe der geschätzte Wert \hat{p}_N^* für frühe Zeitpunkte t immer öfter abgezinst wird und damit eine umso geringere Rolle für den Wert \hat{p}_t^* spielt. Wir sollten also darauf achten, dass die gewählten real-world Zeitreihen genug Perioden enthalten, sodass die Wahl von \hat{p}_N^* keinen zu großen Einfluss auf die geschätzten Fundamentalwerte hat.

Als Fortsetzung des Quellcodes aus Abschnitt 5.2.1 berechnet der folgende R-Quellcode die geschätzten Fundamentalwerte und erzeugt Abbildung 5.5. Da es sich bei den eingelesenen Long Interest Rates um jährliche Zinssätze handelt und wir mit einer monatlichen Zeitskala arbeiten, müssen wir den Exponent 1/12 einfügen. Darüber hinaus ist zu beachten, dass Shiller (2018) Dividenden mit jährlichem Bezug angibt und wir sie somit durch zwölf teilen müssen.

```
#Berechnung der geschätzten Fundamentalwerte, wobei  
berücksichtigt werden muss, dass die Zinssätze und  
Dividendenwerte jährliche Werte sind  
p_hat_star <- vector(length = N)  
p_hat_star[N] <- mean(prices_real_detrended)  
for(t in (N-1):1) {  
  p_hat_star[t] <- (1/(1 + long_interest_rates[t]))^(1/12)  
    * (dividends_real_detrended[t]/12 + p_hat_star[t  
    +1])  
}  
  
#Erzeugung der Abbildung  
plot(x, prices_real_detrended, type = 'l', ylab = '', xlab = '',  
      xaxt = 'n', xaxs = 'i')  
x_text <- c('Februar_1871', 'April_2018')  
axis(1, at = c(1,N), labels = x_text, col.axis = 'black')  
lines(x, p_hat_star, type = 'l', lty = 'dashed')
```

In der erzeugten Abbildung 5.5 sind die trendbereinigten, realen Preise sowie die berechneten, geschätzten Fundamentalwerte für den genannten Zeitraum eingezeichnet. Wir können wie erwartet feststellen, dass die Preise zum Teil lang anhaltend vom Fundamentalwert abweichen, aber immer wieder zum Fundamentalwertniveau zurückkehren.

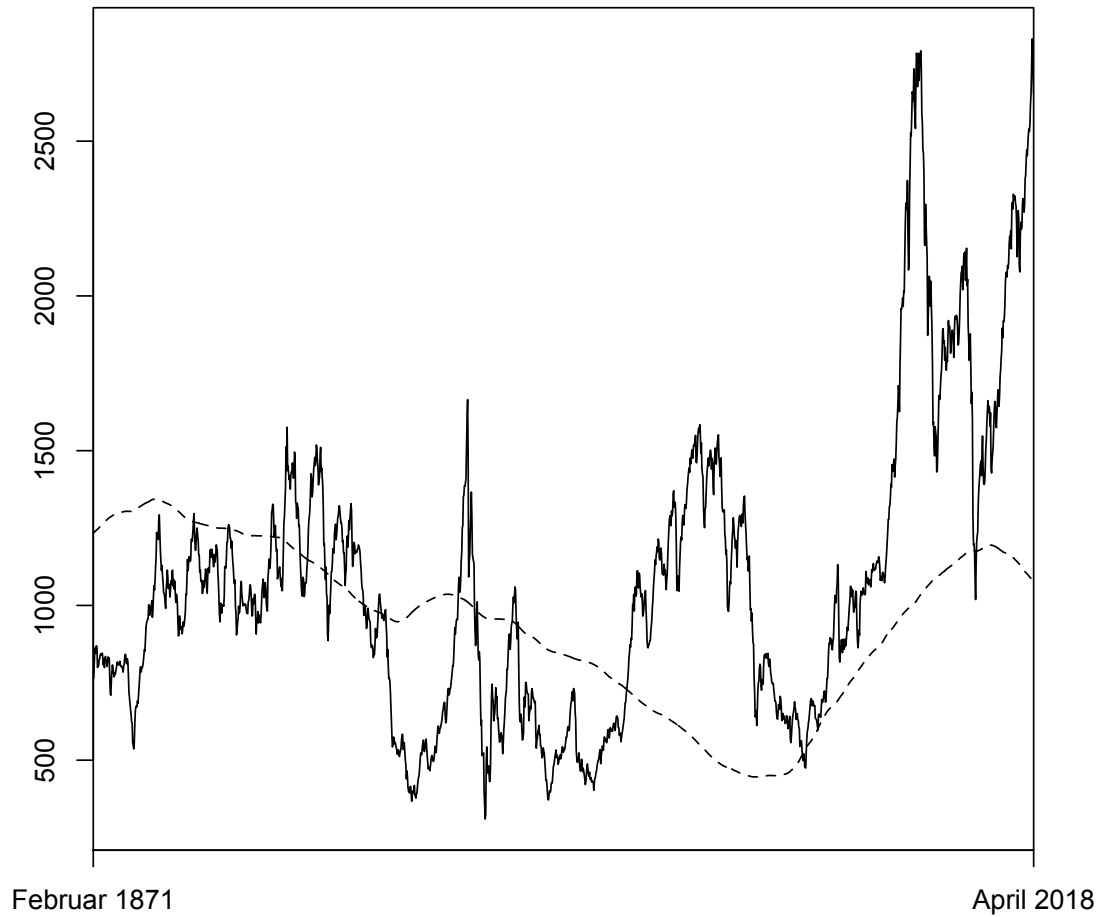


Abbildung 5.5: Die trendbereinigten, realen Preise (durchgezogene Linie) weichen im Zeitraum Februar 1871 bis April 2018 oft über lange Zeit vom geschätzten Fundamentalwert (gestrichelte Linie) ab, kehren jedoch immer wieder zum Fundamentalwertniveau zurück

Kapitel 6

Bootstrappingansatz

In der Realität ist man häufig mit dem Problem konfrontiert, dass nur *eine* real-world Zeitreihe verfügbar ist. Beispielsweise können wir nur ein einziges Mal die tatsächliche Realisation des Preisbildungsprozesses für den S&P 500-Aktienindex zwischen Februar 1871 und April 2018 messen. Uns ist stets unbekannt, welche Preise entstanden wären, wenn sich die zufälligen Einflussfaktoren, wie preisbeeinflussende Informationen, anders entwickelt hätten. Schließlich kann das Zufallsexperiment nicht wiederholt werden.

Das beschriebene Problem haben wir auch bei der Auswertung der Maßzahlen der stilisierten Fakten. Wir können diese jeweils nur einmal an der real-world Zeitreihe auswerten. Hätten wir dagegen die Möglichkeit, die Maßzahl mehrmals an unterschiedlichen Realisationen auszuwerten, so könnten wir den wahren Wert der Maßzahl im real-world DGP präziser schätzen. Genau an diesen Wunsch knüpft die *Bootstrappingmethode* an.

6.1 Bootstrappingmethode von Bradley Efron

Die Bootstrappingidee wurde erstmals von Efron (1979) entwickelt. Es handelt sich um eine *Resampling*-Methode, d. h., eine Möglichkeit, wie wir aus einer Zeitreihe durch wiederholtes Ziehen mit Zurücklegen mehrere Zeitreihen derselben Länge konstruieren können. Efron (1979) betrachtet dabei die Situation, dass X_1, \dots, X_N stochastisch unabhängige, identisch verteilte Zufallsvariablen mit gemeinsamer Verteilungsfunktion F sind und durch $x = (x_1, \dots, x_N)$ Realisationen gegeben seien. Wir interessieren uns für die Verteilung von einer Statistik $T(X, F)$, die von allen Zufallsvariablen $X = (X_1, \dots, X_N)$ sowie der unbekanntem Verteilungsfunktion F abhängen kann. Als Hilfsmittel bei der Schätzung der Verteilung von T bildet Efron (1979) nun sogenannte Bootstrappingzeitreihen.

Zur Konstruktion einer Bootstrappingzeitreihe $x^{(b)}$ wird aus der Menge $\{x_1, \dots, x_N\}$ N -Mal mit Zurücklegen gezogen. Mit anderen Worten: Efron (1979) stellt die diskrete Gleichverteilung mit den Ergebniswerten $\{x_1, \dots, x_N\}$ auf und bildet N stochastisch unabhängige Realisationen. Ist \hat{F} die zugehörige Verteilungsfunktion dieser diskreten Gleichverteilung, dann können wir nach der Konstruktion der Bootstrappingzeitreihe den Wert $T(x^{(b)}, \hat{F})$ berechnen. Der große Vorteil ist nun, dass wir diese Konstruktion beliebig oft wiederholen können. Wir bilden also B Bootstrappingzeitreihen $x^{(1)}, \dots, x^{(B)}$, berechnen anschließend die Werte $T(x^{(b)}, \hat{F})$ für $b = 1, \dots, B$ und nutzen diese zur Schätzung der Verteilung von T . Die Bootstrappingmethode ist somit ein Verfahren, mit dem weitere Pseudorealisationen des real-world DGP generiert werden können.

Nun stellt sich natürlich die Frage, wann das beschriebene Vorgehen sinnvoll ist. Schließlich muss die Verteilungsfunktion \hat{F} der Bootstrappingzeitreihen nicht mit der Verteilungsfunktion F der real-world Zeitreihe übereinstimmen und auch bei der anschließenden Schätzung der Verteilung von T durch die Werte $T(x^{(b)}, \hat{F})$, $b = 1, \dots, B$ entstehen Approximationsfehler. Die theoretische Fehlerabschätzung der Methode ist Thema zahlreicher Literatur (vergleiche z. B. Mammen, 1992).

6.2 Methode und Anwendung von Block Bootstrapping

Wie bereits erwähnt, werden wir den Bootstrappingansatz verwenden, um die in Kapitel 3 beschriebenen Maßzahlen für die stilisierten Fakten zu untersuchen. Genauer interessieren wir uns für die Varianz-Kovarianz-Matrix der Maßzahlen, da wir diese in Abschnitt 7.2 benötigen werden, um die Güte von Modellen zu bewerten.

Sei $m = (m_1, \dots, m_k)$ der Vektor, der die betrachteten Maßzahlen als Komponenten enthält. Dabei können wir alle oder auch nur einen Teil der in Kapitel 3 beschriebenen Kennzahlen wählen. Ich werde mich hier wieder auf die Maßzahlen der vier wichtigsten, stilisierten Fakten keine Autokorrelation in den Erträgen, Volatilitätsclustering, Langzeitgedächtnis und Fat Tails in der Ertragsverteilung einschränken und die in den Abschnitten 3.1 bis 3.5 beschriebenen Kennzahlen verwenden:

- Autokorrelationsfunktion der Erträge für Lag = 1
- Autokorrelationsfunktion der Absoluterträge für Lags = 1, 3, 6, 12, 24
- Korrigierter, empirischer Hurst Exponent der Absoluterträge

- Kurtosis der Erträge
- Hill-Schätzer bezüglich der 2, 5%, 5% und 10% größten, gemessenen Erträge

Es werden also insgesamt $k = 11$ Maßzahlen betrachtet, die sich alle auf die reinen bzw. absoluten Erträge beziehen. Deshalb werden wir die Bootstrappingmethode auf eine real-world Ertragszeitreihe anwenden. Wie gewohnt verwenden wir die Zeitreihe der S&P 500-Erträge zwischen Februar 1871 und April 2018 von Shiller (2018). Ziel ist es, mittels Bootstrapping B viele Pseudo-Ertragszeitreihen $(r_t^{(b)})_{t=1, \dots, N}$, $b = 1, \dots, B$ zu erzeugen, jeweils alle elf Maßzahlen auszuwerten, die Vektoren der Maßzahlen $m^{(b)} = (m_1^{(b)}, \dots, m_{11}^{(b)})$, $b = 1, \dots, B$ aufzustellen und damit einen Schätzer für die gesuchte Varianz-Kovarianz-Matrix zu bestimmen.

Nun haben wir bei der Betrachtung stilisierter Fakten festgestellt, dass eine langfristige Abhängigkeit der Absoluterträge besteht. Damit sind Absoluterträge zu unterschiedlichen Zeitpunkten korreliert und somit stochastisch abhängig, d. h., die Voraussetzung für den Bootstrappingansatz von Efron (1979) ist verletzt. Deshalb werden wir anstatt des klassischen Bootstrappings das sogenannte *Block Bootstrapping* anwenden. Hierbei handelt es sich um eine beliebte Abwandlung der ursprünglichen Bootstrappingmethode, mit der die Abhängigkeitsstruktur der real-world Zeitreihe auf die Bootstrappingzeitreihen übertragen werden kann. Die Idee des Block Bootstrappings ist es, die betrachtete real-world Zeitreihe zunächst in Blöcke fester Länge aufzuteilen. Anschließend führt man je Block die klassische Bootstrappingmethode durch und verkettet die einzelnen Bootstrappingzeitreihen der Blöcke schließlich zu einer N -elementigen Block-Bootstrapping-Zeitreihe. Durch die getrennte Behandlung der Blöcke können wir erreichen, dass die langfristige Abhängigkeit in der Zeitreihe der Absoluterträge beim Bootstrapping nicht verloren geht.

Entscheidend für das Block-Bootstrapping-Verfahren ist die Wahl der Blocklänge. Je größer die Blöcke gewählt werden, desto stärker gleicht das Block Bootstrapping dem klassischen Bootstrapping. Im Extremfall werden Blöcke der Länge N gewählt, d. h., es wird nur ein einziger Block gebildet und die Methode entspricht gerade der klassischen Bootstrappingmethode von Efron (1979). Im gegenteiligen Extremfall haben die Blöcke die Länge eins und damit entspricht jede Block-Bootstrapping-Zeitreihe genau der real-world Zeitreihe und das Bootstrappingverfahren verfehlt seinen Sinn. Die Blocklänge stellt also einen Parameter des Verfahrens dar, der für den konkreten Anwendungsfall möglichst passend bestimmt werden muss. Leider gibt es bislang keine theoretische Analyse für die Wahl einer optimalen Blocklänge. Deshalb werde ich im Folgenden eine eigene, empirische Idee angeben, wie wir die Blocklänge in unserem Fall wählen können.

Gesucht ist also eine Blocklänge, sodass die Block-Bootstrapping-Zeitreihen im Sinne der Maßzahlen ähnlich zu der real-world Zeitreihe sind. Es bezeichnen

$$m^{(b)} = (m_1^{(b)}, \dots, m_{11}^{(b)}), \quad b = 1, \dots, B$$

die Vektoren der Maßzahlen ausgewertet an den Block-Bootstrapping-Zeitreihen $(r_t^{(b)})_{t=1, \dots, N}$, $b = 1, \dots, B$ und

$$m^{\text{emp}} = (m_1^{\text{emp}}, \dots, m_{11}^{\text{emp}})$$

den Vektor der Maßzahlen ausgewertet an den real-world Erträgen $(r_t)_{t=1, \dots, N}$. Ziel ist es, dass die Block-Bootstrapping-Vektoren $m^{(b)}$, $b = 1, \dots, B$ in gewisser Weise mit dem real-world Vektor m^{emp} *verträglich* sind. Dazu werden wir für jede Maßzahl m_i , $i \in \{1, \dots, 11\}$ fordern, dass der empirische Wert m_i^{emp} eine ähnliche Größenordnung wie der *Großteil* der Werte $m_i^{(b)}$, $b = 1, \dots, B$ hat. Zur Veranschaulichung ist in Abbildung 6.1 eine mögliche, beispielhafte Häufigkeitsverteilung von $m_i^{(b)}$, $b = 1, \dots, B$ in Form eines Histogramms skizziert. Liegt nun m_i^{emp} nahe am Zentrum der Häufigkeitsverteilung, dann sind wir mit dem Ergebnis der gewählten Blocklänge zufrieden. Erhalten wir dagegen ein Bild wie im rechten Teil der Abbildung, dann lassen sich m_i^{emp} und $m_i^{(b)}$, $b = 1, \dots, B$ nicht vereinbaren und wir sollten eine andere Blocklänge wählen.

Da wir insgesamt elf Maßzahlen gleichzeitig betrachten, müssen wir eine großzügige Toleranz bei der Abweichung vom Zentrum der Häufigkeitsverteilung akzeptieren. Uns wird es genügen, wenn alle empirischen Maßzahlen m_i^{emp} mindestens genauso nahe am Zentrum der Häufigkeitsverteilung der Werte $m_i^{(b)}$, $b = 1, \dots, B$ sind, wie 95% dieser Werte selbst. Das heißt, m_i^{emp} soll im Bereich der 95% zentralsten Werte $m_i^{(b)}$, $b = 1, \dots, B$ liegen und dies soll wiederum für jede Maßzahl, d. h. jedes $i = 1, \dots, 11$, erfüllt sein. Formal berechnen wir also für jede Maßzahl m_i , $i \in \{1, \dots, 11\}$ die Abstände der Block-Bootstrapping-Maßzahlen $m_i^{(b)}$, $b = 1, \dots, B$ von ihrem eigenen Mittelwert

$$\left| m_i^{(b)} - \frac{1}{B} \sum_{b=1}^B m_i^{(b)} \right|$$

und bilden ein 95%-Quantil dieser Abstände. Definitionsgemäß sind dann 95% der berechneten Abstände kleiner als das Quantil und 5% größer. Damit die empirische Maßzahl m_i^{emp} mindestens so nahe am Mittelwert der Block-Bootstrapping-Maßzahlen ist, wie 95% dieser Werte selbst, müssen wir fordern, dass der Abstand

$$\left| m_i^{\text{emp}} - \frac{1}{B} \sum_{b=1}^B m_i^{(b)} \right|$$

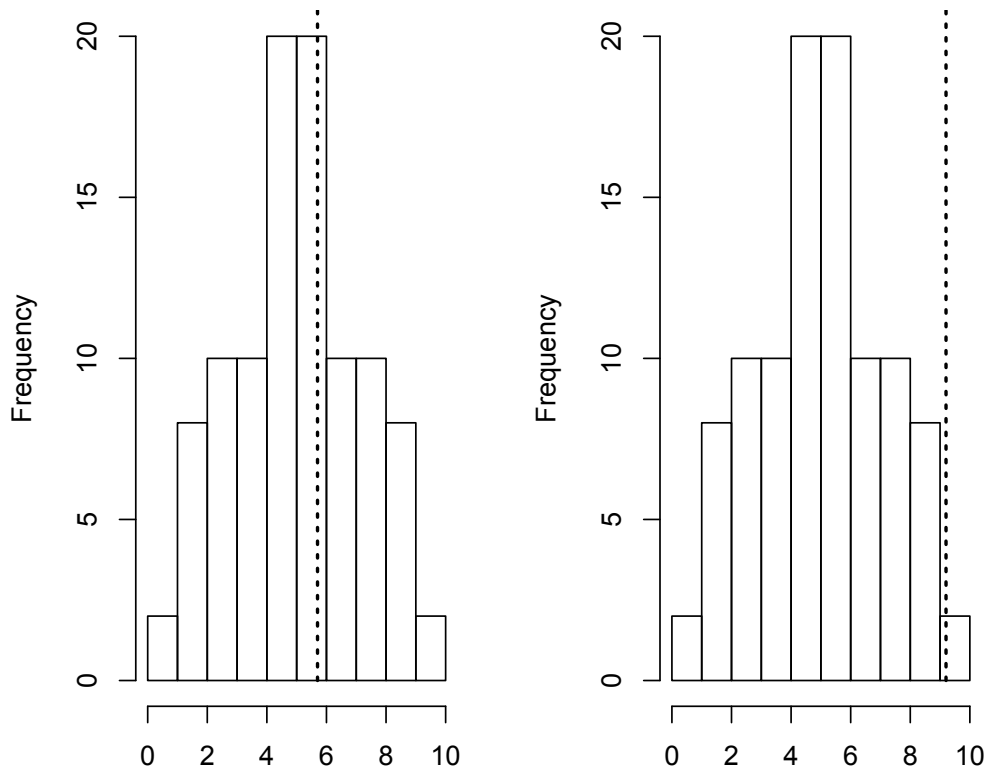


Abbildung 6.1: Liegt die an der real-world Zeitreihe ausgewertete Maßzahl m_i^{emp} (gestrichelte Linie) nahe am Zentrum der Häufigkeitsverteilung der Block-Bootstrapping-Werte $m_i^{(b)}$, $b = 1, \dots, B$, dann sind die Werte verträglich (links), andernfalls lassen sich die Werte nicht miteinander vereinbaren (rechts)

höchstens so groß wie das berechnete 95%-Quantil ist. Die beschriebene Idee und der Akzeptanzbereich ist in Abbildung 6.2 veranschaulicht. Wir akzeptieren eine Blocklänge, wenn die erläuterte Bedingung für jede Maßzahl erfüllt ist.

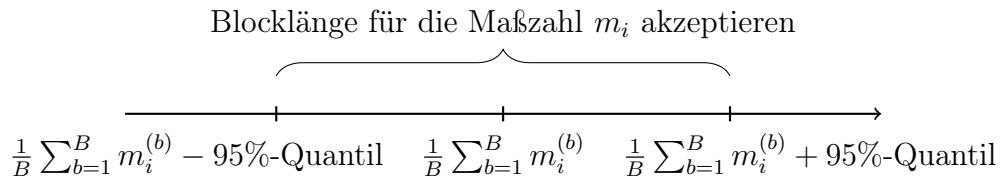


Abbildung 6.2: Hat die empirische Maßzahl m_i^{emp} höchstens den Wert des 95%-Quantils als Abstand vom Mittelwert der Block-Bootstrapping-Maßzahlen, so akzeptieren wir die Blocklänge für die betrachtete Maßzahl, andernfalls nicht

Zu beachten ist, dass die beschriebene Bedingung von den zufällig erzeugten Block-Bootstrapping-Zeitreihen abhängt. Es ist also möglich, dass wir bei einer Durchführung des Verfahrens eine Blocklänge akzeptieren und bei einer weiteren Durchführung lehnen wir dieselbe Blocklänge ab. Durch die groß gewählte Toleranz spielt dieser zufällige Einfluss für hinreichend groß gewählte Werte für B jedoch kaum eine Rolle.

Führt man das beschriebene Verfahren mit den aufgezählten, elf Maßzahlen für die S&P 500-Erträge zwischen Februar 1871 und April 2018 durch und wählt $B = 1000$ als Anzahl der erzeugten Block-Bootstrapping-Zeitreihen, dann erhält man die Akzeptanz der Blocklänge sechs, d. h., für halbjährliche Blöcke halten wir die real-world Zeitreihe und die Block-Bootstrapping-Zeitreihen für verträglich. Abbildung 6.3 zeigt alle elf Häufigkeitsverteilungen der erzeugten Block-Bootstrapping-Maßzahlen. Die gestrichelten Linien skizzieren die jeweiligen Werte der empirischen Maßzahlen, die gemäß unserem Kriterium je hinreichend nahe am Zentrum der Verteilung liegen.

Nachdem wir eine geeignete Blocklänge gefunden haben, können wir die Varianz-Kovarianz-Matrix für die $B = 1000$ berechneten Vektoren $m^{(b)}$, $b = 1, \dots, B$ berechnen und diese als Schätzer für die Varianz-Kovarianz-Matrix des zugrunde liegenden real-world DGP verwenden. Die Matrix ist definiert als

$$\hat{\Sigma} := \frac{1}{B} \sum_{b=1}^B (m^{(b)} - \bar{m}) \cdot (m^{(b)} - \bar{m})^\top, \quad (6.1)$$

$$\bar{m} := \frac{1}{B} \sum_{b=1}^B m^{(b)}$$

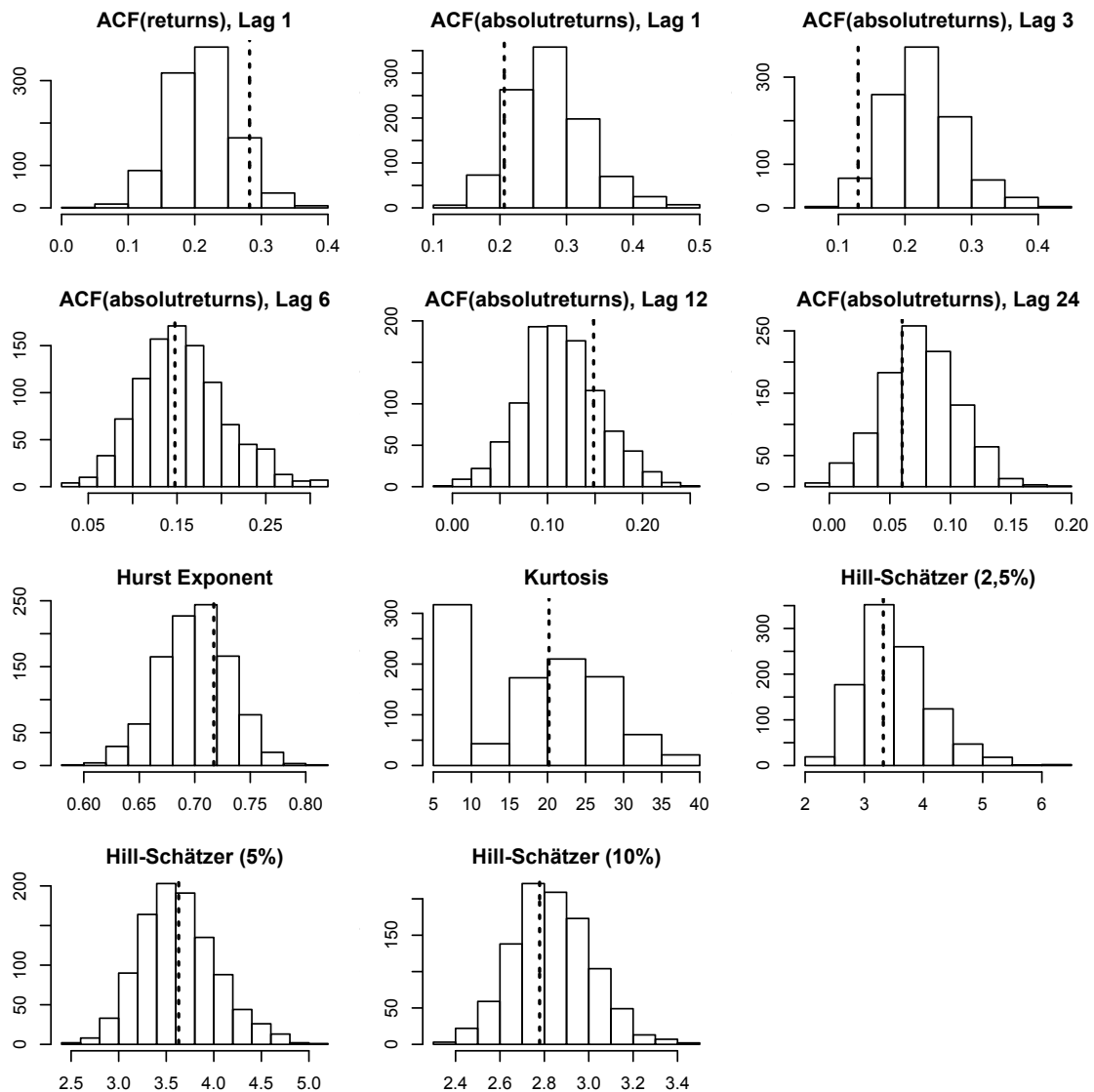


Abbildung 6.3: Histogramme der elf Maßzahlen ausgewertet an den $B = 1000$ Block-Bootstrapping-Zeitreihen und Einzeichnung der Maßzahlen ausgewertet an der real-world Zeitreihe der S&P 500-Erträge zwischen Februar 1871 und April 2018 durch die gestrichelte Linie: Die empirischen Maßzahlen liegen jeweils hinreichend nahe am Zentrum der Häufigkeitsverteilung, sodass wir die gewählte Blocklänge sechs akzeptieren

(vergleiche Franke und Westerhoff, 2011, S. 6). Damit haben wir das Block-Bootstrapping-Verfahren verwendet, um Pseudo-Ertragszeitreihen zu erzeugen, die Maßzahlen an ihnen auszuwerten und schließlich eine geschätzte Varianz-Kovarianz-Matrix der Maßzahlen zu berechnen.

Der folgende Quellcode führt das beschriebene Verfahren durch. In diesem Quellcode verwenden wir erstmals einen Zufallszahlengenerator bei der Erzeugung der Block-Bootstrapping-Zeitreihen. Dies führt natürlich dazu, dass darauf aufbauende Ergebnisse in der weiteren Arbeit von den Ergebnissen der Zufallszahlenziehung abhängen und damit nicht mehr nachvollziehbar sind. Aus diesem Grund wird im Quellcode zunächst der Pseudozufallszahlengenerator initialisiert. Anschließend werden für halbjährliche Blocklängen $B = 1000$ Block-Bootstrapping-Zeitreihen erzeugt und unsere elf Maßzahlen an ihnen ausgewertet. Mit `empirical_moments` sind die an den S&P 500-Erträgen zwischen Februar 1871 und April 2018 ausgewerteten Maßzahlen gemeint. Der hier angegebene Quellcode prüft nun die Akzeptanz der Blocklänge nach dem beschriebenen Kriterium. Außerdem wird die Abbildung 6.3 erzeugt und die Varianz-Kovarianz-Matrix gemäß (6.1) berechnet.

```
#Initialisierung des Pseudozufallszahlengenerators, damit die
Ziehung der Zufallszahlen stets dasselbe Ergebnis liefert,
sodass die Werte nachvollziehbar sind
set.seed(10)

#Wahl der Blocklänge und Berechnung der Blockanzahl
block_length <- 6
block_number <- floor(length(returns)/block_length)

#Festlegung der Anzahl erzeugter Block-Bootstrapping-Zeitreihen
B <- 1000

#Berechnung der B vielen Block-Bootstrapping-Zeitreihen und
Speicherung in den Spalten einer Matrix
matrix_blockbootstrapsamples <- matrix(nrow=block_number*block_
length, ncol=B)
for(b in 1:B) {
  random_numbers <- sample(block_length, block_number*
block_length, replace=TRUE)
  for(j in 1:block_number) {
    for(i in 1:block_length){
      matrix_blockbootstrapsamples[(j-1) *
block_length + i, b] <- returns[(j-1)
* block_length + random_numbers[(j
-1)*block_length+i]]
    }
  }
}
```



```

    }
  }
}

#Berechnung der Maßzahlen je Block-Bootstrapping-Zeitreihe und
  Speicherung in einer Matrix
matrix_moments <- matrix(nrow=11, ncol=B)

for(b in 1:B) {
  matrix_moments[,b] <- evaluate_moments(matrix_
    blockbootstrapsamples[,b])
}

#Prüfung der Akzeptanz der Blocklänge
distance_empirical_moments_from_bootstrapmean <- abs(empirical_
  moments - rowMeans(matrix_moments))
p <- 0.95
vector_p_quantile <- numeric(length=11)
for(j in 1:11) {
  vector_p_quantile[j] <- quantile(abs(matrix_moments[j,]-
    mean(matrix_moments[j,]))) ,p)
}
distance_empirical_moments_from_bootstrapmean < vector_p_
  quantile
#Ergebnis:
#TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

#Erzeugung der Abbildung
par(mfrow=c(4,3),mar=c(3,3,2,1))

hist(matrix_moments[1,],main='ACF(returns)',_Lag_1',xlab='')
lines(rep(empirical_moments[1],par('usr')[4]),c(1:par('usr')[4])
  ,lty='dotted',lwd=2)
hist(matrix_moments[2,],main='ACF(absolutreturns)',_Lag_1',xlab='
  ')
lines(rep(empirical_moments[2],par('usr')[4]),c(1:par('usr')[4])
  ,lty='dotted',lwd=2)
hist(matrix_moments[3,],main='ACF(absolutreturns)',_Lag_3',xlab='
  ')
lines(rep(empirical_moments[3],par('usr')[4]),c(1:par('usr')[4])
  ,lty='dotted',lwd=2)
hist(matrix_moments[4,],main='ACF(absolutreturns)',_Lag_6',xlab='
  ')

```

```

lines(rep(empirical_moments[4], par('usr')[4]), c(1:par('usr')[4])
      , lty='dotted', lwd=2)
hist(matrix_moments[5,], main='ACF(absolutreturns),_Lag_12', xlab=
      '')
lines(rep(empirical_moments[5], par('usr')[4]), c(1:par('usr')[4])
      , lty='dotted', lwd=2)
hist(matrix_moments[6,], main='ACF(absolutreturns),_Lag_24', xlab=
      '')
lines(rep(empirical_moments[6], par('usr')[4]), c(1:par('usr')[4])
      , lty='dotted', lwd=2)
hist(matrix_moments[7,], main='Hurst_Exponent', xlab='')
lines(rep(empirical_moments[7], par('usr')[4]), c(1:par('usr')[4])
      , lty='dotted', lwd=2)
hist(matrix_moments[8,], main='Kurtosis', xlab='')
lines(rep(empirical_moments[8], par('usr')[4]), c(1:par('usr')[4])
      , lty='dotted', lwd=2)
hist(matrix_moments[9,], main='Hill-Schätzer_(2,5%)', xlab='')
lines(rep(empirical_moments[9], par('usr')[4]), c(1:par('usr')[4])
      , lty='dotted', lwd=2)
hist(matrix_moments[10,], main='Hill-Schätzer_(5%)', xlab='')
lines(rep(empirical_moments[10], par('usr')[4]), c(1:par('usr')[4])
      [4], lty='dotted', lwd=2)
hist(matrix_moments[11,], main='Hill-Schätzer_(10%)', xlab='')
lines(rep(empirical_moments[11], par('usr')[4]), c(1:par('usr')[4])
      [4], lty='dotted', lwd=2)

#Berechnung der Varianz-Kovarianz-Matrix für die Maßzahlen
  ausgewertet an den Block-Bootstrapping-Zeitreihen
variance_covariance_matrix <- (matrix_moments[,1] - rowMeans(
  matrix_moments)) %*% t(matrix_moments[,1] - rowMeans(matrix_
  moments))
for(b in 2:B) {
  variance_covariance_matrix <- variance_covariance_matrix
    + (matrix_moments[,b] - rowMeans(matrix_moments)) %*%
      t(matrix_moments[,b] - rowMeans(matrix_moments))
}
variance_covariance_matrix <- (1/B) * variance_covariance_matrix

```

Wir konnten also in diesem Kapitel das Block-Bootstrapping-Verfahren verwenden, um die Varianz-Kovarianz-Matrix der Maßzahlen des zugrunde liegenden real-world DGP zu schätzen. Diese Matrix werden wir bei der Modellbewertung benötigen. Auf diese gehen wir nun im folgenden Kapitel genauer ein.

Kapitel 7

Marktmodellanpassung

Wir haben uns bisher damit beschäftigt, welche Charakteristiken reale Finanzzeitreihen haben, wie wir diese messen können, woher wir real-world Stichproben erhalten und wie man den Fundamentalwert schätzen kann. In Kapitel 6 haben wir außerdem eine Methode kennengelernt, wie mithilfe von Block Bootstrapping ein Schätzer für die Varianz-Kovarianz-Matrix der Maßzahlen stilisierter Fakten berechnet werden kann. Dieses Wissen werden wir nun verwenden, um zu beurteilen, wie gut Modelle den DGP nachahmen. Dabei werden wir uns auf eine bestimmte Art von Modellen einschränken, die strukturelle Parameter besitzen. Ziel ist es, diese Modellparameter möglichst gut zu schätzen, sodass die vom Modell erzeugten Simulationszeitreihen der real-world Zeitreihe in einem bestimmten Sinne ähnlich sind.

Im folgenden Abschnitt 7.1 werden wir zunächst das hier betrachtete, allgemeine Marktmodell einführen. Außerdem werden wir erläutern, wie mithilfe des Modells schrittweise eine Simulationszeitreihe aufgestellt werden kann. Anschließend lernen wir in Abschnitt 7.2 eine Methode kennen, mit der wir eine konkrete Parameterwahl bewerten können. Dort werden wir die geschätzte Varianz-Kovarianz-Matrix aus Kapitel 6 benötigen. In Abschnitt 7.3 stellen wir uns abschließend die Frage, wie die Modellparameter möglichst gut gewählt werden können.

7.1 Aufstellung eines allgemeinen Marktmodells

Wir benötigen ein möglichst allgemeines Marktmodell, das zur Bildung von Simulationszeitreihen verwendet werden kann, die stilisierten Fakten gut nachbildet, aber keine zu komplizierte Struktur aufweist. Das Verhalten des Modells soll durch wählbare Modellparameter bestimmt werden, die wir in Abschnitt 7.3 möglichst gut schätzen wollen. Wir werden uns hier auf die Klasse der *statischen, heterogenen*

Agentenmodelle konzentrieren. Bei dieser Modellart gibt es *mehrere* Händlertypen, die jeweils nach unterschiedlichen Strategien handeln. Jeder einzelne Akteur, oder auch *Agent*, am Finanzmarkt ordnet sich einem der Typen zu und handelt gemäß der entsprechenden Strategie. Bei *statischen* Agentenmodellen geschieht die Zuordnung zum Beginn der ersten Periode und wird danach nicht mehr verändert. Alternativ gibt es auch *dynamische* Modelle, bei denen Akteure im Zeitverlauf ihren Typ ändern können (vergleiche Hommes, 2005, 2006; Franke und Westerhoff, 2011). Der Einfachheit halber werden wir darauf verzichten und gehen von festen Marktanteilen der verschiedenen Händlertypen im Zeitverlauf aus. Heterogene Agentenmodelle haben im Allgemeinen die Eigenschaft, stilisierte Fakten sehr gut nachbilden zu können und erfreuen sich deshalb in den vergangenen Jahrzehnten großer Beliebtheit.

Wir betrachten Modelle mit *zwei* Händlertypen. Einer der beiden Typen ist der sogenannte *Chartist*. Er handelt auf Grundlage der Preistendenz in der vergangenen Periode. Ist der Preis gestiegen, so geht er von einer Fortsetzung der Entwicklung in der kommenden Periode aus und investiert in das betrachtete Finanzprodukt. Bei einem Preisabfall entscheidet er sich entsprechend zur Desinvestition. Die Handelsstrategie gehört zu den Trendfolgestrategien, die wir bereits in Kapitel 2 kennengelernt haben.

Der zweite, für uns wichtige Händlertyp heißt *Fundamentalist*. Im Gegensatz zum Chartisten agiert er unter Berücksichtigung des Fundamentalwertes, den er als wahren Wert des Finanzproduktes ansieht. Fundamentalisten sind davon überzeugt, dass der Preis des Finanzproduktes immer wieder zum Fundamentalwert zurückkehrt. Deshalb investieren sie, wenn der Marktpreis unter dem Fundamentalwert liegt und desinvestiert andernfalls. Um das Verhalten von Fundamentalisten im Marktmodell zu simulieren, werden wir die in Kapitel 5 geschätzten Fundamentalwerte benötigen.

Wir werden nun ein möglichst allgemeines, mathematisches Modell aufstellen, das diesen Vorgaben entspricht und zur Simulation des Marktgeschehens verwendet werden kann. Dabei werden wir uns vor allem an Franke und Westerhoff (2011) orientieren, wobei wir die Dynamik weglassen. Das Modell wird einige strukturelle Parameter enthalten, deren Werte das Modellverhalten beeinflussen. Später werden wir uns dafür interessieren, wie wir die Parameter möglichst gut wählen können, sodass das Modell den DGP in einem gewissen Sinne gut nachbildet.

Unser allgemeines Marktmodell bestimmt zu Beginn jeder Periode t den Marktpreis p_t . Dieser bildet sich aus dem Angebot und der Nachfrage der Vorperiode.

Übersteigt die Nachfrage das Angebot, dann steigt der Preis im Vergleich zur Vorperiode und umgekehrt. Wir werden das Angebot bzw. die Nachfrage auf die beiden Händlertypen aufteilen, sodass $d_{t-1}^C < 0$ für das Angebot und $d_{t-1}^C > 0$ für die Nachfrage der Chartisten und analog d_{t-1}^F für das Angebot bzw. die Nachfrage der Fundamentalisten stehen. Die jeweilige Nachfrage müssen wir natürlich mit den Marktanteilen

$$n^C, n^F \in [0, 1], n^C + n^F = 1 \quad (7.1)$$

der Chartisten bzw. Fundamentalisten gewichten. Außerdem müssen wir sicherstellen, dass die Preise stets nichtnegativ sind. Dies erreichen wir durch den Einsatz der Exponentialfunktion. Im Modell von Franke und Westerhoff (2011) ergibt sich der Preis p_t nun aus der Gleichung

$$p_t := \exp [\log(p_{t-1}) + \theta_1 \cdot (n^C \cdot d_{t-1}^C + n^F \cdot d_{t-1}^F)]$$

beziehungsweise umgestellt

$$\log(p_t) = \log(p_{t-1}) + \theta_1 \cdot (n^C \cdot d_{t-1}^C + n^F \cdot d_{t-1}^F). \quad (7.2)$$

Wir werden annehmen, dass es sich bei den so bestimmten Preisen um reale, trendbereinigte Größen handelt. Dies hat den Vorteil, dass die Vergleichbarkeit mit den in Kapitel 5 geschätzten Fundamentalwerten \hat{p}_t^* sichergestellt ist. Es bezeichnen nun θ_1 , n^C , n^F unsere ersten, strukturellen Parameter des Modells, die wir später möglichst gut schätzen wollen. Der Parameter θ_1 bestimmt, wie sensibel der Preis auf Änderungen in Angebot und Nachfrage reagiert. Aufgrund der beschriebenen Logik fordern wir $\theta_1 \geq 0$. Außerdem wird es natürlich genügen, einen der beiden Marktanteile n^C bzw. n^F zu schätzen und den anderen Anteil durch die Gleichung in (7.1) zu berechnen. Wir werden deshalb nur den Marktanteil der Chartisten n^C als Modellparameter berücksichtigen.

Im nächsten Schritt wollen wir die Angebot- bzw. Nachfragewerte d_t^C und d_t^F spezifizieren. Wir haben bereits die Strategien von Chartisten und Fundamentalisten beschrieben. Für Chartisten sollte demnach d_t^C eine positive Beziehung zur Preisänderung

$$\log(p_t) - \log(p_{t-1})$$

der vergangenen Periode aufweisen. Fundamentalisten bestimmen ihre Nachfrage bzw. ihr Angebot d_t^F dagegen anhand der Abweichung vom Fundamentalwert

$$\log(p_t^*) - \log(p_t).$$

An dieser Stelle wollen wir die geschätzten Fundamentalwerte \hat{p}_t^* aus Kapitel 5 für p_t^* einsetzen. Zusätzlich wollen wir einen zufälligen Einflussfaktor in die Nachfragebestimmung integrieren, denn Händler an Finanzmärkten handeln nur *beschränkt*

rational, d. h., sie lassen sich durch individuelle Stimmungen und Launen beeinflussen, können verfügbare Informationen nur teilweise optimal auswerten und ihre Berechnungsfähigkeiten sind begrenzt (vergleiche Hommes, 2005, Kapitel 1). Aus diesem Grund fügen wir jeweils einen zusätzlichen, zufälligen Störterm ϵ_t^C bzw. ϵ_t^F ein und stellen schließlich die Gleichungen

$$d_t^C := \theta_2 \cdot (\log(p_t) - \log(p_{t-1})) + \epsilon_t^C \quad , \quad \epsilon_t^C \sim \mathcal{N}(0, \sigma_C^2) \quad , \quad (7.3)$$

$$d_t^F := \theta_3 \cdot (\log(\hat{p}_t^*) - \log(p_t)) + \epsilon_t^F \quad , \quad \epsilon_t^F \sim \mathcal{N}(0, \sigma_F^2) \quad (7.4)$$

auf, wobei die strukturellen Parameter θ_2 , θ_3 , σ_C und σ_F wieder nichtnegativ sein sollen. Die Parameter θ_2 bzw. θ_3 bestimmen somit, wie sensibel Chartisten bzw. Fundamentalisten ihre Nachfrage an die jeweilige Preissituation anpassen. Die zufällige Schwankung wird durch die Modellparameter σ_C und σ_F festgelegt.

Damit haben wir unser allgemeines, statisches, heterogenes Agentenmodell mit den sechs strukturellen Parametern θ_1 , θ_2 , θ_3 , n^C , σ_C und σ_F aufgestellt. Das 6-Tupel der Modellparameter werden wir mit

$$\Theta := (\theta_1, \theta_2, \theta_3, n^C, \sigma_C, \sigma_F)$$

bezeichnen.

Sind die Modellparameter Θ sowie ein Startpreis $p_0 > 0$ fest gewählt, dann können wir das Modell verwenden, um einen Preisverlauf für die Zeitskala $t = 1, \dots, N$ zu simulieren, d. h., wir stellen eine Simulationszeitreihe auf. Dazu sind die folgenden Schritte durchzuführen:

1. Berechnung von $n^F = 1 - n^C$ und Initialisierung von $d_0^C = d_0^F = 0$ und $t = 1$
2. Berechnung des Preises $p_t > 0$ gemäß (7.2)
3. Ziehung der Zufallszahlen ϵ_t^C und ϵ_t^F aus $\mathcal{N}(0, \sigma_C^2)$ und $\mathcal{N}(0, \sigma_F^2)$
4. Berechnung der Nachfragen d_t^C und d_t^F gemäß (7.3) und (7.4)
5. Falls $t + 1 \leq N$: Aktualisierung $t \rightarrow t + 1$ und Fortsetzung mit Schritt 2

Die Initialisierung ist so gewählt, dass zu Beginn weder ein Nachfrage- noch ein Angebotsüberschuss existiert. Zu betonen ist außerdem, dass die so berechneten Zeitreihen der Preise $(p_t)_{t=1, \dots, N}$ vom Zufall abhängig sind. Führt man die beschriebene Simulation mehrmals mit identischen Parametern und Initialisierungen durch, dann erhält man wahrscheinlich unterschiedliche Simulationszeitreihen. Dies liegt natürlich an der Ziehung der Zufallszahlen in Schritt drei. Im zweiten Schritt können bei Bedarf zusätzlich die Erträge gemäß ihrer Definition (1.1) bestimmt werden, um simultan die Ertragszeitreihe $(r_t)_{t=1, \dots, N}$ aufzustellen.

7.2 Method of Simulated Moments

Das Ziel dieses Abschnitts ist es, eine Aussage darüber treffen zu können, wie *gut* das in Abschnitt 7.1 beschriebene, allgemeine Modell mit konkreten, strukturellen Parametern Θ und Startpreis p_0 ist. Dazu müssen wir uns natürlich überlegen, auf welche Weise wir die Güte definieren. Es liegt nahe, hierbei die stilisierten Fakten und ihre Messungen zu verwenden. Je *näher* die Maßzahlen ausgewertet an einer Simulationszeitreihe an den Maßzahlen ausgewertet an einer real-world Zeitreihe sind, desto besser bewerten wir das Modell mit den gewählten Parametern. Bei der Definition dieser Nähe besteht wiederum Spielraum. Wir werden uns hier an Franke und Westerhoff (2011) orientieren und im Folgenden die *Method of Simulated Moments (MSM)* erläutern. Bei der Method of Simulated Moments handelt es sich um einen Bewertungsansatz für die Güte eines Marktmodells. Die Idee ist, den Abstand zwischen den Simulationsmaßzahlen und den real-world Maßzahlen durch eine quadratische Verlustfunktion zu messen. Ist der dadurch berechnete Verlust groß, dann liegen die Simulationsmaßzahlen und die real-world Maßzahlen weit voneinander entfernt und wir bewerten das konkrete Marktmodell als schlechtes Modell. Ergibt die Verlustfunktion dagegen einen kleinen Wert, dann sind wir mit dem Simulationsergebnis des Modells zufrieden und bewerten es als gutes Modell.

Sei m^{emp} der Spaltenvektor, der die Maßzahlen ausgewertet an der real-world Stichprobe enthält. Wenn wir also wieder wie in den vorherigen Kapiteln den S&P 500-Datensatz im Zeitraum Februar 1871 bis April 2018 von Shiller (2018) als real-world Stichprobe verwenden, dann entspricht m^{emp} gerade dem Vektor `empirical_moments` aus unseren R-Quellcodes. Es bezeichne außerdem $m_\omega(\Theta, p_0)$ den Spaltenvektor mit den Maßzahlen, die an einem Simulationsergebnis des in Abschnitt 7.1 beschriebenen, allgemeinen Modells für die Parameterwahl Θ und den Startpreis p_0 ausgewertet wurden. Mit der Notation ω soll verdeutlicht werden, dass der Vektor natürlich von den gezogenen Zufallszahlen in Schritt drei des Simulationsverfahrens abhängig ist. Würde man also für die gleichen Parameter und den gleichen Startpreis mehrere Simulationen durchführen, so würde man wahrscheinlich auch unterschiedliche Spaltenvektoren $m_\omega(\Theta, p_0)$ erhalten. Diese Schwankung werden wir später entkräften, indem wir nicht nur eine sondern mehrere Simulationsstichproben erzeugen und einen mittleren Verlust berechnen.

Die Method of Simulated Moments-Verlustfunktion hat nun die folgende, quadratische Form:

$$J[m_\omega(\Theta, p_0)] := [m_\omega(\Theta, p_0) - m^{\text{emp}}]^\top W [m_\omega(\Theta, p_0) - m^{\text{emp}}] \quad (7.5)$$

Dabei bezeichnet

$$W = (w_{ij})_{i,j=1,\dots,k}$$

eine geeignete, quadratische, reellwertige Gewichtungsmatrix der passenden Dimension k . Wenn wir uns auf dieselben Maßzahlen wie in Abschnitt 6.2 beziehen, dann gilt $k = 11$.

Nun stellt sich natürlich die Frage, wie wir die Gewichtungsmatrix definieren können. An dieser Stelle kommt nun unsere geschätzte Varianz-Kovarianz-Matrix aus Abschnitt 6.2 zum Einsatz. Dies lässt sich wie folgt begründen. Wir stellen uns vor, wir könnten beliebig viele Stichproben aus dem real-world DGP ziehen und je Stichprobe den Vektor der ausgewerteten Maßzahlen aufstellen. Wenn nun eine Maßzahl m_i für die real-world Daten stark schwankt, dann scheint sie im DGP eine hohe Varianz zu besitzen. Deshalb kann es sein, dass selbst sehr gute Modelle eine große Differenz $m_\omega(\Theta, p_0)_i - m_i^{\text{emp}}$ ergeben. In diesem Fall sollten wir also die Abweichung in der Verlustfunktion nicht zu stark bestrafen, d. h., wir sollten den entsprechenden Diagonaleintrag w_{ii} dann tendenziell klein wählen. Beobachten wir dagegen anhand der real-world Stichproben eine geringe Varianz von m_i , dann werden wir fordern, dass auch die Differenz $m_\omega(\Theta, p_0)_i - m_i^{\text{emp}}$ gering sein soll. Durch eine große Wahl von w_{ii} können wir dies erreichen. Gleiches gilt natürlich für Einträge außerhalb der Diagonale. Ist die Kovarianz zwischen den m_i - und m_j -Werten für $i \neq j$ groß, dann sollte w_{ij} klein gewählt werden und umgekehrt. Ist also $\hat{\Sigma}$ ein Schätzer für die Varianz-Kovarianz-Matrix der Maßzahlen des real-world DGP, dann wäre

$$W := \hat{\Sigma}^{-1}$$

ein geeigneter Kandidat für die Gewichtungsmatrix, der die beschriebene Logik verfolgt. In Abschnitt 6.2 haben wir uns gerade mit der Frage beschäftigt, wie wir einen Schätzer $\hat{\Sigma}$ für die Varianz-Kovarianz-Matrix der Maßzahlen des real-world DGP mithilfe des Block-Bootstrapping-Verfahrens herleiten können. Setzen wir die Inverse der so erhaltenen Matrix in die Definition der Verlustfunktion (7.5) als Gewichtungsmatrix ein, dann ist die Method of Simulated Moments vollständig definiert.

Damit ist uns nun ein Mittel gegeben, mit dem wir die Güte einer konkreten Parameterwahl $\Theta = (\theta_1, \theta_2, \theta_3, n^C, \sigma_C, \sigma_F)$ bewerten können: Als Startpreis p_0 wählen wir stets den trendbereinigten Preis für Januar 1871 aus der S&P 500-Zeitreihe von Shiller (2018), damit sich die real-world Zeitreihe und die Simulationszeitreihen auf dieselbe Ausgangslage beziehen. Anschließend können wir für die Parameterwahl zum Beispiel $M = 10$ viele Simulationszeitreihen gemäß den in Abschnitt 7.1 angegebenen Schritten berechnen, die Maßzahlen

$$m_{\omega_j}(\Theta, p_0), \quad j = 1, \dots, 10$$

an ihnen auswerten und schließlich den mittleren Wert der Verlustfunktion

$$J[m(\Theta, p_0)] := \frac{1}{M} \sum_{j=1}^M J[m_{\omega_j}(\Theta, p_0)] \quad (7.6)$$

bestimmen.

Mit dem folgenden R-Quellcode können wir eine konkrete Wahl der Modellparameter nach dem beschriebenen Verfahren bewerten und verschiedene Parameterkombinationen vergleichen. Um eine bessere Vergleichbarkeit zu ermöglichen, wird es sinnvoll sein, nicht für jede Parameterkombination je Simulationszeitreihe N Realisationen der Normalverteilungen $\mathcal{N}(0, \sigma_C^2)$ und $\mathcal{N}(0, \sigma_F^2)$ zu bilden, um die Nachfragestörterme der Chartisten und Fundamentalisten aufzustellen. Stattdessen ziehen wir zu Beginn für Chartisten sowie Fundamentalisten jeweils $M \cdot N$ viele Realisationen der *Standardnormalverteilung* $\mathcal{N}(0, 1)$ und verwenden diese Realisationen als Grundlage für *jede* Parameterkombination, d. h., wenn wir die Parameterkombination $\Theta = (\theta_1, \theta_2, \theta_3, n^C, \sigma_C, \sigma_F)$ bewerten wollen, dann bestimmen wir die Nachfragestörterme durch Multiplikation der Realisierungen der Standardnormalverteilung mit den Parametern σ_C bzw. σ_F . Unterschiede in der Bewertung verschiedener Modelle sind damit eindeutig auf die unterschiedlichen, gewählten Parameter und nicht auf den Ausgang der Zufallszahlenziehung zurückzuführen. Nur so haben wir die Möglichkeit, Bewertungsergebnisse zu interpretieren.

```
#Der reale , trendbereinigte Preis für Januar 1871 dient als
  Startpreis
p_0 <- data_Shiller [1,7] *lambda^N

#Anzahl der Simulationsdurchläufe
M <- 10

#Ziehung der standardnormalverteilten Zufallszahlen (für
  verschiedene Parameterwahlen sollten zur besseren
  Vergleichbarkeit die gleichen Zufallszahlen verwendet werden
)
epsilonC_standard_normal_distributed <- rnorm(N*M,0,1)
epsilonF_standard_normal_distributed <- rnorm(N*M,0,1)

#Definition der Gewichtsmatrix
W <- solve(variance_covariance_matrix)
#Ergebnis :
#diag(W) = 615.8992753  497.2279802  431.8564832  588.3765109
           871.3913604  1410.9373008  1496.3630300  0.0303434
           9.9345547   19.0084382   52.7880111
```

```

#Die folgende Funktion berechnet für eine Parameterwahl Theta M
  viele Simulationszeitreihen, bestimmt jeweils den Wert der
  Verlustfunktion und gibt den mittleren Verlust zurück
calculate_loss_function <- function(Theta) {
  theta_1 <- Theta[1]
  theta_2 <- Theta[2]
  theta_3 <- Theta[3]
  nC <- Theta[4]
  sigma_C <- Theta[5]
  sigma_F <- Theta[6]

  #Berechnung des resultierenden Anteils der
    Fundamentalisten
  nF <- 1 - nC

  #M-malige Simulation und Auswertung der Verlustfunktion
    für die gewählten Parameter
  values_loss_function <- vector(length = M)
  prices_simulation <- vector(length = (N+1))
  returns_simulation <- vector(length = N)
  prices_simulation[1] <- p_0
  demandC <- 0
  demandF <- 0
  simulation_moments <- numeric(length = 11)
  for(i in 1:M) {
    #Berechnung der Simulationszeitreihe
    for(t in 1:N) {
      prices_simulation[t+1] <- exp(log(prices_
        _simulation[t]) + theta_1 * (nC*
        demandC + nF*demandF))
      returns_simulation[t] <- (prices_
        simulation[t+1] - prices_simulation[
        t])/prices_simulation[t]
      demandC <- theta_2 * (log(prices_
        simulation[t+1]) - log(prices_
        simulation[t])) + sigma_C * epsilonC
        _standard_normal_distributed[(i-1)*N
        +t]
      demandF <- theta_3 * (log(p_hat_star[t])
        - log(prices_simulation[t+1])) +
        sigma_F * epsilonF_standard_normal_
        distributed[(i-1)*N+t]
    }
  }
}

```

```

    }
    demandC <- 0
    demandF <- 0

    if (any(is.nan(returns_simulation)) | any(!is.numeric(returns_simulation)) | any(is.infinite(returns_simulation))) next

    #Berechnung der Maßzahlen ausgewertet an der Simulationszeitreihe
    simulation_moments <- evaluate_moments(returns_simulation)

    if (any(is.nan(simulation_moments)) | any(!is.numeric(simulation_moments)) | any(is.infinite(simulation_moments))) next

    #Auswertung der Verlustfunktion
    values_loss_function[i] <- t(simulation_moments-empirical_moments) %*% W %*% (simulation_moments-empirical_moments)
  }

  #Rückgabe des mittleren Verlusts, wenn values_loss_function ausschließlich zulässige Werte enthält, ansonsten wird unendlich zurückgegeben
  if ((!any(is.nan(values_loss_function))) & !any(!is.numeric(values_loss_function)) & (!any(is.infinite(values_loss_function)))) {mean(values_loss_function)}
  else Inf
}

```

Die Funktion `calculate_loss_function` kann für eine konkrete Parameterwahl Θ aufgerufen werden und berechnet den entsprechenden Verlustfunktionswert. Es gibt Parameterkombinationen, die zu Fehlern in der beschriebenen Berechnung führen. Solche Fehler sind beispielsweise simulierte Preis- oder Ertragswerte, die außerhalb des darstellbaren Bereichs liegen oder zu einem Fehler bei der Maßzahlenauswertung durch `hurstexp` führen. In diesem Fall gibt die angegebene Funktion `calculate_loss_function` unendlich zurück, d. h., der Parameterkombination wird ein unendlicher Verlust zugeordnet. Anstatt $M = 10$ können natürlich auch mehr Simulationen durchgeführt werden. Dies führt jedoch zu einer Laufzeiterhöhung.

7.3 Schätzung der Modellparameter

Wir stellen uns nun die Frage, wie wir die Modellparameter

$$\Theta = (\theta_1, \theta_2, \theta_3, n^C, \sigma_C, \sigma_F)$$

möglichst gut schätzen können, d. h., wir suchen Parameterwerte, sodass die Verlustfunktion (7.6) einen minimalen Wert annimmt. Leider ist die Lösung des Minimierungsproblems sehr komplex, da wir eine Funktion mit insgesamt sechs Unbekannten betrachten, wir keinerlei Ableitungsinformationen besitzen und die Funktionsauswertung je nach Wahl der Simulationsanzahl M zeitaufwändig ist. Um eine erste, gute Wertkombination zu finden, habe ich zunächst eine systematische Suche durchgeführt, indem die folgenden Parameterwerte betrachtet und je Kombination die Verlustfunktion ausgewertet wird:

$$\begin{aligned}\theta_1, \theta_2, \theta_3 &\in \{0, 1; 0, 5; 1; 1, 5; 2; 2, 5\} \\ n^C &\in \{0, 1; 0, 2; \dots; 0, 9\} \\ \sigma_C, \sigma_F &\in \{0, 1; 0, 5; 1\}\end{aligned}\tag{7.7}$$

Dies ergibt insgesamt 17.496 Wertkombinationen. Der folgende Quellcode führt die beschriebene Rastersuche durch und kommt zum Ergebnis, dass die Parameterwahl

$$\Theta_{\text{Rastersuche}}^* := (0, 1; 2, 5; 1, 5; 0, 5; 0, 1; 0, 1)$$

mit dem mittleren Verlust 29,78677 unter den betrachteten Wertkombinationen bestmöglich ist.

```
theta_values <- c(0.1, 0.5, 1, 1.5, 2, 2.5)
nC_values <- seq(0.1, 0.9, 0.1)
sigma_values <- c(0.1, 0.5, 1)

best_value_loss_function <- Inf
best_theta_1 <- 0
best_theta_2 <- 0
best_theta_3 <- 0
best_nC <- 0
best_sigma_C <- 0
best_sigma_F <- 0

for(a in 1:length(theta_values)) {
  for(b in 1:length(theta_values)) {
    for(c in 1:length(theta_values)) {
      for(d in 1:length(nC_values)) {
        for(e in 1:length(sigma_values)) {
```

```

for(f in 1:length(sigma_values)) {
  result <- calculate_loss_function(c(theta_values[a],
    theta_values[b], theta_values[c], nC_values[d], sigma_
    values[e], sigma_values[f]))
  if(result < best_value_loss_function) {
    best_value_loss_function <- result
    best_theta_1 <- theta_values[a]
    best_theta_2 <- theta_values[b]
    best_theta_3 <- theta_values[c]
    best_nC <- nC_values[d]
    best_sigma_C <- sigma_values[e]
    best_sigma_F <- sigma_values[f]
  }
}
}
}
}
}
}

```

Franke und Westerhoff (2011, S. 7) schlagen die Benutzung des *Nelder-Mead Algorithmus* zur Minimierung der Verlustfunktion vor. Dabei handelt es sich um einen ableitungsfreien Optimierungsalgorithmus zur Minimierung einer Funktion mit mehreren Variablen. Zur Anwendung der numerischen Methode ist allerdings ein hinreichend guter Startpunkt notwendig, da das Verfahren nur in dessen Umgebung nach einem lokalen Minimum sucht. Nach der durchgeführten Rastersuche bietet es sich nun an, $\Theta_{\text{Rastersuche}}^*$ als Startpunkt zu wählen. Der R-Befehl

```

optim(c(0.1, 2.5, 1.5, 0.5, 0.1, 0.1), calculate_loss_function, method
= 'Nelder-Mead', control = list(maxit=100))

```

ruft den Nelder-Mead Algorithmus aus dem Paket `stats` mit dem genannten Startwert auf und führt maximal hundert Iterationsdurchläufe des Verfahrens durch. Dies ergibt eine Verbesserung des Verlustfunktionswertes auf 19,94634 für die Parameterwahl

$$\Theta_{\text{Nelder-Mead}}^* := (0, 07188020; 2, 65214729; 1, 54208764; 0, 51866776; 0, 09583929; 0, 10508706).$$

In Abbildung 7.1 sind für die Parameterwahl $\Theta_{\text{Nelder-Mead}}^*$ zwei modellerzeugte Simulationszeitreihen abgebildet. Wir können feststellen, dass sich die Simulationszeitreihen kaum unterscheiden. Dies liegt an der konkreten Wahl der Parameter. Die geringen Werte $\sigma_C = 0, 09583929$ und $\sigma_F = 0, 10508706$ führen dazu, dass in der Nachfragebildung nur ein geringer, zufälliger Einfluss enthalten ist. Zusätzlich bewirkt der geringe Wert $\theta_1 = 0, 07188020$, dass Preise nur langsam angepasst werden. Ein zufallsbedingter Impuls wirkt sich also kaum auf den Preis aus und

deshalb unterscheiden sich mehrere Simulationszeitreihen nur wenig. Zum Vergleich sind in Abbildung 7.2 für eine andere Parameterwahl

$$\tilde{\Theta} := (0, 05; 4; 1, 5; 0, 95; 0, 25; 0, 1)$$

ebenfalls zwei Simulationszeitreihen abgebildet. Diese unterscheiden sich deutlich stärker voneinander. Das liegt daran, dass Chartisten hier 95% des Marktes einnehmen, ihre Nachfrage durch $\sigma_C = 0, 25$ stärker vom Zufall abhängt und durch $\theta_2 = 4$ auch schneller angepasst wird. Das Modell ist also mit der Parameterwahl $\tilde{\Theta}$ agiler als mit den Modellparametern $\Theta_{\text{Nelder-Mead}}^*$. Nebenbei bemerkt ergibt $\tilde{\Theta}$ einen Verlustfunktionswert von 85, 01822, d. h., wir bewerten diese Parameterwahl wesentlich schlechter als $\Theta_{\text{Nelder-Mead}}^*$.

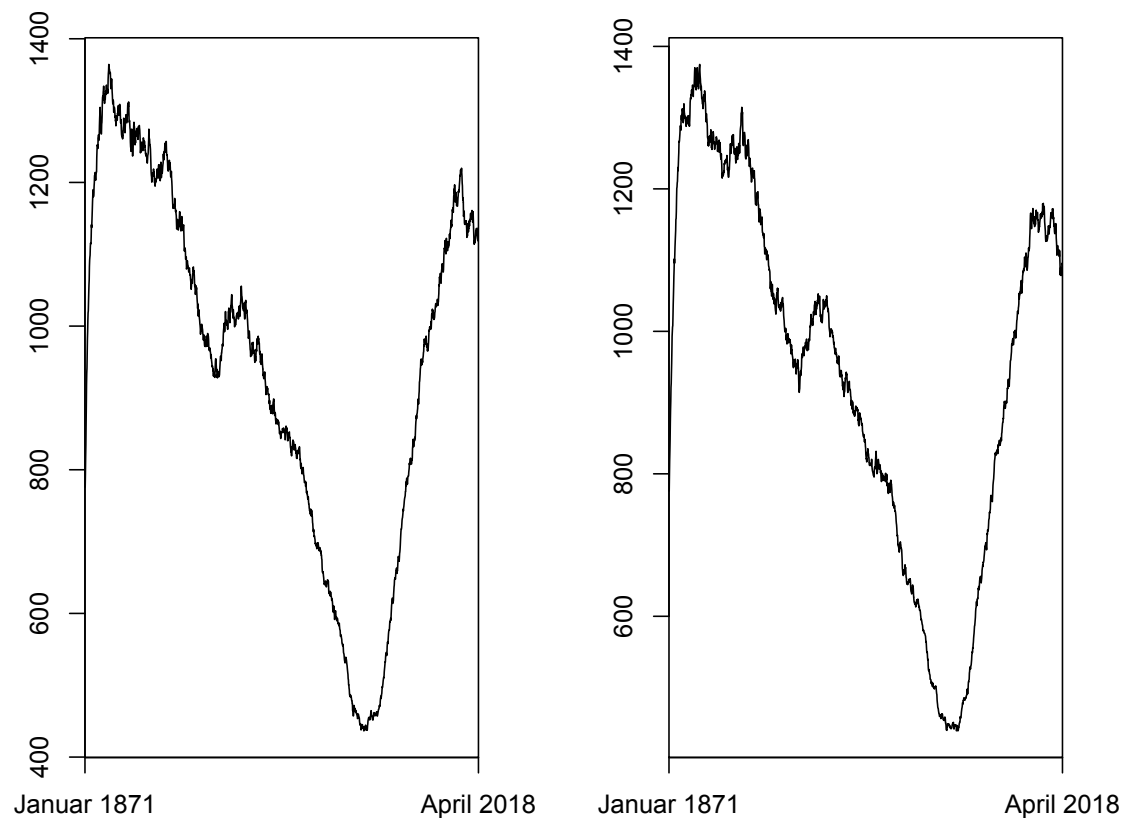


Abbildung 7.1: Für die Modellparameter $\Theta_{\text{Nelder-Mead}}^*$ sind zwei Simulationszeitreihen eingezeichnet, die sich jedoch kaum unterscheiden

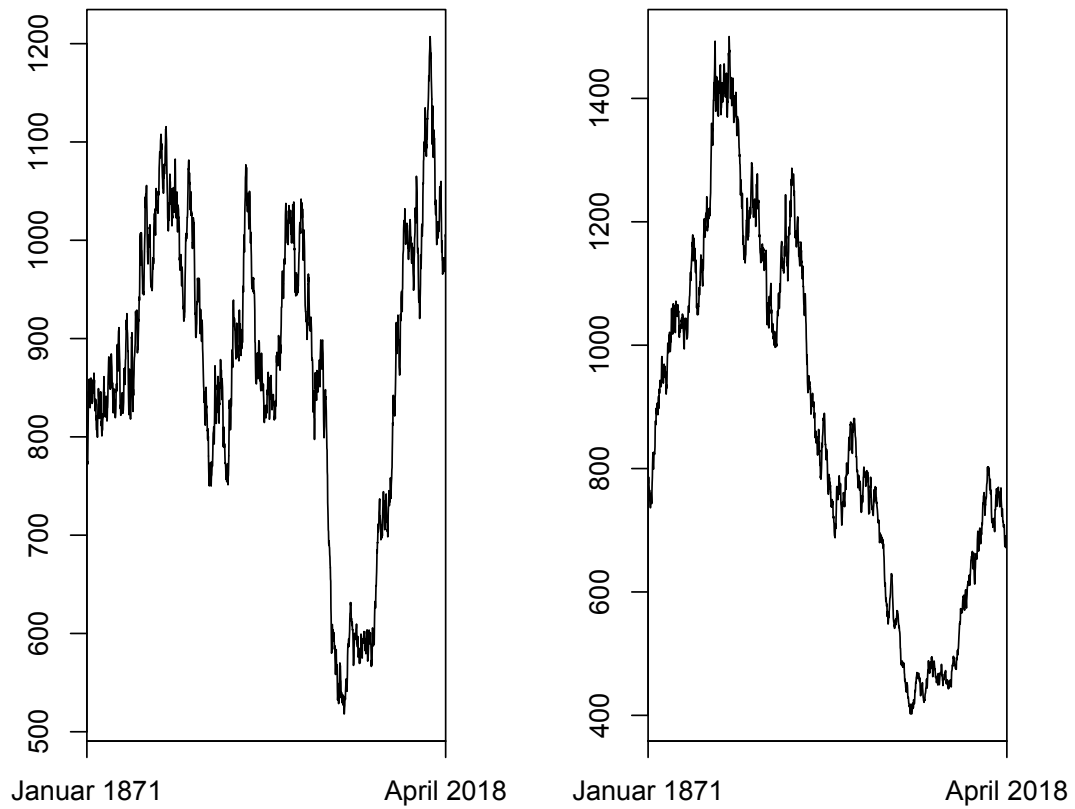


Abbildung 7.2: Für die Parameterwahl $\tilde{\Theta} = (0, 05; 4; 1, 5; 0, 95; 0, 25; 0, 1)$ können sich zwei Simulationszeitreihen stärker unterscheiden als für die Parameterkombination $\Theta_{\text{Nelder-Mead}}^*$

Wenn man Simulationszeitreihen der Parameterwahl $\Theta_{\text{Nelder-Mead}}^*$ betrachtet, fällt außerdem auf, dass sie eine sehr große Ähnlichkeit zum Verlauf der geschätzten Fundamentalwerte besitzen. Dies lässt sich in Abbildung 7.3 gut erkennen. Dort stellt die gepunktete Linie die geschätzten Fundamentalwerte und die durchgezogene Linie eine Simulationszeitreihe der betrachteten Parameterwahl dar. Zum Vergleich sind die realen, trendbereinigten Preise des S&P 500 zwischen Januar 1871 und April 2018 als gestrichelte Linie eingezeichnet. Die Simulationszeitreihe schwankt durchgehend um unsere geschätzten Fundamentalwerte und verlässt, im Gegensatz zu den tatsächlichen Preisen, nie das Niveau über einen längeren Zeitraum. Der Grund hierfür liegt im Marktanteil der Fundamentalisten. Diese nehmen in unserem Modell mit $\Theta_{\text{Nelder-Mead}}^*$ in etwa die Hälfte des Marktes ein und ihre Nachfrage wird kaum vom Zufall beeinflusst. Somit passen sie ihre Nachfrage stets an die aktuelle Abweichung vom Fundamentalwert an und lenken den simulierten Preis dadurch immer wieder zurück zum Fundamentalwertniveau. Wir können also feststellen, dass ein hoher Fundamentalistenanteil dazu führt, dass simulierte Preise nahe am Fundamentalwert sind oder zumindest relativ schnell wieder zu diesem Niveau zurückkehren. Man könnte nun auf die Idee kommen, dass wir mit unserer Modellbewertung Modelle umso besser bewerten, je näher ihre erzeugten Simulationszeitreihen an den geschätzten Fundamentalwerten sind. Um dies zu widerlegen, habe ich die Maßzahlen direkt an unseren geschätzten Fundamentalwerten aus Abschnitt 5.2.2 ausgewertet und den Maßzahlenvektor in die Method of Simulated Moments-Verlustfunktion (7.5) eingesetzt. Dies ergibt den Verlustfunktionswert 1.007,818. Ein Modell, das im Extremfall *exakt* die Fundamentalwerte als Simulationszeitreihe ergibt, würden wir demnach als sehr schlechtes Modell bewerten.

Bei der Durchführung unserer Rastersuche und des Nelder-Mead Algorithmus sollten wir stets im Hinterkopf behalten, dass das Ergebnis natürlich von den gezogenen Zufallszahlen bei der Simulationserzeugung sowie beim Block Bootstrapping abhängt. Somit sollten wir uns mit der Frage beschäftigen, zu welchem Ergebnis wir bei einer anderen Zufallszahlenrealisierung kommen. Wir hoffen, dass unabhängig von der Realisierung dieselben oder zumindest ähnliche Parameterwerte als gut empfunden werden. Die Ziehung der Zufallszahlen sollte also bei unserer Modellbewertung keine entscheidende Rolle spielen. Um dies zu testen, ersetzen wir die Initialisierung des Pseudozufallszahlengenerators in unserem Quellcode aus Abschnitt 6.2 durch:

```
set . seed (123)
```

Anschließend führen wir erneut die Rastersuche für die in (7.7) genannten Parameterwerte durch und kommen auf den kleinstmöglichen Verlust von 36,07525 für

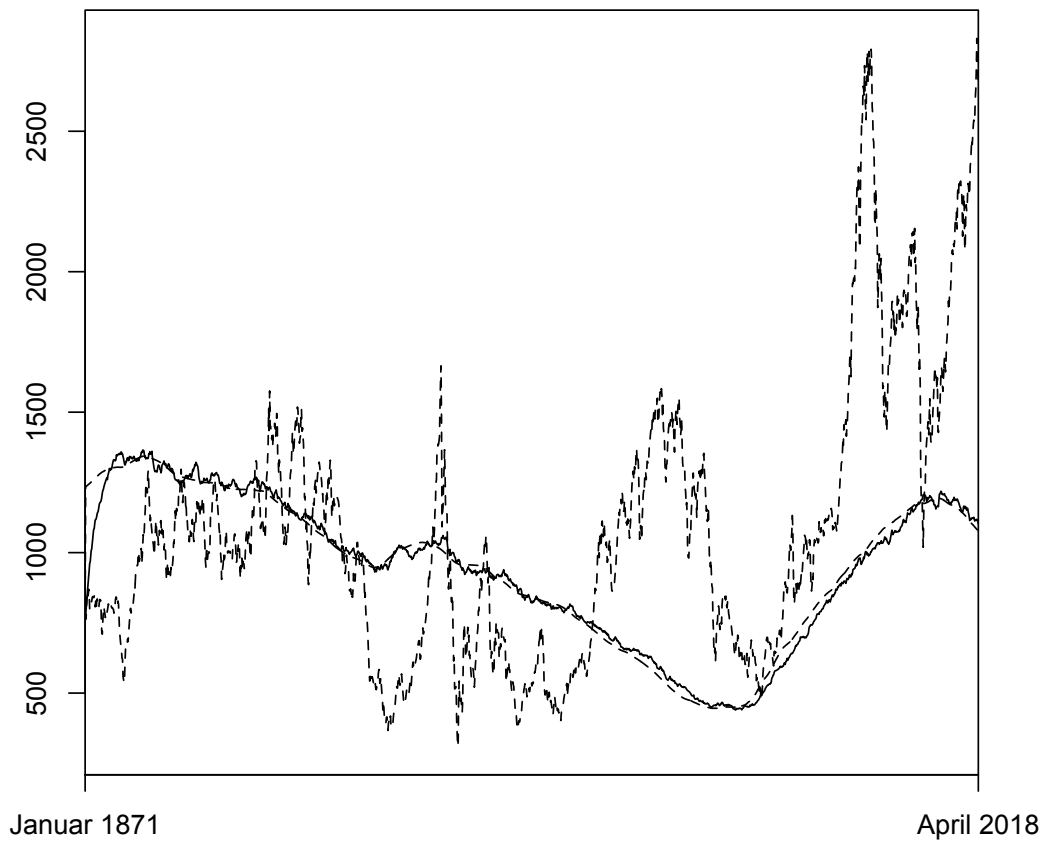


Abbildung 7.3: Vergleich der realen, trendbereinigten S&P 500-Preise (gestrichelte Linie), der geschätzten Fundamentalwerte (gepunktete Linie) und einer Simulationszeitreihe der Parameterwahl $\Theta_{\text{Nelder-Mead}}^*$ (durchgezogene Linie) zwischen Januar 1871 und April 2018: Die simulierten Preise befinden sich stets in der Nähe der geschätzten Fundamentalwerte

die Parameterkombination

$$\Theta_{\text{Rastersuche, seed}(123)}^* := (0, 1; 2, 5; 2; 0, 6; 0, 1; 0, 1).$$

Nutzen wir diese Parameter wieder als Startpunkt des Nelder-Mead Algorithmus, dann kommen wir zu einer Verbesserung des Verlusts auf 15,84329 für die Parameterwahl

$$\Theta_{\text{Nelder-Mead, seed}(123)}^* := (0, 05708742; 2, 73702901; 1, 89806603; \\ 0, 64501615; 0, 11569983; 0, 06238965).$$

In Tabelle 7.1 sind die Verlustfunktionswerte je nach Initialisierung des Pseudozufallszahlengenerators übersichtlich dargestellt. Wir können feststellen, dass die Initialisierung durchaus einen Einfluss auf die Bewertung von Parameterkombinationen hat. Im Wesentlichen bewerten wir aber aufgrund der relativ ähnlichen Verlustfunktionswerte alle vier Kombinationen als *gute* Parameterwahlen. Wir können außerdem feststellen, dass alle vier Parameterkombinationen ähnliche Parameterwerte beinhalten und dadurch auch ähnliche Simulationszeitreihen erzeugen. Insgesamt stellen wir also fest, dass die Initialisierung des Pseudozufallszahlengenerators bei der Modellbewertung durchaus die Bewertungsreihenfolge verschiedener Parameterkombinationen beeinflusst, aber vermutlich keine entscheidende Rolle für unser gesamtes Bewertungsverfahren spielt.

Tabelle 7.1: Verlustfunktionswerte der gefundenen Parameterkombinationen je Initialisierung des Pseudozufallszahlengenerators

Parameter	set.seed(10)	set.seed(123)
$\Theta_{\text{Rastersuche}}^*$	29,78677	38,07978
$\Theta_{\text{Nelder-Mead}}^*$	19,94634	23,51245
$\Theta_{\text{Rastersuche, seed}(123)}^*$	32,82262	36,07525
$\Theta_{\text{Nelder-Mead, seed}(123)}^*$	26,64509	15,84329

Zum Abschluss dieses Abschnitts möchte ich die Wirkung der einzelnen Parameter auf das Aussehen der Simulationszeitreihen untersuchen. Ziel ist es, ein Gefühl dafür zu entwickeln, welcher Parameter die simulierte Preisentwicklung auf welche Weise beeinflusst. Wir werden hierzu Veränderungen in den Parameterwerten vornehmen und anschließend die resultierenden Simulationszeitreihen optisch vergleichen. Dabei sollen die Abbildungen 7.4 und 7.5 helfen. Dort ist links jeweils dieselbe Simulationszeitreihe abgebildet, die mit dem Modellparameter $\Theta_{\text{Nelder-Mead}}^*$ erzeugt wurde. In der jeweils rechten Skizze wurde ein einziger Parameterwert verändert.

Sein veränderter Wert ist in der Überschrift der jeweiligen Skizze angegeben. Natürlich beziehen sich sämtliche Skizzen auf dieselben Zufallszahlenrealisierungen, sodass Unterschiede ausschließlich auf die veränderten Parameterwerte zurückzuführen sind.

Der Parameter θ_1 bestimmt, wie sensibel die Preise auf Nachfrageveränderungen reagieren. Bei einer Erhöhung des Wertes auf 0,2 erwarten wir also, dass die simulierten Preise im Zeitverlauf unruhiger sind. Dies lässt sich in der ersten Skizze der rechten Spalte von Abbildung 7.4 gut erkennen.

Der zweite Parameter θ_2 gibt an, wie stark Chartisten auf vergangene Preisänderungen reagieren und ihre Nachfrage anpassen. Erhöhen wir den Parameterwert auf 15, dann kommt es zu einer starken Nachfrageerhöhung der Chartisten, wenn in der vergangenen Periode die Preise gestiegen sind, und einer starken Angebotserhöhung bei einem Preisabfall in der vergangenen Periode. Ein kurzfristiger Aufwärtstrend im Preisverlauf setzt sich somit tendenziell länger fort und umgekehrt. Kurzfristige *Zacken* im Preisverlauf sollten nun also extremer ausfallen. Dies können wir gut in der zweiten, rechten Skizze in Abbildung 7.4 sehen.

Fundamentalisten haben im Gegensatz zu Chartisten eine eher stabilisierende Wirkung. Sie bewegen durch ihre Nachfrage den Preis in die Richtung des Fundamentalwertes, der im Zeitverlauf in der Regel keine kurzfristigen Schwankungen aufweist. Erhöhen wir also θ_3 wie in der dritten Skizze der rechten Spalte in Abbildung 7.4 auf 15, dann werden simulierte Preise stets schnell wieder zurück zum Fundamentalwertniveau gelenkt, sodass kurzfristige Schwankungen nur sehr schwach ausfallen.

Eine weitere, entscheidende Stellschraube des Modells sind die Marktanteile der Chartisten bzw. Fundamentalisten. Je größer der Chartistenanteil ist, desto stärker fallen kurzfristige Trends aus. Bei einem kleinen Anteil der Chartisten kehrt der simulierte Preis dagegen umso schneller zum Fundamentalwertniveau zurück. Besteht der Markt wie in der ersten, rechten Skizze von Abbildung 7.5 zu 99% aus Chartisten, dann erhalten wir einen unruhigen Preisverlauf, der mit dem Fundamentalwertverlauf kaum noch vereinbar ist.

Die letzten zwei Parameter σ_C und σ_F bestimmen, wie stark die Nachfrage der Chartisten bzw. Fundamentalisten durch den Zufall abgelenkt werden. Erhöht man diese Werte auf 0,3 wie in den unteren zwei Skizzen der rechten Spalte von Abbildung 7.5, dann kommt es in beiden Fällen zu verstärkten *Zacken* im Preisverlauf. Insbesondere führen hohe Werte von σ_C und σ_F dazu, dass sich mehrere

Simulationszeitreihen, die mit denselben Parameterwerten erzeugt wurden, stark unterscheiden können.

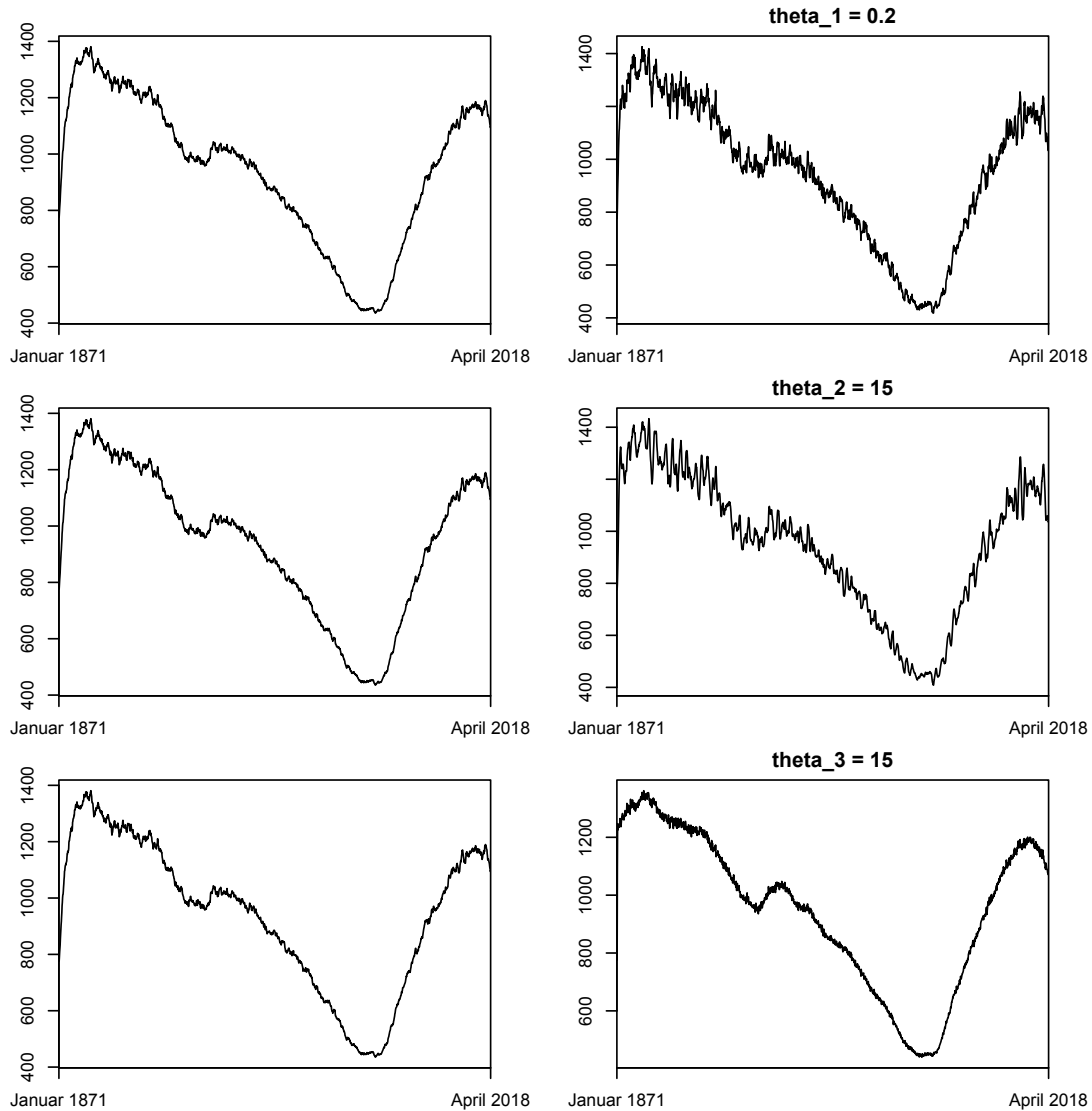


Abbildung 7.4: Veranschaulichung der Wirkung der Parameter θ_1 , θ_2 und θ_3 : In der linken Spalte ist dieselbe Simulationszeitreihe für die Parameterkombination $\Theta_{\text{Nelder-Mead}}^*$ abgebildet. Rechts wurde ein einziger Parameter verändert. Sein neuer Wert wird jeweils in der Abbildungsüberschrift angegeben. Bei allen Simulationszeitreihen werden dieselben Zufallszahlenrealisierungen verwendet, sodass Unterschiede ausschließlich auf verschiedenen Parameterwerte zurückzuführen sind

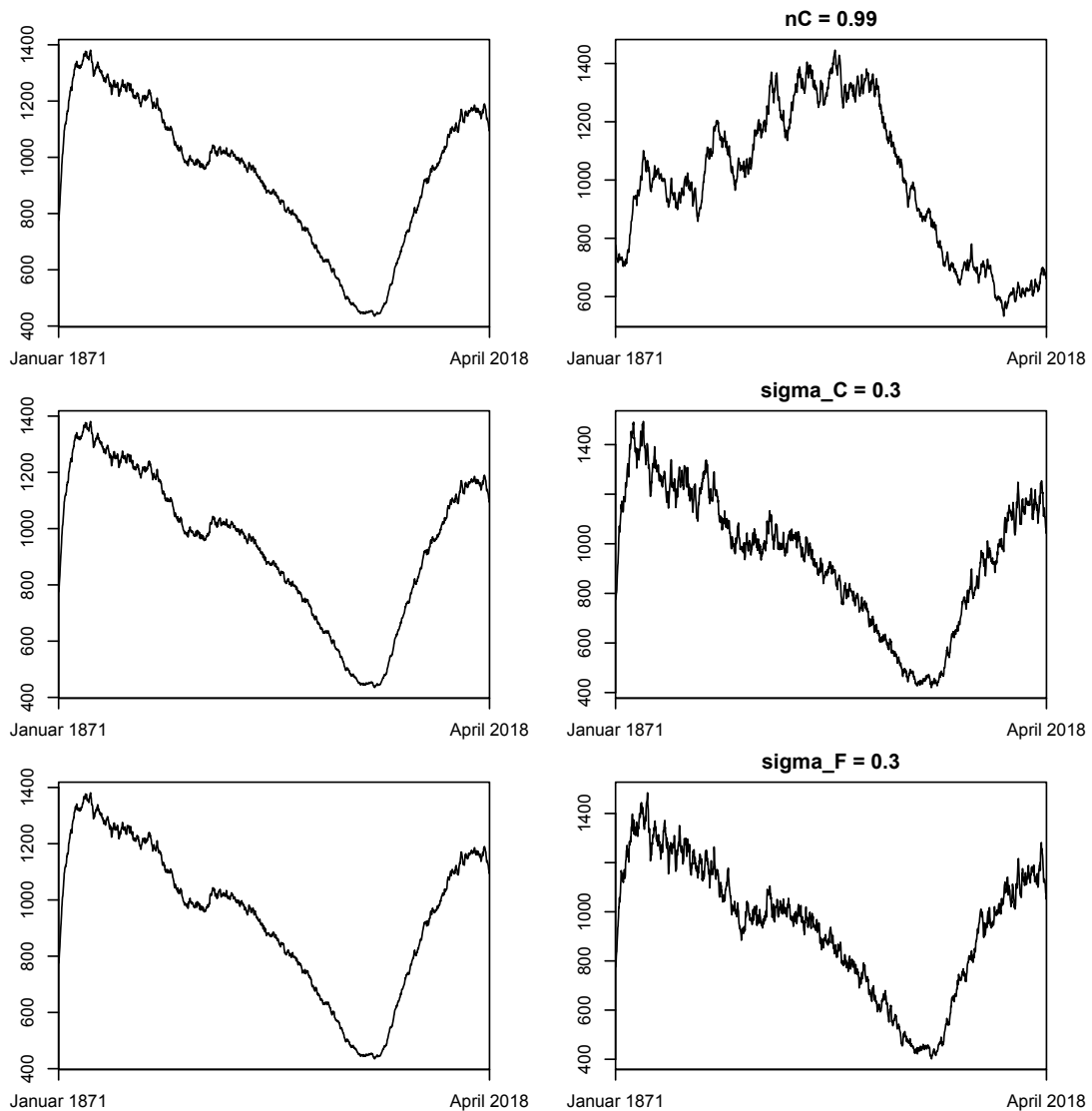


Abbildung 7.5: Veranschaulichung der Wirkung der Parameter n^C , σ_C und σ_F : In der linken Spalte ist dieselbe Simulationszeitreihe für die Parameterkombination $\Theta_{\text{Nelder-Mead}}^*$ abgebildet. Rechts wurde ein einziger Parameter verändert. Sein neuer Wert wird jeweils in der Abbildungsüberschrift angegeben. Bei allen Simulationszeitreihen werden dieselben Zufallszahlenrealisierungen verwendet, sodass Unterschiede ausschließlich auf verschiedenen Parameterwerte zurückzuführen sind

Wir haben nun ein Gefühl dafür entwickelt, welchen Einfluss die einzelnen Parameter auf das Aussehen der Simulationszeitreihe haben. Außerdem haben wir betrachtet, welche Rolle die Zufallszahlenziehung in unserer Modellbewertung spielt und mit $\Theta_{\text{Rastersuche}}^*$ und $\Theta_{\text{Nelder-Mead}}^*$ konnten wir zwei Parameterkombinationen finden, die zu einem geringen Verlustfunktionswert führen. Die damit erzeugten Simulationszeitreihen weichen nur gering voneinander ab und haben einen ähnlichen Verlauf wie unser geschätzter Fundamentalwert. Es bleibt offen, ob wir beispielsweise eine Maßzahl für den stilisierten Fakt *lang anhaltende Abweichung vom Fundamentalwert* in die Modellbewertung aufnehmen sollten, sodass Simulationszeitreihen guter Modelle, ähnlich wie unsere real-world Zeitreihe, über längere Zeiträume vom Fundamentalwertniveau abweichen. Eine weitere offene Fragestellung ist, ob mit einem Optimierungsalgorithmus andere Parameterkombinationen gefunden werden können, die zu einem geringeren, mittleren Verlust führen.

Kapitel 8

Fazit

Wir können nun zum Beispiel aus Motivation 0.1 zurückkommen. Es geht also um die Frage, wie wir die abgebildete Simulationszeitreihe aus Abbildung 1 bewerten können. Mit den in dieser Arbeit erläuterten Maßzahlen stilisierter Fakten und der Method of Simulated Moments-Verlustfunktion haben wir nun geeignete Mittel zur Bewertung kennengelernt. Wir werten demnach zunächst unsere insgesamt elf Maßzahlen an der betrachteten Simulationszeitreihe aus. Die Ergebniswerte sind in Tabelle 8.1 angegeben. Zum Vergleich sind dort nochmals die Maßzahlen ausgewertet an unserer real-world Zeitreihe des S&P 500-Aktienindex genannt. Wir können bereits bei der Betrachtung der Werte feststellen, dass diese stark voneinander abweichen.

Tabelle 8.1: Vergleich der ausgewerteten Maßzahlen an der real-world Zeitreihe und an der Simulationszeitreihe aus Abbildung 1

Maßzahl	real-world Wert	Simulationswert
ACF(returns), Lag 1	0,28	0,77
ACF(absolutreturns), Lag 1	0,21	0,55
ACF(absolutreturns), Lag 3	0,13	0,11
ACF(absolutreturns), Lag 6	0,15	0,01
ACF(absolutreturns), Lag 12	0,15	-0,01
ACF(absolutreturns), Lag 24	0,06	-0,02
Hurst Exponent	0,72	0,57
Kurtosis	20,20	3,07
Hill-Schätzer (2,5%)	3,32	5,44
Hill-Schätzer (5%)	3,63	4,30
Hill-Schätzer (10%)	2,78	3,49

Wenn wir nun die Method of Simulated Moments-Verlustfunktion (7.5) an den Maßzahlenvektor der Simulationszeitreihe auswerten, dann erhalten wir den Verlust 188,6394. Verglichen mit den berechneten Verlusten in Abschnitt 7.3 ist dieser Wert relativ hoch. Im Sinne der hier betrachteten Maßzahlen und deren mittels Bootstrapping ermittelten Gewichtung bewerten wir die betrachtete Simulationszeitreihe also als eher schlechte Simulation, obwohl sie bei einer rein optischen Betrachtung durchaus als gut empfunden werden kann. Die Simulationszeitreihe wurde im Übrigen durch das Marktmodell aus Abschnitt 7.1 mit den Parametern

$$\Theta = (0, 19; 4; 1, 5; 0, 99; 0, 1; 0, 1)$$

erzeugt.

Wir können an diesem Beispiel sehen, dass das Ergebnis unsers Bewertungsverfahrens nicht immer mit der optischen Einschätzung übereinstimmt. Das Verfahren bewertet Marktmodelle, die verglichen mit Zeitreihen aus der Realität recht stabile Simulationszeitreihen mit wenig kurzfristigen Schwankungen erzeugen, als gute Modelle und umgekehrt. Dies liegt unter anderem daran, dass wir uns bei der Bewertung auf eine konkrete Auswahl stilisierter Fakten und bestimmte Messungen für diese beschränken. Bei dieser Auswahl besteht natürlich Spielraum. Beispielsweise könnte es sinnvoll sein, eine Maßzahl für die lang anhaltende Abweichung vom Fundamentalwert ins Bewertungsverfahren mit aufzunehmen, damit wir nur diejenigen Modelle als gut bewerten, deren Simulationszeitreihen auch über längere Zeiträume vom Fundamentalwert abweichen.

Außerdem möchte ich an dieser Stelle darauf hinweisen, dass wir uns in dieser Arbeit auf statische Marktmodelle eingeschränkt haben. Es ist durchaus vorstellbar, dass durch die Hinzunahme von Dynamik weitaus bessere Ergebnisse erzielt werden können. Bei dynamischen Modellen können Akteure am Finanzmarkt im Zeitverlauf ihre Strategie ändern. Dadurch kann möglicherweise ein besserer Einklang zwischen lang anhaltenden Abweichungen vom Fundamentalwert und Rückkehr zum Fundamentalwertniveau erreicht werden, da die dynamischen Verschiebung der Marktanteile von Chartisten und Fundamentalisten gerade zu diesem Wechsel führen könnte.

Auch wenn es offenbar noch Verbesserungs- und Erweiterungspotential gibt, haben wir in dieser Arbeit ein allgemeines Verfahren kennengelernt, mit dem Marktmodelle bewertet werden können. Wir haben eine Idee entwickelt, wie Marktmodellanpassung funktionieren kann und können dieses Wissen als gute Grundlage für weiterführende Arbeiten in diesem Bereich nutzen.

Quellcodes

Der Vollständigkeit halber sollen in diesem Kapitel die Quellcodes zu den verwendeten Abbildungen und Tabellen bereitgestellt werden. Es handelt sich um Quellcodes für die Softwareumgebung R. Mit `prices`, `long_interest_rates`, `prices_real`, `dividends_real` und `returns` sind die eingelesenen Zeitreihen aus dem S&P 500-Datensatz von Shiller (2018) gemeint. Wie die Datensätze eingelesen werden können, wurde in Abschnitt 4.2 erläutert. Die folgenden Quellcodes bilden teilweise Fortsetzungen der in der Arbeit angegebenen Codes. Darauf werde ich in den jeweiligen Quellcodeüberschriften hinweisen. Außerdem werde ich dort vermerken, welche speziellen R-Pakete notwendig sind.

Listing 8.1: Quellcode zu Abbildung 1 und Tabelle 8.1 (Fortsetzung des Codes aus Abschnitt 7.2)

```
#Parameterwahl der abgebildeten Simulationszeitreihe aus der  
Motivation  
theta_1 <- 0.19  
theta_2 <- 4  
theta_3 <- 1.5  
nC <- 0.99  
sigma_C <- 0.1  
sigma_F <- 0.1  
  
#Berechnung des resultierenden Anteils der Fundamentalisten  
nF <- 1 - nC  
  
prices_simulation <- vector(length = (N+1))  
returns_simulation <- vector(length = N)  
prices_simulation[1] <- p_0  
demandC <- 0  
demandF <- 0  
  
#Berechnung der Simulationszeitreihe  
for(t in 1:N) {  
  prices_simulation[t+1] <- exp(log(prices_simulation[t])
```

```

+ theta_1 * (nC*demandC + nF*demandF))
returns_simulation[t] <- (prices_simulation[t+1] -
  prices_simulation[t])/prices_simulation[t]
demandC <- theta_2 * (log(prices_simulation[t+1]) - log(
  prices_simulation[t])) + sigma_C * epsilonC_standard
  _normal_distributed[t]
demandF <- theta_3 * (log(p_hat_star[t]) - log(prices_
  simulation[t+1])) + sigma_F * epsilonF_standard_
  normal_distributed[t]
}

#Berechnung der Maßzahlen ausgewertet an der
  Simulationszeitreihe
simulation_moments <- evaluate_moments(returns_simulation)

#Erzeugung der Abbildung
par(mfrow = c(1,2), mar = c(2,3,2,3))
plot(c(1:(N+1)), c(p_0,prices_real_detrended), type = 'l', xlab
  = '', ylab = '', xaxt = 'n', xaxs = 'i', ylim = c(250,2900),
  main = 'Realität')
axis(1, at = c(1,N+1), labels = c('_Januar_1871_', '_April_2018_'
  ))
plot(c(1:(N+1)), prices_simulation, type = 'l', xlab = '', ylab
  = '', xaxt = 'n', xaxs = 'i', ylim = c(250,2900), main = '
  Simulation')
axis(1, at = c(1,N+1), labels = c('_Januar_1871_', '_April_2018_'
  ))

#Auswertung der Verlustfunktion
t(simulation_moments-empirical_moments) %*% W %*% (simulation_
  moments-empirical_moments)
#Ergebnis: 188.6394

#Ausgabe der real-world Maßzahlen und der simulierten Maßzahlen
round(empirical_moments,2)
#Ergebnis: 0.28 0.21 0.13 0.15 0.15 0.06 0.72 20.20 3.32
  3.63 2.78
round(simulation_moments,2)
#Ergebnis: 0.77 0.55 0.11 0.01 -0.01 -0.02 0.57 3.07 5.44
  4.30 3.49

```

Listing 8.2: Quellcode zu Abbildung 2.1

```
#Betrachtung der letzten 97 Preise, d.h. Preise für April 2010
bis April 2018
pricesApril2010toApril2018 <- prices[(length(prices)-96):length(
  prices)]
x <- c(1:97)
plot(x, pricesApril2010toApril2018, type='l', main='', xlab='', ylab=
  '', xaxt='n', xaxs='i')
z <- seq(1,97,12)
x_text <- c('April_2010', 'April_2011', 'April_2012', 'April_2013',
  'April_2014', 'April_2015', 'April_2016', 'April_2017', 'April_
  2018')
axis(1, at=z, labels=x_text, col.axis='black', las=2)
```

Listing 8.3: Quellcode zu Abbildung 2.2

```
plot(density(returns), xlim = c(-0.2,0.2), main = '', xlab = '',
  ylab = '')
x <- seq(-0.3, 0.3, 0.001)
y <- dnorm(x, mean(returns), sd(returns))
lines(x, y, lty = 'dashed')
```

Listing 8.4: Quellcode zu Abbildung 3.1 (Paket `stats` notwendig)

```
acorr <- acf(returns, lag = 120, type = 'correlation', plot =
  TRUE, main = '')
#der folgende Befehl liefert den Wert 0,282 der
Autokorrelationsfunktion für das Lag 1
acorr[1]
```

Listing 8.5: Quellcode zu Abbildung 3.2 und den Werten aus Tabelle 3.1 (Paket `stats` notwendig)

```
absolutreturns <- abs(returns)
squaredreturns <- returns * returns
par(mfrow = c(1,2))
acorr_abs <- acf(absolutreturns, lag = 120, type = 'correlation',
  plot = TRUE, main = '')
acorr_square <- acf(squaredreturns, lag = 120, type = '
  correlation', plot = TRUE, main = '')
#um die Werte der Autokorrelationsfunktion der Absoluterträge
für die Lags 1,3,6,12 und 24 abzulesen, kann acorr_abs
ausgegeben werden:
acorr_abs
```

Listing 8.6: Quellcode zu Abbildung 3.3

```

logn <- c(1:5)
logRSn <- c(1.9, 3.2, 3.8, 4.9, 6.1)
x <- c(0:6)
y <- c(1:7)
plot(logn, logRSn, xlab = 'log(n)', ylab = 'log((R/S)n)')
lines(x, y, type = 'l')

```

Listing 8.7: Quellcode zu Abbildung 3.4

```

expectationAnisLloyd <- function(n) {
  prefactor <- (n-0.5)/n
  sum <- 0
  for(i in 1:(n-1)) {
    sum <- sum + sqrt((n-i)/i)
  }
  if(n <= 340)
    return(prefactor * (gamma((n-1)/2)/(sqrt(pi) *
      gamma(n/2))) * sum)
  else
    return(prefactor * (1/sqrt(n * pi/2)) * sum)
}
n <- seq(10,10000,1)
exact <- vector(length = length(n))
for(i in 1:length(n)) {
  exact[i] <- expectationAnisLloyd(n[i])
}
asymptotic <- sqrt(0.5 * pi * n)
plot(log(n), log(asymptotic), type = 'l', lty = 2, ylab = 'log(E
  (R/S)n)', xlim = c(2.6, 8.9))
text(6.1, 3.4, 'asymptotisch, Steigung 0.5', srt = 37)
lines(log(n), log(exact))
text(6.3, 3.2, 'exakt', srt = 37)

```

Listing 8.8: Quellcode zu Abbildung 3.5 (Paket `pracma` notwendig)

```

#Definition der lokalen Variablen der Funktion hurstexp
x <- abs(returns)
d <- 50 #entspricht dmin aus hurstexp

#Sourcecode von hurstexp aus dem Paket pracma:
d <- max(2, floor(d[1]))
N <- length(x)
if (N%%2 != 0) {

```

```

        x <- c(x, (x[N-1] + x[N])/2)
        N <- N + 1
    }
    rssimple <- function(x) {
        n <- length(x)
        y <- x - mean(x)
        s <- cumsum(y)
        rs <- (max(s) - min(s))/sd(x)
        log(rs)/log(n)
    }
    rscal <- function(z, n) {
        m <- length(z)/n
        y <- matrix(x, n, m)
        e <- apply(y, 2, mean)
        s <- apply(y, 2, std)
        for (i in 1:m) y[, i] <- y[, i] - e[i]
        y <- apply(y, 2, cumsum)
        mm <- apply(y, 2, max) - apply(y, 2, min)
        return(mean(mm/s))
    }
    divisors <- function(n, n0 = 2) {
        n0n <- n0:floor(n/2)
        dvs <- n0n[n%%n0n == 0]
        return(dvs)
    }
    N <- length(x)
    dmin <- d
    N0 <- min(floor(0.99 * N), N - 1)
    N1 <- N0
    dv <- divisors(N1, dmin)
    for (i in (N0 + 1):N) {
        dw <- divisors(i, dmin)
        if (length(dw) > length(dv)) {
            N1 <- i
            dv <- dw
        }
    }
    OptN <- N1
    d <- dv
    x <- x[1:OptN]
    N <- length(d)
    RSe <- numeric(N)
    for (i in 1:N) RSe[i] <- rscal(x, d[i])

```

```

Pe <- polyfit(log10(d), log10(RSe), 1)

#Erzeugung der Zeichnung:
plot(log10(d), log10(RSe), xlab='log10(n)', ylab='log10((R/S)n)'
)
d_extended <- c(50, d, 1000)
lines(log10(d_extended), Pe[2] + Pe[1] * log10(d_extended), lty=
'dashed')
text(2.4, 1.55, 'lineare_Regressionsgerade', srt=42)

```

Listing 8.9: Quellcode zu Tabelle 3.3 mit einer Funktion zur Auswertung der empirischen Konfidenzintervalle von Weron (2002)

```

confidenceintervalWeron <- function(N) {
  K <- log2(N)
  lowerbound10 <- 0.5 - exp(-7.35 * log(log(K)) + 4.06)
  upperbound10 <- exp(-7.07 * log(log(K)) + 3.75) + 0.5
  lowerbound5 <- 0.5 - exp(-7.33 * log(log(K)) + 4.21)
  upperbound5 <- exp(-7.20 * log(log(K)) + 4.04) + 0.5
  lowerbound1 <- 0.5 - exp(-7.19 * log(log(K)) + 4.34)
  upperbound1 <- exp(-7.51 * log(log(K)) + 4.58) + 0.5
  cat("10%:_(", lowerbound10, ", ", upperbound10, ")", "\n")
  cat("5%:_(", lowerbound5, ", ", upperbound5, ")", "\n")
  cat("1%:_(", lowerbound1, ", ", upperbound1, ")", "\n")
}
#nach der Verkürzung hat die Zeitreihe der Absoluterträge die
  Länge 1764
confidenceintervalWeron(1764)

```

Listing 8.10: Quellcode zu Abbildung 3.6

```

x <- seq(-8, 8, 0.01)

par(mfrow=c(1,2))

plot(x, dt(x,3), type='l', ylim=c(0,0.4), ylab='Dichte', xlab='')
text(0, 0.4, 'steilgipflig')
lines(x, dnorm(x,sd=sqrt(3)), lty='dashed')

plot(x, dunif(x,-3,3), type='l', ylim=c(0,0.4), ylab='Dichte',
xlab='')
text(0, 0.4, 'flachgipflig')
lines(x, dnorm(x,sd=sqrt(3)), lty='dashed')

```

Listing 8.11: Quellcode zu Abbildung 3.7 (Paket `evd` notwendig)

```
x <- seq(-10, 10, 0.1)
par(mfrow=c(2,3), oma=c(2,2,2,2), mar=c(5,2,2,2))
plot(x, pgumbel(x), type='l', ylab='', xlab='', main='Gumbel')
plot(x, pfrechet(x), type='l', ylab='', xlab='', main='Frechet',
      alpha=1')
plot(x, prweibull(x), type='l', ylab='', xlab='', main='Weibull',
      alpha=1')
plot(x, dgumbel(x), type='l', ylab='', xlab='', main='Gumbel')
plot(x, dfrechet(x), type='l', ylab='', xlab='', main='Frechet',
      alpha=1')
plot(x, drweibull(x), type='l', ylab='', xlab='', main='Weibull',
      alpha=1')
mtext('Verteilungsfunktionen', side=3, line=0, outer=TRUE)
mtext('Dichtefunktionen', side=3, line=-24, outer=TRUE)
```

Listing 8.12: Quellcode zu Abbildung 4.1

```
plot(c(1,2), c(48.17,48.93), type='l', xlab='', ylab='Dividende',
      xaxt='n', yaxt='n', xaxs='i')
x <- c(1, 1+1/3, 1+2/3,2)
x_text <- c('Sep_2017', 'Okt_2017', 'Nov_2017', 'Dez_2017')
axis(1, at=x, labels=x_text, col.axis='black', las=2)
y<-c(48.17 * 2/3 + 48.93 * 1/3, 48.17 * 1/3 + 48.93 * 2/3)
axis(2, at=y, labels=round(y,2), col.axis='black', las=2)
lines(c(1+1/3,1+1/3), c(48.17*2/3+48.93*1/3,0), lty=2)
lines(c(1+2/3,1+2/3), c(48.17*1/3+48.93*2/3,0), lty=2)
lines(c(0,1+1/3), c(48.17*2/3+48.93*1/3,48.17*2/3+48.93*1/3),
      lty=2)
lines(c(0,1+2/3), c(48.17*1/3+48.93*2/3,48.17*1/3+48.93*2/3),
      lty=2)
```

Listing 8.13: Quellcode zu Abbildung 5.2 und Tabelle 5.1 (Paket `pracma` notwendig)

```
N <- length(prices_real)
x <- seq(1:N)

calculate_R_squared <- function(lin_polynom, regressor,
  observation_dependent_variable) {
  ESS <- 0
  for(i in 1:length(regressor)) { ESS <- ESS + (lin_polynom[2]
    + lin_polynom[1] * regressor[i] - mean(observation_
    dependent_variable))^2 }
}
```

```

TSS <- 0
for(i in 1:length(observation_dependent_variable)) { TSS <-
  TSS + (observation_dependent_variable[i] - mean(
    observation_dependent_variable))^2 }
R_squared <- ESS/TSS
R_squared
}

#Lineare Regression der realen Preise ergibt R_squared=0.5922143
polynom_prices <- polyfit(x, prices_real, 1)
calculate_R_squared(polynom_prices, x, prices_real)

#Schätzung des exponentiellen Zusammenhangs der realen Preise
  durch lineare Regression der logarithmierten Gleichung
  ergibt R_squared=0.7822638
polynom_log_prices <- polyfit(x, log(prices_real), 1)
calculate_R_squared(polynom_log_prices, x, log(prices_real))

#Lineare Regression der realen Dividenden ergibt R_squared
  =0.739724
polynom_dividends <- polyfit(x, dividends_real, 1)
calculate_R_squared(polynom_dividends, x, dividends_real)

#Schätzung des exponentiellen Zusammenhangs der realen
  Dividenden durch lineare Regression der logarithmierten
  Gleichung ergibt R_squared=0.8518733
polynom_log_dividends <- polyfit(x, log(dividends_real), 1)
calculate_R_squared(polynom_log_dividends, x, log(dividends_real
))

#Erstellung der Abbildung
par(mfrow = c(1,2))

plot(x, prices_real, type='l', main='reale_Preise', ylab='',
  xlab='', xaxt='n', xaxs='i')
z <- c(1,N)
x_text <- c('Februar_1871', 'April_2018')
axis(1, at = z, labels = x_text, col.axis = 'black')
lines(x, polynom_prices[2] + polynom_prices[1] * x, type = 'l',
  lty = 'dotted')
lines(x, exp(polynom_log_prices[2] + polynom_log_prices[1] * x),
  type = 'l', lty = 'dashed')

```



```

plot(x, dividends_real, type = 'l', main = 'reale_Dividenden',
     ylab = '', xlab = 'l', xaxt = 'n', xaxs = 'i')
axis(1, at = z, labels = x_text, col.axis = 'black')
lines(x, polynom_dividends[2] + polynom_dividends[1] * x, type =
      'l', lty = 'dotted')
lines(x, exp(polynom_log_dividends[2] + polynom_log_dividends[1]
            * x), type = 'l', lty = 'dashed')

```

Listing 8.14: Quellcode zu Abbildung 5.4 (Fortsetzung des Codes aus Abschnitt 5.2.1)

```

par(mfrow = c(1,2))

plot(x, prices_real, type = 'l', main = 'reale_Preise', ylab = '
', xlab = '', xaxt = 'n', xaxs = 'i')
z <- c(1,N)
x_text <- c('Februar_1871', 'April_2018')
axis(1, at = z, labels = x_text, col.axis = 'black')
lines(x, rep(mean(prices_real),N), type = 'l', lty = 'dashed')

plot(x, prices_real_detrended, type = 'l', main = 'reale_
trendbereinigte_Preise', ylab = '', xlab = '', xaxt = 'n',
xaxs = 'i')
axis(1, at = z, labels = x_text, col.axis = 'black')
lines(x, rep(mean(prices_real_detrended),N), type = 'l', lty = '
dashed')

```

Listing 8.15: Quellcode zu Abbildung 6.1

```

x <- c(seq(0,1.5,0.5), seq(1.6,4,0.1), seq(4.05,6,0.05), seq
      (6.1,8.5,0.1), seq(9,10,0.5))

par(mfrow=c(1,2))

hist(x, main='', xlab='')
lines(rep(5.7,par('usr')[4]+2), c(0:(par('usr')[4]+1)), lty='
dotted', lwd=2)

hist(x, main='', xlab='')
lines(rep(9.2,par('usr')[4]+2), c(0:(par('usr')[4]+1)), lty='
dotted', lwd=2)

```

Listing 8.16: Quellcode zu Abbildung 7.1 und 7.2 (Fortsetzung des Codes aus Abschnitt 7.2)

```

par(mfrow=c(1,2))
show_two_simulations <- function(Theta) {
  theta_1 <- Theta[1]
  theta_2 <- Theta[2]
  theta_3 <- Theta[3]
  nC <- Theta[4]
  sigma_C <- Theta[5]
  sigma_F <- Theta[6]

  #Berechnung des resultierenden Anteils der
  Fundamentalisten
  nF <- 1 - nC

  prices_simulation <- vector(length = (N+1))
  prices_simulation[1] <- p_0
  demandC <- 0
  demandF <- 0
  for(i in 1:2) {
    #Berechnung der Simulationszeitreihe
    for(t in 1:N) {
      prices_simulation[t+1] <- exp(log(prices_
        _simulation[t]) + theta_1 * (nC*
        demandC + nF*demandF))
      demandC <- theta_2 * (log(prices_
        simulation[t+1]) - log(prices_
        simulation[t])) + sigma_C * epsilonC
        _standard_normal_distributed[(i-1)*N
        +t]
      demandF <- theta_3 * (log(p_hat_star[t])
        - log(prices_simulation[t+1])) +
        sigma_F * epsilonF_standard_normal_
        distributed[(i-1)*N+t]
    }
    demandC <- 0
    demandF <- 0

    #Erzeugung der Abbildung
    plot(c(1:(N+1)), prices_simulation, type = 'l',
      xlab = '', ylab = '', xaxt = 'n', xaxs = 'i'
    )
    axis(1, at = c(1,N+1), labels = c('Januar_1871',

```

```

        'April_2018'))
    }
}

show_two_simulations(c(0.07188020, 2.65214729, 1.54208764,
  0.51866776, 0.09583929, 0.10508706)
show_two_simulations(c(0.05,4.00,1.50,0.95,0.25,0.10))

```

Listing 8.17: Quellcode zu Abbildung 7.3 (Fortsetzung des Codes aus Abschnitt 7.2)

```

show_simulation_and_fundamental_value_and_real_prices <-
function(Theta) {
  theta_1 <- Theta[1]
  theta_2 <- Theta[2]
  theta_3 <- Theta[3]
  nC <- Theta[4]
  sigma_C <- Theta[5]
  sigma_F <- Theta[6]

  #Berechnung des resultierenden Anteils der
  Fundamentalisten
  nF <- 1 - nC

  prices_simulation <- vector(length = (N+1))
  prices_simulation[1] <- p_0
  demandC <- 0
  demandF <- 0
  for(t in 1:N) {
    prices_simulation[t+1] <- exp(log(prices_
      simulation[t]) + theta_1 * (nC*demandC + nF*
      demandF))
    demandC <- theta_2 * (log(prices_simulation[t
      +1]) - log(prices_simulation[t])) + sigma_C
      * epsilonC_standard_normal_distributed[(i-1)
      *N+t]
    demandF <- theta_3 * (log(p_hat_star[t]) - log(
      prices_simulation[t+1])) + sigma_F *
      epsilonF_standard_normal_distributed[(i-1)*N
      +t]
  }

  #Plot der berechneten Simulationszeitreihen im Vergleich
  mit den realen, trendbereinigten Preisen und den

```

```

        geschätzten Fundamentalwerten
    par(mfrow=c(1,1))
    plot(c(1:(N+1)), c(p_0,prices_real_detrended), type = 'l',
        , xlab = '', ylab = '', xaxt = 'n', xaxs = 'i',lty=
        'dashed')
    axis(1, at = c(1,N+1), labels = c('Januar_1871','April_
        2018'))
    lines(c(1:(N+1)), c(p_0,p_hat_star), lty = 'dashed')
    lines(c(1:(N+1)),prices_simulation ,lty='solid')
}

show_simulation_and_fundamental_value_and_real_prices(c
(0.07188020, 2.65214729, 1.54208764, 0.51866776, 0.09583929,
0.10508706))

```

Listing 8.18: Quellcode zu Abbildung 7.4 und 7.5 (Fortsetzung des Codes aus Abschnitt 7.2)

```

show_one_simulation <- function(Theta) {
    theta_1 <- Theta[1]
    theta_2 <- Theta[2]
    theta_3 <- Theta[3]
    nC <- Theta[4]
    sigma_C <- Theta[5]
    sigma_F <- Theta[6]

    #Berechnung des resultierenden Anteils der
        Fundamentalisten
    nF <- 1 - nC

    prices_simulation <- vector(length = (N+1))
    prices_simulation[1] <- p_0
    demandC <- 0
    demandF <- 0
    #Berechnung der Simulationszeitreihe
    for(t in 1:N) {
        prices_simulation[t+1] <- exp(log(prices_
            simulation[t]) + theta_1 * (nC*demandC + nF*
            demandF))
        demandC <- theta_2 * (log(prices_simulation[t
            +1]) - log(prices_simulation[t])) + sigma_C
            * epsilonC_standard_normal_distributed[(i-1)
            *N+t]
        demandF <- theta_3 * (log(p_hat_star[t]) - log(

```

```

        prices_simulation[t+1])) + sigma_F *
        epsilonF_standard_normal_distributed[(i-1)*N
        +t]
    }
    plot(c(1:(N+1)),prices_simulation,type='l', xlab = '',
        ylab = '', xaxt = 'n', xaxs = 'i')
    axis(1, at = c(1,N+1), labels = c('Januar_1871','
        April_2018'))
}
par(mfrow=c(3,2),mar=c(2,3,2,3))
#Standardparameterwahl ist Theta^*_Nelder-Mead
Theta_star_NelderMead <- c(0.07188020, 2.65214729, 1.54208764,
    0.51866776, 0.09583929, 0.10508706)
show_one_simulation(Theta_star_NelderMead)
show_one_simulation(c(0.2,2.75,1.5,0.5,0.1,0.1))
title('theta_1=0.2')
show_one_simulation(Theta_star_NelderMead)
show_one_simulation(c(0.1,15,1.5,0.5,0.1,0.1))
title('theta_2=15')
show_one_simulation(Theta_star_NelderMead)
show_one_simulation(c(0.1,2.75,15,0.5,0.1,0.1))
title('theta_3=15')
show_one_simulation(Theta_star_NelderMead)
show_one_simulation(c(0.1,2.75,1.5,0.99,0.1,0.1))
title('nC=0.99')
show_one_simulation(Theta_star_NelderMead)
show_one_simulation(c(0.1,2.75,1.5,0.5,0.3,0.1))
title('sigma_C=0.3')
show_one_simulation(Theta_star_NelderMead)
show_one_simulation(c(0.1,2.75,1.5,0.5,0.1,0.3))
title('sigma_F=0.3')

```


Literaturverzeichnis

- M. Abramowitz und I. A. Stegun. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. United States Department Of Commerce, 10 edition, 1972.
- A. A. Anis und E. H. Lloyd. The expected value of the adjusted rescaled Hurst range of independent normal summands. *Biometrika*, 63(1):111–116, 1976.
- Board of Governors of the Federal Reserve System. Webseite. <https://www.federalreserve.gov/releases/h15/>, 30. Mai 2018.
- Bureau of Labor Statistics. Handbook of Methods, Chapter 17. The Consumer Price Index. 2018.
- A. Christmann. Skript zur Vorlesung Einführung in die Statistik. Universität Bayreuth, 2016.
- R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.
- A. Cowles. *Common Stock Indexes*. Principia Press, 1939.
- B. Efron. Bootstrap methods: Another look at the Jackknife. *The annals of statistics*, 7(1):1–26, 1979.
- R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 50(4):987–1007, 1982.
- E. F. Fama. Mandelbrot and the stable Paretian hypothesis. *The journal of business*, 36(4):420–429, 1963.
- Federal Reserve Economic Data (FRED). Webseite FRED. <https://fred.stlouisfed.org/series/GS10>, 30. Mai 2018.

- R. A. Fisher und L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press, 1928.
- R. Franke und F. Westerhoff. Structural stochastic volatility in asset pricing dynamics: Estimation and model contest. BERG working paper series on government and growth, No. 78, ISBN 978-3-931052-88-1, 2011.
- B. Gnedenko. Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of mathematics*, pages 423–453, 1943.
- Gruppe Deutsche Boerse. Webseite deutsche-boerse.com Suchbegriff Dividende. <http://deutsche-boerse.com/dbg-de/ueber-uns/services/know-how/boersenlexikon/boersenlexikon-article/Dividende/2561844>, 7. Juni 2018a.
- Gruppe Deutsche Boerse. Webseite deutsche-boerse.com Suchbegriff Marktkapitalisierung. <http://deutsche-boerse.com/dbg-de/ueber-uns/services/know-how/boersenlexikon/boersenlexikon-article/Marktkapitalisierung/2561970>, 7. Juni 2018b.
- K. H. Hamed. Improved finite-sample Hurst exponent estimates using rescaled range analysis. *Water resources research*, 43(4), 2007.
- J. Hartung, B. Elpelt, und K.-H. Klösener. *Statistik Lehr- und Handbuch der angewandten Statistik*, volume 9. R. Oldenbourg Verlag München Wien, 1993.
- B. M. Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 3(5):1163–1174, 1975.
- S. Homer und R. E. Sylla. *A history of interest rates*. Rutgers University Press, 1996.
- C. H. Hommes. Heterogeneous Agent Models in Economics and Finance. Tinbergen Institute Discussion Paper, No. 05-056/1, 2005.
- C. H. Hommes. Interacting Agents in Finance. Tinbergen Institute Discussion Paper, No. 06-029/1, 2006.
- H. E. Hurst. Long term storage capacity of reservoirs. *ASCE Transactions*, 116 (776):770–808, 1951.

- Y. Ivanova, C. Neely, D. Rapach, und P. Weller. Can risk explain the profitability of technical trading in currency markets? *Federal Reserve Bank of St. Louis Working Paper Series*, 2014.
- C. M. C. Lee, J. Myer, und B. Swaminathan. What is the Intrinsic Value of the Dow? *The Journal of Finance*, 54(5):1693–1741, October 1999.
- T. Lux. Stochastic behavioral asset pricing models and the stylized facts. Technical report, Economics Working Paper, Kiel, 2008.
- T. Lux und M. Ausloos. Market Fluctuations I: Scaling, Multiscaling, and Their Possible Origins. In *The Science of Disasters*, chapter 13. Springer, Berlin, Heidelberg, 2002.
- E. Mammen. *When Does Bootstrap Work?: Asymptotic Results and Simulations*, volume 77. Springer-Verlag New York, 1 edition, 1992.
- B. B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419, October 1963.
- B. B. Mandelbrot und J. R. Wallis. Noah, Joseph, and operational hydrology. *Water resources research*, 4(5):909–918, 1968.
- E. E. Peters und D. Peters. *Fractal market analysis: applying chaos theory to investment and economics*, volume 24. John Wiley & Sons, 1994.
- R. Schlittgen und B. H. J. Streitberg. *Zeitreihenanalyse*, volume 8. Oldenbourg Verlag, 1999.
- R. J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, 71(3):421–436, June 1981.
- R. J. Shiller. *Market volatility*, volume 3. MIT press, 1989.
- R. J. Shiller. Website. <http://www.econ.yale.edu/shiller/data.htm>, 23. April 2018.
- J. J. Siegel. The Shiller CAPE ratio: A new look. *Financial Analysts Journal*, 72(3):41–50, 2016.
- M. E. Solari und A. A. Anis. The mean and variance of the maximum of the adjusted partial sums of a finite number of independent normal variates. *The Annals of Mathematical Statistics*, 28(3):706–716, 1957.

Statistisches Bundesamt. Webseite. <https://www.destatis.de/DE/ZahlenFakten/GesamtwirtschaftUmwelt/Preise/Verbraucherpreisindizes/Methoden/verbraucherpreisindex.html>, 29. Mai 2018.

G. F. Warren und F. A. Pearson. *Gold and Prices*. New York: John Wiley and Sons, 1935.

R. Weron. Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A: Statistical Mechanics and its Applications*, 312(1–2): 285–299, 2002.

J. B. Williams. *The theory of investment value*, volume 36. Harvard university press Cambridge, 1938.

Finance Yahoo. Webseite. <https://finance.yahoo.com/>, 24. Mai 2018.

Selbständigkeitserklärung Hiermit erkläre ich, dass ich die Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die Arbeit nicht bereits zur Erlangung eines akademischen Grades eingereicht habe.

A handwritten signature in blue ink, appearing to read 'A. Schuppel', written in a cursive style.

Bayreuth, 27. November 2018