

2ND
EDITION

Quantitative **Trading**

HOW TO BUILD YOUR OWN ALGORITHMIC TRADING BUSINESS

ERNEST P. CHAN

WILEY

Quantitative Trading

Founded in 1807, John Wiley & Sons is the oldest independent publishing company in the United States. With offices in North America, Europe, Australia, and Asia, Wiley is globally committed to developing and marketing print and electronic products and services for our customers' professional and personal knowledge and understanding.

The Wiley Trading series features books by traders who have survived the market's ever changing temperament and have prospered—some by reinventing systems, others by getting back to basics. Whether a novice trader, professional, or somewhere in-between, these books will provide the advice and strategies needed to prosper today and well into the future.

For a list of available titles, visit our Web site at www.WileyFinance.com.

Quantitative Trading

*How to Build Your Own
Algorithmic Trading Business*

Second Edition

ERNEST P. CHAN

WILEY

Copyright © 2021 by Ernest P. Chan. All rights reserved.

Revised Second Edition © 2026

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

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, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

The manufacturer's authorized representative according to the EU General Product Safety Regulation is Wiley-VCH GmbH, Boschstr. 12, 69469 Weinheim, Germany, e-mail: Product_Safety@wiley.com.

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and the authors have used their best efforts in preparing this work, including a review of the content of the work, neither the publisher nor the authors make any 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.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data is Available:

ISBN 9781394378043 (Revised Second Ed pbk)

ISBN 9781394428236 (Revised Second Ed ePDF)

ISBN 9781394428229 (Revised Second Ed ePub)

ISBN 9781119800064 (Second Ed hardback)

Cover Design: Wiley

Cover Image: © Jobalou/Getty Images

Author Photo: Courtesy of the Author

To my late parents, Hung Yip and Ching, and to Ben, Sarah, and Ethan.

Contents

Foreword by Rishi K. Narang	xi
Preface to the 2nd Edition	xv
Preface	xix
Acknowledgments	xxv
 CHAPTER 1: The Whats, Whos, and Whys of Quantitative Trading	 1
Who Can Become a Quantitative Trader?	2
The Business Case for Quantitative Trading	4
Scalability	5
Demand on Time	5
The Nonnecessity of Marketing	7
The Way Forward	8
 CHAPTER 2: Fishing for Ideas	 11
How to Identify a Strategy that Suits You	14
Your Working Hours	14
Your Programming Skills	15
Your Trading Capital	15
Your Goal	19
A Taste for Plausible Strategies and Their Pitfalls	20
How Does It Compare with a Benchmark, and How Consistent Are Its Returns?	20
How Deep and Long Is the Drawdown?	23
	vii

How Will Transaction Costs Affect the Strategy?	24
Does the Data Suffer from Survivorship Bias?	26
How Did the Performance of the Strategy Change over the Years?	27
Does the Strategy Suffer from Data-Snooping Bias?	28
Does the Strategy “Fly under the Radar” of Institutional Money Managers?	30
Summary	30
References	31

CHAPTER 3: Backtesting **33**

Common Backtesting Platforms	34
Excel	34
MATLAB	34
Python	36
R	38
QuantConnect	40
Blueshift	40
Finding and Using Historical Databases	40
Are the Data Split and Dividend Adjusted?	41
Are the Data Survivorship-Bias Free?	44
Does Your Strategy Use High and Low Data?	46
Performance Measurement	47
Common Backtesting Pitfalls to Avoid	57
Look-Ahead Bias	58
Data-Snooping Bias	59
Transaction Costs	72
Strategy Refinement	77
Summary	78
References	79

CHAPTER 4: Setting Up Your Business **81**

Business Structure: Retail or Proprietary?	81
Choosing a Brokerage or Proprietary Trading Firm	85
Physical Infrastructure	87
Summary	89
References	91

CHAPTER 5: Execution Systems 93

What an Automated Trading System Can Do for You	93
Building a Semiautomated Trading System	95
Building a Fully Automated Trading System	98
Minimizing Transaction Costs	101
Testing Your System by Paper Trading	103
Why Does Actual Performance Diverge from Expectations?	104
Summary	107

CHAPTER 6: Money and Risk Management 109

Optimal Capital Allocation and Leverage	109
Risk Management	120
Model Risk	124
Software Risk	125
Natural Disaster Risk	125
Psychological Preparedness	125
Summary	130
Appendix: A Simple Derivation of the Kelly Formula when Return Distribution Is Gaussian	131
References	132

CHAPTER 7: Special Topics in Quantitative Trading 133

Mean-Reverting versus Momentum Strategies	134
Regime Change and Conditional Parameter Optimization	137
Stationarity and Cointegration	147
Factor Models	160
What Is Your Exit Strategy?	169
Seasonal Trading Strategies	174
High-Frequency Trading Strategies	186
Is it Better to Have a High-Leverage versus a High-Beta Portfolio?	188
Summary	190
References	192

CHAPTER 8: Conclusion	193
Next Steps	197
References	198
Appendix: A Quick Survey of MATLAB	199
Bibliography	205
About the Author	209
Index	211

Foreword

Since its initial publication in 2021, Dr. Ernest Chan's book, *Quantitative Trading*, has served as a crucial reference for independent systematic researchers and traders. In my mind, it has served two key purposes. First, and most obviously, it provides budding practitioners with information that is instrumental to the successful development of trading strategies. The second reason is subtler, but no less important: Chan explains why it's a total myth that only the largest, most well-resourced quant firms can succeed.

What's great about Chan is that he has modeled (forgive the pun) these virtues himself: he has been building successful models for many years, and he has been doing so without the enormous resources of the firms that have hundreds of millions, if not billions, of revenue per year to spend on their tooling and resources. This makes him uniquely qualified to lead readers on this journey.

A key lesson in this book is one that is close to my own heart: the simplest useful solution is probably the best solution. It echoes Occam's razor, which is directly translated as: "You should not multiply entities unnecessarily." And Einstein reportedly said: "Everything should be as simple as possible, but not simpler." The even more terse, "necessary and sufficient" bears here, too. All of this is to say that, while this principle is easily enough understood, Chan does the work to show the reader why this is true, and how to achieve this level of parsimony in their own research.

He also outlines a number of other useful elements of developing and implementing quant strategies: how to backtest while avoiding the myriad pitfalls thereof, how to think about the capacity for a strategy (and the related topic of transaction cost modeling), how to

think about bet sizing, various prediction models (this is the bread and butter for a soon-to-be practitioner, by the way), and even some practical guidance on coding. These may seem mundane at first glance, but please be on notice: the devil is very much in the details in this world. The key strategy types are well known to pretty much everyone: mean reversion, value, momentum, sentiment, technicals, etc. The key to actually making profits is buried deep in detailed decisions around how one implements these ideas.

The second vital purpose *Quantitative Trading* serves is to highlight that smaller practitioners can not only survive, but actually thrive, even in today's competitive landscape. I have personally seen one-person quant shops put up statistically significant positive results trading large capitalization tickers in the US market, with data no more sophisticated than what is readily available to pretty much anyone. My experience is both theoretical and anecdotal.

The idea that a firm with tens of billions under management is playing the same game as a practitioner with maybe \$500,000 or \$10,000,000 of positions is ridiculous at its face. Yes, these firms have so much more to play with, as far as resources go. They have all the data they can ingest, all the compute to process it, and all the talent to make it happen. But in the end, their biggest challenge is that they're, well, *big*. It turns out that this is no small problem (again, sorry for the pun). The markets don't care how much edge you have in your bet. The size of your trades (and the detectability of the patterns through which you put these trades into the market), in the context that you place them, is a huge determinant of the outcome of these trades.

The raw truth is that much of the resource advantage that larger firms have is spent on simply managing that larger size. The fee structures associated with these large businesses, with all their economies of scale in areas like operations, compliance, data, execution and the like make it beneficial to get large. But the real reason is not that they have a huge edge over smaller firms. It's that the resources (specifically, management fees) that come with their size allow them to manage their size. Smaller traders don't need as many resources in part because they don't need to put on so much size – they move the markets less than larger players typically do.

Chan supplies readers with the ammunition to take this hopeful idea and make it into something real. Is this book going to give you a step-by-step recipe to generate profits? Please do not be so naive. Do you need your own excellent ideas? Of course you do! You are embarking on an endeavor to predict the future in a marketplace that is much more noise than signal, which is constantly evolving, and which is unbelievably competitive (with many of the brightest minds in the world as your competitors). But will it arm you to compete, if you heed its lessons and do a great deal of your own learning and gaining hard-won experience? I believe it can.

—RISHI K. NARANG

Preface to the 2nd Edition

When I first started thinking about writing the 2nd edition, I had a measure of dread. What could I have added that would be new and interesting? After writing the first draft, I was relieved, and incredibly excited, at the prospect of sharing with you my latest knowledge, techniques, and insights, ranging from the addition of some new functions that make our PCA example run more than 10x faster, to a novel application of machine learning.

In the 1st edition of this book, published more than a decade ago, I maintained that independent quantitative traders can beat institutional managers at their own game. Many of you have taken that advice to heart, and many retail quantitative trading communities and platforms have been built to serve just such an ambition. But does the premise still hold?

Over the years, many readers reached out and told me how successful they have been in improving and trading the strategies I discussed in my books, and others told me how they have simply been inspired by my books to become successful traders. Our fund is invested in some of these readers, some of whom have been managing many millions more dollars than we are. So, the answer to the above question is a resounding “YES!”

I also exhorted retail traders new to quantitative trading to start with the simplest strategies (examples of which are described in this and my previous books). Do simple strategies still work? Or do we all have to become mathematicians or machine learning experts?

My colleagues and I traded some of the strategies described in this book live since it was first published in 2009, and ran true out-of-sample backtests on others, and I was as surprised as they are that

many still work after all these years. But the issues of “alpha decay,” and the even-more-dreaded “regime change,” are ever threatening. I will talk more about that below.

Speaking of machine learning and artificial intelligence, I didn’t really think much of those techniques in my first book. In fact, the only artificial intelligence platform that I described there has gone out of business. But you may hear that AI is everywhere nowadays, and many fundamental advances in AI have been made since then. For example, the dropout technique that gave birth to deep learning achieved fame in 2012 (Gershgorn, 2017). Should retail traders still avoid AI/ML?

It is as difficult to apply AI/ML to finance in 2021 as it was in 2009, but you may be surprised to hear that we have finally succeeded (Chan, 2020). We have benefited from other giants in the industry who graciously share their insights and knowledge with everyone (López de Prado, 2018). We, in turn, tried to make it easier for every retail trader (even those who are not programmers) to benefit from this technology by launching *predictnow.ai*. Here is the spoiler: The key to successfully apply AI/ML to finance is to focus on *metalabeling* – i.e., finding the probability of profit of your own simple basic trading strategy, and not to use it to predict the market directly. Why? Your own trading strategy’s past track record is private; no one else is trying to predict its success. Meanwhile, millions of people around the world are watching the same public market, and everyone is trying to predict where it will go. Competition and arbitrage naturally mean that signal-to-noise ratio is very low and predictive successes are few and far in between.

But that’s not all. There is another novel use of machine learning that I will discuss in a completely revised Example 7.1 of this book.

Despite our luck with the longevity of some of the strategies I described, most arbitrage opportunities eventually fade away—the notorious *alpha decay* that professionals like to lament. Alpha decay can be due to competition—too many people trading the same strategy, but equally often it is due to regime shift caused by market structure or macroeconomic changes. Adapt and evolve your strategies, or watch them die (Lo, 2019). The market is not stationary; why

should your strategies be? The most agonizing decision a quantitative trader needs to make is to decide when to abandon a strategy during a prolonged drawdown, despite repeated efforts to evolve it. It is ultimately a discretionary decision—you have to judge based on your market knowledge whether there is a fundamental reason your strategy stopped working. To gain this market knowledge, you have to constantly absorb public knowledge disseminated on social media. That is the reason I set aside an hour each day to go through my Twitter feed (@chanep). I have highlighted some of the Twitterers I follow in Chapter 2. More so than providing specific strategy examples, I hope my books will also improve your market intuition in making these discretionary decisions.

One major addition to this edition is the inclusion of Python and R codes to every example. Even though MATLAB is still my favorite backtesting language, there is no reason to exclude the other two most popular languages. Other things that I added and changed in the 2nd edition:

- Chapter 1: A bit more about fully automated trading and marketing your strategies to investors. Also, a scare episode during Covid-19.
- Chapter 2: Updated the educational and trading resources for budding quant traders, including the new URL for my own blog. Also, a good word for Millennium Partners' founder (not that he needs it).
- Chapter 3: Extensive changes on MATLAB code that remove a major bug, and new commentary and codes for Python and R. Description of some new quant trading platforms. One item of particular interest: I discuss a mathematically rigorous way to decide how much backtest data and how long a paper trading period is needed. Another mathematical technique was referenced that determines how data snooping will affect your live Sharpe ratio.
- Chapter 4: Much has changed in the world of brokers and infrastructure providers for algorithmic traders since the first edition. Even the name of the US regulator for brokers has changed. You will find them all updated.

- Chapter 5: It is now much easier than before to build a fully automated trading system. The new ways are described in this chapter.
- Chapter 6: New insights on the Kelly formula and its practical impact. Python and R codes for demonstrating capital allocation using the Kelly formula are added. Also included is a discussion on why loss aversion is *not* a behavioral bias, which is opposite to what I previously believed. It stems from a profound mathematical insight that threatens to upend the economics profession.
- Chapter 7: This chapter is extensively updated. I describe a novel machine learning technique we invented called Conditional Parameter Optimization that can be used to optimize the trading parameters of a strategy based on market regimes. Also added are new high-performance MATLAB/Python/R codes on using PCA, new Python/R codes on checking for stationarity and cointegration, and some surprising out-of-sample results on seasonal trading strategies. I also clarified the difference between time-series and cross-sectional factors.
- Chapter 8: Conclusions remain largely the same. Yes, a retail trader can beat the professionals. But a retail trader can also hire a professional to help generate alpha and diversify.

REFERENCES

- Chan, Ernest. 2020. "What Is the Probability of Your Profit?" PredictNow .ai. <https://www.predictnow.ai/blog/what-is-the-probability-of-profit-of-your-next-trade-introducing-predictnow-ai/>
- Gershgorin. 2017. "The data that transformed AI research—and possibly the world." *Qz*. <https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world/>
- Lo, Andrew. 2019. *Adaptive Markets: Financial Evolution at the Speed of Thought*. Princeton University Press.
- López de Prado, Marcos. 2018. *Advances in Financial Machine Learning*. Wiley.

Preface

By some estimates, quantitative or algorithmic trading now accounts for over 80 percent of the equity trading volume (*Economist*, 2019). There are, of course, innumerable books on the advanced mathematics and strategies utilized by institutional traders in this arena. However, can an independent, retail trader benefit from these algorithms? Can an individual with limited resources and computing power backtest and execute strategies over thousands of stocks, and come to challenge the powerful industry participants at their own game?

I will show you how this can, in fact, be achieved.

WHO IS THIS BOOK FOR?

I wrote this book with two types of readers in mind:

1. Aspiring independent (“retail”) traders who are looking to start a quantitative trading business.
2. Students of finance or other technical disciplines (at the undergraduate or MBA level) who aspire to become quantitative traders and portfolio managers at major institutions.

Can these two very different groups of readers benefit from the same set of knowledge and skills? Is there anything common between managing a \$100 million portfolio and managing a \$100,000 portfolio? My contention is that it is much more logical and sensible

for someone to become a profitable \$100,000 trader before becoming a profitable \$100 million trader. This can be shown to be true on many fronts.

Many legendary quantitative hedge fund managers such as Dr. Edward Thorp of the former Princeton-Newport Partners (Poundstone, 2005) and Dr. Jim Simons of Renaissance Technologies Corp. (Lux, 2000) started their careers trading their own money. They did not begin as portfolio managers for investment banks and hedge funds before starting their own fund management business. Of course, there are also plenty of counterexamples, but clearly this is a possible route to riches as well as intellectual accomplishment, and for someone with an entrepreneurial bent, a preferred route.

Even if your goal is to become an institutional trader, it is still worthwhile to start your own trading business as a first step. Physicists and mathematicians are now swarming Wall Street. Few people on the Street are impressed by a mere PhD from a prestigious university anymore. What is the surest way to get through the door of the top banks and funds? To show that you have a systematic way to profits—in other words, a track record. Quite apart from serving as a stepping stone to a lucrative career in big institutions, having a profitable track record as an independent trader is an invaluable experience in itself. The experience forces you to focus on simple but profitable strategies, and not get sidetracked by overly theoretical or sophisticated theories. It also forces you to focus on the nitty-gritty of quantitative trading that you won't learn from most books: things such as how to build an order entry system that doesn't cost \$10,000 of programming resources. Most importantly, it forces you to focus on risk management—after all, your own personal bankruptcy is a possibility here. Finally, having been an institutional as well as a retail quantitative trader and strategist at different times, I only wish that I had read a similar book before I started my career at a bank—I would have achieved profitability many years sooner.

Given these preambles, I won't make any further apologies in the rest of the book in focusing on the entrepreneurial, independent traders and how they can build a quantitative trading business on their own, while hoping that many of the lessons would be useful on their way to institutional money management as well.

WHAT KIND OF BACKGROUND DO YOU NEED?

Despite the scary-sounding title, you don't need to be a math or computer whiz in order to use this book as a guide to start trading quantitatively. Yes, you do need to possess some basic knowledge of statistics, such as how to calculate averages, standard deviations, or how to fit a straight line through a set of data points. Yes, you also need to have some basic familiarity with Excel. But what you don't need is any advanced knowledge of stochastic calculus, neural networks, or other impressive-sounding techniques.

Though it is true that you can make millions with nothing more than Excel, it is also true that there are tools that, if you are proficient with them, will enable you to backtest trading strategies much more efficiently, and may also allow you to retrieve and process data much more easily than you otherwise can. Best among these tools are MATLAB, Python, and R, and they are the most common research platforms that many institutional quantitative strategists and portfolio managers use. Therefore, I will demonstrate how to backtest the majority of strategies using all three languages. In fact, I have included a brief tutorial in the appendix on how to do some basic programming in MATLAB, which is my favorite among the three. For a tutorial on how to use R for finance, I recommend Regenstein (2018). For Python, I like the book by the inventor of its Pandas package, McKinney (2017). MATLAB for home use costs about as much as Microsoft Office, while Python and R are free.

WHAT WILL YOU FIND IN THIS BOOK?

This book is definitely not designed as an encyclopedia of quantitative trading techniques or terminologies. It will not even be about specific profitable strategies (although you can refine the few example strategies embedded here to make them quite profitable). Instead, this is a book that teaches you how to find a profitable strategy yourself. It teaches you the characteristics of a good strategy, how to refine and

backtest a strategy to ensure that it has good historical performance, and, more importantly, to ensure that it will remain profitable in the future. It teaches you a systematic way to scale up or wind down your strategies depending on their real-life profitability. It teaches you some of the nuts and bolts of implementing an automated execution system in your own home. Finally, it teaches you some basics of risk management, which is critical if you want to survive over the long term, and also some psychological pitfalls to avoid if you want an enjoyable (and not just profitable) life as a trader.

Even though the basic techniques for finding a good strategy should work for any tradable securities, I have focused my examples on an area of trading I personally know best: statistical arbitrage trading in stocks. While I discuss sources of historical data on stocks, futures, and foreign currencies in the chapter on backtesting, I did not include options because those are quite complicated to backtest for someone new to algorithmic trading. If you are really keen on learning that, please read Chan (2017).

The book is organized roughly in the order of the steps that traders need to undertake to set up their quantitative trading business. These steps begin at finding a viable trading strategy (Chapter 2), then backtesting the strategy to ensure that it at least has good historical performance (Chapter 3), setting up the business and technological infrastructure (Chapter 4), building an automated trading system to execute your strategy (Chapter 5), and managing the money and risks involved in holding positions generated by this strategy (Chapter 6). I will then describe in Chapter 7 a number of important advanced concepts with which most professional quantitative traders are familiar, and finally conclude in Chapter 8 with reflections on how independent traders can find their niche and how they can grow their business. I have also included an appendix that contains a tutorial on using MATLAB.

You'll see two different types of boxed material in this book:

- Sidebars containing an elaboration or illustration of a concept
- Examples, accompanied by Excel (for some), MATLAB, Python, and R codes (for all)

Readers who want to learn more and keep up to date with the latest news, ideas, and trends in quantitative trading are welcome to visit my blog predictnow.ai/blog, where I will do my best to answer their questions. You will find that the website contains articles and presentations of many aspects of quantitative trading. Readers of this book will have free access to the software codes contained in this book located at epchan.com/book and will find the password in a later chapter to enter that website.

—ERNEST P. CHAN
October 2020

REFERENCES

Economist. 2019. “March of the Machines. The stockmarket is now run by computers, algorithms, and passive managers.” October 5. www.economist.com/briefing/2019/10/05/the-stockmarket-is-now-run-by-computers-algorithms-and-passive-managers.

Lux, Hal. 2000. “The Secret World of Jim Simons.” *Institutional Investor Magazine*, November 1.

Poundstone, William. 2005. *Fortune’s Formula*. New York: Hill and Wang.

