

# Applied Economic Forecasting using Time Series Methods

Eric Ghysels  
University of North Carolina, Chapel Hill, United States

Massimiliano Marcellino  
Bocconi University, Milan, Italy

**OXFORD**  
UNIVERSITY PRESS

**OXFORD**  
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries.

Published in the United States of America by Oxford University Press  
198 Madison Avenue, New York, NY 10016, United States of America.

© Oxford University Press 2018

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, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization. Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer.

CIP data is on file at the Library of Congress  
ISBN 978-0-19-062201-5

9 8 7 6 5 4 3 2 1

Printed by Sheridan Books, Inc., United States of America

# Contents

Preface

xv

## Part I Forecasting with the Linear Regression Model

<b>1</b>	<b>The Baseline Linear Regression Model</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	The basic specification . . . . .	4
1.3	Parameter estimation . . . . .	6
1.4	Measures of model fit . . . . .	8
1.5	Constructing point forecasts . . . . .	10
1.6	Interval and density forecasts . . . . .	13
1.7	Parameter testing . . . . .	15
1.8	Variable selection . . . . .	18
1.9	Automated variable selection procedures . . . . .	21
1.9.1	Forward selection (FWD) . . . . .	22
1.9.2	Least angle regressions (LARS) . . . . .	22
1.9.3	LASSO and elastic net estimator (NET) . . . . .	23
1.10	Multicollinearity . . . . .	24
1.11	Example using simulated data . . . . .	26
1.11.1	Data simulation procedure . . . . .	26
1.12	Empirical examples . . . . .	32
1.12.1	Forecasting Euro area GDP growth . . . . .	32
1.12.2	Forecasting US GDP growth . . . . .	43
1.13	A hint of dynamics . . . . .	47
1.13.1	Revisiting GDP forecasting . . . . .	48

1.13.2	Forecasting default risk . . . . .	49
1.14	Concluding remarks . . . . .	55
<b>2</b>	<b>Model Mis-Specification</b>	<b>57</b>
2.1	Introduction . . . . .	57
2.2	Heteroskedastic and correlated errors . . . . .	57
2.2.1	The Generalized Least Squares (GLS) estimator and the feasible GLS estimator . . . . .	60
2.3	HAC estimators . . . . .	64
2.4	Some tests for homoskedasticity and no correlation . . . . .	66
2.5	Parameter instability . . . . .	69
2.5.1	The effects of parameter changes . . . . .	70
2.5.2	Simple tests for parameter changes . . . . .	71
2.5.3	Recursive methods . . . . .	74
2.5.4	Dummy variables . . . . .	76
2.5.5	Multiple breaks . . . . .	78
2.6	Measurement error and real-time data . . . . .	80
2.7	Instrumental variables . . . . .	82
2.8	Examples using simulated data . . . . .	84
2.9	Empirical examples . . . . .	94
2.9.1	Forecasting Euro area GDP growth . . . . .	94
2.9.2	Forecasting US GDP growth . . . . .	107
2.9.3	Default risk . . . . .	114
2.10	Concluding remarks . . . . .	117
<b>3</b>	<b>The Dynamic Linear Regression Model</b>	<b>119</b>
3.1	Introduction . . . . .	119
3.2	Types of dynamic linear regression models . . . . .	120
3.3	Estimation and testing . . . . .	125
3.4	Model specification . . . . .	126
3.5	Forecasting with dynamic models . . . . .	128
3.6	Examples with simulated data . . . . .	130
3.7	Empirical examples . . . . .	132
3.7.1	Forecasting Euro area GDP growth . . . . .	133
3.7.2	Forecasting US GDP growth . . . . .	138

3.7.3	Default risk . . . . .	142
3.8	Concluding remarks . . . . .	142
<b>4</b>	<b>Forecast Evaluation and Combination</b>	<b>145</b>
4.1	Introduction . . . . .	145
4.2	Unbiasedness and efficiency . . . . .	146
4.3	Evaluation of fixed event forecasts . . . . .	149
4.4	Tests of predictive accuracy . . . . .	150
4.5	Forecast comparison tests . . . . .	152
4.6	The combination of forecasts . . . . .	154
4.7	Forecast encompassing . . . . .	156
4.8	Evaluation and combination of density forecasts . . . . .	157
4.8.1	Evaluation . . . . .	158
4.8.2	Comparison . . . . .	159
4.8.3	Combination . . . . .	159
4.9	Examples using simulated data . . . . .	160
4.10	Empirical examples . . . . .	164
4.10.1	Forecasting Euro area GDP growth . . . . .	164
4.10.2	Forecasting US GDP growth . . . . .	167
4.10.3	Default risk . . . . .	169
4.11	Concluding remarks . . . . .	170

## Part II Forecasting with Time Series Models

<b>5</b>	<b>Univariate Time Series Models</b>	<b>173</b>
5.1	Introduction . . . . .	173
5.2	Representation . . . . .	174
5.2.1	Autoregressive processes . . . . .	176
5.2.2	Moving average processes . . . . .	180
5.2.3	ARMA processes . . . . .	181
5.2.4	Integrated processes . . . . .	183
5.2.5	ARIMA processes . . . . .	184
5.3	Model specification . . . . .	185
5.3.1	AC/PAC based specification . . . . .	185
5.3.2	Testing based specification . . . . .	186
5.3.3	Testing for ARCH . . . . .	187

5.3.4	Specification with information criteria . . . . .	187
5.4	Estimation . . . . .	189
5.5	Unit root tests . . . . .	190
5.6	Diagnostic checking . . . . .	195
5.7	Forecasting, known parameters . . . . .	196
5.7.1	General formula . . . . .	196
5.7.2	Some examples . . . . .	197
5.7.3	Some additional comments . . . . .	201
5.8	Forecasting, estimated parameters . . . . .	202
5.9	Multi-steps (or direct) estimation . . . . .	204
5.10	Permanent-transitory decomposition . . . . .	207
5.10.1	The Beveridge Nelson decomposition . . . . .	207
5.10.2	The Hodrick-Prescott filter . . . . .	209
5.11	Exponential smoothing . . . . .	209
5.12	Seasonality . . . . .	211
5.12.1	Deterministic seasonality . . . . .	212
5.12.2	Stochastic seasonality . . . . .	213
5.13	Examples with simulated data . . . . .	214
5.13.1	Stationary ARMA processes . . . . .	214
5.13.2	Estimation: Full sample analysis . . . . .	218
5.13.3	Estimation: Subsample analysis . . . . .	220
5.13.4	Forecasting . . . . .	227
5.14	Empirical examples . . . . .	228
5.14.1	Modeling and forecasting the US federal funds rate . . . . .	228
5.14.2	Modeling and forecasting US inventories . . . . .	240
5.15	Concluding remarks . . . . .	251
<b>6</b>	<b>VAR Models</b> . . . . .	<b>253</b>
6.1	Representation . . . . .	254
6.2	Specification of the model . . . . .	256
6.3	Estimation . . . . .	257
6.4	Diagnostic checks . . . . .	260
6.5	Forecasting . . . . .	261
6.6	Impulse response functions . . . . .	263
6.7	Forecast error variance decomposition . . . . .	268
6.8	Structural VARs with long-run restrictions . . . . .	269
6.9	VAR models with simulated data . . . . .	271

6.10	Empirical examples . . . . .	282
6.10.1	GDP growth in the Euro area . . . . .	282
6.10.2	Monetary transmission mechanism . . . . .	289
6.11	Concluding remarks . . . . .	300
<b>7</b>	<b>Error Correction Models</b>	<b>301</b>
7.1	Introduction . . . . .	301
7.2	Spurious regressions . . . . .	302
7.3	Cointegration and error correction models . . . . .	303
7.4	Engle and Granger cointegration test . . . . .	305
7.5	The Johansen cointegration test . . . . .	306
7.6	MA representations of cointegrated processes . . . . .	308
7.7	Forecasting in the presence of cointegration . . . . .	311
7.8	The effects of stochastic trends on forecasts . . . . .	312
7.9	Example using simulated series . . . . .	313
7.10	Empirical examples . . . . .	323
7.10.1	Term structure of UK interest rates . . . . .	323
7.10.2	Composite coincident and leading indexes . . . . .	334
7.10.3	Predicting coincident and leading indices in the period before the crisis . . . . .	334
7.10.4	Predicting coincident and leading indices during the financial crisis . . . . .	340
7.11	Concluding remarks . . . . .	342
<b>8</b>	<b>Bayesian VAR Models</b>	<b>345</b>
8.1	Introduction . . . . .	345
8.2	A primer on Bayesian econometrics . . . . .	346
8.2.1	Parameters as random variables . . . . .	347
8.2.2	From prior to posterior distribution . . . . .	347
8.2.3	An example: The posterior of the mean when the variance is known . . . . .	348
8.2.4	The Bayesian linear regression model . . . . .	350
8.3	Baseline Bayesian VAR case . . . . .	353

8.3.1	Baseline BVAR specification . . . . .	353
8.3.2	Prior parameterizations . . . . .	354
8.4	Forecasting with the BVAR . . . . .	356
8.4.1	The proper method . . . . .	356
8.4.2	The pseudo-iterated approach . . . . .	357
8.4.3	The direct forecasting approach . . . . .	358
8.5	Example with simulated data . . . . .	359
8.6	Empirical examples . . . . .	361
8.6.1	GDP growth in the Euro area . . . . .	361
8.6.2	Multi-country inflation rates . . . . .	362
8.7	Concluding remarks . . . . .	364

### Part III TAR, Markov Switching, and State Space Models

<b>9</b>	<b>TAR and STAR Models</b>	<b>369</b>
9.1	Introduction . . . . .	369
9.2	Specification . . . . .	371
9.3	Estimation . . . . .	373
9.4	Testing for STAR and TAR . . . . .	374
9.5	Diagnostic tests . . . . .	376
9.6	Forecasting . . . . .	376
9.6.1	Point forecasts . . . . .	376
9.6.2	Interval forecasts . . . . .	378
9.7	Artificial neural networks . . . . .	379
9.8	Example: Simulated data . . . . .	382
9.8.1	Testing for linearity versus non-linearity . . . . .	382
9.8.2	Estimation . . . . .	384
9.8.3	Forecasting . . . . .	386
9.8.4	Two-steps ahead forecasting . . . . .	388
9.9	Example: Forecasting industrial production growth . . . . .	392
9.10	Concluding remarks . . . . .	397
<b>10</b>	<b>Markov Switching Models</b>	<b>399</b>
10.1	Introduction . . . . .	399

10.2	Markov chains . . . . .	400
10.3	Mixture of i.i.d. distributions . . . . .	403
10.4	Markov switching dynamic models . . . . .	406
10.5	Empirical example: Simulated data . . . . .	410
10.6	Empirical example: Industrial production . . . . .	413
10.7	Concluding remarks . . . . .	417
<b>11</b>	<b>State Space Models</b>	<b>419</b>
11.1	Introduction . . . . .	419
11.2	Models in state space form . . . . .	420
11.2.1	An ARMA(2,2) model . . . . .	421
11.2.2	A time-varying parameters (TVP) regression model . . . . .	422
11.2.3	A model with stochastic volatility . . . . .	423
11.2.4	A dynamic factor model . . . . .	424
11.2.5	Unobserved component models . . . . .	425
11.3	The Kalman filter . . . . .	426
11.3.1	The key equations of the Kalman filter . . . . .	427
11.3.2	The iterative procedure . . . . .	429
11.3.3	Some derivations and additional results . . . . .	430
11.4	Example with simulated data: The TVP regression model . . . . .	432
11.5	Empirical examples . . . . .	440
11.5.1	Forecasting US GDP growth with a TVP model . . . . .	440
11.5.2	Forecasting GDP growth with dynamic factor models . . . . .	444
11.6	Concluding remarks . . . . .	449

## Part IV Mixed Frequency, Large Datasets, and Volatility

<b>12</b>	<b>Models for Mixed-Frequency Data</b>	<b>453</b>
12.1	Introduction . . . . .	453
12.2	Bridge equations . . . . .	455
12.3	MIDAS Regressions . . . . .	459
12.3.1	The basic MIDAS regression model . . . . .	460
12.3.2	The AR-MIDAS and ADL-MIDAS models . . . . .	463
12.3.3	Slope or no slope? . . . . .	463
12.3.4	U-MIDAS and MIDAS with step functions . . . . .	465

12.3.5	Extensions of the MIDAS regression model . . . . .	466
12.4	Mixed-frequency VAR . . . . .	470
12.4.1	Parameter-driven approach . . . . .	471
12.4.2	Observation-driven approach . . . . .	474
12.5	Example: Nowcasting with simulated data . . . . .	476
12.5.1	The bridge equation approach . . . . .	477
12.5.2	The MIDAS regression approach . . . . .	480
12.6	Example: Nowcasting US GDP growth . . . . .	482
12.6.1	Predicting the worst trough of the crisis . . . . .	483
12.6.2	Predicting mild positive and negative growth rates . . . . .	493
12.6.3	Predicting GDP growth rates, 2007-2012 . . . . .	496
12.7	Concluding remarks . . . . .	501
<b>13</b>	<b>Models for Large Datasets</b>	<b>503</b>
13.1	Introduction . . . . .	503
13.2	Factor models . . . . .	504
13.3	The three pass regression filter . . . . .	509
13.4	Large unbalanced datasets . . . . .	512
13.4.1	Mixed-frequency factor models . . . . .	512
13.4.2	Mixed-frequency three pass regression filter . . . . .	516
13.4.3	Missing observations and ragged edges . . . . .	519
13.5	Example with simulated data . . . . .	520
13.6	Empirical example: Forecasting GDP growth . . . . .	521
13.7	Concluding remarks . . . . .	527
<b>14</b>	<b>Forecasting Volatility</b>	<b>531</b>
14.1	Introduction . . . . .	531
14.2	ARCH-type models . . . . .	532
14.2.1	Model specifications . . . . .	533
14.2.2	Estimation . . . . .	537
14.3	MIDAS regressions and volatility forecasting . . . . .	539
14.3.1	Realized volatility . . . . .	540
14.3.2	Realized volatility and MIDAS regressions . . . . .	541
14.3.3	HAR models . . . . .	541
14.3.4	Direct versus iterated volatility forecasting . . . . .	542
14.3.5	Variations on the theme of MIDAS regressions . . . . .	544

<i>CONTENTS</i>	xiii
14.4 GARCH models again . . . . .	545
14.5 Volatility forecasting evaluation . . . . .	546
14.6 Forecasting S&P 500 index volatility . . . . .	548
14.6.1 ARCH-type models . . . . .	549
14.6.2 Realized volatility . . . . .	551
14.7 Concluding remarks . . . . .	555
<b>Bibliography</b>	<b>559</b>
<b>Subject Index</b>	<b>587</b>
<b>Author Index</b>	<b>592</b>

# Preface

Economic forecasting is a key ingredient of decision making both in the public and in the private sector. Governments and central banks, consumers and firms, banks and financial institutions base many of their decisions on future expected economic conditions or on the predictions of specific indicators of interest such as income growth, inflation, unemployment, interest rates, exchange rates, earnings, wages, oil prices, and so on. Unfortunately, economic outcomes are the realization of a vast, complex, dynamic and stochastic system, which makes forecasting very difficult and forecast errors unavoidable. However, forecast precision and reliability can be enhanced by the use of proper econometric models and methods, like those we present in this book.

We wish to satisfy an audience including both undergraduate and graduate students willing to learn basic and advanced forecasting techniques, and researchers in public and private institutions interested in applied economic forecasting. For this reason, the book is divided into four parts. In Part I, the first chapter provides a brief review of basic regression analysis, followed by two chapters dealing with specific regression topics relevant for forecasting, such as model mis-specification, including structural breaks, and dynamic models and their predictive properties. The fourth chapter, on the topic of forecast evaluation and combination, is the first exclusively dedicated to forecasting. The material in Part I could be used in the senior year of an undergraduate program for students who have a strong interest in applied econometrics, or as an introduction to economic forecasting for graduate students and researchers from other disciplines.

Part II of the book is devoted to time series models, in particular univariate autoregressive integrated moving average (ARIMA) models and vector autoregressive (VAR) models. Specifically, Chapter 5 deals with ARIMA models, Chapter 6 with VAR models, Chapter 7 with cointegration and er-

ror correction (ECM) models, and Chapter 8 with Bayesian methods for VAR analysis. The material progressively becomes better suited for master level and/or PhD level classes, in particular the chapters on cointegration and Bayesian VARs. Models considered in Part II are also common tools for practical forecasting in public and private institutions.

Parts III and IV of the book contain a collection of special topics chapters. Each chapter is self-contained, and therefore an instructor or researcher can pick and choose the topics she or he wants to cover. Part III deals mainly with modeling parameter time-variation. Specifically, Chapter 9 presents Threshold and Smooth Transition Autoregressive (TAR and STAR) models, Chapter 10 Markov switching regime models, and Chapter 11 state space models and the Kalman filter, introducing, for example, models with random coefficients and dynamic factor models.

Part IV deals with mixed frequency data models and their use for now-casting in Chapter 12, forecasting using large datasets in Chapter 13 and, finally, volatility models in Chapter 14.

Each chapter starts with a review of the main theoretical results to prepare the reader for the various applications. Examples involving simulated data follow, to make the reader familiar with application using at first stylized settings where one knows the true data generating process and one learns how to apply the techniques introduced in each chapter. From our own teaching experience we find this to be extremely useful as it creates familiarity with each of the topics in a controlled environment. The simulated examples are followed by real data applications - focusing on macroeconomic and financial topics. Some of the examples run across different chapters, particularly in the early part of the book. All data are public domain and cover Euro area, UK, and US examples, including forecasting US GDP growth, default risk, inventories, effective federal funds rates, composite index of leading indicators, industrial production, Euro area GDP growth, UK term structure of interest rates, to mention the most prominent examples.

The book is mostly software neutral. However, for almost all the simulated and empirical examples we provide companion *EViews*<sup>®</sup> and *R* code. The former is mostly - but not exclusively - a menu driven licensed software package whereas the latter is an open source programming language and software environment for statistical computing and graphics supported by the

*R* Foundation for Statistical Computing.<sup>1</sup> We provide code and data for all the simulated and empirical examples in the book on the web.<sup>2</sup> Moreover, all the tables and figures appearing in the book were produced using *EViews*<sup>®</sup>.

As the title suggests, this book is an *applied* time series forecasting book. Hence, we do not have the pretense to provide the full theoretical foundations for the forecasting methods we present. Indeed, there are already a number of books with thorough coverage of theory. The most recent example is the excellent book by Elliott and Timmermann (2016) which provides an in-depth analysis of the statistical theory underlying predictive models, covering a variety of alternative forecasting methods in both classical and Bayesian contexts, as well as techniques for forecast evaluation, comparison, and combination. Equally useful and complimentary are the many standard econometrics textbooks such as Hendry (1995), Judge, Hill, Griffiths, Lütkepohl, and Lee (1988), Stock and Watson (2009), Wooldridge (2012), and the many graduate time series textbooks such as Box and Jenkins (1976), Clements and Hendry (1998), Diebold (2008), Hamilton (1994), or Lütkepohl (2007), among others.

We should also highlight that this book is an applied *time series* forecasting book. As such, we will not discuss forecasting using economic theory-based structural models, such as Dynamic Stochastic General Equilibrium (DSGE) models, see, e.g., Del Negro and Schorfheide (2013) for an exhaustive overview and references. Typically, these structural models are better suited for medium- and long-horizon forecasting, and for policy simulations, while time series methods work better for nowcasting and short-term forecasting (say, up to one-year ahead).

This book would not have been completed without the invaluable and skillful help of a number of our current and former TAs and RAs. Special thanks go to Cristina Angelico, Francesco Corsello, Francesco Giovanardi, Novella Maugeri, and Nazire Özkan who helped with all the simulation and empirical examples throughout the entire book, as well as Hanwei Liu who wrote the *R* codes. Their contributions to the book were invaluable. We would also like to thank many cohorts of undergraduate, master, and PhD

---

<sup>1</sup>Please visit the web page <http://www.eviews.com> for further information on *EViews*<sup>®</sup> and *R* programming language Wikipedia web page [https://en.wikipedia.org/wiki/R\\_\(programming\\_language\)](https://en.wikipedia.org/wiki/R_(programming_language)).

<sup>2</sup>Please visit our respective webpages [www.unc.edu/~eghsels](http://www.unc.edu/~eghsels) or [www.igier.unibocconi.it/marcellino](http://www.igier.unibocconi.it/marcellino) to download the code and data.

students who took courses based on the materials covered in the book at Bocconi University and at the University of North Carolina, Chapel Hill and the Kenan-Flagler Business School.

We are also grateful to a number of coauthors of papers on the topics covered in the book. In particular, Eric Ghysels would like to thank Lucia Alessi, Elena Andreou, Jennie Bai, Fabio Canova, Xilong Chen, Riccardo Colacito, Robert Engle, Claudia Foroni, Lars Forsberg, Patrick Gagliardini, René Garcia, Clive Granger, Andrew Harvey, Jonathan Hill, Casidhe Horan, Lynda Khalaf, Andros Kourtellos, Virmantas Kvedaras, Emanuel Moench, Kaiji Motegi, Denise Osborn, Alberto Plazzi, Eric Renault, Mirco Rubin, Antonio Rubia, Pedro Santa-Clara, Pierre Siklos, Arthur Sinko, Bumjean Sohn, Ross Valkanov, Cosmé Vodounou, Fangfang Wang, Jonathan Wright, and Vaidotas Zemlys (and Massimiliano of course!). Massimiliano Marcellino would like to thank Knut-Are Aastveit, Angela Abbate, Elena Angelini, Mike Artis, Anindya Banerjee, Guenter Beck, Stelios Bekiros, Ralf Burggemann, Andrea Carriero, Roberto Casarin, Todd Clark, Francesco Corielli, Sandra Eickmeier, Carlo Favero, Laurent Ferrara, Claudia Foroni, Giampiero Gallo, Ana Galvão, Pierre Guérin, Jerome Henry, Christian Hepenstrick, Kristin Hubrich, Oscar Jorda, George Kapetanios, Malte Knüppel, Hans-Martin Krolzig, Danilo Leiva-Leon, Wolfgang Lemke, Helmut Lütkepohl, Igor Masten, Gian Luigi Mazzi, Grayham Mizon, Matteo Mogliani, Alberto Musso, Chiara Osbat, Fotis Papailias, Mario Porqueddu, Esteban Prieto, Tommaso Proietti, Francesco Ravazzolo, Christian Schumacher, Dalibor Stevanovic, Jim Stock, Fabrizio Venditti, and Mark Watson (and Eric of course!).

Several colleagues provided useful comments on the book, in particular Frank Diebold, Helmut Lütkepohl, Serena Ng, Simon Price, Frank Schorfheide, and Allan Timmermann. Of course, we remain responsible for any remaining mistakes or omissions.

Finally, we would like to thank our families for their continuous support and understanding. We forecast a quieter and more relaxed “Life after the Book,” but as all forecasters, we could be proven wrong.

Eric Ghysels, Chapel Hill, United States

Massimiliano Marcellino, Milan, Italy