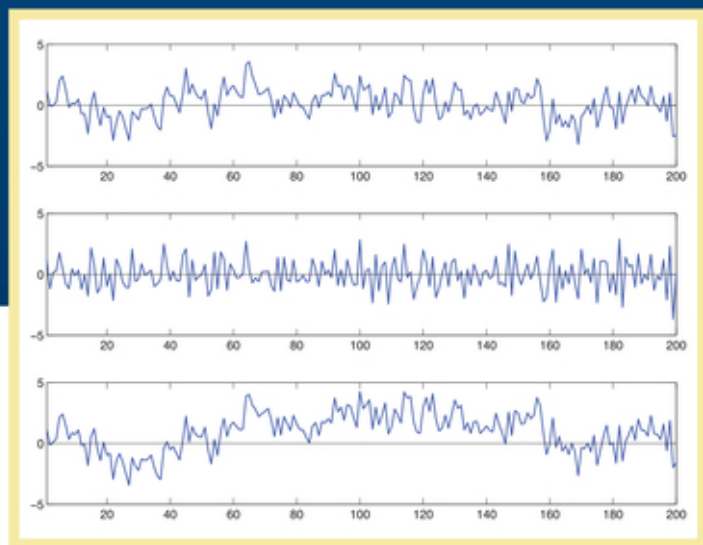


Wiley Series in Probability and Statistics

TIME SERIES ANALYSIS WITH LONG MEMORY IN VIEW

UWE HASSLER



WILEY

Time Series Analysis with Long Memory in View

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by *Walter A. Shewhart* and *Samuel S. Wilks*

Editors: *David J. Balding, Noel A. C. Cressie, Garrett M. Fitzmaurice, Geof H. Givens, Harvey Goldstein, Geert Molenberghs, David W. Scott, Adrian F. M. Smith, and Ruey S. Tsay*

Editors Emeriti: *J. Stuart Hunter, Iain M. Johnstone, Joseph B. Kadane, and Jozef L. Teugels*

The **Wiley Series in Probability and Statistics** is well established and authoritative. It covers many topics of current research interest in both pure and applied statistics and probability theory. Written by leading statisticians and institutions, the titles span both state-of-the-art developments in the field and classical methods.

Reflecting the wide range of current research in statistics, the series encompasses applied, methodological and theoretical statistics, ranging from applications and new techniques made possible by advances in computerized practice to rigorous treatment of theoretical approaches. This series provides essential and invaluable reading for all statisticians, whether in academia, industry, government, or research.

A complete list of titles in this series can be found at
<http://www.wiley.com/go/wsp>

Time Series Analysis with Long Memory in View

Uwe Hassler

Goethe University Frankfurt

WILEY

This edition first published 2019

© 2019 John Wiley & Sons, Inc

Edition History

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Uwe Hassler be identified as the author of the material in this work has been asserted in accordance with law.

Registered Office(s)

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

Editorial Office

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty

In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data

Names: Hassler, Uwe, author.

Title: Time series analysis with long memory in view / Uwe Hassler.

Description: 1st edition. | Hoboken, NJ : John Wiley & Sons, 2019. | Series:

Wiley series in probability and statistics |

Identifiers: LCCN 2018023760 (print) | LCCN 2018036671 (ebook) | ISBN

9781119470281 (Adobe PDF) | ISBN 9781119470427 (ePub) | ISBN 9781119470403 (hardcover)

Subjects: LCSH: Time-series analysis.

Classification: LCC QA280 (ebook) | LCC QA280 .H385 2019 (print) | DDC 519.5/5—dc23

LC record available at <https://lcn.loc.gov/2018023760>

Cover Design: Wiley

Cover Image: Courtesy of Uwe Hassler

Set in 10/12pt WarnockPro by SPi Global, Chennai, India

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

PROMETHEUS

“Then I invented arithmetic for them,
the most ingenious acquired skill,
and joining letters to write down words,
so they could store all things in Memory,
the working mother of the Muses’ arts.”

AESCHYLUS, *Prometheus Bound*

Quoted from the translation by Ian Johnston, *Richer Resources Publications*,
2012

Contents

	List of Figures	<i>xi</i>
	Preface	<i>xiii</i>
	List of Notation	<i>xv</i>
	Acronyms	<i>xvii</i>
1	Introduction	<i>1</i>
1.1	Empirical Examples	<i>1</i>
1.2	Overview	<i>6</i>
2	Stationary Processes	<i>11</i>
2.1	Stochastic Processes	<i>11</i>
2.2	Ergodicity	<i>14</i>
2.3	Memory and Persistence	<i>22</i>
2.4	Technical Appendix: Proofs	<i>25</i>
3	Moving Averages and Linear Processes	<i>27</i>
3.1	Infinite Series and Summability	<i>27</i>
3.2	Wold Decomposition and Invertibility	<i>32</i>
3.3	Persistence versus Memory	<i>37</i>
3.4	Autoregressive Moving Average Processes	<i>47</i>
3.5	Technical Appendix: Proofs	<i>51</i>
4	Frequency Domain Analysis	<i>57</i>
4.1	Decomposition into Cycles	<i>57</i>
4.2	Complex Numbers and Transfer Functions	<i>62</i>
4.3	The Spectrum	<i>63</i>
4.4	Parametric Spectra	<i>68</i>
4.5	(Asymptotic) Properties of the Periodogram	<i>72</i>
4.6	Whittle Estimation	<i>76</i>
4.7	Technical Appendix: Proofs	<i>81</i>

5	Differencing and Integration	89
5.1	Integer Case	89
5.2	Approximating Sequences and Functions	91
5.3	Fractional Case	95
5.4	Technical Appendix: Proofs	99
6	Fractionally Integrated Processes	103
6.1	Definition and Properties	103
6.2	Examples and Discussion	108
6.3	Nonstationarity and Type I Versus II	114
6.4	Practical Issues	118
6.5	Frequency Domain Assumptions	120
6.6	Technical Appendix: Proofs	123
7	Sample Mean	127
7.1	Central Limit Theorem for $I(0)$ Processes	127
7.2	Central Limit Theorem for $I(d)$ Processes	129
7.3	Functional Central Limit Theory	132
7.4	Inference About the Mean	139
7.5	Sample Autocorrelation	141
7.6	Technical Appendix: Proofs	145
8	Parametric Estimators	149
8.1	Parametric Assumptions	149
8.2	Exact Maximum Likelihood Estimation	150
8.3	Conditional Sum of Squares	154
8.4	Parametric Whittle Estimation	156
8.5	Log-periodogram Regression of FEXP Processes	161
8.6	Fractionally Integrated Noise	164
8.7	Technical Appendix: Proofs	165
9	Semiparametric Estimators	169
9.1	Local Log-periodogram Regression	169
9.2	Local Whittle Estimation	175
9.3	Finite Sample Approximation	182
9.4	Bias Approximation and Reduction	184
9.5	Bandwidth Selection	188
9.6	Global Estimators	193
9.7	Technical Appendix: Proofs	195
10	Testing	197
10.1	Hypotheses on Fractional Integration	197
10.2	Rescaled Range or Variance	199

10.3	The Score Test Principle	204
10.4	Lagrange Multiplier (LM) Test	205
10.5	LM Test in the Frequency Domain	210
10.6	Regression-based LM Test	213
10.7	Technical Appendix: Proofs	218
11	Further Topics	223
11.1	Model Selection and Specification Testing	223
11.2	Spurious Long Memory	226
11.3	Forecasting	229
11.4	Cyclical and Seasonal Models	231
11.5	Long Memory in Volatility	234
11.6	Fractional Cointegration	236
11.7	R Packages	240
11.8	Neglected Topics	241
	Bibliography	245
	Index	267

List of Figures

- Figure 1.1** Annual minimal water levels of the Nile river. 2
- Figure 1.2** Monthly opinion poll in England, 1960–1996. 3
- Figure 1.3** Monthly US inflation, 1966–2008. 4
- Figure 1.4** Daily realized volatility, 1993–2007. 5
- Figure 1.5** Monthly unemployment rate, 1972–2008. 6
- Figure 2.1** Processes with (top to bottom) (a) moderate persistence, (b) antipersistence, and (c) strong persistence. 24
- Figure 3.1** White noise and differences thereof $(\epsilon_t - \epsilon_{t-1})$. 39
- Figure 3.2** Antipersistent process with $c_j = -(1/2)^j$. 40
- Figure 3.3** Antipersistent process with $c_j = -j/(j+1)!$. 41
- Figure 3.4** Long memory under antipersistence with $c_j = (-1)^j/(j \ln 2)$. 43
- Figure 3.5** Long memory under moderate persistence, $c_j = (-1)^{j+1}/j$. 44
- Figure 3.6** Long memory under strong persistence. $c_j = 1/(j+1)$. 45
- Figure 3.7** Long memory process under strong persistence with $c_j = 1/j^{0.8}$. 46
- Figure 4.1** AR(1) and EXP(1) with $\phi = 0.8$ and $\psi = 1.6$. 71
- Figure 6.1** Simulated $I(2)$ and $I(1)$ series. 110
- Figure 6.2** Simulated $I(d)$ series and their impulse responses. 111
- Figure 6.3** FIN(d) processes in the frequency and time domain. 112
- Figure 6.4** FIMA($d, 1$) processes with $d = 0.45$ and θ . 113
- Figure 6.5** Spectra of ARFI(1, d) processes. 113
- Figure 7.1** $\text{eff}(d) = C_{s,d}^2 / (C_{s,d}^{\text{II}})^2$. 138
- Figure 7.2** Autocorrelation bias for $d = 0.2$. 143
- Figure 7.3** Autocorrelation bias for $d = 0.45$. 144
- Figure 7.4** Quantiles and averages of $\hat{\rho}(h)$ for FIN(d), $T = 1000$. 145

Preface

Scope of the Book

Since the book by Box and Jenkins (1970), autoregressive moving average (ARMA) models integrated of order d are a standard tool for time series analysis, where typically $d \in \{0, 1, 2\}$. The integrated ARMA (ARIMA) model of order d means that a time series has to be differenced d times in order to obtain a stationary and invertible ARMA representation. The papers by Granger and Joyeux (1980) and Hosking (1981) extended the ARIMA model with integer d to the so-called fractionally integrated model, where d takes on noninteger values, often restricted to $|d| < 1/2$. In particular, the case of $0 < d < 1/2$ corresponds to a stationary model with long memory, where the latter means that the autocorrelations die out so slowly that they are not absolutely summable. For $1/2 \leq d < 1$, the fractionally integrated model bridges the gap from stationarity to the so-called unit root behavior ($d = 1$), where past shocks have a permanent effect on the present and values of $d > 1$ allow for even more extreme persistence.

This book grew out of lecture notes from which I taught PhD courses on time series analysis and in particular on time series with long memory. Long memory and fractional integration have become key concepts in time series analysis over the last decades. For instance, the updated edition of Box and Jenkins (1970), i.e. Box et al. (2015), contains a section on long memory and fractional integration, and so do Kirchgässner et al. (2013), Pesaran (2015), or Palma (2016). Also, previous textbooks like Brockwell and Davis (1991, Section 13.2) and Fuller (1996, Section 2.11) include short sections on this topic. Contrary to these books on general times series analysis containing only short digressions into the realm of long memory, there are nowadays specialized monographs dedicated to this topic exclusively, most recently by Giraitis et al. (2012) and Beran et al. (2013); see also the earlier books by Beran (1994) and Palma (2007). The approach of the present book differs from both routes, from the general interest track and from the specialized long memory track. I rather attempt to introduce into the theory of univariate time series analysis, and the foundations thereof, in

such a way that long memory and fractional integration arise as a special case, naturally embedded into the general theory. This is reflected by the title: Time Series Analysis with Long Memory in View. This view is largely directed by the author's research agenda in this field over the last 25 years.

Acknowledgment

Twenty-five years ago I wrote my doctoral thesis on time series with long memory under the supervision of Professor Wolters at the Freie Universität Berlin. Jürgen Wolters passed away in November 2015. I take this opportunity to commemorate his enthusiasm, generosity, and open-mindedness from which I profited so much not only during my doctoral studies but also later on as his coauthor. Since my thesis, I have written a sequence of papers on long memory. I am indebted to many anonymous referees for writing, in many cases, critical and constructive reports on my papers before publication. Most papers were written with coauthors. I thank them for sharing their knowledge and endurance with me. All of them I owe insights that influenced my research agenda and hence this book. In particular, I wish to mention Matei Demetrescu and Mehdi Hosseinkouchack with whom the collaboration was especially fruitful. The intense discussions we had on a daily basis when they held postdoc positions at Goethe University Frankfurt shaped my view not only on how to address long memory but also on time series analysis in general. Christoph Hanck, Paulo Rodrigues, and Verena Werkmann have proofread an earlier draft of this book, and their many comments and corrections are gratefully acknowledged. Finally, I am grateful to the Volkswagen Stiftung for financing a year of sabbatical leave in 2014/2015 by an *opus magnum* grant; without this support it would not have been possible to write this book.

October 2017

Uwe Hassler

List of Notation

\mathbb{C}	set of complex numbers
\mathbb{N}	set of natural numbers
\mathbb{N}_0	set of natural numbers including 0
\mathbb{R}	set of real numbers
\mathbb{Z}	set of integers
$\lfloor x \rfloor$	largest integer smaller or equal to $x \geq 0, x \in \mathbb{R}$
$\ln x$	natural logarithm of $x > 0$, also $\ln(x)$
I_k	identity matrix of dimension k
γ	Euler's constant
$P(\cdot)$	probability
$E(\cdot)$	expectation operator
$\text{Var}(\cdot)$	variance operator
$\text{Cov}(\cdot, \cdot)$	covariance operator
$\gamma(h)$	autocovariance at lag h
$\rho(h)$	autocorrelation at lag h
ω^2	long-run variance
L	lag operator
Δ	difference operator
$\mathcal{E}\mathcal{X}\mathcal{P}(\theta)$	exponential distribution with parameter θ
$\mathcal{N}(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$\chi^2(n)$	chi-square distribution with n degrees of freedom
\rightarrow	convergence
$\xrightarrow{a.s.}$	almost sure convergence
$\xrightarrow{2}$	convergence in mean square
\xrightarrow{p}	convergence in probability
\xrightarrow{D}	convergence in distribution
\Rightarrow	weak convergence
\approx	approximately equal

$a := b$	a is defined to equal b
$A \Rightarrow B$	A implies B
$A \Leftrightarrow B$	A and B are equivalent
$a_n \sim b_n$	a_n/b_n converges to 1
$y_t \sim I(d)$	y_t is integrated of order d
$\sim \mathcal{N}(0, 1)$	follows a standard normal distribution

Acronyms

AIC	Akaike information criterion
AR	autoregressive
ARCH	autoregressive conditional heteroskedasticity
ARFIMA	autoregressive fractionally integrated moving average
ARMA	autoregressive moving average
BIC	Bayesian information criterion
CIR	cumulated impulse response
CLT	central limit theorem
CMT	continuous mapping theorem
CSS	conditional sum of squares
DCT	dominated convergence theorem
DFT	discrete Fourier transform
EXP	exponential model
fBm	fractional Brownian motion
FCLT	functional central limit theorem
FEXP	fractional EXP
FIN	fractionally integrated noise
GARCH	generalized autoregressive conditional heteroskedasticity
LLN	law of large numbers
LM	Lagrange multiplier
MA	moving average
MAC	memory and autocorrelation consistent
MDS	martingale difference sequence
ML	maximum likelihood
MSE	mean squared error
OLS	ordinary least squares
WLLN	weak law of large numbers

1

Introduction

1.1 Empirical Examples

Figure 1.1 displays 663 annual observations of minimal water levels of the Nile river. This historical data is from Beran (1994, Sect. 12.2) and ranges from the year 622 until 1284. The second panel contains the sample autocorrelations $\hat{\rho}(h)$ at lag $h \in \{1, 2, \dots, 30\}$. The maximum value, $\hat{\rho}(1) = 0.57$, is not particularly large, but the autocorrelogram dies out only very slowly with $\hat{\rho}(30) = 0.15$ still being significantly positive. Such a slowly declining autocorrelogram is characteristic of what we will define as long memory or strong persistence. It reflects that the series exhibits a very persistent behavior in that we observe very long cyclical movements or (reversing) trends. Note, e.g. that from the year 737 until 805, there are only three data points above the sample average ($=11.48$), i.e. there are seven decades of data below the average. Then the series moves above the average for a couple of years, only to swing down below the sample mean for another 20 years from the year 826 on. Similarly, there is a long upward trend from 1060 on until about 1125, followed again by a long-lasting decline. Such irregular cycles or trends due to long-range dependence, or persistence, have first been discovered and discussed by Hurst, a British engineer who worked as hydrologist on the Nile river; see in particular Hurst (1951). Mandelbrot and Wallis (1968) coined the term *Joseph effect* for such a feature; see also Mandelbrot (1969). This alludes to the biblical seven years of great abundance followed by seven years of famine, only that *cycles* in Figure 1.1 do not have a period of seven years, not even a constant period.

Long memory in the sense of strong temporal dependence as it is obvious in Figure 1.1 has been reported in many fields of science. Hipel and McLeod (1994, Section 11.5) detected long memory in hydrological or meteorological series like annual average rainfall, temperature, and again river flow data; see also Montanari (2003) for a survey. A further technical area beyond geophysics with long memory time series is the field of data network traffic in computing; see Willinger et al. (2003).

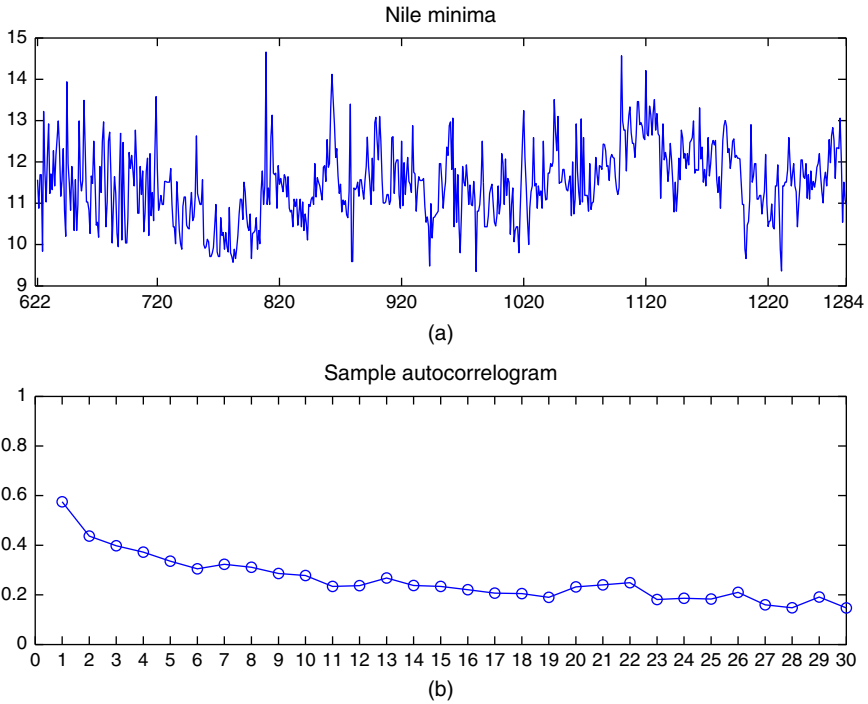


Figure 1.1 Annual minimal water levels of the Nile river.

The second data set that we look into is from political science. Let p_t denote the poll data on partisanship, i.e. the voting intention measured by monthly opinion polls in England. More precisely, p_t is the portion of people supporting the Labor Party. The sample ranges from September 1960 until October 1996 and has been analyzed by Byers et al. (1997).¹ Figure 1.2 contains the logit transformation of this poll data,

$$y_t = \ln \left(\frac{p_t}{1 - p_t} \right),$$

such that $y_t = 0$ for $p_t = 50\%$; here $\ln(x)$ stands for the natural logarithm of x . We observe long-lasting upswings followed by downswings amounting to a pseudocyclical pattern or reversing trends. This is well reflected and quantified by the sample autocorrelations in the lower panel, decreasing from $\hat{\rho}(1) = 0.9$ quite slowly to $\hat{\rho}(24) \approx 0.2$. Independently of Byers et al. (1997), Box-Steffensmeier and Smith (1996) detected long memory in US opinion poll data on partisanship. Long memory in political popularity has been confirmed

¹ We downloaded the data from James Davidson's homepage on May 5, 2016. The link is <http://people.exeter.ac.uk/jehd201/bdpdata.txt>.

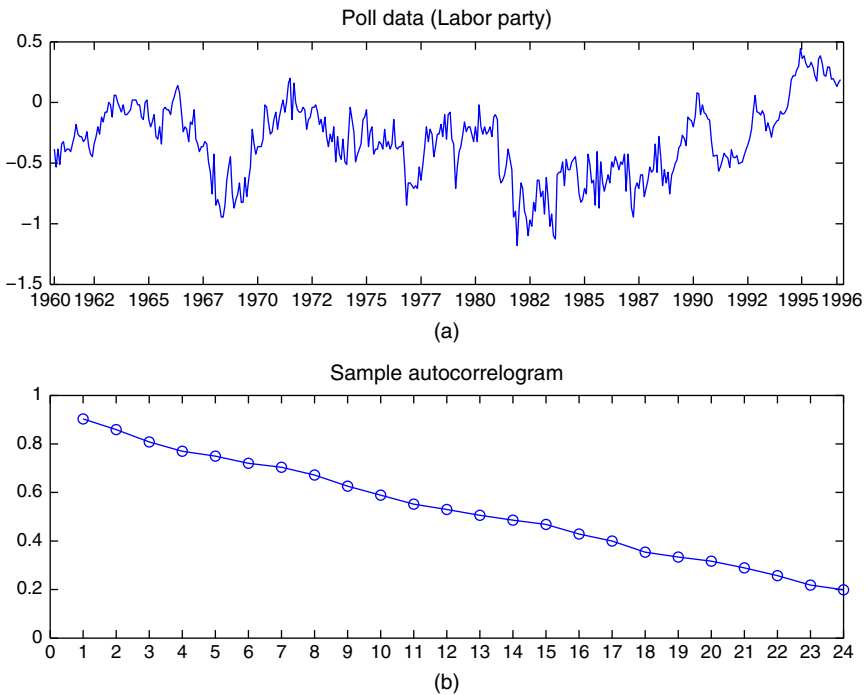


Figure 1.2 Monthly opinion poll in England, 1960–1996.

in a sequence of papers; see Byers et al. (2000, 2007), and Dolado et al. (2003); see also Byers et al. (2002) for theoretical underpinning of long memory in political popularity. Further evidence on long memory in political science has been presented by Box-Steffensmeier and Tomlinson (2000); see also the special issue of *Electoral Studies* edited by Lebo and Clarke (2000).

Since Granger and Joyeux (1980), the fractionally integrated autoregressive moving average (ARMA) model gained increasing popularity in economics. The empirical example in Granger and Joyeux (1980) was the monthly US index of consumer food prices. Granger (1980) had shown theoretically how the aggregation of a large number of individual series may result in an index that is fractionally integrated, which provided theoretical grounds for long memory as modeled by fractional integration in price indices. A more systematic analysis by Geweke and Porter-Hudak (1983) revealed long memory in different US price indices. These early papers triggered empirical research in long memory in inflation rates in independent work by Delgado and Robinson (1994) for Spain and by Hassler and Wolters (1995) and Baillie et al. (1996) for international evidence. Since then, there has been offered abundant evidence in favor of long memory in inflation rates; see, e.g. Franses and Ooms (1997),

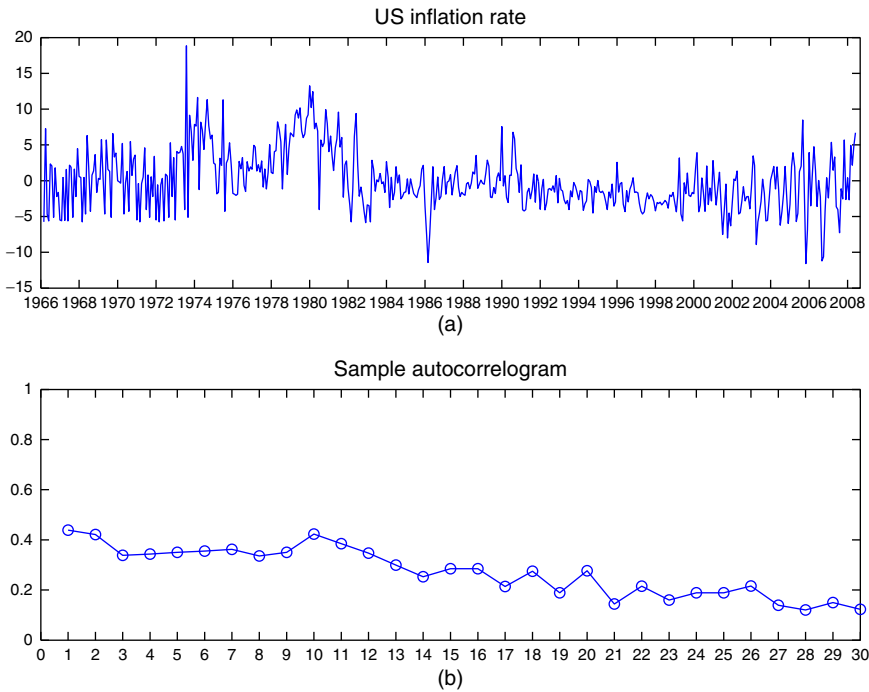


Figure 1.3 Monthly US inflation, 1966–2008.

Baum et al. (1999), Franses et al. (1999), Hsu (2005), Kumar and Okimoto (2007), Martins and Rodrigues (2014), and Hassler and Meller (2014), where the more recent research focused on breaks in persistence, i.e. in the order of fractional integration. For an early survey article on further applications in economics, see Baillie (1996).

Figure 1.3 gives a flavor of the memory in US inflation. The seasonally adjusted and demeaned data from January 1966 until June 2008 has been analyzed by Hassler and Meller (2014). The autocorrelations fall from $\hat{\rho}(1) = 0.44$ to a minimum of $0.12 = \min\{\hat{\rho}(h)\}$, $h = 1, 2, \dots, 30$. Again, this slowly declining autocorrelogram mirrors the reversing trends in inflation, although Hassler and Meller (2014) suggested that the persistence may be superimposed by additional features like time-varying variance.

The fourth empirical example is from the field of finance. Figure 1.4 displays daily observations from January 4, 1993, until May 31, 2007. This sample of 3630 days consists of the logarithm of realized volatility of International Business Machines Corporation (IBM) returns computed from underlying five-minutes data; see Hassler et al. (2016) for details. Although the dynamics of the series

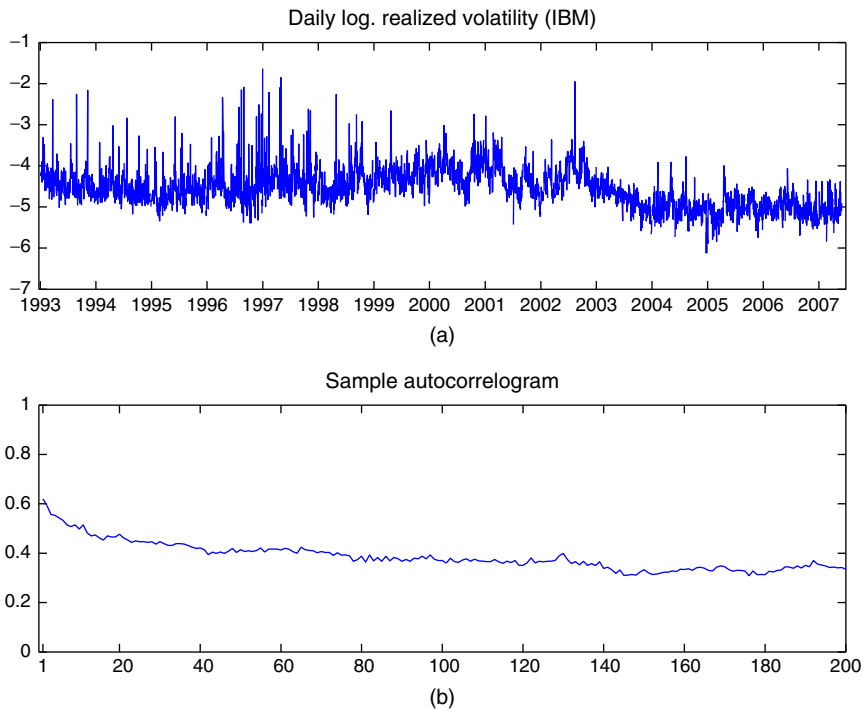


Figure 1.4 Daily realized volatility, 1993–2007.

is partly masked by extreme observations, one clearly may distinguish periods of weeks where the data tends to increase, followed by long time spans of decrease. The high degree of persistence becomes more obvious when looking at the sample autocorrelogram. Starting off with $\hat{\rho}(1) = 0.62$, the decline is extremely slow with $\hat{\rho}(400)$ still being well above 0.2. Long memory in realized volatility is sometimes considered to be a stylized fact since the papers by Andersen et al. (2001, 2003). Such a view is supported by the special issue in *Econometric Reviews* edited by Maasoumi and McAleer (2008).

Finally, with the last example we return to economics. Figure 1.5 shows 435 monthly observations from 1972 until 2008. The series is the logarithm of seasonally adjusted US unemployment rates (number of unemployed persons as a percentage of the civilian labor force); see Hassler and Wolters (2009) for details. The sample average of log-unemployment is 1.7926; compare the straight line in the upper panel of Figure 1.5. Here, the trending behavior is so strong that the sample average is crossed only eight times over the period of 35 years. The deviations from the average are very pronounced and very long relative to the sample size. In that sense the series from Figure 1.5 seems to be most persistent of all the five examples considered in this introduction.

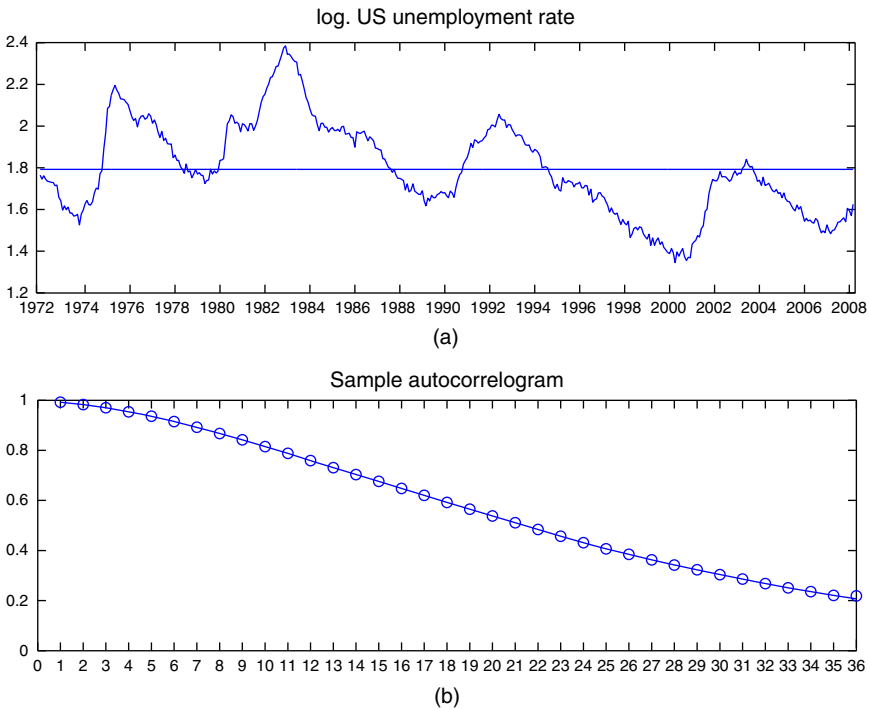


Figure 1.5 Monthly unemployment rate, 1972–2008.

This is also expressed by the sample autocorrelogram virtually beginning at one ($\hat{\rho}(1) = 0.992$) and $\hat{\rho}(h) > 0.2$ for $h \in \{1, 2, \dots, 36\}$. What is more, the autocorrelations decline almost linearly in h , which is indicative of an $I(1)$ process or an $I(d)$ process with even $d > 1$; see Hassler (1997, Corollary 3) and Section 7.5. Hence, the log-unemployment data seems to be most persistent, or most strongly trending, among our empirical examples.

1.2 Overview

There are two natural approaches to long memory modeling by fractional integration. The first one takes the nonstationary $I(1)$ model as starting point, i.e. processes integrated of order 1. Such processes are often labeled as unit root processes in econometrics, where they play a major role within the cointegration framework; see, for instance, Johansen (1995), Lütkepohl (2005), or Pesaran (2015). The extension from the $I(1)$ model to the more general $I(d)$ model might be considered as a nearby approach from an econometric point of view. The second approach starts off with the classical stationary time series

model, where the moving average coefficients from the Wold decomposition are assumed to be absolutely summable and to sum to a value different from 0. For this model, which may be called integrated of order 0, $I(0)$ (see Chapter 6), it holds true that the scaled sample average converges with the square root of the sample size to a nondegenerate normal distribution. This $I(0)$ model underlying the major body of time series books from Anderson (1971) over Brockwell and Davis (1991) and Hamilton (1994) to Fuller (1996) may be generalized to the stationary $I(d)$ process for $d < 1/2$. The latter can be further extended to the region of nonstationarity ($d \geq 1/2$). Here, we follow this second route starting with the $I(0)$ case. More precisely, the outline of the book is as follows.

A definition of stationarity of stochastic processes is given in the next chapter. Moreover, Chapter 2 contains a discussion of ergodicity that corrects expositions found in some books (see Example 2.2). Next, we show that a familiar sufficient condition for ergodicity in the mean (defined in Definition 2.3) is also necessary; see Proposition 2.2. Then we distinguish between (short and long) memory (Definition 2.4) and different degrees of persistence on statistical grounds: Short memory is separated from long memory to characterize under what circumstances the variance of the sample average is of order $1/T$, where T denotes the sample size; see Proposition 2.3. Persistence is defined (Definition 2.5) to characterize the absence or presence and strength of a trend component in a process; see also Eq. (4.22).

Chapter 3 focuses on moving average processes of infinite order, sometimes called linear processes. This is motivated by Wold's theorem in Section 3.2. We thus have a unified framework to embed the classical process of moderate persistence as well as processes with antipersistence or strong persistence, which may or may not display long memory at the same time. The discussion of memory vs. persistence is picked up again in Section 3.3. The discussion of Examples 3.2 through 3.5 shows that the series from Figures 1.1 to 1.5 display both long memory and strong persistence, which motivates the model of fractional integration in Chapter 6. Before leaving Chapter 3, we provide some interesting results on the summability of the classical ARMA process (Proposition 3.5) established with a sequence of technical lemmata.

Chapter 4 introduces to the frequency domain where much of the long memory analysis is settled. The frequency domain is not only useful for data analysis, but it also allows for a deeper theoretical study. For instance, the classical concept of invertibility can be recast following Bloomfield (1985) and Bondon and Palma (2007) in a way (Proposition 4.6) that extends the region of invertibility of fractionally integrated processes; see Proposition 6.2. Next, we introduce the so-called exponential model formulated in the frequency domain. This exponential model is typically not treated in time series books, although it is particularly convenient in the context of long memory as modeled by fractional integration. Similarly, time series books typically do not deal with so-called

Whittle estimation, which is a frequency domain approximation to maximum likelihood that we present in Section 4.6, thus laying the foundation for memory estimation in Chapters 8 and 9.

Chapter 5 opens the route to fractional integration. It is a short chapter on the fractional difference and integration operator, respectively. We provide four technical lemmata that will be used repeatedly in subsequent chapters. Chapter 6 defines the stationary fractionally integrated process (of type I), building on a precise definition of $I(0)$ processes; see Assumption 6.2. Conditions for (different degrees of) persistence follow under minimal restrictions from Lemma 5.4, while Proposition 6.1 translates this into the frequency domain. Corollary 6.1 and Proposition 6.3 reflect the persistence as (short or long) memory in the time domain. After a discussion of parametric fractionally integrated models in Section 6.2, two different types of nonstationarity are discussed in Section 6.3: First, type II fractionally integrated processes are only asymptotically stationary if $d < 1/2$. Second, the case $d \geq 1/2$ covers nonstationarity for both type I and type II processes. Proposition 6.6 shows that classical parametric models imply frequency domain assumptions often entertained in the literature. For the rest of the book, we assume the fractionally integrated models as introduced in Chapter 6.

Chapter 7 sets off with what seems to be the most general central limit theorem currently available for moving average processes. It is applied to the sample average of fractionally integrated processes, closing in particular the gap at $d = -1/2$ in the literature; see Corollary 7.1. Section 7.3 extends the central limit theorem to a functional central limit theory, where fractional Brownian motions show up in the limit. Two seemingly different representations of type II fractional Brownian motion are shown to be identical in Lemma 7.2. Finally, this chapter contains in Section 7.5 an exposition on the behavior of the sample autocorrelations under fractional integration.

The eighth chapter is dedicated to the estimation of all other parameters except for the mean, assuming a fully parametric model of fractional integration. Theorem 8.1 gives the general structure of the limiting covariance matrix of the asymptotic normal distribution for (different approximations to) maximum likelihood, while Corollary 8.1 focuses in particular on the integration parameter d . Approximations to maximum likelihood may be settled in the time domain (Proposition 8.2) or in the frequency domain (Whittle estimation, Proposition 8.3). In particular, we find that the nonstationarity-extended Whittle estimator (Proposition 8.4) overcomes all pitfalls of exact maximum likelihood, except for being parametric of course. Section 8.5 paves the way to semiparametric estimation in that it studies the log-periodogram regression in the presence of a so-called exponential model for the short memory component. While consistency is established in Proposition 8.5, we learn that the estimator is less efficient than corresponding estimators rooted in the maximum likelihood principle.

Chapter 9 begins with the already familiar log-periodogram regression, however, now in the presence of short memory, which is not parametrically modeled. The whole chapter is dedicated to procedures that are semiparametric in the sense that they are robust with respect to short memory. This comes at the price of reduced efficiency. Indeed, we obtain a slower rate of convergence compared with parametric estimators. Within the class of semiparametric estimators, there exist differences in efficiency, too, and the local Whittle estimator (Proposition 9.4 or 9.5), respectively its versions allowing for nonstationarity (Proposition 9.6 or 9.7), turn out to be superior.

Since semiparametric estimators are burdened with large variances, it is interesting to have powerful tests that allow to discriminate statistically, e.g. between short memory and long memory or between stationarity and nonstationarity. This issue is addressed in Chapter 10. The first test builds on a classical rescaled range analysis that can be traced back to Hurst (1951). It has been improved by the rescaled variance test that is designed to provide a better balance of power and size in finite samples. A different approach is adopted in Section 10.4 dedicated to Lagrange multiplier (LM) tests. In Section 10.6, the original LM test is recast in a convenient lag-augmented regression framework (Proposition 10.8), and it has the nice property of robustness against conditional and even unconditional heteroskedasticity (Proposition 10.9). At the same time it is asymptotically most powerful against local alternatives.

Long memory is a still rapidly growing field of applied and theoretical research. Therefore, we close the book with a collection of further topics in the final chapter.

All chapters contain a final section called “Technical Appendix: Proofs” (except for the last chapter “Further Topics”). There, we give the mathematical proofs of results provided and discussed in the main text. Some proofs just accomplish or spell out simple steps to adapt proofs from the literature to our context. Other proofs are truly original in that they establish new results that cannot be drawn from the literature. By separating the proofs from the propositions in the main text, we hope to improve the readability of the book. Finally, it should be stressed that the book is not fully self-contained. While in some propositions we spell out all required assumptions, there are many cases where we refer to the literature. For brevity and convenience one finds in the latter case formulations like “...satisfying Assumption 6.3 [...] and some further restrictions by Robinson (1995b);” such that the reader is expected to read up details from the provided reference, namely, Robinson (1995b), for this example from Proposition 9.4.