

# Introduction to Time Series Using Stata

SEAN BECKETTI



A Stata Press Publication  
StataCorp LP  
College Station, Texas



Copyright © 2013 by StataCorp LP  
All rights reserved. First edition 2013

Published by Stata Press, 4905 Lakeway Drive, College Station, Texas 77845

Typeset in L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub>

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN-10: 1-59718-132-3

ISBN-13: 978-1-59718-132-7

Library of Congress Control Number: 2012951897

No part of this book may be reproduced, stored in a retrieval system, or transcribed, in any form or by any means—electronic, mechanical, photocopy, recording, or otherwise—without the prior written permission of StataCorp LP.

Stata, **STATA**, Stata Press, Mata, **MATA**, and NetCourse are registered trademarks of StataCorp LP.

Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations.

L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub> is a trademark of the American Mathematical Society.

# Contents

	List of tables	xiii
	List of figures	xv
	Preface	xxi
	Acknowledgments	xxvii
<b>1</b>	<b>Just enough Stata</b>	<b>1</b>
1.1	Getting started . . . . .	2
1.1.1	Action first, explanation later . . . . .	2
1.1.2	Now some explanation . . . . .	6
1.1.3	Navigating the interface . . . . .	7
1.1.4	The gestalt of Stata . . . . .	13
1.1.5	The parts of Stata speech . . . . .	15
1.2	All about data . . . . .	19
1.3	Looking at data . . . . .	29
1.4	Statistics . . . . .	49
1.4.1	Basics . . . . .	49
1.4.2	Estimation . . . . .	53
1.5	Odds and ends . . . . .	60
1.6	Making a date . . . . .	62
1.6.1	How to look good . . . . .	63
1.6.2	Transformers . . . . .	65
1.7	Typing dates and date variables . . . . .	68
1.8	Looking ahead . . . . .	69
<b>2</b>	<b>Just enough statistics</b>	<b>71</b>
2.1	Random variables and their moments . . . . .	72

2.2	Hypothesis tests . . . . .	73
2.3	Linear regression . . . . .	74
2.3.1	Ordinary least squares . . . . .	74
2.3.2	Instrumental variables . . . . .	77
2.3.3	FGLS . . . . .	77
2.4	Multiple-equation models . . . . .	78
2.5	Time series . . . . .	79
2.5.1	White noise, autocorrelation, and stationarity . . . . .	80
2.5.2	ARMA models . . . . .	82
<b>3</b>	<b>Filtering time-series data</b>	<b>85</b>
3.1	Preparing to analyze a time series . . . . .	87
3.1.1	Questions for all types of data . . . . .	87
	How are the variables defined? . . . . .	87
	What is the relationship between the data and the phenomenon of interest? . . . . .	88
	Who compiled the data? . . . . .	90
	What processes generated the data? . . . . .	90
3.1.2	Questions specifically for time-series data . . . . .	91
	What is the frequency of measurement? . . . . .	91
	Are the data seasonally adjusted? . . . . .	91
	Are the data revised? . . . . .	91
3.2	The four components of a time series . . . . .	92
	Trend . . . . .	93
	Cycle . . . . .	95
	Seasonal . . . . .	97
3.3	Some simple filters . . . . .	100
3.3.1	Smoothing a trend . . . . .	103
3.3.2	Smoothing a cycle . . . . .	109
3.3.3	Smoothing a seasonal pattern . . . . .	114
3.3.4	Smoothing real data . . . . .	115

3.4	Additional filters . . . . .	121
3.4.1	ma: Weighted moving averages . . . . .	123
3.4.2	EWMAAs . . . . .	125
	exponential: EWMAAs . . . . .	126
	dexponential: Double-exponential moving averages . . . . .	130
3.4.3	Holt–Winters smoothers . . . . .	131
	hwinters: Holt–Winters smoothers without a seasonal component . . . . .	131
	shwinters: Holt–Winters smoothers including a seasonal component . . . . .	137
3.5	Points to remember . . . . .	138
<b>4</b>	<b>A first pass at forecasting</b>	<b>141</b>
4.1	Forecast fundamentals . . . . .	141
4.1.1	Types of forecasts . . . . .	142
4.1.2	Measuring the quality of a forecast . . . . .	144
4.1.3	Elements of a forecast . . . . .	144
4.2	Filters that forecast . . . . .	146
4.2.1	Forecasts based on EWMAAs . . . . .	148
4.2.2	Forecasting a trending series with a seasonal component . . . . .	159
4.3	Points to remember . . . . .	165
4.4	Looking ahead . . . . .	166
<b>5</b>	<b>Autocorrelated disturbances</b>	<b>167</b>
5.1	Autocorrelation . . . . .	168
5.1.1	Example: Mortgage rates . . . . .	169
5.2	Regression models with autocorrelated disturbances . . . . .	172
5.2.1	First-order autocorrelation . . . . .	173
5.2.2	Example: Mortgage rates (cont.) . . . . .	175
5.3	Testing for autocorrelation . . . . .	176
5.3.1	Other tests . . . . .	177
5.4	Estimation with first-order autocorrelated data . . . . .	178

5.4.1	Model 1: Strictly exogenous regressors and autocorrelated disturbances . . . . .	179
	The OLS strategy . . . . .	181
	The transformation strategy . . . . .	183
	The FGLS strategy . . . . .	185
	Comparison of estimates of model 1 . . . . .	188
5.4.2	Model 2: A lagged dependent variable and i.i.d. errors . . . . .	189
5.4.3	Model 3: A lagged dependent variable with AR(1) errors . . . . .	192
	The transformation strategy . . . . .	193
	The IV strategy . . . . .	195
5.5	Estimating the mortgage rate equation . . . . .	196
5.6	Points to remember . . . . .	198
<b>6</b>	<b>Univariate time-series models</b>	<b>201</b>
6.1	The general linear process . . . . .	202
6.2	Lag polynomials: Notation or prestidigitation? . . . . .	203
6.3	The ARMA model . . . . .	205
6.4	Stationarity and invertibility . . . . .	208
6.5	What can ARMA models do? . . . . .	210
6.6	Points to remember . . . . .	214
6.7	Looking ahead . . . . .	215
<b>7</b>	<b>Modeling a real-world time series</b>	<b>217</b>
7.1	Getting ready to model a time series . . . . .	218
7.2	The Box–Jenkins approach . . . . .	226
7.3	Specifying an ARMA model . . . . .	228
	7.3.1 Step 1: Induce stationarity (ARMA becomes ARIMA) . . . . .	228
	7.3.2 Step 2: Mind your p’s and q’s . . . . .	233
7.4	Estimation . . . . .	243
7.5	Looking for trouble: Model diagnostic checking . . . . .	253
	7.5.1 Overfitting . . . . .	253
	7.5.2 Tests of the residuals . . . . .	254

7.6	Forecasting with ARIMA models . . . . .	257
7.7	Comparing forecasts . . . . .	262
7.8	Points to remember . . . . .	266
7.9	What have we learned so far? . . . . .	267
7.10	Looking ahead . . . . .	269
<b>8</b>	<b>Time-varying volatility</b>	<b>271</b>
8.1	Examples of time-varying volatility . . . . .	272
8.2	ARCH: A model of time-varying volatility . . . . .	277
8.3	Extensions to the ARCH model . . . . .	285
8.3.1	GARCH: Limiting the order of the model . . . . .	286
8.3.2	Other extensions . . . . .	292
	Asymmetric responses to “news” . . . . .	293
	Variations in volatility affect the mean of the observable series . . . . .	295
	Nonnormal errors . . . . .	296
	Odds and ends . . . . .	296
8.4	Points to remember . . . . .	298
<b>9</b>	<b>Models of multiple time series</b>	<b>299</b>
9.1	Vector autoregressions . . . . .	300
9.1.1	Three types of VARs . . . . .	302
9.2	A VAR of the U.S. macroeconomy . . . . .	303
9.2.1	Using Stata to estimate a reduced-form VAR . . . . .	305
9.2.2	Testing a VAR for stationarity . . . . .	309
	Other tests . . . . .	312
9.2.3	Forecasting . . . . .	316
	Evaluating a VAR forecast . . . . .	325
9.3	Who’s on first? . . . . .	329
9.3.1	Cross correlations . . . . .	330
9.3.2	Summarizing temporal relationships in a VAR . . . . .	335
	Granger causality . . . . .	336

	How to impose order . . . . .	339
	FEVDs . . . . .	343
	Using Stata to calculate IRFs and FEVDs . . . . .	343
9.4	SVARs . . . . .	356
9.4.1	Examples of a short-run SVAR . . . . .	359
9.4.2	Examples of a long-run SVAR . . . . .	368
9.5	Points to remember . . . . .	371
9.6	Looking ahead . . . . .	372
<b>10</b>	<b>Models of nonstationary time series</b>	<b>375</b>
10.1	Trends and unit roots . . . . .	376
10.2	Testing for unit roots . . . . .	380
10.3	Cointegration: Looking for a long-term relationship . . . . .	385
10.4	Cointegrating relationships and VECMs . . . . .	387
10.4.1	Deterministic components in the VECM . . . . .	391
10.5	From intuition to VECM: An example . . . . .	392
	Step 1: Confirm the unit root . . . . .	397
	Step 2: Identify the number of lags . . . . .	399
	Step 3: Identify the number of cointegrating relationships . . . . .	400
	Step 4: Fit a VECM . . . . .	404
	Step 5: Test for stability and white-noise residuals . . . . .	414
	Step 6: Review the model implications for reasonableness . . . . .	415
10.6	Points to remember . . . . .	422
10.7	Looking ahead . . . . .	422
<b>11</b>	<b>Closing observations</b>	<b>425</b>
11.1	Making sense of it all . . . . .	425
11.2	What did we miss? . . . . .	426
11.2.1	Advanced time-series topics . . . . .	427
11.2.2	Additional Stata time-series features . . . . .	429
	Data management tools and utilities . . . . .	429
	Univariate models . . . . .	430



Multivariate models . . . . .	430
11.3 Farewell . . . . .	431
<b>References</b>	<b>433</b>
<b>Author index</b>	<b>437</b>
<b>Subject index</b>	<b>439</b>

*(Pages omitted)*

# Preface

Welcome.

Time-series analysis is a relatively new branch of statistics. Most of the techniques described in this book did not exist prior to World War II, and many of the techniques date from just the last few decades. The novelty of these techniques is somewhat surprising, given the importance of forecasting in general and of predicting the future consequences of today's policy actions in particular. The explanation lies in the relative difficulty of the statistical theory for time series. When I was in graduate school, one of my econometrics professors admitted that he had switched his focus from time series when he realized he could produce three research papers a year on cross-section topics but only one paper per year on time-series topics.

## Why another book on time series?

The explosion of research in recent decades has delivered a host of powerful and complex tools for time-series analysis. However, it can take a little while to become comfortable with applying these tools, even for experienced empirical researchers. And in industry, these tools sometimes are applied indiscriminately with little appreciation for their subtleties and limitations. There are several excellent books on time-series analysis at varying levels of difficulty and abstraction. But few of those books are linked to software tools that can immediately be applied to data analysis.

I wrote this book to provide a step-by-step guide to essential time-series techniques—from the incredibly simple to the quite complex—and, at the same time, to demonstrate how these techniques can be applied in the Stata statistical package.

Why Stata? There are, after all, a number of established, powerful statistical packages offering time-series tools. Interestingly, the conventions adopted by these programs for describing and analyzing time series vary widely, much more widely than the conventions used for cross-section techniques and classical hypothesis testing. Some of these packages focus primarily on time series and can be used on non-time-series questions only with a bit of difficulty. Others have to twist their time-series procedures into a form that fits the rest of the structure of their package.

I helped out in a small way when Stata was first introduced. At that time, the most frequent question posed by users (and potential users) was, “When will time series be available?” For a long time, we would tell users (completely sincerely) that these techniques would appear in the next release, in six to twelve months. However, we

repeatedly failed to deliver on this promise. Version after version appeared with many new features, but not time series. I moved on to other endeavors, remaining a Stata user but not a participant in its production. Like other users, I kept asking for time-series features—I needed them in my own research. I finally became frustrated and, using Stata’s programming capabilities, cobbled together some primitive Stata functions that helped a bit.

Why the delay? Part of the reason was other, more time-critical demands on what was, at the beginning, a small company. However, I think the primary reason was StataCorp’s commitment to what they call the “human-machine interface”. There are lots of packages that reliably calculate estimates of time-series models. Many of them are difficult to use. They present a series of obstacles that must be overcome before you can test your hypotheses on data. Frequently, it is challenging to thoroughly examine all aspects of your data. And they make it onerous to switch directions as the data begin to reveal their structure.

Stata makes these tasks easy—at least, easy by comparison to the alternatives. I find that the facility of Stata contributes to better analyses. I attempt more, I look more deeply, because it is easy. The teams that work for me use several different packages, not just Stata, depending on the task at hand. I find that I get better, more thorough analyses from the team members using Stata. I do not think it is a coincidence.

When Stata finally gained time-series capabilities, it incorporated a design that retains the ease of use and intuitiveness that has always been the hallmark of this package. That is why I use Stata rather than any of the other candidate packages.

Despite the good design poured into Stata, time-series analysis is still tough. That is just the nature of the time-series inference task. I tend to learn new programs by picking up the manual and playing around. I certainly have learned a lot of the newer, more complex features of Stata that way. However, I do not think it is easy to learn the time-series techniques of Stata just from reading the Stata *Time-Series Reference Manual*—and it is a very well-written manual. I know—I tried. For a long time, I stuck with my old, home-brew Stata functions to avoid the task of learning something different, even after members of my staff had adopted the new Stata tools.

Writing this book provided me with the opportunity to break out of my bad habits and make the transition to Stata’s powerful time-series features. And I am glad I did. Once you come up the learning curve, I think these tools will knock your socks off. They certainly lower the barrier to many ambitious types of empirical research.

I hope you are the beneficiary of my learning process. I have attempted in these pages to link theory with tools in a way that smooths the path for you. Please let me know if I have succeeded. Contact the folks at Stata Press with your feedback—good or bad—and they will pass it along to me.

## Who should read this book?

**Stata users trying to figure out Stata's time-series tools.** You will find detailed descriptions of the tools and how to apply them combined with detailed examples and an intuitive explanation of the theory underlying each tool.

**Time-series researchers considering Stata for their work.** Each commercial time-series package takes a different approach to characterizing time-series data and models. Stata's unique approach offers distinct advantages that this book highlights.

**Researchers who know a bit about time series but want to know more.** The gestalt of time-series analysis is not immediately intuitive, even to researchers with a deep background in other statistical techniques.

**Researchers who want more extensive help than the manual can provide.** It is clear and well written, but, at the end of the day, it is a manual, not a tutorial.

## How is this book organized?

Like Gaul, this book is divided into three parts.

**Preliminaries.** Preparation for reading the rest of the book.

**Chapter 1: Just enough Stata.** A quick and easy introduction for the complete novice. Also useful if you have not used Stata for a while.

**Chapter 2: Just enough Statistics.** A cheat sheet for the statistical knowledge assumed in later chapters.

**Filtering and Forecasting.** A nontechnical introduction to the basic ways to analyze and forecast time series. Lots of practical advice.

**Chapter 3: Filtering time-series data.** A checklist of questions to answer before your analysis. The four components of a time series. Using filters to suppress the random noise and reveal the underlying structure.

**Chapter 4: A first pass at forecasting.** Forecast fundamentals. Filters that forecast.

**Time-series models.** Modern approaches to time-series models.

**Chapter 5: Autocorrelated disturbances.** What is autocorrelation? Regression models with autocorrelation. Testing for autocorrelation. Estimation with first-order autocorrelated data.

**Chapter 6: Univariate time-series models.** The general linear process. Notation conventions. The mixed autoregressive moving-average model. Stationarity and invertibility.

- Chapter 7: Modeling a real-world time series: The example of U.S. gross domestic product.** Getting ready to model a time series. The Box–Jenkins approach. How to specify, estimate, and test an autoregressive moving-average model. Forecasting with autoregressive integrated moving-average models. Comparing forecasts.
- Chapter 8: Time-varying volatility: Autoregressive conditional heteroskedasticity and generalized autoregressive conditional heteroskedasticity models.** Examples of time-varying volatility. A model of time-varying volatility. Extensions to the autoregressive conditional heteroskedasticity model.
- Chapter 9: Models of multiple time series.** Vector autoregressions. A vector autoregression of the U.S. macroeconomy. Cross correlations, causality, impulse–response functions, and forecast-error decompositions. Structural vector autoregressions.
- Chapter 10: Models of nonstationary time series.** Trends and unit roots. Cointegration. From intuition to vector error-correction models.
- Chapter 11: Closing observations.** Making sense of it all. What did we miss?

## Ready, set, . . .

I am a reporter. I am reporting on the work of others. Work on the statistical theory of time-series processes. Work on the Stata statistical package to apply this theory. As a reporter, I must give you an unvarnished view of these topics. However, as we are frequently reminded in this postmodern world, none of us can be completely objective, try as we will. Each of us has a perspective, a slant informed by our life experiences.

Here is my slant. I was trained as an academic economist. I became a software developer to pay my way through graduate school and found I liked the challenges of good software design as much as I liked economic research. I began my postgraduate career in academics, transitioned to the Federal Reserve System, and eventually ended up in research in the financial services industry, where I have worked for a number of leading firms (some of them still in existence). I believe I have learned something valuable at each stage along the way.

For the purposes of this book, the most important experience has been to see how statistical research, good and bad, is performed in academics, the Fed, and industry. Good academic research applies cutting-edge research to thorny problems. Bad academic research gets caught up in footnotes and trivia and loses sight of real-world phenomena. The Federal Reserve produces high-quality research, frequently published in the best academic journals. A signature characteristic of research within the Fed is a deep knowledge of the institutional details that can influence statistical relationships. However, Fed research occasionally exhibits an oversimplified perspective of the workings of the financial services industry. Industry has to make decisions in real time. Accordingly, industry research has to generate answers quickly. Good industry research

makes wise tactical choices and selects reasonable shortcuts around technical obstacles. Bad industry research is “quick and dirty”.

Embrace the good, avoid the bad. Perhaps because the latter half of my career has been spent in industry, my personal bent is to recognize the limitations of the tools I use without becoming distressed over them. I am more interested in intuition than in proofs.

Here are three articles that sum up the approach I try to emulate:

- Diaconis, P. 1985. Theories of data analysis: From magical thinking through classical statistics. In *Exploring Data Tables, Trends, and Shapes*, ed. D. C. Hoaglin, F. Mosteller, and J. W. Tukey, 1–36. New York: Wiley.
- Ehrenberg, A. S. C. 1977. Rudiments of numeracy. *Journal of the Royal Statistical Society, Series A* 140: 277–297.
- Wainer, H. 1984. How to display data badly. *American Statistician* 38: 137–147.

Do not say I did not warn you. Now get cracking and learn some stuff.

*(Pages omitted)*



# 4 A first pass at forecasting

## Chapter map

- 4.1 Forecast fundamentals.** The types of forecasts covered in this book. Other types of forecasts. Measuring the quality of a forecast. The elements of a forecast.
- 4.2 Filters that forecast.** And ones that do not. Trailing moving averages. Forecasts based on exponentially weighted moving averages (EWMAs). Forecasting a trending series with a seasonal component. Matching the filter to the type of forecast needed. Pros and cons of each approach.
- 4.3 Points to remember.** Three models. Different filters for different situations.
- 4.4 Looking ahead.** Moving from an intuitive approach to forecasting to a formal statistical approach. Short-horizon versus long-horizon forecasts.

Forecasting is perhaps the most familiar and certainly one of the most important uses of time-series models. Accordingly, much of this book is concerned with aspects of forecasting. We have not introduced modern time-series models yet; nonetheless, some useful forecasting can be done with tools we have already discussed. Indeed a great deal of forecasting in industry uses these methods.

This chapter provides a gentle introduction to forecasting. First, we introduce some of the fundamentals of forecasting—the different types of forecasts, the elements of a forecast, and the criteria for choosing one forecasting method over another. Next we explain how some of the filters discussed in the previous chapter can be used for forecasting and the situations to which they are suited. Most of the information we need was discussed in detail in the last chapter, so this first pass at practical forecasting will be brief. More advanced methods for forecasting will be introduced—and compared with these simpler methods—in later chapters.

## 4.1 Forecast fundamentals

A forecast is a statement about the outcome, timing, or probability distribution of some event whose realization is not known currently. That said, there are many different kinds of forecasts, and the time-series techniques discussed in this book apply to only a subset of them.

### 4.1.1 Types of forecasts

For the most part, this book explores stochastic, linear, dynamic time-series models for forecasting continuous variables. Let's consider the import of these adjectives one at a time.

**Stochastic:** We cannot forecast perfectly because the variables we are interested in are influenced by random disturbances in addition to any systematic determinants.

**Linear:** Linear models provide adequate, tractable time-series representations of many variables. As a practical matter, nonlinear models are much more difficult to analyze and understand, and there are fewer established techniques for nonlinear models.

**Dynamic:** This feature is the defining characteristic of time-series models—the influence of the past on the present and future values of a variable. Analyzing the dynamic relationships among variables comprises the bulk of this book.

**Continuous:** Modeling discrete variables such as the outcomes of elections or sporting events requires specialized statistical techniques, even in a cross-section setting. Dynamic relationships—the essence of time-series models—already add significant complexity to statistical analysis. Indeed this complexity justifies the writing of a book such as this one. Adding discrete variables to the mix increases the complexity beyond the scope of this book.

The time-series models we consider can do a lot. The most common application is the generation of point forecasts, that is, the likeliest values at each future date. However, these models also can produce confidence intervals and complete probability distributions of future events. As with static (non-time-series) regression models, they also can be used for policy experiments: to calculate the likely impact of changes in an explanatory variable on a dependent variable. Moreover, these time-series models are well suited to analyzing the relationships among jointly dependent time series. These models also can produce forecasts of the timing of events, although the forecasts typically are expressed in terms of rates of decay of random disturbances to the time path of a variable or the expected time until a variable crosses a specified threshold.

Before we start generating forecasts, let's review some of the types of forecasts that arise in practice and point out whether the time-series models we discuss are well adapted to producing these types of forecasts.

Although we are restricting our attention to models of continuous variables, predictions of events with discrete, qualitative outcomes may be the most familiar.<sup>1</sup> Which candidate will win the presidential election? Which football team will win this Sunday's contest? If I ask Cindy to the prom, will she say yes? Some of these forecasts concern unique, one-time events. In these cases, time-series models, which emphasize the influence of the past on the present and future, generally are not applicable. In these situations, handicapping can be a useful approach. Instead of attempting to construct a model for the probability distribution of an event—say, the winner of this year's Kentucky Derby—we may base our forecast on the opinion of an expert (the racing form), the empirical distribution of the opinions of a panel of experts (a group of professional handicappers), or the opinions of the public at large (the betting odds that reflect the volume of betting on each horse).

Some forecasts concern events that have occurred already but whose outcome is unknown at present. The most familiar examples arise in gambling. Does my opponent's hole card give him a higher-valued poker hand than mine? Many government and industry statistics are published only with a significant lag, and participants in financial markets forecast these not-yet-disclosed measurements of past events to anticipate their likely impacts on the prices of stocks, bonds, and other financial assets. Time-series models can be used for these types of forecasts as long as dynamic relationships play a role in the determination of the variable of interest.

Some forecasts are focused primarily on the timing of an event whose eventual occurrence is assumed to be nearly certain. For example, market participants generally agreed that the rapid increases in house prices in the years prior to 2007 were unsustainable and that an eventual collapse of this housing bubble was inevitable. However, there was no consensus on the timing (or, as it turns out, the magnitude) of the housing collapse. Time-series models can be used for this type of forecast, but frequently some other type of analysis is better adapted to this task. Survival analysis and the failure analysis techniques of statistical quality control often provide convenient methods.

A final class of forecasts—clearly beyond the scope of this book—includes speculations about alternative histories (how would the United States be different today if the South had won the Civil War?) and the forecasts of futurists (what will computers be like in 30 years, or how will human society adapt to climate change in the 21st century?) These types of forecasts are fascinating. However, time-series models are of little use in producing and evaluating forecasts with such global and open-ended scope.

---

1. Stata provides a rich set of tools for models with discrete outcomes (see the Stata Press book *Regression Models for Categorical Dependent Variables Using Stata* [2nd ed., 2006] or type `help logistic` from within Stata if you are in a hurry). Survival analysis—also well covered by Stata—provides models for the timing of discrete events. See the Stata 12 manual *Survival Analysis and Epidemiological Tables Reference Manual* or the Stata Press book *An Introduction to Survival Analysis Using Stata* (3rd ed., 2010) for more information.

### 4.1.2 Measuring the quality of a forecast

As we develop different approaches to forecasting, we need some way to measure the quality of each approach to choose among them. The primary measure of the quality of a forecast is a loss function, that is, a function that assigns a value—the loss—to the difference between the forecast of a variable,  $\hat{y}_t$ , and its realization,  $y_t$ :

$$\text{Loss} = L(y_t - \hat{y}_t)$$

A model that produces forecasts with smaller expected loss is preferred, other things being equal, to a model that produces forecasts with larger expected loss. Except in special cases, we measure loss by the square of the forecast error,

$$L(y_t - \hat{y}_t) = (y_t - \hat{y}_t)^2$$

Expected loss is not the only statistical criterion for choosing a forecasting method. Unbiasedness generally is a desirable characteristic, and the forecast method with minimum expected loss may not be unbiased, forcing us to choose between objectives.

Pragmatic criteria can be as important as statistical ones. Some forecast methods are more difficult to apply and understand, more computationally intensive, or more expensive to use than others. These costs may outweigh any statistical advantages of these methods. Fitting and interpreting modern time-series models requires both a significant degree of statistical sophistication and a considerable amount of practical experience—in other words, a firm may have to employ an experienced statistician to use these techniques. In industry, the additional precision of advanced forecast techniques may not justify the cost of employing a statistician—a simpler method that can be applied by an analyst may be adequate. Moreover, it often is challenging for statisticians to communicate effectively to business leaders the advantages and limitations of various statistical approaches.

The forecast methods covered in this chapter are relatively simple to implement and understand and do not require the assistance of a statistician. As a consequence, these techniques are the workhorses of a great deal of forecasting in industry. Moreover, their short-term to medium-term accuracy often is quite reasonable. In settings where large numbers of forecasts are needed routinely—for example, in a company that must forecast the inventory of hundreds or thousands of specialized parts—these methods may be the only feasible ones. For now, we will evaluate the performance of forecast methods on pragmatic and intuitive grounds. We will introduce more sophisticated techniques in later chapters. At that point, we will focus on expected loss and similar measures of forecast quality.

### 4.1.3 Elements of a forecast

The elements of a forecast consist of an information set; a projection date or time (usually the latest date or time for which the information set is available); a forecast horizon (the amount of time between the projection date and the event being predicted);

and a model of some sort that relates the information set available at the projection date to the event at the forecast horizon. These elements are not fixed by the phenomenon under analysis. You can choose to invest in an extensive information set in an effort to maximize the accuracy of your forecast. Alternatively, you may choose to save money and effort by limiting the information set at the cost of a somewhat greater error variance in your forecast. The forecast horizon is determined solely by your requirements. You may need to forecast the value of a variable next period but may not care about longer-horizon forecasts. Alternatively, you may need to predict the value at a later period—for example, you may need to forecast your available cash at a distant date when you are required to make a large payment—but you may not care about the value of the variable in the intervening periods. To a certain extent, even the model is your choice. All models are approximations to reality, and your choice of model may reflect your trade-off between the closeness of the approximation and other objectives. Of course, there are limits: not all possible model choices will be supported by the data.

The minimum-loss information set includes all the information that can improve the forecast. A reduced information set will produce forecasts with a higher-expected loss. However, it may be costly to construct a model that includes every potentially useful predictor. It also may be costly to collect all the variables in the ideal information set, especially if the forecast is updated frequently. As a result, forecasts frequently are based on reduced information sets.

Univariate time-series models are based on the most commonly used reduced information set. Univariate time-series models predict a variable based solely on its past values (and, potentially, past values of random disturbances). In fact, the defining characteristic of a time-series model is the inclusion of past values of the dependent variable.

The univariate time-series model is a benchmark. More complex models are gauged by their improvement (loss reduction) relative to a univariate time-series model (in much the same way that  $R^2$  measures regression performance by the reduction in error variance relative to a model based solely on the sample mean). It can be surprisingly difficult to improve on the forecasts of a well-specified univariate model. As a result, univariate time-series models are used frequently in practical forecasting. All the filter-based forecasts illustrated in this chapter are types of univariate time-series models.

The most common forecast horizon is one-step-ahead, that is, the first unknown value, and models often are compared based on their one-step-ahead performance alone. The longer the forecast horizon, the greater the number of random disturbances that affect the outcome. As a result, the error variance of a forecast generally increases with the forecast horizon.

**Note:** A  $K$ -step-ahead forecast is a prediction of events in period  $t + K$  conditional on information in period  $t$ . When  $K > 1$ , we have a multistep-ahead forecast.

It is important to distinguish between in-sample and out-of-sample forecasts and between static forecasts and dynamic forecasts. Consider the forecast equation

$$y_{t+1} = \mu + \rho y_t + \epsilon_{t+1}$$

An in-sample forecast of  $y_{t+1}$  is given by

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

where the “hats” ( $\hat{\phantom{x}}$ ) indicate estimates, and both  $y_{t+1}$  and  $y_t$  are part of the sample used to calculate  $\hat{\mu}$  and  $\hat{\rho}$ . Out-of-sample forecasts apply the same formula to  $y$ 's that were not included in the estimation sample. In-sample forecasts do not provide the most rigorous test of a model, because the model is fit to optimize in-sample performance.

Static forecasts are based on the actual values of period  $t$  information. Sticking to the same forecast equation, we write a static, one-period-ahead forecast of  $y_{t+1}$  as

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

that is, we use the observed value of  $y_t$  in forecasting  $y_{t+1}$ . In contrast, a dynamic,<sup>2</sup> one-period-ahead forecast of  $y_{t+1}$  is

$$\hat{y}_{t+1} = \mu + \rho \hat{y}_t$$

In this case, we do not observe the actual value of  $y_t$ , so we must base our forecast of  $y_{t+1}$  on a prior forecast of  $y_t$ . Dynamic forecasts generally are less accurate than static forecasts because they include the error in predicting  $y_t$  in the forecast.

## 4.2 Filters that forecast

In the previous chapter, we introduced a wide variety of filters that help in highlighting patterns in time series. Some of these filters also can be used for forecasting. These methods have the advantage of low cost, both in the time and effort required for estimation and in the amount of computation required for forecasting. This low cost comes at the expense of some of the sophistication, flexibility, and (sometimes) accuracy provided by formal time-series models.

To generate a forecast, a filter must estimate future values as a function of current and past information (or, equivalently, current values as a function of past information).<sup>3</sup>

- 
2. We use the word “dynamic” in two senses in this book. Sometimes “dynamic” indicates the calculation of distant predictions based on earlier predictions, as described here. Earlier in the chapter, we used “dynamic” to denote a formula or model that describes the evolution of a variable over time. Related, but not identical, ideas. Unfortunately, accepted terminology in time-series analysis can be confusing, as you will see in later chapters. We will do our best to make our meaning clear from context.
  3. When trying to forecast the value of a variable that has already been realized but is not yet known, we may include any information available when the forecast is calculated. Some of this information may postdate the variable we are forecasting.

This condition eliminates all the smoothers provided by the `tssmooth nl` command because they combine information from the past, present, and future.

Filters that can produce forecasts are provided by the `tssmooth` command. The four flavors of EWMA<sub>s</sub>—`tssmooth exponential`, `tssmooth dexponential`, `tssmooth hwinters`, and `tssmooth shwinters`—include a `forecast(#)` option that appends predicted values to the end of the smoothed series.<sup>4</sup> Each of these methods operates differently and is suitable for a specific type of forecasting task. The table below lists the types of forecasts produced by each of these `tssmooth` subcommands.

Table 4.1. Forecasting with `tssmooth`

Subcommand	Able to generate forecasts?	Nature of forecast	Uses
<code>nl</code>	No	—	—
<code>ma</code>	No	—	—
<code>exponential</code>	Yes	All forecasts are equal to the last in-sample EWMA	Noisy but nontrending series
<code>dexponential</code>	Yes	Forecasts follow the last estimated trend	Series with a linear secular trend
<code>hwinters</code>	Yes	Forecasts follow the last estimated trend	Series with a linear secular trend
<code>shwinters</code>	Yes	Forecasts follow the last estimated trend and seasonal pattern	Series with a regular seasonal pattern on top of a linear secular trend

To compare the forecasts produced by these methods, we will construct several forecasts of the U.S. civilian unemployment rate, a series we introduced in the prior chapter on filters.

4. Technically, the `tssmooth ma` command can be used to forecast as long as the `window` option includes only past values, that is, as long as we specify a trailing moving average. However, this smoother offers no advantage over `tssmooth exponential`, so Stata does not provide a `forecast` option for `tssmooth ma`.

### 4.2.1 Forecasts based on EWMA

Let's start by reexamining the time series of unemployment rates.

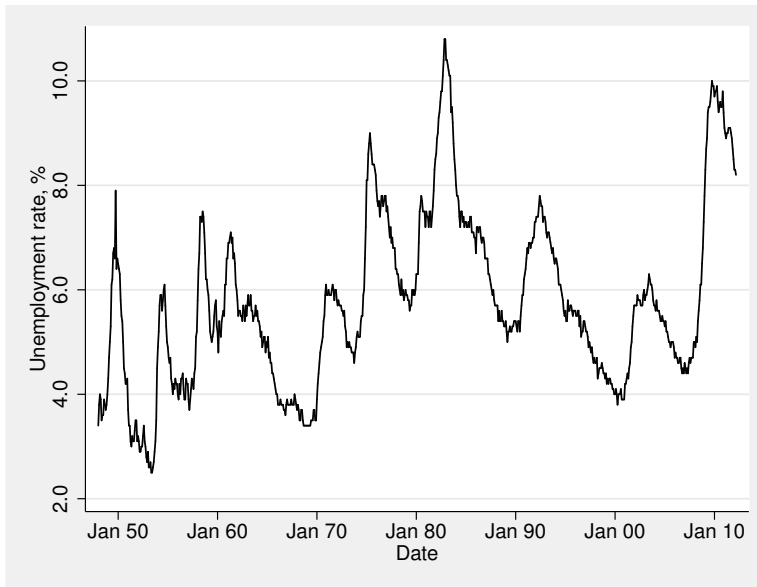


Figure 4.1. U.S. civilian unemployment rate, January 1948 through March 2012

In chapter 3, we used the unemployment rate as an example of a series with a prominent cyclical component. Unlike the artificial cyclical series we constructed to demonstrate the performance of alternative smoothers, the cycle in the unemployment rate is irregular in frequency and shape. It exhibits sharp, inverted, V-shaped peaks associated with recessions and slightly skewed, U-shaped troughs associated with economic expansions.

From this inspection, it is clear that none of our smoothers can produce useful long-run forecasts. Simple EWMA sets all future values to a constant equal to the final estimate of the local mean. The other smoothers project constant trends. None of these smoothers project future oscillations, a key feature of the unemployment rate series. Nonetheless, these forecasts may be useful for short periods, especially when the unemployment rate is hovering near a peak or trough. And, in practice, we usually are most interested in fairly short-term forecasts of the unemployment rate—often just the forecast of next month's rate.



The U.S. unemployment rate began increasing sharply in May 2008 and peaked in October 2009 at 10%. To highlight the strengths and weaknesses of these forecasting approaches, we will generate forecasts from two different starting dates—April 2009, when unemployment was 8.9% and still trending up, and December 2009, just after the peak. We will begin with the `tssmooth exponential` command. We have set  $\alpha$ , the weight on the current observation, to 0.9, higher than Granger's recommended maximum of 0.7 but lower than Stata's estimate of 0.9998.<sup>5</sup> We could generate these two forecasts by running the `tssmooth exponential` command twice, once through April 2009 and once through December 2009, and use the `forecast()` option to append as many projections as we need.<sup>6</sup> For this example, though, there is an easier way. The forecasts of `tssmooth exponential` all are equal to the final smoothed value, so we can smooth the entire series once and generate two forecast series equal to the smoothed values in April and December 2009.

```
. use ${ITSUS_DATA}/monthly, clear
  (Monthly data for ITSUS)
. keep date unrate
. keep if unrate!=.
  (420 observations deleted)
. tsset
      time variable:  date, Jan 48 to Mar 12
                delta:  1 month
. tssmooth exponential ewma = unrate, parms(0.9)
exponential coefficient =      0.9000
sum-of-squared residuals =     39.882
root mean squared error =     .22744
. label variable ewma "EWMA(0.9)"
. list date if date==tm(2009m4) | date==tm(2009m12)
```

	date
736.	Apr 09
744.	Dec 09

```
. generate apr = ewma[736] if _n>736
  (736 missing values generated)
. label variable apr "April 2009 forecast"
. generate dec = ewma[744] if _n>744
  (744 missing values generated)
. label variable dec "December 2009 forecast"
```

5. The Stata estimate suggests that  $\alpha \approx 1$ . In other words, the best estimate of tomorrow's unemployment rate is today's rate. That random walk view of the unemployment rate clearly does not hold up over an extended horizon, but, as we shall see below, it works pretty well for a one-step-ahead forecast.

6. The `forecast()` option will extend the current dataset if necessary.

```
. twoway (conn unrate date, msymbol(o)) (line apr dec date, lpattern(solid dash))
> if date>tm(2005m12), ytitle(Percent)
```

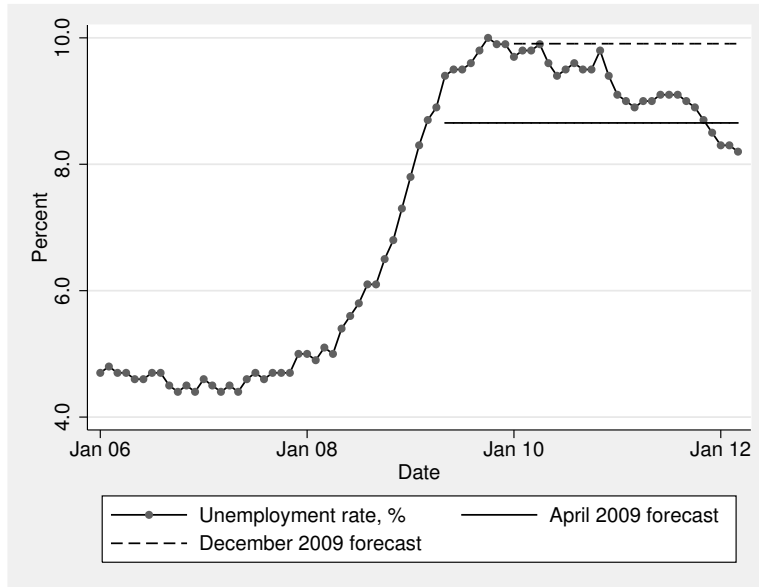


Figure 4.2. EWMA forecasts with different projection dates

The April forecast—when unemployment is still rising—performs very poorly. The December forecast fares better, at least for a few months, because the unemployment rate remained within 0.2% points of 9.9% (the December forecast) until May 2010. Of course, it was impossible to be certain in April 2009 that unemployment was still trending upward or to be certain a few months later in December that the rate was flattening out for a while.

The double exponential smoother projects future values along the last estimated local trend. We will repeat this exercise with the `tssmooth dexponential` command to see how that approach performs.<sup>7</sup>

```
. tssmooth dexponential dewma = unrate if date<tm(2009m5), parms(0.5)
> forecast(35)

double-exponential coefficient =      0.5000
sum-of-squared residuals      =      32.311
root mean squared error       =      .20952

. generate dapr = dewma if date>tm(2009m4)
(736 missing values generated)
```

7. Now we need the `forecast()` option because the projections of the double exponential smoother are not constant.

```

. label variable dapr "April 2009 forecast"
. drop dewma
. tssmooth dexponential dewma = unrate if date<tm(2010m1), parms(0.5)
> forecast(27)

double-exponential coefficient =      0.5000
sum-of-squared residuals      =      32.703
root mean squared error       =      .20966

. generate ddec = dewma if date>tm(2009m12)
(744 missing values generated)
. label variable ddec "December 2009 forecast"
. twoway (conn unrate date, msymbol(o)) (line dapr ddec date, lpattern(solid dash))
> if date>tm(2005m12), ytitle(Percent)

```

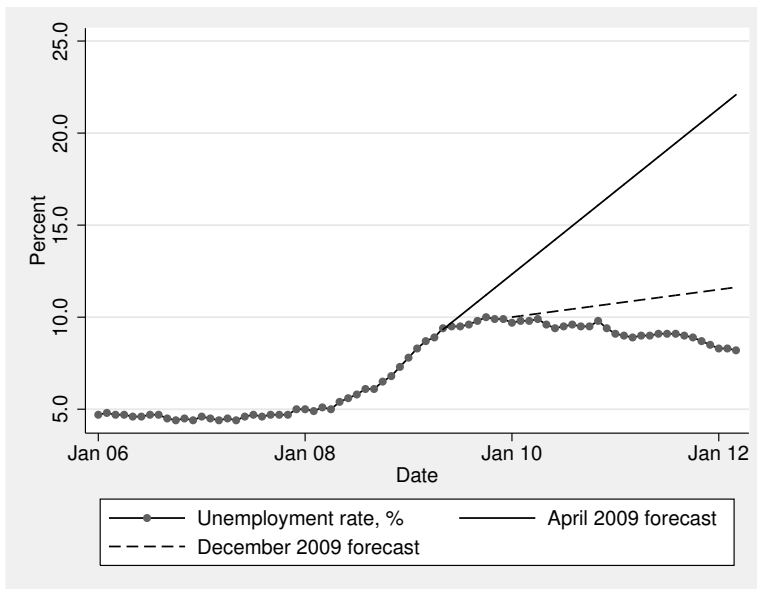


Figure 4.3. DEWMA forecasts with different projection dates

By projecting a fixed trend, both forecasts end up overpredicting by increasing amounts as the forecast horizon increases. Again the December forecast is a little luckier because the forecast begins after the turning point in the unemployment rate.