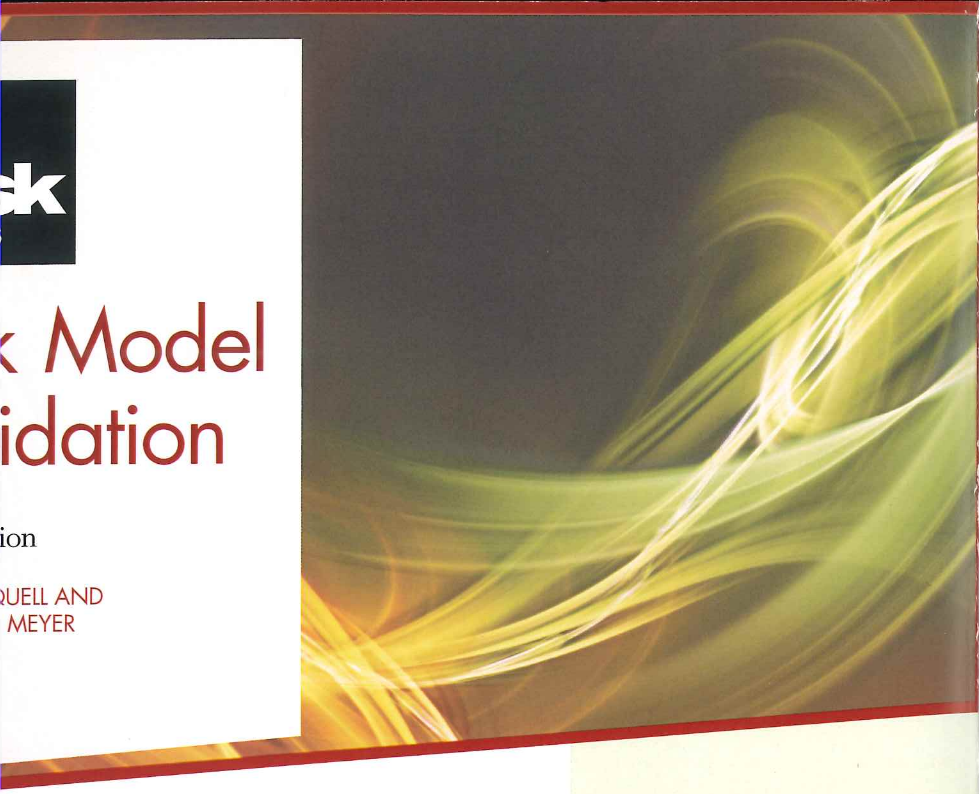


itative risk models have been presented as f the causes of the financial crisis that started 7. In this fully updated second edition, authors ian Meyer and Peter Quell give a holistic view models: their construction, appropriateness, tion and why they play such an important role financial markets.

w edition provides financial institutions with a to raise the key questions when it comes to ting the results of quantitative risk models into ss decisions.

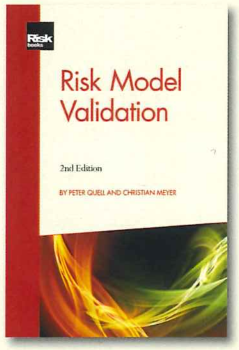
rs will be able to:

- uate the validity of a model;
- e the model's quality, consistency egulatory compliance;
- ove a framework for validation; and
- a model-risk approach for their institution

rs include:

- s of Quantitative Risk Models
- Can a Risk Model Fail?
- Regulatory Perspective on Risk Model Validation
- ation Toolbox 1: Focus on Model Results
- ation Toolbox 2: Focus on Model Assumptions
- ation Toolbox 3: Focus on Data and Software
- ementing a Model Risk Framework

oks.com/risk-model-validation-2nd



Book: £145
ISBN: 9781782722632

eBook: £110
ISBN: 9781782722991

For more information and to order:

Tel:
+44 (0) 870 240 8859

Email:
books@incisivemedia.com

The Journal of Investment Strategies
EDITORIAL BOARD

Editor-in-Chief

ARTHUR M. BERD General Quantitative LLC

Advisory Board

ROBERT ENGLE NYU Stern School of Business
KENNETH A. FROOT Harvard Business School
ROBERT JARROW Cornell University, Johnson School of Business

Editorial Board

JESPER ANDREASEN Danske Bank
MARCO AVELLANEDA NYU Courant Institute
VINEER BHANSALI PIMCO
JEAN-PHILIPPE BOUCHAUD Capital Fund Management
PETER CARR Morgan Stanley
JOHN CHISHOLM Acadian Asset Management
DEAN CURNUTT Macro Risk Advisors
VLADIMIR FINKELSTEIN Blue Mountain Capital
ROSS GARON Cubist Systematic Strategies
LISA GOLDBERG UC Berkeley
PHILIPPE GOUGENHEIM Gougenheim Investments
ALI HIRSA Sauma Capital LLC
PHILIPPE ITHURBIDE Amundi Asset Management
PETTER KOLM NYU Courant Institute
CHRIS LIMBACH PGGM
ALEX LIPTON Bank of America Merrill Lynch
MARCOS LOPEZ DE PRADO Guggenheim Partners and Lawrence Berkeley National Laboratory
DILIP MADAN University of Maryland
ATTILIO MEUCCI SYMMYS
JACQUES PEZIER University of Sussex
VLADIMIR PITERBARG Barclays Capital
JEFFREY ROSENBERG Blackrock
PIERRE SARRAU Blackrock
EUAN SINCLAIR Bluefin Trading
DIDIER SORNETTE ETH Zurich
CYRILLE URFER Gonet & Cie
PETER ZANGARI MSCI Barra

SUBSCRIPTIONS

The Journal of Investment Strategies (Print ISSN 2047-1238 | Online ISSN 2047-1246) is published quarterly by Infopro Digital, Haymarket House, 28–29 Haymarket, London SW1Y 4RX, UK. Subscriptions are available on an annual basis, and the rates are set out in the table below.

	UK	Europe	US
Risk.net Journals	£2045	€2895	$3225
Print only	£775	€1075	$1265
Risk.net Premium	£2895	€4145	$4545

Academic discounts are available. Please enquire by using one of the contact methods below.

All prices include postage. All subscription orders, single/back issues orders, and changes of address should be sent to:

UK & Europe Office: Infopro Digital, Haymarket House, 28–29 Haymarket, London SW1Y 4RX, UK. Tel: +44 (0) 207 316 9300

US & Canada Office: Infopro Digital, 55 Broad Street, Floor 22, New York, NY 10005, USA. Tel: +1 646 736 1850

Asia & Pacific Office: Infopro Digital, Unit 1704-05 Berkshire House, Taikoo Place, 25 Westlands Road, Hong Kong. Tel: +852 3411 4888

Website: www.risk.net/journals **E-mail:** info@risk.net

The Journal of Investment Strategies

GENERAL SUBMISSION GUIDELINES

The Journal of Investment Strategies welcomes submissions from practitioners as well as academics. Manuscripts and research papers submitted for consideration must be original work that is not simultaneously under review for publication in another journal or other publication outlets. All papers submitted for consideration should follow strict academic standards in both theoretical content and empirical results. Papers should be of interest to a broad audience of sophisticated practitioners and academics.

Submitted papers should follow *Webster's New Collegiate Dictionary* for spelling, and *The Chicago Manual of Style* for punctuation and other points of style. Papers should be submitted electronically via our online submissions site:

https://editorialexpress.com/cgi-bin/e-editor/e-submit_v15.cgi?dbase=risk

Please clearly indicate which journal you are submitting to.

Papers should be submitted as either a LaTeX file or a Word file ("source file"). The source file must be accompanied by a PDF file created from the version of the source file that is submitted. LaTeX files need to have an explicitly coded bibliography included or be sent with a BBL file. All files must be clearly named and saved by author name and date of submission.

A concise and factual abstract of between 150 and 200 words is required and it should be included in the main document. Four to six keywords should be included after the abstract. Submitted papers must also include an Acknowledgements section and a Declaration of Interest section. Authors should declare any funding for the paper or conflicts of interest. In-text citations should follow the author-date system as outlined in *The Chicago Manual of Style*. Reference lists should be formatted in APA style.

The number of figures and tables included in a paper should be kept to a minimum. Figures and tables must be included in the main PDF document and also submitted as clearly numbered editable files (please see the online submission guidelines for guidance on editable figure files). Figures will appear in color online, but will be printed in black and white. Footnotes should be used sparingly. If footnotes are necessary then these should be included at the end of the page and should be no more than two sentences. Appendixes will be published online as supplementary material.

Before submitting a paper, authors should consult the full author guidelines at:

http://www.risk.net/static/risk-journals-submission-guidelines

Queries may also be sent to:

The Journal of Investment Strategies, Infopro Digital, Haymarket House,
28–29 Haymarket, London SW1Y 4RX, UK
Tel: +44 (0)20 7004 7531; Fax: +44 (0)20 7484 9758
E-mail: journals@infopro-digital.com

The Journal of

Investment Strategies

The journal

The Journal of Investment Strategies is an international refereed journal focusing on the rigorous treatment of modern investment strategies.

Investment strategy research is a distinct subject matter, combining insights from various areas of financial economics and utilizing techniques from econometrics and statistics, among other fields. As an applied field of research, it directly impacts the practice of asset management, a large and diverse industry with many constituents, including traditional and alternative buy-side investment managers as well as the sell-side and independent advisors. As an academic topic, it presents unique and interesting challenges for understanding the sources of investment returns and for formulating consistent and systematic methodologies for portfolio management in a dynamic context.

Much research in this area, particularly the studies going beyond classical portfolio theory, is not currently accessible to a wider audience. In particular, many research papers originating from the industry are not well distributed. *The Journal of Investment Strategies*, therefore, has three fundamental aims:

(1) to foster high-quality, original and innovative work on investment strategies;
(2) to provide practitioners and academics with access to the resulting technical research; and
(3) to serve as an educational forum on timely issues concerning investment strategies.

Content Guidelines

Topics considered for publication in the journal include:

- Fundamental strategies;
- Relative value strategies;

- Tactical strategies;
- Event-driven strategies;
- Algorithmic trading strategies;
- Principal investment strategies;
- Portfolio management and asset allocation;
- Econometric and statistical methods.

The Editorial Board will consider two types of submission:

- Full-length original papers covering both theory and practice; and
- Short expository or discussion papers for publication in the Investment Strategy Forum.

Research Papers

Research papers are typically full-length academic papers of between 8,000 and 10,000 words in length. Topics for research can be of interest to both academic and industry professionals. Papers submitted for publication must be original and should not have been published or considered for publication elsewhere.

Investment Strategy Forum

The Forum is intended to provide rapid communication on findings and ideas about investment strategies that are topical, expository and educational in nature. The mission of the Forum is to promote active discussions of current issues. Articles should not exceed 6,000 words. The main goal of these submitted papers is to educate a wider audience and to increase understanding of the issues and topics that may not be easily accessible to either academics or practitioners.

CONTENTS

Editor-in-Chief: Arthur M. Berd
Publisher: Nick Carver
Journals Manager: Sarah Campbell
Editorial Assistant: Ciara Smith
Subscription Sales Manager: Aaraa Javed
Global Key Account Sales Director: Michelle Godwin
Composition and copyediting: T&T Productions Ltd
Printed in UK by Printondemand-Worldwide

© Infopro Digital Risk (IP) Limited, 2017. All rights reserved. No parts of this publication may be reproduced, stored in or introduced into any retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the copyright owners.

Risk journals

LETTER FROM THE EDITOR-IN-CHIEF

Arthur M. Berd
Founder and CEO, General Quantitative LLC

Welcome to the first issue of the seventh volume of *The Journal of Investment Strategies*, in which we present four papers that I believe will be of special interest to active investors and portfolio managers while also hopefully being of interest to academic and industry researchers.

In the issue's first paper, "Efficient trading in taxable portfolios", Sanjiv R. Das, Dennis Yi Ding, Vincent Newell and Daniel N. Ostrov tackle the very important but relatively underexplored topic of tax-efficient portfolio management. The vast majority of standard portfolio models ignore this effect, and they are strictly speaking only applicable to nontaxable or tax-deferred investments such as those from tax-exempt institutions or from individual retirement accounts. As soon as one is dealing with taxable portfolios, the nonlinear and path-dependent nature of taxes makes this problem very nontrivial. Its impact, though, is difficult to overestimate as taxes may account for a very sizeable portion of returns – especially in the case of active management, which can generate short-term capital gains – and the difference between the pre- and after-tax returns could be extremely important.

The authors focus on a specific problem setting related to life-cycle investment management, ie, the investment strategy that changes the allocation between risky and low-/no-risk assets as the investor ages and transitions from the accumulation phase to the decumulation phase of life after retirement. To make the problem tractable, they assume that the risky asset is a non-dividend-yielding stock index and that the low-risk asset is cash. Solving the problem via a Monte Carlo simulation and optimization approach, they are able to retain the full complexity of the US tax code, including limited capital gains loss deductions, differences between short-term and long-term capital gains taxes, and the treatment of taxes upon the death of a taxpayer. They are able to demonstrate the sensitivity of the solutions to various important input parameters, including capital gains tax rates, the current cost basis, the expected rates of stock return and interest rates, the volatility of the stock, the level of transaction costs, and the (very long) time horizon of the strategy. They uncover many interesting and novel results, including a relatively counterintuitive finding that rising tax rates make the larger stock allocation more preferred rather than less so.

I believe this paper will find interested readers among the many industry practitioners working in the wealth and retirement management areas. The paper's findings may have a significant impact on how they approach the practical management of their portfolios.

Our second paper, "Portfolio concentration and equity market contagion: evidence on the 'flight to familiarity' across indexing methods" by Lars Kaiser, explores the intriguing concept of "flight to familiarity" as an explanation for contagion effects in equity portfolios. The concept helps explain both similarities and differences between the large number of alternative index weighting schemes that have been discussed in the literature, and used by practitioners, over the past two decades. Some of the most prominent alternative schemes, deviating from the market capitalization weights, are the $1/N$ equal weight approach, the diversity weighted model, the inverse beta model, the minimum-variance model, the equal risk contribution model and the maximum-diversification model, to name but a few. The authors explore metrics including absolute and relative holdings overlap, Euclidean distance and the Bray–Curtis dissimilarity measure, and they demonstrate that the holding overlap is not sufficient to explain the correlations between the portfolios based on different weighting schemes.

Instead, a different regime-dependent pattern emerges, where one can see that in high-volatility, low-correlation states of the market, portfolios tend to become more concentrated and more similar, while in low-volatility, high-correlation states they become less similar (hence, the term "flight to familiarity"). This seems natural and intuitive and is explained by the fact that in low-risk states of the market the perceived accuracy of return forecasts is high and portfolio managers are therefore willing to go further afield to harvest alpha, while in high-risk states the "unfamiliar" assets are hard to predict and a more concentrated portfolio of assets that are easier to risk manage therefore makes more sense.

I think readers will benefit from the empirical evidence presented in this paper to help them understand how best to handle portfolio construction across different market regimes.

In the third paper in this issue, "Tail protection for long investors: trend convexity at work", Tung-Lam Dao, Trung-Tu Nguyen, Cyril Deremble, Yves Lemperiere, Jean-Philippe Bouchaud and Marc Potters explore the importance of trend convexity for the tail risk management of long market exposure investment strategies. The authors provide an algorithmic representation of trend-following strategies and demonstrate that their performance can be largely explained by the difference between the long-term and short-term variance (measured under an assumption of zero mean) of the underlying asset returns. The trend is essentially an increase in the long-run variance because the mean, if it is nonzero, contributes to the price differential, while the short-term variance remains relatively unchanged because of the $1/T$ scale that enters the computation.

Furthermore, the paper proves what some (but not all) researchers and practitioners instinctively know to be true: that trend-following strategies have strongly positive convexity and potentially negative alpha, which is entirely intuitive if one understands that these strategies essentially have option-like payoffs. See, for example, Berd (2010), in which I explained that the trend-following investor essentially bets on the wings of the prospective returns distribution realizing more than statistically expected while the middle of the distribution realizes less, which is the same expectation that an investor trading a straddle would have. Unsurprisingly, trend-following trading strategy dynamically replicates such a payoff. Dao *et al*'s paper demonstrates this thesis in full detail and quite unambiguously. Furthermore, it goes beyond second-order (volatility/convexity) exposure and also shows a dependence on the third-order (skewness) exposure of the strategy, which is highly nontrivial and has a lot to do with the long-run accumulation of tail risks.

As a corollary to the paper's findings regarding the statistical signatures of the trending strategies, the authors show that commodity trading advisor strategies provide a natural hedge for downside tail risk for both traditional and risk-parity-weighted long-only investment strategies. This is true even in the absence of positive expected performance. It is even more true, and more attractive from an investor's point of view, in the light of the well-documented long-term positive performance of trend-following strategies. I concur with the authors that having a reasonably large allocation to such strategies is highly advisable for most investors.

In the issue's fourth and final paper, "Speed and dimensions of trading", Boris Gnedenko and Igor Yelnik look at another important aspect of trading strategies. While much attention has been devoted to the risk-based analysis of portfolio composition, there has been much less focus on the risk-based study of trading and turnover. However, for any active strategy these are critical issues, and having only a nominal description of the turnover is clearly an oversight. The authors give a fairly comprehensive treatment of this problem and describe an elegant framework for analysis, deriving metrics such as effective number of trades and effective number of trading dimensions.

The authors' approach is very compelling: it is certainly natural to think about turnover and trading costs in the same terms as the current composition of the portfolio. And for most quantitative portfolio managers the latter is best described in terms of risk bets, not in terms of asset weights. The turnover of risk bets is therefore also the proper measure for describing trading activity. It allows us to compare, on a like-for-like basis, the predicted transaction costs and the projected excess returns within each effectively independent risk bet. In the process of doing this, the authors naturally develop the invariant metrics that describe the effective number of these risk bets, and the effective number of trades within each risk bet. They represent, respectively, the dimension and the speed of trading; indeed, they represent a very coherent and clear description of the trading activity in a complex portfolio. I found this paper very illuminating and I am sure many of our readers will too.

In conclusion, I hope that this issue will prove to be a useful reference for practitioners and will inspire researchers in industry and academia both to pose more probing questions and to look at those questions from nonstandard angles. Our job at *The Journal of Investment Strategies* is to promote such diffusion of ideas and a fresh take on problems.

REFERENCES

Berd, A. M. (2010). Investment strategy returns: volatility, asymmetry, fat tails and the nature of alpha. In *Lessons from the Financial Crisis*, A. M. Berd (ed.). Risk Books, London.

Journal of Investment Strategies **7**(1), 1–40
DOI: 10.21314/JOIS.2017.094

Risk journals

Research Paper

Efficient trading in taxable portfolios

Sanjiv R. Das, Dennis Yi Ding, Vincent Newell and Daniel N. Ostrov

Department of Mathematics and Computer Science, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053, USA; emails: srdas@scu.edu, dennis.yi.ding@gmail.com, vincent.newell@gmail.com, dostrov@scu.edu

(Received February 7, 2017; accepted October 23, 2017)

ABSTRACT

We determine life-cycle trading strategies for portfolios subject to the US tax system. Our method employs Monte Carlo optimization. It accommodates long horizons (between forty and sixty years) and large numbers of trading periods (eg, 480), while accounting for the full cost basis history of the portfolio's stock holdings, thus sidestepping the curse of dimensionality. We present many new results that provide insights into questions about taxable portfolio investing which were previously unexplorable. Some of our results challenge current conventional wisdom. For instance, we establish circumstances where raising the allocation of stock is optimal though counterintuitive and demonstrate the suboptimality of the 5/25 rebalancing rule, even as a rule of thumb.

Keywords: taxes; capital losses; portfolio optimization; simulation; path dependence; expected utility.

1 INTRODUCTION

> Death and taxes and childbirth! There's never any convenient time for any of them.
>
> Margaret Mitchell, *Gone with the Wind*

Taxes make financial problems both more complicated and more interesting. The optimal portfolio problem is no exception to this general statement. Due to the fact

Corresponding authors: S. R. Das and D. N. Ostrov

Print ISSN 2047-1238 | Online ISSN 2047-1246

© 2017 Infopro Digital Risk (IP) Limited

that capital losses may be deducted from an investor's ordinary income – which is taxed at a higher rate than (long-term) capital gains – the US tax code injects two options into portfolio analysis. First, Constantinides (1983) argues that, ceteris paribus, when tax rates on losses are the same as (or greater than) rates on gains, portfolio policies become tilted toward opting to immediately realize any losses, ie, exercising a "tax put" option. This immediate exercise rule for the tax put option is easy to implement. Second, Constantinides (1984) points to a more subtle tax-timing (or "restart") option, where reaping the gains resets the cost basis, and therefore resets the tax put to be at the money, which allows investors to better exploit the higher rate for losses. Both of these tax options are important, but it is far more challenging to determine how to exercise the second option optimally.

We consider a dynamic portfolio problem, where, at each time t, the investor determines the fraction of their wealth, $f \in [0, 1]$, that will be placed into a risky asset, subject to capital gains taxation and capital losses tax deductions. The remaining fraction $(1 - f)$ is placed into a risk-free asset. (For simplicity, we will refer to the risky asset as "stock" and the risk-free asset as "cash" for the remainder of this paper.) Our choice of f is made in a tax-optimized manner, where the complexity of the US tax code is considered over a large number of potential trading periods (~ 500). This optimal choice of f depends on the cost basis of the stock holdings in the portfolio, which, in turn, depends on the path of the stock's price history as well as the investor's purchasing history. Thus the problem is heavily path dependent and high dimensional. In particular, the full cost basis of all the stock in the portfolio can be described by a vector of random, evolving length. Following this method, each stock purchase still held in the portfolio corresponds to an entry in the vector. Attempting to keep track of every possibility for the basis can quickly render the equation for determining the optimal f intractable, as discussed in Ostrov and Wong (2011), for example. Therefore, many papers assume that it is approximately acceptable to use a weighted-average cost basis encompassing all stock in the portfolio, instead of using the actual cost basis for each stock purchase. This assumption allows us to replace the vector describing the full basis with a scalar representing the average cost basis. This change can make solving a (dynamic programming) Bellman equation to determine the optimal f tractable.[1] Examples of papers that have used the average cost basis include Dammon *et al* (2001, 2004), Gallmeyer *et al* (2006), Tahar *et al* (2010) and Dai *et al* (2015).

[1] The state space for the dynamic programming problem comprises the current wealth of the investor and the tax basis. If the weighted-average tax basis is used, then the state space will be two-dimensional. If the complete tax basis is used, then the dimension of the state space has no definite upper bound.

Dybvig and Koo (1996) implemented a model using the full cost basis, but they were limited to four periods before the problem became intractable. DeMiguel and Uppal (2005) subsequently used nonlinear programming, namely the Sparse Nonlinear OPTimizer (SNOPT) algorithm devised by Gill *et al* (2002). By adding linear constraints, they were able to extend the use of the model to ten periods. For their ten-period model, they found that using the full cost basis provided a 1% certainty equivalent advantage over using the average cost basis, which is robust to various parameterizations.

In contrast with these previous models, we use a simulation algorithm with optimization that runs very fast. Our algorithm uses a nonlinear optimizer in the language of R programming that calls a simulation run in compiled C for 50 000 paths of portfolio wealth and policy. We generally run our algorithm for forty-year portfolio horizons with quarterly trading, which equates to 160 potential trading periods. However, we also implement the model for longer horizons, eg, sixty years, or for monthly trading, ie, 480 periods, which takes marginally longer. Our approach has the advantage of being both simple and flexible. This yields a number of benefits, such as allowing us to

(1) determine solutions using the full cost basis, even if the basis is large and complicated;

(2) address a number of complex features of the US tax code that are generally difficult to accommodate; and

(3) work with any desired stochastic processes to simulate the stock movement (although, for simplicity, we use a geometric Brownian motion in this paper).

Our model keeps track of the entire tax basis, however, we can also implement a simplification to accommodate the average basis. This will allow us to compare the effect of using the average cost basis with that of the full cost basis. This is key for understanding whether or not it is justified to use the average cost basis approximation that current dynamic programming approaches (eg, the Bellman equation) implement. DeMiguel and Uppal (2005) argue that using the average tax basis is reasonable, since, in their simulations, most stock holdings comprise a single basis. They state that only 4% of holdings comprise additional bases. However, their conclusion is based on a ten-period model, where there are few periods in which to make new purchases. With many more periods and more complex tax code features, as in our model, there is greater potential for the difference in optimality between the full tax basis problem and the average tax basis problem to increase. Having a large number of periods is more like a Bellman model, where potential sales and purchases occur in continuous time or over multiple periods in discrete time.

A comprehensive paper by Dammon *et al* (1989) examines both the tax put option and the restart option (see also a review by Dammon and Spatt (2012)). Their paper shows that the value of these options depends on the pattern of gains, the length of the portfolio's time horizon (T) and whether, at this horizon, the investor is alive and liquidates their portfolio or is deceased. (In the case of a deceased investor, taxes will be lower, since all capital gains are forgiven upon death and the basis is reset.) We will explore the effect of the investor's status – alive or deceased – at the horizon T and show how this may lead to substantially different optimal strategies, even when T is large. A difference between our paper and that of Dammon *et al* (1989) is that they do not seek to determine an optimal trading policy. Instead, they examine three trading rules suggested by Constantinides (1984) and compare the performance of these rules with a buy-and-hold strategy. For high volatility stocks, they find that the value of the restart option is lower than that suggested by Constantinides (1984) because such stocks also tend to have higher capital gains. They conclude that offset rules matter. Hence, in our model, we account for tax loss offsets and carryforwards exactly as per the US tax code.

It is well established (see, for example, Davis and Norman (1990), Dai *et al* (2011) and Dai *et al* (2015); or, in the asymptotic case, see Shreve and Soner (1994), Whaley and Wilmott (1997), Atkinson and Mokkhavesa (2002), Janaček and Shreve (2004), Rogers (2004) or Goodman and Ostrov (2010)) that, in the absence of taxes, the optimal trading strategy is to maintain the stock fraction within a specific "no-transaction region" (also referred to as a "hold region"). When the portfolio's stock fraction is in the interior of this no-transaction region, there is no trading. Trading only occurs to rebalance the portfolio so that it does not escape the no-transaction region. The same logic applies to the optimal strategy for taxable portfolios, but with one tweak: the optimal strategy, as per Constantinides (1983) and Ostrov and Wong (2011), requires realizing any capital losses as soon as possible (that is, exercising the tax put option), no matter where we are in the region, and then rebuying an essentially equivalent stock or index to avoid wash sale restrictions, as discussed further in Section 2.1. This process collects the tax advantages of realized losses without rebalancing the portfolio. As before, we only rebalance to prevent the portfolio from escaping our region. Because we transact to realize losses in the interior of our region, we refer to any region using this optimal strategy in the presence of taxes as a "no-rebalancing region" instead of a "no-transaction region". In the case of a single stock, as explored in this paper, the region is an interval.

Because our simulation approach works forward in time, whereas Bellman equation approaches either work backward in time or employ specific models for which the value function's dependence on time is removed, we are able to obtain new results that complement those previously obtained by Bellman equation approaches. By working forward in time, we are not restricted to using either a single cost basis or an average

cost basis. We are not restricted to approximating the tax code for losses, either by requiring all losses to be immediately claimed or only allowing losses to cancel gains. We are not encumbered by the "curse of dimensionality", which may impose less realistic model assumptions, such as an exponential distribution for the investor's time of death or a restricted class of investor utility function. Nonetheless, our approach has limitations, and there are certain cases for which the Bellman approach appears better suited. For example, in Dai *et al* (2015), the Bellman approach explores cases where it is better to sell stock even if this triggers short-term capital gains taxes. Our approach cannot be generalized to determine such cases. However, our algorithm can be generalized, as we will later show, to allow the values of the lower and upper bounds of the no-rebalancing interval to depend upon the ratio of stock basis in the portfolio to current stock price, as in Dai *et al* (2015).

Working forward in time allows us to derive an optimization algorithm that reveals many new findings about the optimal no-rebalancing interval for the stock fraction. Some of our results confirm statements in the current literature, while others contradict previous claims, but many of our results provide an insight into questions hitherto unaddressed in the extant literature. The main findings of this paper are as follows.

First, we provide a simulation-based method that quickly generates an optimal portfolio trading strategy. Our method utilizes the full cost basis history, a complex, realistic tax model, and many more trading periods (we implemented 480) than have ever been considered in the literature when using the full cost basis history (Haugh *et al* (2016), for instance, consider twenty periods). Our model can be extended in a number of directions, making it possible to accommodate just about any stochastic model for stock price evolution and any utility function that is applied to the portfolio worth at the horizon time, T. Our model can also be applied to multiple stocks; however, this would require certain assumptions about the geometric nature of the no-rebalancing region. By contrast, the approach applied by DeMiguel and Uppal (2005) to the multiple stock question required no a priori knowledge about the geometry of the no-rebalancing region, though it was computationally limited to seven trading periods for two stocks and four trading periods for four stocks.

Second, our portfolio rule comprises an optimal no-rebalancing interval for the stock fraction that can be specified by the interval's center, denoted by f^*, and the interval's width, denoted by Δf. We show that optimality is reduced much more by movement away from the optimal center than movement away from the optimal width. When the optimal interval width is positive, as opposed to zero, the reduction in optimality, due to perturbations of the interval width from its optimal value, is particularly small.

Third, as in Dammon *et al* (1989), we find that materially different optimal portfolio choices can be made depending on whether the investor is assumed to continue living or to expire upon liquidation of the portfolio at the horizon time, T. Moreover, our results indicate that this difference in optimal strategy does not vanish as T increases, even though the tax treatment only varies at time T, suggesting that these strategies have long memory.

Fourth, counterintuitively, the optimal stock fraction in a taxable portfolio (or, more specifically, f^*, the interval's center, if the interval width $\Delta f > 0$) is often higher than the optimal stock fraction for a portfolio without taxes, rather than lower. This is the case whether or not the investor is assumed to be alive or dead when the portfolio is liquidated at time T. Because there are generally more capital gains than capital losses, it is intuitive to think that taxes should make the stock less desirable. However, the tax rate used to credit losses is generally significantly higher than the tax rate for gains, a fact that may often be exploited to make the taxable stock more, not less, desirable. We analyze when this is the case.

Fifth, our analysis shows that the static band rules commonly used to determine the no-rebalancing region are generally problematic. The "5/25" rule of thumb, for example, suggests that for our stock–cash scenario we should use a no-rebalancing interval with a width of $\Delta f = 0.10$, since this corresponds to ± 5 percentage points for the portfolio's stock fraction.[2] Our results, however, show that the optimal width of the no-rebalancing interval can vary dramatically (and often turns out to be zero), depending on a number of parameters. Specifically, the optimal width increases as we increase the stock's expected return, the capital gains tax rate, or the size of the portfolio, and decreases as we increase the risk-free return, the capital losses tax rate, or the time horizon, T, of the portfolio. The optimal width also increases if the investor is assumed to expire at time T. (Surprisingly, in our results, the optimal width does not appear to be influenced by the volatility of the stock, when intuitively one might expect it to be. See, for example, the trading range in Section 3.2.1.) Further, we explain why, when we roughly average the effect of all these parameters, it is better to choose $\Delta f = 0$ (that is, to be continually rebalancing) than to choose $\Delta f = 0.10$ as suggested by the 5/25 rule of thumb.

Sixth, also counterintuitively, in a case where the investor is assumed to expire at the portfolio horizon and the tax rate on gains is increasing, more, rather than less, investment in the stock is generally optimal. This is because the amount of rebalancing (along with the associated capital gains) needed to maintain the desired fraction, f, of the portfolio in stock decreases as f increases from $f = 0.5$ to $f = 1$. Indeed, to maintain $f = 1$, no rebalancing is needed, so no capital gains are generated.

[2] The 5/25 rule of thumb recommends that a portfolio be rebalanced if the actual fraction of the portfolio in an asset class deviates from the desired fraction by five percentage points (in cases where the desired fraction is 25% or higher), or if the actual fraction deviates by more than 25% from the desired fraction (in cases where the desired fraction is less than 25%).

Seventh, we assessed monthly, quarterly and semiannual trading schemes, and found that the choice between these three trading frequencies has no material effect on the optimal allocation strategy. This assessment is novel in comparison with the extant literature, which does not assess this issue due to limitations in the number of periods it has been possible to consider.

Eighth, using a large collection of scenarios, mostly with forty-year time horizons, we find that, on average, using the full cost basis provides only a 0.65% certainty equivalent advantage over using the average cost basis if the investor is assumed to be alive at the portfolio horizon time, T. This advantage reduces even further to 0.27% when the investor is assumed to expire at time T. These numbers are a bit lower than the 1% figure reported by DeMiguel and Uppal, further supporting the validity of using the average cost basis approximation approach outlined in optimization models, such as the Bellman models that require the approximation for tractability.

Ninth, as in Dai *et al* (2015), we allow the no-rebalancing interval's upper and lower bounds to depend on the ratio of the basis price of the portfolio's stock to the current stock price. Dai *et al* (2015) employ the average cost basis price. However, since we allow for the full cost basis, we select the highest cost basis price in the portfolio, since that corresponds to the first stock traded. We find that allowing the no-rebalancing region's bounds to depend on this ratio of the highest cost basis to the current price yields no discernible certainty equivalent advantage.

Tenth, for completeness and comparison with Leland (2000), we also reoptimize the taxable portfolio in the presence of proportional transactions costs. We confirm Leland's result that transaction costs reduce portfolio churn (ie, they increase the optimal width Δf). However, unlike Leland, we do not find that transaction costs materially reduce the optimal fraction of the portfolio in the stock position. That is, they do not reduce the optimal f^*.

In Section 2, we present our modeling approach, explicitly accounting for various features of the tax code and for the evolving cost basis of all the stock in the portfolio. Section 3 follows with numerical simulations of the model. We report and explain the effect on the optimal trading strategy caused by varying the values of the stock's expected return, the cash interest rate, the stock's volatility, the investor's risk aversion, the tax rate on losses, the tax rate on gains, the initial portfolio worth, the portfolio's time horizon and the periods between trading opportunities. We then report and explain the effect on the optimal trading strategy of letting the strategy change halfway to liquidation (at time $T/2$), of using the average cost basis in place of the full cost basis, of allowing the optimal strategy to depend on the ratio of the portfolio's highest cost basis to the current stock price, and of incorporating transaction costs into the model. Section 4 concludes.

2 MODEL

In this section, we present the assumptions behind our model and the details of our simulation of the model.

2.1 Assumptions and notation

We have two assets in our portfolio model: stock and cash. The return on the stock is risky; the return on the cash is certain. Our model applies two sets of assumptions: those for the stock and cash positions, and those for the tax model.

For the stock and cash positions, we make the following assumptions.

(1) We assume the stock evolves via a geometric Brownian motion with a constant expected return, μ, and a constant volatility, σ.

(2) The tax-free continuously compounded interest rate for the cash position, r, is assumed to be constant. We note that cash positions with constant interest taxed at a constant rate can be converted to an equivalent tax-free rate, r, and, of course, periodically compounded interest at a constant rate can also be converted to a continuously compounded rate, r.

(3) Except where otherwise specified, we assume that stock and cash can be bought and sold in any quantity, including non-integer amounts, with negligible transaction costs.

In portfolios subject to tax, the impact of taxes on the optimal strategy is generally much stronger than the impact of transactions costs. At the same time, tax codes are generally more complex than models for transactions costs. For tax-free portfolios like 401(k)s and Roth individual retirement accounts (IRAs), where transaction costs take center stage, there is an extensive literature on portfolio optimization (see Akian *et al* 1996; Atkinson and Ingpochai 2006; Bichuch 2012; Goodman and Ostrov 2010; Leland 2000; Liu 2004; Muthuraman and Kumar 2006).

(4) For simplicity, we do not consider dividends for the stock, although the model can easily be altered to approximate the effect of dividends by adjusting μ, the growth rate of the stock.

For the tax model, we make the following assumptions.

(1) As stipulated by the tax code in the United States, we assume a limit of no more than US$3000 in net losses can be claimed at the end of each year. Net losses in excess of this amount are carried over to subsequent years. Should the investor expire when the portfolio is liquidated at time T, all remaining carried-over

capital losses are lost. We encode all these features into our portfolio simulation program.

(2) For simplicity, we assume that for all times prior to the portfolio being liquidated, the capital gains tax rate, τ_g, is constant and applies to both long- and short-term gains. According to the US tax system, when the portfolio is liquidated, we employ one of two capital gains rates ($\tau^{liq} = \tau_g$ if the investor is alive, or $\tau^{liq} = 0$ if the investor is dead) to reflect the fact that capital gains are forgiven when an investor expires. Similarly, for both short- and long-term capital losses that are claimed prior to the liquidation of the portfolio, we assume a constant rate, τ_l, which corresponds to the marginal income tax rate of the investor. Our model can be altered to accommodate different rates for short- and long-term gains and losses. However, optimizing with this short-term/long-term model poses difficulties, unless one restricts the strategy to something reasonable but possibly suboptimal, such as assuming that short-term gains will never be realized.

(3) We allow for wash sale rules, but we make an additional assumption so that they have no effect. Specifically, we assume the presence of other stocks or stock indexes in our market with essentially the same value of μ and σ. For a loss to qualify for tax credit, wash sale rules in the US stipulate that the investor must wait at least thirty-one days before repurchasing a stock that was sold at a loss. But if the investor sells all of a stock that is at a loss and then immediately buys a different stock with the same μ and σ, the investor will avoid triggering a wash sale, while still taking full advantage of the loss for tax purposes. This strategy of selling and repurchasing lowers the cost basis of any stock with losses. Yet, because it allows for earlier use of the losses for taxes, it is always superior to the strategy of buying and holding when transaction costs are negligible (see, for example, Constantinides 1983; Ostrov and Wong 2011).

2.2 Trading strategy

We implement the following trading strategy. For simplicity, we start with a portfolio that is strictly cash, although initial portfolios containing stock positions with various cost bases could just as easily be accommodated. We then immediately buy stock so that the portfolio's stock allocation attains a selected fraction, f^{init}, of the total portfolio's worth.

After each time period of length h years passes in the simulation, we consider three types of trade. First, we sell and repurchase any stock with a loss, avoiding wash sale restrictions as described above, to generate money from these capital losses.[3] Second, if f (the fraction of the portfolio's value in the stock position) is below a selected lower threshold, f^{l} (which is constant over time), we purchase stock until $f = f^{\mathrm{l}}$. Third, if f is above a selected upper threshold, f^{u} (which is also constant over time), then the stock with the highest cost basis is sold until $f = f^{\mathrm{u}}$. Choosing to sell stock with the highest cost basis is optimal, since it minimizes capital gains.

At the end of each year, taxes on any net capital gains are paid from the cash position, and any net losses that can be realized are used to purchase additional stock.

We keep track of the cost basis of all stock purchases. This makes both the problem and the portfolio value path dependent, ruling out any easy dynamic programming approach to determining the optimal stock proportions at each point in time for the portfolio, as explained in the introduction.

Our goal is to determine the three values, f^{init}, f^{l} and f^{u}, that may optimize the expected utility of the portfolio at a specified final portfolio liquidation time, T. Our model can easily accommodate any utility function; however, for our own simulations, we have chosen power law utility functions (ie, constant relative risk aversion). Normally, power law utilities lead to results that are independent of the initial portfolio worth. However, this will not be the case here due to the US\$3000 limit on losses that can be claimed per year.

2.3 The algorithm for simulating a single run over T years

For given values of f^{init}, f^{l} and f^{u}, our algorithm works via Monte Carlo simulation over a large number of runs. For each run, we proceed with the following algorithm.

We start by purchasing stock so that our initial cash-only portfolio attains the given initial stock fraction, f^{init}, at $t = 0$.

We simulate the market over each time period h. From time t to time $t + h$, the stock price, S, advances via a geometric Brownian motion, so

$$S_{t+h} = S_t \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)h + \sigma\sqrt{h}Z\right],$$

where Z is a standard normal random variable. The cash position, C, advances by $C_{t+h} = C_t \exp(rh)$.

As we buy and sell stocks over time, we develop a portfolio with stock positions purchased at various prices. We keep track of each of these separately to determine the correct basis for computing gains or losses when we sell the stock. We let J be the number of different purchase prices (corresponding to the various purchase times) of the stock. Define B_j, where $j = 1, 2, \dots, J$, to be these J purchase prices in order of purchase time, and let N_j be the number of stocks purchased at each of these prices.[4] Since we sell and repurchase any stock that has incurred a loss, we guarantee that $B_1 \leqslant B_2 \leqslant \cdots \leqslant B_J$. For the initial stock purchase at $t = 0$, for example, we have $J = 1$, since there is only one stock position; $B_1 = S_0$ is the initial stock price; and the number of shares bought is $N_1 = f^{\mathrm{init}} W_0 / S_0$, where W_0 is the initial worth of the portfolio.

After each time period h has elapsed, we consider our three possible trading actions in the following order.

(1) Collect any losses: define k such that $k, k+1, \dots, J$ are the only positions with losses. That is, $B_k, B_{k+1}, \dots, B_J$ are the only values of B_j that are greater than the current stock price, S. We sell all of these positions and then immediately buy them back, recalling that we are actually buying back other stocks or stock indexes that have the same μ and σ to avoid wash sale restrictions. We subtract these losses from our current gains, so the new value for the gains, denoted by G, is

$$G = G^{\mathrm{old}} - \sum_{j=k}^{J} N_j (B_j - S).$$

Since positions k through J have all been repurchased at the current stock price, it is now the case that

$$N_k = N_k^{\mathrm{old}} + N_{k+1}^{\mathrm{old}} + \cdots + N_J^{\mathrm{old}},$$

where $B_k = S$ and we reduce J to equal k, since positions $k + 1$ through J^{old} no longer exist.

(2) If the current stock fraction, f, is below f^{l}, we buy stock until $f = f^{\mathrm{l}}$. Since we are buying stock, we must create a new position at the current purchase price. Hence we increase J by one. For this new J position, we have $B_J = S$, and N_J equals the number of stocks needed to ensure that $f = f^{\mathrm{l}}$.

(3) If the current stock fraction, f, is above f^{u}, we sell stock, beginning with position J, then position $J - 1$, etc, until we have $f = f^{\mathrm{u}}$. We add these gains to G. They must be gains, because all losses have already been collected by this stage. If this procedure results in the liquidation of any positions, we reduce the value of J to reflect this.

[3] This assumes that the investor has an income not derived from their investment, so the losses can be used to lower the tax on the income in a way that is equivalent to generating money from tax shields.

[4] Should our model be altered to distinguish between short- and long-term gains and losses, the time of purchase, t_j, must also be recorded.

This trading strategy of buying and selling stock so as to stay just within the interval $[f^{\mathrm{l}}, f^{\mathrm{u}}]$ is a standard optimizing strategy when there are either no transaction costs or proportional transaction costs, as is the case here (see, for example, Davis and Norman 1990).

At the end of each year, we either pay taxes or collect tax credits, depending on the sign of G. If $G > 0$, we have incurred gains; so, we remove $G\tau_{\mathrm{g}}$ from the cash account and then set $G = 0$. If $G < 0$, we have incurred losses that we use to buy stock. Since we are buying stock, we increase J by one and then set $B_J = S$, the current stock price. If the losses, $-G$, are less than the annual limit of US$3000, they generate $-G\tau_{\mathrm{l}}$ dollars, which purchases $N_J = (-G\tau_{\mathrm{l}})/S$ shares of stock. We may then set $G = 0$. If the losses are more than US$3000, we purchase $N_J = 3000\tau_{\mathrm{l}}/S$ shares of stock, and then we set $G = G^{\mathrm{old}} + 3000$, so the excess losses are carried over to the next year.

At the end of T years, we liquidate all the stock positions in the portfolio. If we assume the portfolio owner is alive, capital gains from this liquidation are paid. If we assume the owner has expired, no capital gains are paid. Any carried-over losses are lost.

As detailed in the next subsection, we seek to average the utility of the final portfolio worth over all of our Monte Carlo runs to determine an approximation for the expected utility. We use the same simulated stock runs for comparing the expected utilities for different f^{init}, f^{l} and f^{u} combinations. This allows us to converge to the values of f^{init}, f^{l} and f^{u} that optimize the expected utility for this fixed set of simulations. We then use different sets of simulations to check for consistency in the optimal values determined for f^{init}, f^{l} and f^{u}.

2.4 The expected utility estimator and the optimization program

Working with our model is quite computationally intensive. We generally chose a base case liquidation time, T, of forty years. Each forty-year run was computed using the complex algorithm from the previous section. We also needed to average the utility at liquidation over a high number (up to 50 000) of these runs to estimate the expected utility of the terminal wealth, $E[U(W_T)]$, corresponding to any specific values of f^{init}, f^{l} and f^{u}. Further, the optimization algorithm requires several calls to this expected utility estimator for various values of f^{init}, f^{l} and f^{u}.

Given this approach, we required very fast computation from the expected utility estimator as well as an efficient optimizer. We therefore programmed the expected utility estimator in C programming language. We also compiled and linked it in a way that made it callable from the R programming language, and therefore suitable for use with the optimizer in R, a language that is stable, fast and accurate. The optimization function in R is `constrOptim`; we chose this constrained optimizer because we needed to restrict the values of f^{init}, f^{l} and f^{u} so that $0 \leqslant f^{\mathrm{l}} \leqslant f^{\mathrm{init}} \leqslant f^{\mathrm{u}} \leqslant 1$.

We applied a power utility function,

$$U(W) = \frac{W^{1-\alpha}}{1-\alpha}, \tag{2.1}$$

to the terminal wealth, although we can easily accommodate any other utility function in our simulation program. The power law utility, however, is particularly suited to our model – where f^{l} and f^{u} are constant with respect to time – because, in the absence of taxes, Merton (1992) showed that, for the power law utility, the optimal fraction of the portfolio held as stocks is

$$f_{\text{Merton}} = \frac{\mu - r}{\alpha\sigma^2},$$

which is also constant with respect to time.

Given our choice of the power law utility, our estimate of the expected utility at time T becomes

$$E[U(W_T)] \approx \frac{1}{M}\sum_{m=1}^{M}\frac{1}{1-\alpha}(W_T(m))^{1-\alpha},$$

where α is the coefficient of relative risk aversion for the investor, m indexes each simulated run, M is the total number of runs and $W_T(m)$ is the terminal wealth for run m generated in the simulation.

Optimization is run in two stages. In the first stage, we find the values of f^{init}, f^{l} and f^{u} that are optimum over a case with only $M = 1000$ runs. This stage is fast, since M is small. We then use these three values as starting guesses for the optimum f^{init}, f^{l} and f^{u} in the second stage, where we optimize over $M = 50\,000$ runs. This two-step process has the benefits of speed from the first stage and precision from the second stage. We used standard computing hardware for this simulated optimization procedure, and each optimization runs in under five minutes. We obtained similar results when using other sets of 50 000 sample paths generated using independent sets of random numbers.

3 RESULTS AND ANALYSIS

In this section, we present and analyze numerical results obtained from the simulation model described in Section 2. We begin in Section 3.1 with a discussion of the optimal stock fraction range, $[f^{\mathrm{l}}, f^{\mathrm{u}}]$ (ie, the optimal no-rebalancing interval) for the "base case", in which we assign the below values to the following nine parameters:

(1) the stock growth rate, $\mu = 7\% = 0.07$ (per annum);

(2) the risk-free rate, $r = 3\% = 0.03$ (per annum);

(3) the stock volatility, $\sigma = 20\% = 0.20$ (per annum);

(4) the risk-aversion parameter, $\alpha = 1.5$ in our utility function in (2.1);

(5) the tax rate on losses, $\tau_{\mathrm{l}} = 28\% = 0.28$;

(6) the tax rate on gains, $\tau_{\mathrm{g}} = 15\% = 0.15$;

(7) the initial portfolio worth, $W_0 =$ US\$100 000;

(8) the time horizon before portfolio liquidation, $T = 40$ years;

(9) the period between potential trades, $h = 0.25$ years (ie, quarterly trading and rebalancing).

We will consider the base case in two scenarios: when the investor expires at $T = 40$, so that any remaining capital gains are forgiven, and when the investor is alive at $T = 40$, so that taxes on any remaining net capital gains must be paid.

In Section 3.2, we show the sensitivity of this optimal stock fraction range by varying these nine parameter values, one at a time, from their base case values. Then, in Section 3.3, we consider the effects of changing the model in four ways:

(1) letting the stock fraction range, $[f^{\mathrm{l}}, f^{\mathrm{u}}]$, change values at time $T/2 = 20$ years;

(2) using the average cost basis instead of the full cost basis;

(3) allowing the bounds of the no-rebalancing interval to depend upon the ratio of the highest cost basis to the current stock price;

(4) incorporating proportional transaction costs when we buy and sell stock.

A no-rebalancing interval can be defined by specifying f^{l} and f^{u} or by specifying the components

$$f^* = \frac{f^{\mathrm{l}} + f^{\mathrm{u}}}{2} \quad \text{(the center (or midpoint) of the interval)}$$

and

$$\Delta f = f^{\mathrm{u}} - f^{\mathrm{l}} \quad \text{(the width of the interval)}\,.$$

We often prefer to use f^* and Δf, since it makes more sense to analyze how changes affect the center and the width of the trading strategy, instead of how they affect f^{l} and f^{u}. In our simulations, we found the effect of varying f^{init} to be quite small. Therefore, we do not discuss or present the optimal values of f^{init}, even though our algorithm determines and uses them.

3.1 Optimal strategy for the base case

For the base case specified above, with a deceased investor at $T = 40$, our analysis shows that the optimal trading strategy would be to set $f^* = 0.764$ and $\Delta f = 0.168$. With a living investor at $T = 40$, the optimal trading strategy is to set $f^* = 0.711$ and $\Delta f = 0$. Note that $\Delta f = 0$ means that the investor is best off continually rebalancing.

The effect of living as opposed to dying on f^* and Δf will hold in general, and not just in the base case, as we will see throughout Section 3.1. When the investor is alive at time T and must pay capital gains, there are two key effects.

(1) Having to pay capital gains taxes at time T makes the stock less desirable, so f^* gets smaller when the investor is alive at liquidation.

(2) Having to pay capital gains taxes at time T makes having capital gains less desirable, so Δf also gets smaller.

Further, even though the status of the investor (living or deceased) at time T only affects the tax treatment at time T, it has a considerable effect on our optimal long-term investing strategy, especially on the optimal Δf, as we see here in the base case and throughout Section 3.1.

As stated previously, in the absence of taxes, Merton (1992) shows that it is optimal to keep the stock fraction equal at all times to

$$f_{\mathrm{Merton}} = \frac{\mu - r}{\alpha\sigma^2}, \tag{3.1}$$

which, for our base case, corresponds to $f_{\mathrm{Merton}} = \frac{2}{3}$. Note, however, that this Merton stock fraction is smaller than either $f^* = 0.764$ or $f^* = 0.711$. Though it may seem counterintuitive, the optimal stock fraction in the taxable accounts is higher here than in the account with no taxes.

Why? Since τ_{l}, the refund tax rate for capital losses, is higher than τ_{g}, the tax rate for capital gains, the advantage of culling capital losses from stock in a taxable account can, on average, outweigh the disadvantage of paying capital gains taxes. This is the case here. We emphasize that only the stock positions require different tax treatments. The cash positions, both in our algorithm and in the Merton expression (3.1), use the same tax-free rate, $r = 0.03$. That is, we are not choosing a higher stock fraction in the taxable portfolio because it is better to keep the cash position shielded from taxes, as investors are often advised when considering a taxable account versus a 401(k) or Roth account.

Though our algorithm determines the optimal f^* and Δf, it can easily be simplified to determine the loss incurred by an investor should they employ a suboptimal f^*

and/or Δf. We quantify this loss using the certainty equivalent, C, which is defined by

$$U(C) = E[U(W_T)],$$

where W_T is the portfolio worth at time T. Further, the investor has no preference between starting at $t = 0$ with Ce^{-rT} dollars that must be invested as cash at the risk-free rate r until time T (leading to C dollars at $t = T$) or starting at $t = 0$ with W_0 dollars invested in our stock and cash portfolio until time T (leading to a random variable for the value of W_T).

To quantify the disadvantage of using a suboptimal strategy rather than the optimal one, we employ $C_{\text{subopt}}/C_{\text{opt}} < 1$, the ratio of certainty equivalents under these two strategies, to define c, the certainty equivalent difference, via the following equation:

$$1 + c = \frac{C_{\text{subopt}}}{C_{\text{opt}}} = \frac{U^{-1}(E_{\text{subopt}}[U(W)])}{U^{-1}(E_{\text{opt}}[U(W)])}.$$

Since $U(W) = (W^1 - \alpha)/(1 - \alpha)$, this can be reexpressed as

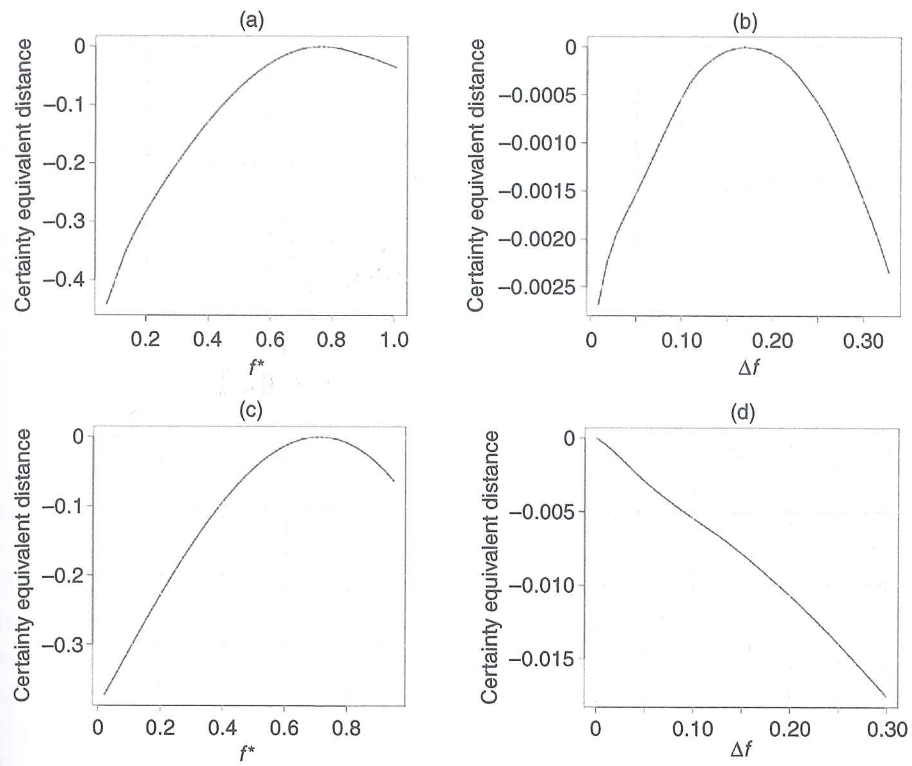
$$c = \left(\frac{E_{\text{subopt}}[U(W)]}{E_{\text{opt}}[U(W)]}\right)^{1/(1-\alpha)} - 1 \leqslant 0. \tag{3.2}$$

Note that the optimal strategy corresponds to $c = 0$, and c becomes progressively negative as the strategy becomes more suboptimal.

Figure 1 shows the effect on c when we deviate from the optimal strategy. For parts (a) and (c), we fix Δf at its optimal value and let f^* vary. For parts (b) and (d), we fix f^* at its optimal value and let Δf vary. Noting the scale on the vertical axis, we see that the sensitivity of c to suboptimal values of f^* is far larger than its sensitivity to suboptimal values of Δf. That is, it is much more important for investors to determine their optimal portfolio stock fraction (or, more precisely, the center of their optimal no-rebalancing interval) than it is to determine the precise optimal width of this interval.

Should the investor get the precise value of the optimal f^* wrong, however, there is also considerable forgiveness built in, due to the fact that – at least in the base cases for the investor dying (Figure 1(a)) or living (Figure 1(c)) – the optimal f^* is on the interior of the domain $f \in [0, 1]$. Because of this and the smoothness of the graphs, the derivative of c is zero at the optimum value of f^*. This leads to a small loss in the certainty equivalent difference for even moderate deviations from the optimal f^*. In the two base cases shown in Figure 1, for example, we see no more than a 1% loss in the certainty equivalent difference over forty years, until the value chosen for f^* differs from the optimal f^* value by around ten percentage points.

FIGURE 1 The loss to the investor, measured as the certainty equivalent difference, resulting from employing a suboptimal strategy for the base case.



In parts (a) and (b), the investor expires during liquidation of the portfolio, whereas in parts (c) and (d) the investor is still alive at this time. Parts (a) and (c) examine the effect of using a suboptimal f^*, showing it to be stronger than the effect of using a suboptimal Δf, as depicted in parts (b) and (d).

The same effect occurs for Δf in the base case with a deceased investor (Figure 1(b)). Because the optimal Δf is an interior value in the domain of possible Δf values, the derivative of c must be zero at the optimal Δf. Thus we see little sensitivity to moderate deviations from this value. This is not the case, however, in the base case when the investor is still alive (Figure 1(d)). Since the optimal Δf is zero, the endpoint of the domain of possible Δf values, the derivative of c need not be zero at the optimal Δf. This means the sensitivity of c to deviations from $\Delta f = 0$ will, in general, be greater, as is the case here.

There is considerable debate about whether taxable portfolios should be rebalanced continually (only selling stock with long-term capital gains, of course) or only when the portfolio deviates too much from its optimal stock fraction, as in the 5/25 rule of thumb discussed in the introduction. In this paper, we will present numerous examples in which it is optimal to rebalance continually (ie, the optimal $\Delta f = 0$) and almost equally numerous examples where it is optimal to rebalance only when the portfolio deviates too much (ie, the optimal $\Delta f > 0$). It might therefore be assumed that this paper suggests that either rule of thumb is equally valid. However, this is not the case. If the optimal Δf is not too large, the observation in the previous paragraph tells us that, as a rule of thumb, it is better to continually rebalance, because the loss incurred by wrongly choosing $\Delta f = 0$ when the optimal Δf is a small positive value will generally be less than the loss caused by wrongly choosing a small positive value for Δf when the optimal $\Delta f = 0$. That said, using the wrong rule of thumb is unlikely to have significant consequences. After all, as we can see in parts (b) and (d) of Figure 1, a wide range of suboptimal Δf may be used in the base case where the loss over forty years in the certainly equivalent difference remains under 1%.

3.2 The effect of varying parameters on the optimal strategy

In the remainder of Section 3, f^* and Δf will denote the f^* and Δf of the optimal no-rebalancing interval.

In the figures of this subsection, the parameters being varied are displayed on the horizontal axes, while stock fractions, f, are displayed on the vertical axes. Recall that stock positions are constrained to be long only, so $f \in [0, 1]$. Three curves are shown on the graphs, which represent the upper boundary, f^{u} (red line), the lower boundary, f^{l} (blue line), and the midpoint between f^{u} and f^{l}, f^* (green line). It is reasonable to think that f^* and the average value of f over time are almost equal, since, for example, they are equal to leading order as the interval width $\Delta f = f^{\mathrm{u}} - f^{\mathrm{l}}$ gets small in the related continuous time scenario considered in Goodman and Ostrov (2010). We will sometimes refer to f^* as the average stock fraction for this reason and, of course, because f^* is the average of the two values f^{u} and f^{l}.

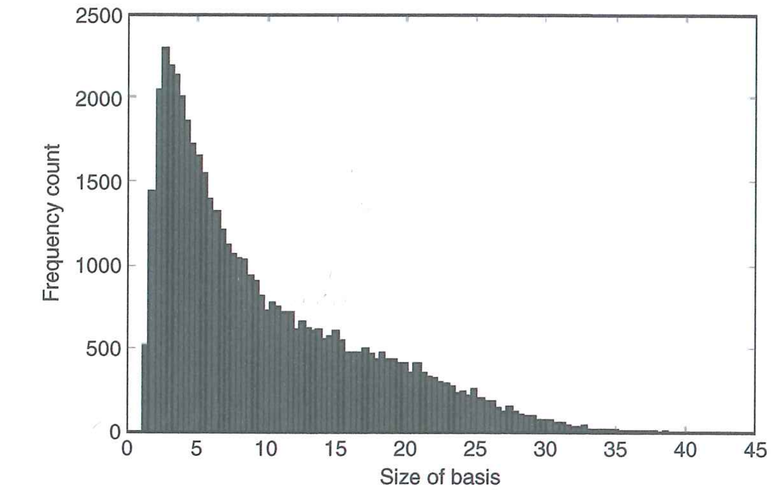
In many of the figures, we will see how changes to our parameter values will cause Δf to become bigger or smaller, often to the point of causing a transition between Δf being positive and being zero. These changes in the size of Δf are determined by shifting balances among a number of opposing factors.

Two factors may push Δf to be bigger.

(1) The bigger Δf, the more useful deferring capital gains becomes, even when the tax on gains must be paid at time T due to the investor being alive.

(2) If the investor is deceased at time T, then Δf will be even bigger to make it more likely that larger amounts of capital gains will be forgiven at time T.

FIGURE 2 Distribution of the number of levels in the tax basis on average in one life cycle.



A histogram of the average basis levels for each path, across 50 000 paths. We assume the standard US\$3000 annual limit for claiming losses and that all parameter values, except the varying parameter, take their base case values.

In opposition to these effects are the factors that push Δf to be smaller.

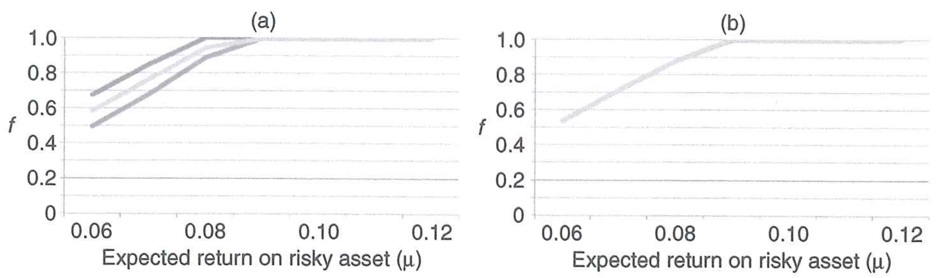
(1) The smaller Δf, the more useful claiming capital losses becomes, since $\tau_{\mathrm{l}} > \tau_{\mathrm{g}}$ according to current US tax law. That is, the smaller Δf, the more we can take advantage of the two tax options discussed in Constantinides (1983, 1984).

(2) The smaller Δf, the more control we have in keeping the portfolio at or near the stock fraction that optimizes the investor's expected utility.

In this subsection, we demonstrate and then explain how f^* and Δf are affected by altering our model parameters, one at a time, from their base case values. We present results for variations in each parameter using two graphs per figure: (a), on the left, will correspond to the case where the investor is assumed to expire at the portfolio horizon time T and (b), on the right, will correspond to the case where they are assumed to be alive.

First, we keep track of the number of levels in the tax basis of a simulated path. Earlier papers by DeMiguel and Uppal (2005) and Dai *et al* (2015) argue that there is rarely more than one level in the basis. We find otherwise, with an average of between nine and ten levels in the basis over its life cycle. The distribution of basis levels across the 50 000 simulated paths is shown in Figure 2. This is for the optimal strategy under base case parameters.

FIGURE 3 Variation in the optimal stock fraction range, caused by a varying expected stock growth rate, μ.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon. In this figure, and those below, it is optimal to rebalance just enough so that the portfolio stock fraction stays between the blue curve, f^{l}, and the red curve, f^{u}. The green curve, f^*, represents the center (ie, the midpoint) of this interval. We assume the standard US$3000 annual limit for claiming losses and that all parameter values, except the varying parameter, take their base case values.

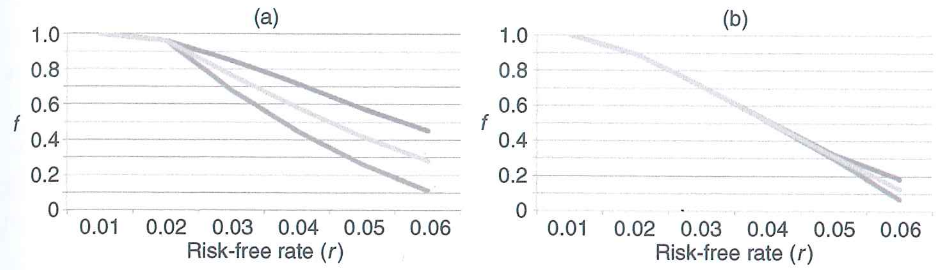
3.2.1 Varying the stock and cash growth rates, μ and r

The stock growth rate, μ. Critical to any long-run portfolio strategy are the assumed average growth rates of the portfolio's financial instruments. Financial planning tools make assumptions about these as a critical part of their processes. Here, we first examine the effect of changing the expected stock growth rate, μ, over the values $0.06, 0.07, 0.08, \dots, 0.12$ per annum, while holding all other parameters at their base case values (given above).

In Figure 3, we see, as expected, that the optimal average stock fraction, f^* (represented by the green line), increases as μ increases, until the point when the investor is best off placing the entire portfolio in stocks, ie, $f^* = 1$. Once this happens, of course, the optimal stock fraction interval, $[f^{\mathrm{l}}, f^{\mathrm{u}}]$ (represented by the blue and red lines), collapses to point $f^{\mathrm{l}} = f^{\mathrm{u}} = 1$, so that there is no cash. We also note, as expected from our discussion at the beginning of Section 3.1, that both f^* and Δf are smaller in Figure 3(b), where the investor is alive at the portfolio's liquidation time $T = 40$. Finally, we note that the values seen in Figure 3, cases (a) and (b), for the base case, $\mu = 0.07$, correspond to the numerical values given in Section 3.1.

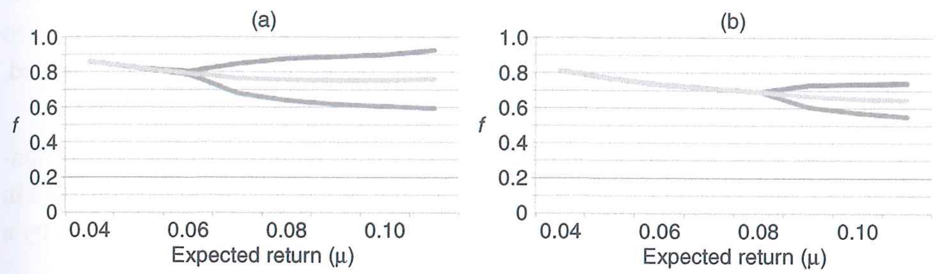
The risk-free rate, r. Essentially, increasing r, the cash interest rate (ie, the risk-free growth rate) has the opposite effect upon the optimal policy to increasing μ. We examine this effect in Figure 4, as we change the interest rate, r, over the values $0.01, 0.02, \dots, 0.06$ per annum. Again, note that the base case value in Figure 4, $r = 0.03$, corresponds, as it must, to the results shown in Figure 3 at the base case value, $\mu = 0.07$.

FIGURE 4 Optimal stock fraction range when varying the risk-free rate, r.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon.

FIGURE 5 Optimal stock fraction range when varying the stock growth rate, μ, while keeping the excess return, $\mu - r$,constant.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon.

The stock growth rate, μ, and the risk-free rate, r, while holding $\mu - r$ constant. As a related analysis, we also examine the effect of increasing the stock growth rate, μ, to the values $0.04, 0.05, 0.06, \dots, 0.16$ per annum, while equally increasing the risk-free rate, so as to keep the equity risk premium constant at $\mu - r = 0.04$.

Recall that in the absence of taxes, it is optimal to keep the stock fraction equal to $f_{\mathrm{Merton}} = (\mu - r)/\alpha\sigma^2$ at all times. If we change μ while keeping the equity risk premium, $\mu - r$, constant, f_{Merton} does not change. That is, so long as $\mu - r = 0.04$, the optimal Merton strategy stays unchanged from the base case, namely $f^* = \frac{2}{3}$ and $\Delta f = 0$.

But what happens to the optimal strategy in the presence of taxes? In Figure 5, we let μ vary while holding the equity risk premium constant at $\mu - r = 0.04$. We see that as μ increases, f^* decreases slightly and Δf increases.

Because $\mu - r$ remains constant, both μ and r increase equally in value, but the effect of tax law on the stock is quite different from its effect on the cash. As μ increases, there are fewer stock losses and more stock gains, which means fewer opportunities to take advantage of the tax break created by culling losses due to τ_l being greater than τ_g. Because culling losses is less advantageous, the stock becomes less useful relative to the cash, and so f^* decreases as μ (and r) increase.

As the probability of loss decreases, the ability of a small Δf to reap losses diminishes. At the same time, since there are more gains, the advantage of deferring paying tax on these gains increases. Both of these effects push Δf to increase as μ (and r) increase.

3.2.2 Varying risk and risk aversion, σ and α

The stock volatility, σ. Portfolio risk is a function of the volatility of the stock, σ. We examine the effect of changing the stock risk, σ, over the values 0.15, 0.20, 0.25 and 0.30. Not surprisingly, we see in Figure 6 that the optimal average stock fraction, f^*, decreases as risk increases, starting from an all-stock position when σ is very low. The effect of σ on Δf, however, is surprisingly small. As is always the case when the investor is alive at time T, we observe smaller values for f^* and for Δf. Indeed, Δf shrinks to zero when the investor is alive.

The risk-aversion parameter, α. We also looked at changing the risk-aversion parameter, α, over the values $0.7, 1.1, 1.5, \dots, 3.9$. The results of changing α, as seen in Figure 7, are similar to the results of increasing σ, and they can be explained by a similar line of reasoning.

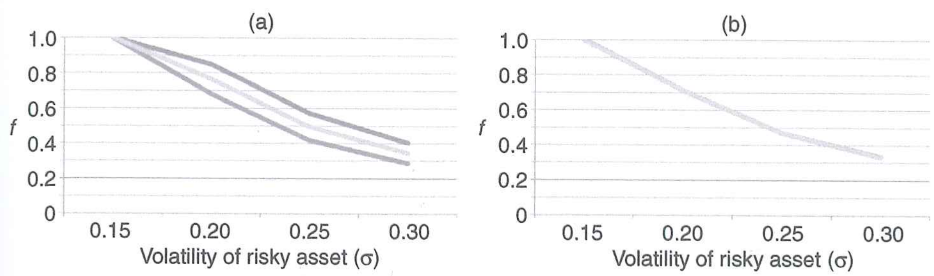
3.2.3 Varying tax rates, τ_l and τ_g

Since the US tax code has different rates for capital losses (τ_l) and capital gains (τ_g), we examine the impact of varying each rate separately.

The tax rate on losses, τ_l. We study the effect of changing the tax rate on capital losses, τ_l, over the values of the marginal tax rate in the current US tax code: 0.1, 0.15, 0.25, 0.28, 0.33, 0.35 and 0.396. From a tax point of view, losses may be seen as beneficial; the higher τ_l, the greater the tax shielding experienced by the investor. As a consequence, in Figure 8, we see that the optimal average stock fraction, f^*, increases with τ_l, since the additional tax shielding mitigates the risk of holding stocks – the downside to increasing τ_l – thereby making the stock more desirable.

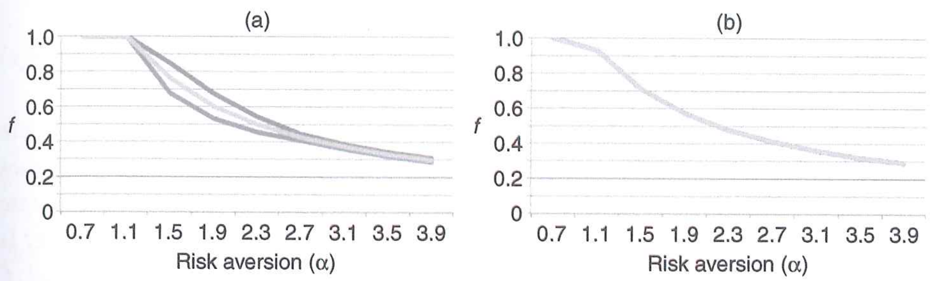
As expected, Δf decreases as τ_l increases, since the investor optimally rebalances more often as the benefits of taking losses increase.

FIGURE 6 Optimal stock fraction range when varying the stock volatility, σ.



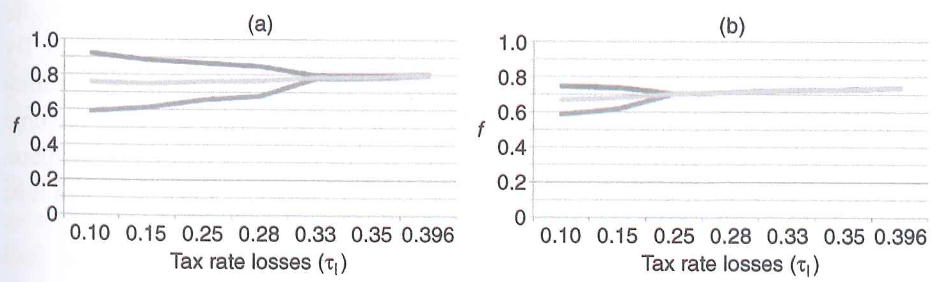
(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon.

FIGURE 7 Optimal stock fraction range when varying the risk aversion coefficient, α.



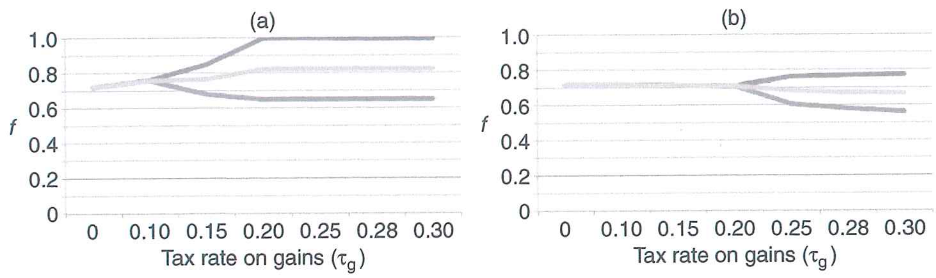
(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon.

FIGURE 8 Optimal stock fraction range when varying the tax rate on losses, τ_l.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon. Note that, since the differences between our experimental τ_l values are not uniform, neither are the points on the horizontal axes.

FIGURE 9 Optimal stock fraction range when varying the tax rate on gains, τ_g.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon. Note that, since the differences between our experimental τ_g values are not uniform, neither are the points on the horizontal axes.

The tax rate on gains, τ_g. We examined the effect of changing the capital gains tax rate, τ_g, over the values 0, 0.1, 0.15, 0.20, 0.25, 0.28 and 0.30. Given our analysis for τ_l, it seems intuitive to expect the stock to become less desirable as τ_g increases, so our desired stock fraction, f^*, should decrease. Yet a glance at the plot in Figure 9(a) shows that this intuition is wrong!

Why should f^* increase? To provide some simple intuition, consider the case where the stock fraction $f = 1$. Whether the stock goes up or down in worth, f continues to equal one. Thus there is never a need to realize capital gains, whatever the increase in tax rate, τ_g, until liquidation time T. Therefore, in the case where the investor is deceased and capital gains are forgiven at time T, this highest possible value of f is clearly more desirable than, say, $f = 0.9$ or $f = 0.8$, where capital gains will occur every time period h. Further, the desirability of this higher f increases as τ_g increases.

We can quantify and expand this intuition with a simple calculation over a year's time horizon. Let f be the fraction of the portfolio we want in stock, μ be the annual return on the stock and r be the annual return on the cash. Assume we have a portfolio worth US\$1 at $t = 0$, so we have f dollars of stock and $(1 - f)$ dollars in cash. By $t = 1$, we have $(1 + \mu)f$ dollars of stock and $(1 + r)(1 - f)$ dollars in cash. Adding these gives a total portfolio worth of $(1 + r) + (\mu - r)f$ dollars at $t = 1$. Assume we choose $\Delta f = 0$. We then need to rebalance so that the stock fraction is f again. This means we want to have $((1 + r) + (\mu - r)f)f$ dollars of stock; to attain this, we must sell

$$[(1 + \mu)f] - [((1 + r) + (\mu - r)f)f] = (\mu - r)f(1 - f)$$

dollars of stock as capital gains. This capital gains function, $(\mu - r)f(1 - f)$, is parabolic in f. It equals zero at $f = 0$, increases to its maximum value at $f = \frac{1}{2}$

and then decreases back to zero at $f = 1$. Thus, we may establish that, in a region where $f > \frac{1}{2}$, increasing f actually lowers capital gains.

This calculation corresponds to the case in which the investor expires at the portfolio liquidation time, T, and so capital gains taxes are not paid at liquidation. When the investor is alive and capital gains taxes must be paid at liquidation, there are two opposing effects to be balanced: the higher deferral of capital gains that using a higher f provides (as explained in the previous paragraph) and the higher capital gains taxes that must be paid at liquidation when a higher f is used. We see in Figure 9(b) that f^* slightly decreases as τ_g increases, so, in this case, the latter effect clearly outweighs the former.

We also see in the plot on the right that Δf increases as τ_g increases. This is to be expected, since the higher τ_g is, the less advantageous it is to realize gains in order to reset the cost basis and, thereby, increase the likelihood of reaping the losses, which initially have a higher rate than the gains. Once τ_g surpasses $\tau_l = 0.28$, the situation is flipped, and capital gains become more destructive than capital losses are advantageous. This makes rebalancing progressively less desirable as τ_g increases, causing Δf to increase further.

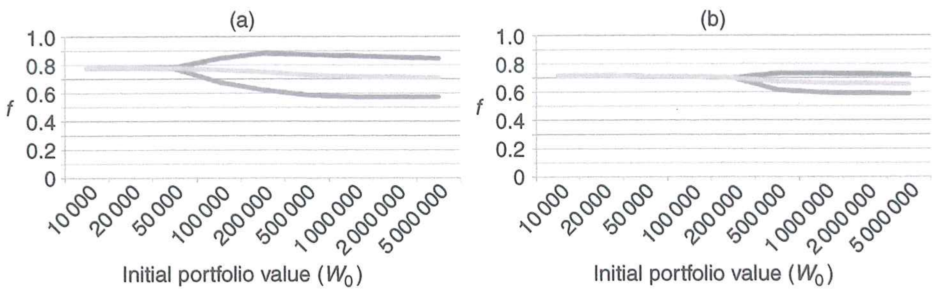
3.2.4 Varying the initial portfolio worth, W_0

Our power law utility function, outlined in (2.1), has the property of constant relative risk aversion. In the absence of taxes, this implies that the optimal stock fraction is independent of the portfolio size, as reflected by the absence of W_0 in the Merton expression given in (3.1).

If tax policy were strictly dictated by proportional factors such as τ_g and τ_l, then we would also expect the optimal policy with taxes to be independent of W_0 (see, for example, Dammon *et al* 2001). However, the US\$3000 limit on annual claimed losses is a constant factor rather than a proportional one. Therefore, the optimal strategy will indeed be affected by W_0.

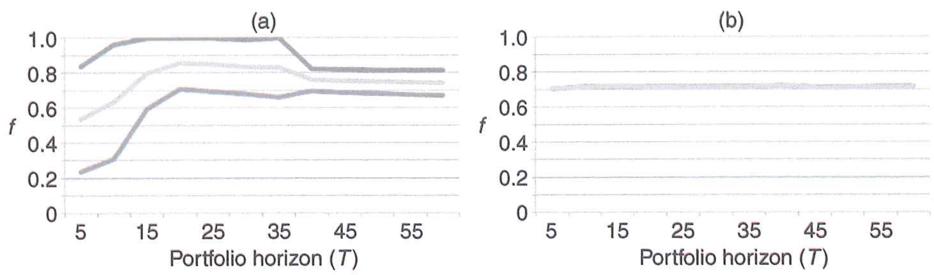
We see this effect in Figure 10. We study the effect of changing the initial size of the portfolio, W_0, over the values (in US\$): 10 000, 20 000, 50 000, 100 000, 200 000, 500 000, 1 000 000, 2 000 000 and 5 000 000. As W_0 increases, there is a mild decline in f^* due to the fact that more and more losses must be carried over to subsequent years, making the stock less valuable. With a small portfolio, keeping Δf smaller corresponds, in general, to more collectable losses: a desirable outcome, since the US\$3000 limit rarely interferes. As the portfolio becomes larger, however, the US\$3000 limit on losses is more easily reached, and the advantage of keeping Δf small is diminished. When this happens, the tax deferral provided by a larger Δf becomes a more dominant factor, and therefore Δf grows as W_0 increases, as shown in Figure 10.

FIGURE 10 Optimal stock fraction range when varying the initial wealth, W_0.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon. Note that, since the differences between our experimental τ_g values are not uniform, neither are the points on the horizontal axes. In fact, the values of W_0 have been chosen so that the horizontal axes are close to logarithmic.

FIGURE 11 Optimal stock fraction range when varying the portfolio horizon, T (in years).



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon.

3.2.5 Varying the time horizon before portfolio liquidation, T

Next, we examine the effect of the portfolio horizon, T, on the optimal strategy by changing T (measured in years) over the values 5, 10, 15, ..., 60.

First, we consider Figure 11(a) for the case where the investor expires at time T. As T initially begins to increase, we see f^{l}, f^{u} – and therefore f^* – increasing too. This is no surprise since the growth of T means the advantages of deferring capital gains taxes and of having capital gains forgiven at liquidation also increase, which makes the stock's worth relative to the cash increase. By $T = 15$, $f^{u} = 1$ and the advantages of having capital gains forgiven at liquidation now completely outweigh both the incentive to position the portfolio at its optimal stock fraction and the incentive to sell stock so that more losses can be generated for tax credits.

However, by $T = 40$, this no longer holds, and we see f^{u} decrease. There are two effects behind this.

(1) When we use an interval of the form $[f^{l}, 1]$, we do not sell stock. In this case, because $\mu > r$, the longer the time horizon, the more likely the stock fraction f is to drift into the high end of the $[f^{l}, 1]$ interval. The closer the stock fraction drifts toward one, the farther it drifts from its optimal fraction, and so the portfolio becomes exposed to too much risk unless we reduce f^{u}.

(2) If $f^{u} = 1$, we do not sell stock over long time horizons. Thus, in cases where the stock does well, the portfolio does particularly well, due, in part, to the significant forgiveness of particularly big gains at time T when the investor expires. However, in cases where the stock does not do well overall, the portfolio will do even worse, since stock is not sold when it reaches an $f^{u} < 1$. Such early sales could potentially generate losses over a long time horizon, cushioning the damage caused by bad returns. This greater "wealth disparity" over longer time horizons is penalized by the concavity of the utility function, eventually forcing f^{u} to decrease as T increases.

In the case where the investor does not expire at liquidation time T, we see in Figure 11(b) that f^* is a constant just above 0.7 and $\Delta f = 0$. The fact that Δf is reduced when the investor does not expire at time T is, of course, expected. Because it is reduced to $\Delta f = 0$, the advantages of deferring taxes are not capitalized upon and so f^* becomes independent of T, as in the Merton case when $\Delta f = 0$. In this instance, $f_{\text{Merton}} = \frac{2}{3}$ regardless of the value of T.

3.2.6 Varying the period between potential trading, h

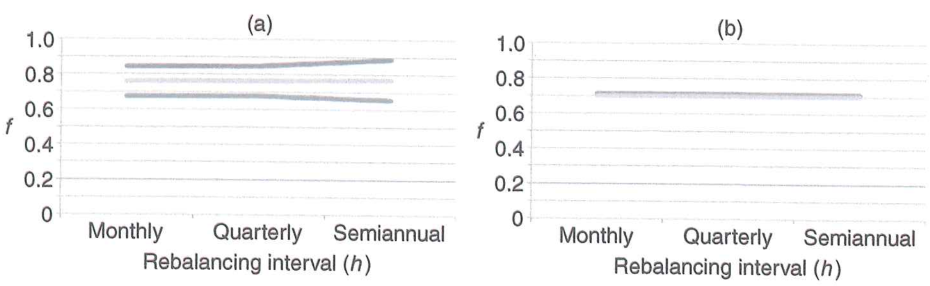
We study the effect of changing the time period, h, between potential trades within the portfolio (ie, culling losses and/or rebalancing) over the values $1/12 = 0.8333$ (monthly), $1/4 = 0.25$ (quarterly) and $1/2 = 0.5$ (semiannually).

In Figure 12(b), we see that increasing h appears to have no real effect on f^* and only slightly increases Δf in the case of Figure 12(a), where the investor expires at time $T = 40$.

Why is there a slight increase in Δf in this case? One of the advantages of a smaller Δf is that it increases the likelihood of capital gains and capital losses, which is desirable since $\tau_l > \tau_g$. When we increase h, however, this advantage is reduced, since it becomes more likely that the losses will be canceled by gains before they can be realized. With the advantage of a smaller Δf reduced, the factors that push Δf to expand become more dominant, and so Δf increases a little as h increases.

From a practical standpoint, however, it is more important to note that this increase in Δf is small. That is, over the reasonable range of h values considered here, the

FIGURE 12 Optimal stock fraction range when varying the potential trading interval, h.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon. Note that, since the differences between our experimental τ_g values are not uniform, neither are the points on the horizontal axes.

frequency of potential trading is not a particularly important factor in the choice of optimal strategy.

3.3 The effect of changing the model on the optimal strategy

In this subsection, we consider the effect of changing the model in four ways:

(1) Letting the optimal stock fraction range, $[f^{\mathrm{l}}, f^{\mathrm{u}}]$, change values when the portfolio is halfway to liquidation (ie, at $T/2 = 20$ years), instead of remaining constant.

(2) Using the average cost basis instead of the full cost basis to provide a quantitative measure of the suboptimality generated by using the average cost basis.

(3) Letting f^{l} and f^{u} each depend on the ratio of the highest cost basis in the portfolio to the current stock price to quantify any advantage this may generate.

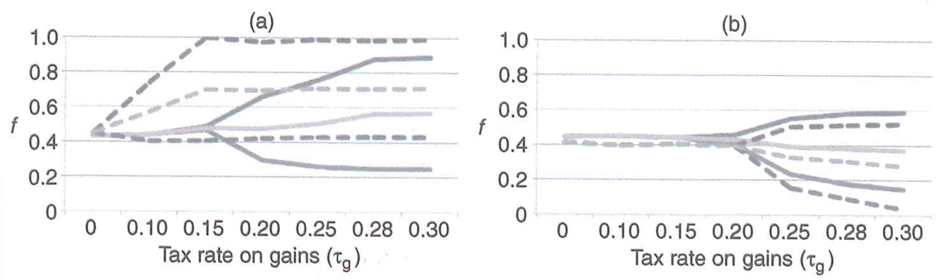
(4) Incorporating transaction costs when we buy and sell stock to understand their effect on the optimal stock fraction range, $[f^{\mathrm{l}}, f^{\mathrm{u}}]$, in the presence of taxes.

3.3.1 The effect of a time-dependent no-rebalancing region

We study the effect of allowing f^{l} and f^{u} to change values when we transition from the initial twenty years, during which there is a long time until liquidation, to the final twenty years, during which there is a short time until liquidation.

We must now optimize over five variables instead of three: f^{init} (the initial stock fraction), $f^{\mathrm{l},0\text{–}20}$ and $f^{\mathrm{u},0\text{–}20}$ (the values of f^{l} and f^{u} in the initial twenty years), and $f^{\mathrm{l},20\text{–}40}$ and $f^{\mathrm{u},20\text{–}40}$ (the values of f^{l} and f^{u} in the final twenty years). As a

FIGURE 13 Optimal stock fraction range for the first and second halves of the portfolio horizon when varying the tax rate on gains, τ_g.



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon. The entire horizon of the portfolio is forty years. For the first twenty years, stock fractions are shown in bold lines; for the last twenty years they are shown in dashed lines. As before, the upper bound on the stock fraction is shown by a red line, the lower bound is shown by a blue line and the average of these two bounds, f^*, is given by a green line.

typical example of our results, in Figure 13 we show the values of $f^{\mathrm{l},0\text{–}20}$, $f^{\mathrm{u},0\text{–}20}$, $f^{\mathrm{l},20\text{–}40}$ and $f^{\mathrm{u},20\text{–}40}$ in the context of changing τ_{g}.

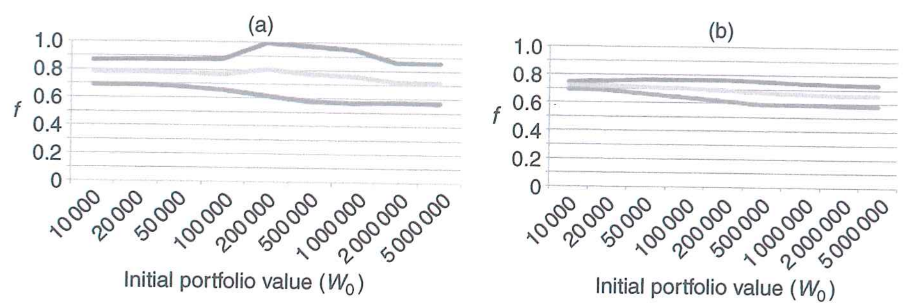
Unsurprisingly, when the investor expires at liquidation time T, the ability to change stock proportions makes a significant difference, as shown in Figure 13(a). Specifically, since capital gains are forgiven at time T, the stock becomes more valuable in the final twenty years, so $f^{*,20\text{–}40} > f^{*,0\text{–}20}$. To create more forgivable capital gains, we also see that $\Delta f^{20\text{–}40} > \Delta f^{0\text{–}20}$.

Equally unsurprisingly, we find that when the investor is alive at time T, the ability to change stock proportions makes little difference, so $f^{\mathrm{l},0\text{–}20} \approx f^{\mathrm{l},20\text{–}40}$ and $f^{\mathrm{u},0\text{–}20} \approx f^{\mathrm{u},20\text{–}40}$, as shown in Figure 13(b). The only difference over time is at the portfolio horizon, when liquidation is required and the associated capital gains tax hits. This required liquidation diminishes the advantage of having stock, because it restricts our ability to defer capital gains taxes. Thus, $f^{*,20\text{–}40}$, the optimal stock proportion held when we are closer to liquidation, dips a little from $f^{*,0\text{–}20}$, the optimal proportion when we are further from liquidation.

3.3.2 The effect of the average cost basis versus the full cost basis

As discussed in the introduction, this paper's simulation approach allows us to use the entire cost basis, resulting in a comprehensive solution to the problem of portfolio optimization with taxes. Preceding works have attempted to account for the full tax basis but have been limited to very few allowable trading times. Therefore, they are unsuitable for modeling the entire long-horizon life cycle of an investor. Instead, for computational convenience in multiperiod optimization, most past works have used

FIGURE 14 Optimal stock fraction range when the algorithm uses the average cost basis rather than the full cost basis, when varying the initial wealth, W_0.



(a) Investor deceased, average cost basis. (b) Investor alive, average cost basis. Note that, since the differences between our experimental τ_g values are not uniform, neither are the points on the horizontal axes. The values of W_0 have been chosen so that the horizontal axis is close to logarithmic.

the average tax basis instead of the full tax basis. Using the average tax basis keeps the state space small, allowing a numerical solution of the dynamic programming problem. Of course, using the average cost basis, which is allowed in US tax law for mutual funds, is guaranteed to be suboptimal compared with using the full cost basis.

But how suboptimal is using the average cost basis?

Our large-scale algorithm enables us to directly compare the effect of using the full cost basis with the average cost basis for the long-horizon problem. To make this comparison, we reran each of the scenarios used to generate Figures 3 to 10 (with the exception of Figure 5), using the average cost basis instead of the full cost basis. This allows us to compare, for example, the graphs in Figure 10, using the full cost basis, with the graphs in Figure 14, using the average cost basis. The observed changes between these figures (where W_0 is varied) are quite similar to those observed when other parameters are varied.

Across all scenarios, we found that using the average cost basis had very little effect on f^* but tended to increase Δf. The average increase in Δf was 6.0 percentage points for cases of both living and deceased investors at time T; the maximum increase in Δf was 16.8 percentage points when the investor was alive at time T, and 18.6 percentage points when the investor expired at time T. The increase in Δf can be explained by the fact that using the average cost basis makes losses in the portfolio less likely to occur, thereby reducing the advantage of keeping Δf small.

Our main concern in switching to the average cost basis, however, is not the change in the optimal policy. It is the loss to the investor created by this switch. To quantify this loss, we again use the certainty equivalent difference, which for our current

circumstance is

$$c = \left(\frac{E_{\text{avg}}[U(W)]}{E_{\text{full}}[U(W)]}\right)^{1/(1-\alpha)} - 1. \tag{3.3}$$

Applying (3.3) across each of our scenarios, we found that when the investor expires at time T, the average certainty equivalent difference, c, was only -0.27%, with a maximum certainty equivalent difference of -1.12%. When the investor is alive at time T, the average difference increases to -0.65%, with a maximum difference of -1.73%. For comparison, if we quantify the loss resulting from living at time T (with no gains forgiven) versus expiring at time T (with all gains forgiven) over the same set of scenarios (using the full cost basis for both), the cost equivalent difference

$$\hat{c} = \left(\frac{E_{\text{alive}}[U(W)]}{E_{\text{dead}}[U(W)]}\right)^{1/(1-\alpha)} - 1$$

averages -9.97% with a maximum difference of -23.6%.

The data in the above paragraph alone gives considerable justification for past works that have employed the average cost basis to determine trading strategies in taxable portfolios, since the loss generated by using the average cost basis in place of the full cost basis is clearly not that great. Yet the case for justifying these average cost basis models is even stronger: these models will generate optimal values for f^* and Δf for the average cost basis. Yet, in practice, these values would always be used with the full cost basis strategy. So, how much of the value of c in (3.3) is due to the f^* and Δf from the average cost basis model being suboptimal when an investor uses the full cost basis (as they would in practice)? How much is due to the effect of an investor actually using the average cost basis method rather than the full cost basis?

Consider the base case again. If the investor expires at $T = 40$, then, for the full cost basis optimization, we have that $f^* = 0.764$ and $\Delta f = 0.168$, while for the average cost basis optimization we have that $f^* = 0.770$ and $\Delta f = 0.228$, which, using (3.3), corresponds to $c = -0.0019$. But recall from Section 3.1 that small changes to f^* and large changes to Δf, as we have here, usually have little impact on the certainty equivalent difference. In fact, if we compute

$$\tilde{c} = \left(\frac{E^{\text{a}}_{\text{full}}[U(W)]}{E_{\text{full}}[U(W)]}\right)^{1/(1-\alpha)} - 1,$$

where $E^{\text{a}}_{\text{full}}$ means using the full cost basis with the values $f^* = 0.770$ and $\Delta f = 0.228$ that came from the average cost basis optimization, we get $\tilde{c} = -0.0003$. That is, the actual loss incurred by using the optimal strategy for the average cost basis model is much smaller than indicated above; specifically, it is only $\tilde{c}/c = 3/19$ of the loss indicated above, as long as the full cost basis is employed for actual trading, as it would be by any investor interested in minimizing taxes.

The base case where the investor is alive at $T = 40$ shows less dramatic results. In this case, for the full cost basis optimization, we have that $f^* = 0.711$ and $\Delta f = 0$, while for the average cost basis optimization we have that $f^* = 0.701$ and $\Delta f = 0.127$, which corresponds with $c = -0.0090$ and $\tilde{c} = -0.0060$. Therefore, $\tilde{c}/c = 2/3$, where before it was only $3/19$. The higher value for $\tilde{c}/c$ is a result of the optimal Δf being zero for the full cost basis optimization, which – as seen in Figure 1(b) and (d), and explained in Section 3.1 – leads to a higher sensitivity to changes in Δf.

3.3.3 The effect of allowing the no-rebalancing region to vary with the cost basis to current stock price ratio

Papers such as Dai *et al* (2015) allow f^{l} and f^{u} to depend on b_{avg}, the ratio of the average cost basis to the current stock price, B_{avg}/S. Dai *et al* (2015) give examples in which optimizing the no-rebalancing interval (and specifically its dependence on b_{avg}) leads to strategies that produce certainty equivalent advantages of between 0.84% and 5.20% over the strategy of deferring all short-term gains and realizing all losses and long-term gains. They do not consider the question explored here: the certainty equivalent advantage created by allowing the no-rebalancing region to depend on the aforementioned ratio of the cost basis to the stock price (as opposed to not allowing such dependence).

Since our algorithm allows for the full cost basis, it makes little sense for us to employ b_{avg}. Instead, we define $b = B_J/S$, where, as before, B_J corresponds to the highest cost basis of the stock in the portfolio, which is the first stock that should be traded. We then define

$$f^{\mathrm{l}} = \max[0, c_1 - c_2 * \max[0, 1 - b]],$$
$$f^{\mathrm{u}} = \min[1, c_3 + c_4 * \max[0, 1 - b]],$$

where we restrict all four c_i values to be positive as well as $0 \leqslant c_1 \leqslant c_3 \leqslant 1$. This model keeps $0 \leqslant f^{\mathrm{l}} \leqslant f^{\mathrm{u}} \leqslant 1$. It also allows Δf to grow as b grows, creating the potentially desirable effect of discouraging the sale of stock as the capital gains-related implications of selling stock increase. Unlike that of Dai *et al* (2015), our model restricts f^{l} and f^{u} to being linear functions of b. Thus, the optimal dependence of f^{l} and f^{u} on b is only approximated here. In our new model, we optimize over five variables – f_{init}, c_1, c_2, c_3 and c_4 – in place of optimizing our normal three variables, f_{init}, f^{l} and f^{u}. Since the subcase $c_2 = c_4 = 0$ in our five-variable optimization yields our normal three-variable optimization, we are guaranteed that the five-variable optimization will be superior to our normal three-variable optimization.

We considered the base cases for both deceased and living investors at the portfolio horizon T. We found that using the five-variable optimization versus the three-variable

optimization made no discernible difference. Specifically, in both cases, the certainty equivalent difference between allowing (linear) dependence of f^{l} and f^{u} on b versus not allowing such dependence on b was less than a basis point over the course of $T = 40$ years, which is the limit of our model's ability to discern certainty equivalent differences. This evidence suggests that it is not necessary to use models where the boundaries of the no-rebalancing region, f^{l} and f^{u}, depend on b, the ratio of the (highest) cost basis to the current stock price.

3.3.4 The effect of incorporating transaction costs

Finally, we investigate the effect of incorporating transaction costs into our model. This issue was also explored in Leland (2000), who found that transaction costs reduce optimal portfolio churn by 50%. He also found that capital gains taxes lead to lower investment in stock. Yet we find that this is not always the case.

Assume that, for the current time t, the current number of shares of stock in the portfolio is N_t, the current stock price is S_t, the current worth of the portfolio's cash position is C_t and the number of shares to be transacted at time t is n. Let e denote the proportion of the worth of a trade lost to transaction costs. Based on Domowitz *et al* (2001) and Pollin and Heintz (2011), we assume transaction costs range from 0 to 50bps of the value of the transaction, ie, we consider the e values 0 (our base case), 0.0005, 0.0010, ..., 0.0050. These proportional transaction costs can be incurred at the following four points in the simulation.

(1) When any stock has a capital loss, we sell it and buy back the same value of an equivalent stock. Thus, our cash balance must be reduced by $2(nS_te)$, a value that corresponds to the transaction costs incurred in both selling and repurchasing the stock. We note that, while it is always optimal to sell and buy back stock with a capital loss when there are no transaction costs, this is no longer guaranteed to be the optimal strategy when transaction costs apply. However, for the small transaction costs considered here, the investor will, generally, still be better off selling and buying back stock with a capital loss. So we continue to implement this strategy in our transaction cost model.

(2) When the stock fraction falls below f^{l}, we buy stock so that the new stock fraction equals f^{l} after transaction costs. This must satisfy the equation

$$f^{\mathrm{l}} = \frac{(N_t + n)S_t}{(N_t + n)S_t + C_t - nS_t - nS_te},$$

which, after rearrangement, leads to the following expression for the number of shares that will be bought:

$$n = \frac{N_tS_t(f^{\mathrm{l}} - 1) + C_tf^{\mathrm{l}}}{S_t(f^{\mathrm{l}}e + 1)}.$$

Using the subscript (t^+) to denote values just after rebalancing at time t, we then update the total number of shares of stock

$$N_{(t^+)} = N_t + n$$

and the cash balance

$$C_{(t^+)} = C_t - nS_t - nS_t e.$$

(3) When the stock fraction rises above f^{u}, we sell stock so that the new stock fraction equals f^{u} after transaction costs. This must satisfy the equation

$$f^{\mathrm{u}} = \frac{(N_t - n)S_t}{(N_t - n)S_t + C_t + nS_t - nS_t e},$$

which, after rearrangement, leads to the following expression for the number of shares that will be sold:

$$n = \frac{N_t S_t (f^{\mathrm{u}} - 1) + C_t f^{\mathrm{u}}}{S_t (f^{\mathrm{u}} e - 1)}.$$

We then update the total number of shares of stock

$$N_{(t^+)} = N_t - n$$

and the cash balance

$$C_{(t^+)} = C_t + nS_t - nS_t e.$$

(4) Finally, at the end of the year, if there are capital losses, then the tax break generated by these losses is used to buy additional shares. This additional number of shares (after transaction costs) will be

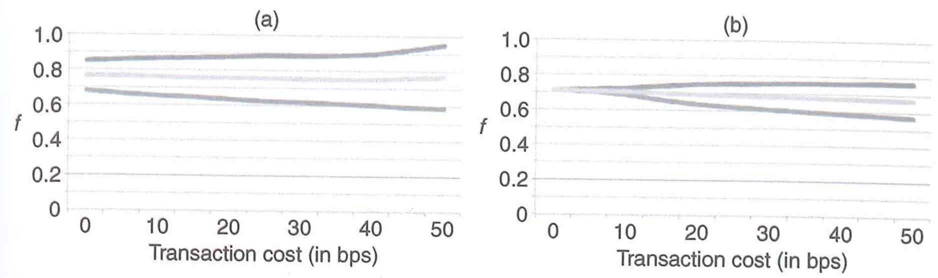
$$n = \frac{\tau_1 \min[3000, \max(0, -G)]}{S_t(1+e)},$$

since US\$3000 is the annual limit allowed for taking tax losses and G, as before, represents the realized gains in the portfolio. Therefore, shares are only bought when $G < 0$, ie, when there are losses.

As expected, we see from Figure 15 that rebalancing occurs less often as transaction costs increase. That is, Δf increases as e increases. What is more surprising is that the optimal average stock fraction, f^*, is essentially unaffected by e, at least over the range of values for e considered here.

Considerable attention has been devoted to understanding the asymptotic effect of small proportional transaction costs in portfolios that are not subject to taxes: see, for

FIGURE 15 Optimal stock fraction range when varying transactions costs (in bps).



(a) Investor deceased at portfolio horizon. (b) Investor alive at portfolio horizon. These costs are stated in terms of the percentage of the value of each transaction lost to costs.

example, Atkinson and Mokkhavesa (2002), Goodman and Ostrov (2010), Janaček and Shreve (2004), Rogers (2004), Shreve and Soner (1994) and Whaley and Wilmott (1997), all of whom conclude that the order of growth of Δf is given by $\Delta f \sim O(e^{1/3})$ when e is small. What happens to this asymptotic expression in portfolios subject to taxes? The expression has no relevance to the case in Figure 15(a), where the investor expires at time T, since, in this case, $\Delta f \neq 0$ when $e = 0$. Figure 15(b), in which the investor is alive at time T, has more potential to be connected to the asymptotic expression, since $\Delta f = 0$ when $e = 0$, and Δf grows as e grows. However, even in this case, it is clear from Figure 15(b) that Δf does not grow by order $e^{1/3}$, so the asymptotic formula for cases without taxes is not relevant to cases with taxes.

4 CONCLUSIONS

In this paper, we have modeled the optimal trading strategy for a taxable portfolio with a stock position and a cash position. In this model, we consider the effects of the US tax system, including keeping track of the cost basis of stocks whenever they are bought and sold, utilizing the tax benefits from stocks with losses as well as the annual US\$3000 limit on claimable losses and the practice of carrying over losses above this limit. We develop a large-scale simulated optimization program that offers facile computation on standard hardware, extending the number of trading periods into the hundreds ($\sim$500), whereas earlier models managed many fewer ($\sim$20 periods).

For our model, we have determined the optimal static interval $[f^{\mathrm{l}}, f^{\mathrm{u}}]$ within which to maintain f, the fraction of the portfolio invested in stock. When the portfolio is inside this interval, it is optimal to trade only to reap capital losses, and not to rebalance. When the portfolio strays outside this interval, it is optimal to rebalance the portfolio back to the nearest endpoint, f^{l} or f^{u}. We note that a generalization of the static interval is feasible by making f^{l} and f^{u} functions of a state variable such as the tax basis.

Our experiments to determine this optimal interval provide a number of insights concerning the best strategies for investors to pursue in their taxable portfolios. A number of our conclusions differ from conventional wisdom and the intuition of many investors.

- The optimal value of f, the fraction of the portfolio in stock, is often higher for taxable accounts than for tax-free accounts, such as the Roth IRA (as also noted in Dammon *et al* 2004).

  This is the case even if the cash position is not taxed and if we assume that the investor is alive at liquidation, meaning that taxes on all capital gains must be paid. This is explained by the fact that τ_{l}, the refund tax rate for capital losses, is equal to the marginal income tax rate and higher than τ_{g}, the (long-term) tax rate for capital gains. Therefore, the benefit of culling capital losses from stock can, on average, outweigh the disadvantage of paying capital gains taxes, making the stock more useful in the taxable portfolio than in the tax-free portfolio. For example, in Section 3.1, we considered a base case where the investor is alive at the time of liquidation. We found that the optimal value of $f = f^{\mathrm{l}} = f^{\mathrm{u}}$ was about four percentage points higher than the value of f_{Merton}, the optimal constant stock fraction in a tax-free portfolio.

- If the capital gains tax rate increases, then f, the fraction of the portfolio in stock, should be raised, not lowered.

  In the previous section, we substantiate this both experimentally and with an intuitive explanation. This conclusion assumes that the portfolio is designed to be given to a beneficiary, so that gains are forgiven upon the investor's death. It also assumes that the stock position is larger than the cash position.

- The 5/25 rule for rebalancing taxable portfolios is less than ideal, even as a rule of thumb.

  The "5" part of the 5/25 rule corresponds to an interval width, $\Delta f = f^{\mathrm{u}} - f^{\mathrm{l}} = 2 \times 0.05 = 0.10$. That is, it recommends using an interval for the stock fraction with a width of ten percentage points. (The "25" part of the rule is irrelevant for our stock–cash model if f^*, the center of the optimal interval, is between 0.20 and 0.80.)

  However, we find that there are common circumstances in which the optimal $\Delta f = 0$, and almost equally common circumstances in which the optimal $\Delta f > 0$. Moreover, because the effect of using a suboptimal Δf will generally be more harmful when the optimal $\Delta f = 0$ than when the optimal $\Delta f > 0$ (for reasons explained in Section 3.1), our analysis suggests that, for a rule of thumb, it makes more sense to adopt $\Delta f = 0$. In other words, it is preferable to adopt a strategy of continual rebalancing, provided transaction costs are small.

- The optimal interval – in which the stock fraction, f, should not be rebalanced – is not improved by allowing it to get bigger as the stock price increases further from its purchase price.

  It is intuitive to think that, since the capital gains taxes incurred by selling stock increase as the difference between the stock price and the cost basis grows, it should be optimal to let Δf increase as this difference grows, so as to avoid these progressively costly sales. However, as we saw in Section 3, allowing f^{l} and f^{u} to linearly depend on b, the ratio of the highest basis price to the current stock price, leads to no discernible certainty equivalent advantage, suggesting that the attention paid to the dependence of the optimal f^{l} and f^{u} on b in previous papers (see Dai *et al* 2015) is likely unnecessary.

Overall, our model shows that the optimal width of the stock fraction interval, $\Delta f = f^{\mathrm{u}} - f^{\mathrm{l}}$, reacts in the following ways.

- It significantly increases when the stock's expected return, μ, increases or the risk-free interest rate, r, decreases; when the tax rate for capital gains, τ_{g}, increases or the tax rate for capital losses, τ_{l}, decreases; when the initial portfolio size, W_0, increases; or when the lifetime of the portfolio, T, decreases.

- It slightly increases when the time interval between potential transactions, h, increases or when the utility function risk-aversion parameter, α, in (2.1) decreases.

- It is essentially unaffected when the stock's volatility, σ, changes.

Further, our model demonstrates that the optimal portfolio stock fraction, or, more specifically, the optimal center, $f^* = (f^{\mathrm{l}} + f^{\mathrm{u}})/2$, of the stock fraction interval reacts as follows.

- It significantly increases when μ increases or when r, σ or α decrease.

- It slightly increases when τ_{l} increases; when W_0 decreases; when τ_{g} increases, assuming the investor expires at time T; or when τ_{g} decreases, assuming the investor is alive at time T.

- It is essentially unaffected when h changes, or, in the case where the investor is alive at time T, when T changes.

- It increases when T initially increases in the case where the investor expires at time T. However, as T increases further, it levels off, then slowly decreases and finally levels off again.

Finally, our model provides a number of other insights about optimal investing in taxable portfolios.

- Using a suboptimal value for f^* is generally far more detrimental to the investor than using a suboptimal value for Δf.

- The optimal strategy is almost completely unaffected by the timescale of our considered trading actions: whether monthly, quarterly or semiannual.

- When the optimal stock fraction interval, $[f^{\mathrm{l}}, f^{\mathrm{u}}]$, is allowed to depend on time, we see far more dynamic behavior for this interval when the investor is deceased at time T than when the investor is alive. Unsurprisingly, when the investor expires at time T, both f^* and Δf increase over time. When the investor is alive at time T, f^* decreases slightly over time, while Δf stays more or less constant.

- While using the full cost basis history as opposed to the average cost basis has a significant effect on the optimal stock fraction interval, $[f^{\mathrm{l}}, f^{\mathrm{u}}]$, it has surprisingly little effect on the outcome for the investor, generally leading to a certainty equivalent difference of less than 1% over forty years. This justifies the use of the average cost basis approximation, employed out of necessity by many previous papers in the form of Bellman equation approaches as a means of investigating optimal taxable portfolio strategies.

- As the magnitude of proportional transaction costs increases, the width of the optimal stock fraction interval, Δf, also increases. However, the center, f^*, for this interval remains surprisingly constant.

There are a number of ways in which this research may be extended to generalize the setting outlined in this paper, by capitalizing on our ability to solve an optimization problem over simulated portfolios with one risk-free and one risky asset, while tracking the full tax basis. First, we may extend this simulation approach to the case of multiple risky assets, in which it may be able to clarify the geometry of the optimal no-rebalancing region. Second, the model may be extended to account for the difference in tax treatment between short- and long-term capital gains. Third, the optimal evolution of the shape of the rebalancing region over time can be further explored. We leave these interesting avenues open for further research.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

We thank Ananth Madhavan and other participants at the 2015 *Journal of Investment Management* conference, San Diego, CA, for their comments.

REFERENCES

Akian, M., Menaldi, J. L., and Sulem, A. (1996). On an investment-consumption model with transactions costs. *SIAM Journal of Control and Optimization* **34**(1), 329–364.

Atkinson, C., and Ingpochai, P. (2006). The influence of correlation on multi-asset portfolio optimization with transaction costs. *The Journal of Computational Finance,* **10**(2), 53–96.

Atkinson, C., and Mokkhavesa, S. (2002). Perturbation solution of optimal portfolio theory with transaction costs for any utility function. *IMA Journal of Management Mathematics* **13**(2), 131–151.

Bichuch, M. (2012). Asymptotic analysis for optimal investment in finite time with transaction costs. *SIAM Journal on Financial Mathematics* **3**(1), 433–458.

Constantinides, G. M. (1983). Capital market equilibrium with personal taxes. *Econometrica* **51**(3), 611–636.

Constantinides, G. M. (1984). Optimal stock trading with personal taxes. *Journal of Financial Economics* **13**, 65–89.

Dai, M., Liu, H., and Zhong, Y. (2011). Optimal consumption and investment with differential long-term/short-term capital gain tax rates. Working Paper, Washington University, St. Louis, MO.

Dai, M., Liu, H., Yang, C., and Zhong, Y. (2015). Optimal tax-timing with asymmetric long-term/short-term capital gains tax. *Review of Financial Studies* **28**(9), 2687–2721.

Dammon, R. M., and Spatt, C. S. (1996). The optimal trading and pricing of securities with asymmetric capital gains taxes and transactions costs. *Review of Financial Studies* **9**(3), 921–952.

Dammon, R. M., and Spatt, C. S. (2012). Taxes and investment choice. *Annual Review of Finance and Economics* **4**, 411–429.

Dammon, R. M., Dunn, K. B., and Spatt, C. S. (1989). A reexamination of the value of tax options. *Review of Financial Studies* **2**(3), 341–372.

Dammon, R. M., Spatt, C. S., and Zhang, H. (2001). Optimal consumption and investment with capital gains taxes. *Review of Financial Studies* **14**(3), 583–616.

Dammon, R. M., Spatt, C. S., and Zhang, H. (2004). Optimal asset location and allocation with taxable and tax-deferred investing. *Journal of Finance* **59**(3), 999–1037.

Davis, M. H. A., and Norman, A. (1990). Portfolio selection with transaction costs. *Mathematics of Operations Research* **15**(4), 676–713.

DeMiguel, A., and Uppal, R. (2005). Portfolio investment with the exact tax basis via nonlinear programming. *Management Science* **51**(2), 277–290.

Domowitz, I., Glen, J., and Madhavan, A. (2001). Global equity trading costs. Working Paper, Ivestment Technology Group, Inc.

Dybvig, P., and Koo, H. (1996). Investment with taxes. Working Paper, Washington University, St. Louis, MO.

Gallmeyer, M. F., Kaniel, R., and Tompaidis, S. (2006). Tax management strategies with multiple risky assets. *Journal of Financial Economics* **80**, 243–291.

Gill, P., Murray, W., and Saunders, M. (2002). SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM Review* **47**(1), 99–131

Goodman, J., and Ostrov, D. N. (2010). Balancing small transaction costs with loss of optimal allocation in dynamic stock trading strategies. *SIAM Journal on Applied Mathematics* **70**(6), 1977–1998.

Haugh, M., Iyengar, G., and Wang, C. (2016). Tax-aware dynamic asset allocation. *Operations Research* **64**(4), 849–866 (https://doi.org/10.1287/opre.2016.1517).

Huang, J. (2008). Taxable and tax-deferred investing: a tax-arbitrage approach. *Review of Financial Studies* **21**(5), 2173–2207.

Janaček, K., and Shreve, S. E. (2004). Asymptotic analysis for optimal investment and consumption with transaction costs. *Finance and Stochastics* **8**(2), 181–206.

Leland, H. E. (2000). Optimal portfolio implementation: asset management with transactions costs and capital gains taxes. Working Paper, University of California, Berkeley.

Liu, H. (2004). Optimal consumption and investment with transactions costs and multiple risky assets. *Journal of Finance* **59**(1), 289–338.

Merton, R. C. (1992). *Continuous-Time Finance*, revised edn. Blackwell Publishing, Oxford.

Muthuraman, K., and Kumar, S. (2006). Multidimensional portfolio optimization with proportional transaction costs. *Mathematical Finance* **16**(2), 301–335.

Ostrov, D. N., and Wong, T. G. (2011). Optimal asset allocation for passive investing with capital loss harvesting. *Applied Mathematical Finance* **18**(4), 291–329.

Pollin, R., and Heintz, J. (2011). Transaction costs, trading elasticities, and the revenue potential of financial transaction taxes for the United States. Working Paper, Political Economy Research Institute.

Rogers, L. C. G. (2004). Why is the effect of proportional transaction costs $O(\delta^{2/3})$? In *Mathematics of Finance*, Yin, G., and Zhang, Q. (eds), pp. 303–308. American Mathematical Society, Providence, RI.

Shreve, S. E., and Soner, H. M. (1994). Optimal investment and consumption with transaction costs. *Annals of Applied Probability* **4**(3), 609–692.

Tahar, I. B., Soner, H. M., and Touzi, N. (2010). Merton problem with taxes: characterization, computation, and approximation. *SIAM Journal of Financial Mathematics* **1**, 366–395.

Whaley, A. E., and Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance* **7**(3), 307–324.