CFA INSTITUTE RESEARCH FOUNDATION / **MONOGRAPH**

LIFETIME FINANCIAL ADVICE

A Personalized Optimal Multilevel Approach

THOMAS M. IDZOREK, CFA
PAUL D. KAPLAN, CFA

CFA Institute Research Foundation

LIFETIME FINANCIAL ADVICE: A PERSONALIZED OPTIMAL MULTILEVEL APPROACH

THOMAS M. IDZOREK, CFA, AND PAUL D. KAPLAN, CFA

WITH A FOREWORD BY ROGER G. IBBOTSON

CFA Institute
Research
Foundation

Statement of Purpose

The CFA Institute Research Foundation is a not-for-profit organization established to promote the development and dissemination of relevant research for investment practitioners worldwide.

Photo credit: Getty Images/guowei ying

ISBN: 978-1-952927-37-9

BIOGRAPHIES

Thomas M. Idzorek, CFA, is chief investment officer, retirement, for Morningstar Investment Management LLC. He also serves as a member of Morningstar's 401(k) committee, Public Policy Council, Global Investment Committee, US Investment Policy Committee, and the editorial board of *Morningstar* magazine. He is a former member of the editorial board of the *Financial Analysts Journal*. Idzorek previously served as the president of Morningstar's global investment management division, where he oversaw the firm's global investment management entities, investment management consulting businesses, retirement solutions, Morningstar's broker/dealer (Morningstar Investment Services, LLC), global index business, and financial wellness business. Additionally, he has served as president of Ibbotson Associates, president of Morningstar Associates, board member/responsible officer for a number of Morningstar Investment Management subsidiaries, global chief investment officer for Morningstar Investment Management, chief investment officer for Ibbotson Associates, and director of research and product development for Ibbotson. Before joining Ibbotson (acquired by Morningstar in 2006), he worked for Zephyr Associates. Idzorek has written numerous journal articles, contributed to three CFA Institute Research Foundation books, and is a co-creator of the popularity asset pricing model (PAPM). He is an expert on multi-asset-class strategic asset allocation, the Black–Litterman model, target date funds, retirement income solutions, fund-of-funds optimization, risk budgeting, and performance analysis. Idzorek is the key methodological creator of Morningstar's target date and retirement managed account (robo-advice) solutions. He holds a bachelor's degree in marketing from Arizona State University and a master's degree in business administration from Thunderbird School of Global Management.

Paul D. Kaplan, CFA, was director of research for Morningstar Canada and a senior member of Morningstar's global research team. He led the development of many of the quantitative methodologies behind Morningstar's fund analysis, indexes, adviser tools, and other services. Kaplan conducted research on asset allocation, retirement income planning, portfolio construction, index methodologies, and other investment topics. Many of his research papers have been published in professional books and publications, such as the *Financial Analysts Journal* and the *Journal of Portfolio Management*, and he has served on the editorial board of the *Financial Analysts Journal*. Kaplan has received a Graham & Dodd Top Award and a Graham & Dodd Award of Excellence. Many of his works appear in his book *Frontiers of Modern Asset Allocation*. Kaplan has contributed to three CFA Institute Research Foundation books and is the lead co-creator of the popularity asset pricing model (PAPM). Previously, he has served as quantitative research director for Morningstar Europe in London, director of quantitative research in the United States, and chief investment officer of Morningstar Associates, LLC, where he developed and managed the investment methodology for Morningstar's retirement planning and advice services. Previously, Kaplan was a vice president of Ibbotson Associates and served as the firm's chief economist and director of research. Prior to that, he served on the economics faculty of Northwestern University, where he taught international finance and statistics. Kaplan holds a bachelor's degree in mathematics, economics, and computer science from New York University, and doctorate in economics from Northwestern University. Now retired, he volunteers at Probability Management, a nonprofit dedicated to making uncertainty actionable through tools, standards, applications, and training.

CONTENTS

CFA Institute

PROFESSIONAL LEARNING QUALIFIED ACTIVITY
This publication qualifies for 7.5 PL credits under the guidelines of the CFA Institute Professional Learning Program.

FOREWORD

All of us face the question of how to consume and invest over our lifetimes. We make these decisions periodically as we progress through our lives. We decide in the context of many other related decisions: how much and what type of education to get, what kind of jobs we intend to have, what type of family we plan to support, when might we plan to retire, and how strong a bequest desire we have. The investment products available to us have various risk and expected return trade-offs. Insurance products are also available that can protect our lifetime earnings, insure us against outliving our assets, or protect our investments against downside events.

An optimal solution to these questions is extremely complex, but all of us must plunge ahead. We have to consider how much we expect to earn, not only in the immediate future but also over the course of our lifetime with changing jobs, promotions, eventual retirement, and the probability of survival during each period along the way. We have to decide how much to spend and how much to invest at each of these points in time. Our investment asset allocation is dependent not only on our risk tolerance but also on the growth potential and riskiness of our employment as these change over our lifetime. We have to determine how our financial allocations should be located, that is, which assets should be invested in taxable accounts and which should be invested in tax-deferred tax accounts.

If we believe we have an edge in some areas, we may wish to be active investors, attempting to achieve alpha but also ending up with tracking error (i.e., active risk). We also may have preferences or needs for various characteristics that go beyond risk. These could include a need for liquidity, a preference for recognizable brands, or an interest in environmental, social, and governance (ESG) causes. Finally, along the way, we have to consider our bequest desires and how flexible they are, considering our changing circumstances.

These questions may seem impossible to answer. Yet we all make these decisions throughout our lives. For the most part we do this in ad hoc ways. We plunge ahead because we have to. Each day or each year, we have to decide what to spend, what to save, and how to invest. We usually do not consider the whole picture but rather compartmentalize each decision. This practice leads to suboptimal decisions because we have not integrated the different sources of our wealth, the different stages of our life, or our personal preferences along the way.

Fortunately, rich theoretical approaches are available to handle advice over our lifetimes. These involve the concepts of human capital and financial capital. Human capital is the present value of all that we earn over our lifetime. It includes wages, salary, bonuses, medical care, and other perks, along with retirement benefits such as Social Security, defined benefit plans, and Medicare. Financial capital includes investments in stocks, bonds, and real estate. In addition, individuals usually purchase protection products, such as life insurance, retirement annuities, and other types of insurance. The problem we face is knowing how to integrate the different types of wealth—human and financial—with the different ways that one can invest during the various stages of life, while also making use of insurance products.

One way to cope with this complexity is to stay in the present: even though we are making decisions in the context of our whole life, at each point in time, we are making only today's decisions, albeit with an estimate of what the future may look like. When our circumstances change in a significant and potentially unexpected manner, we can adapt and make new decisions that can take us off the path that we had envisioned earlier and start us on a new path instead.

Lifetime Financial Advice, 2007

In the early 2000s, many of us at Ibbotson Associates worked on models to solve these problems. Ibbotson Associates was sold to Morningstar Inc. in 2006, but the work continued at Morningstar. In 2007, CFA Institute Research Foundation® published *Lifetime Financial Advice: Human Capital, Asset Allocation, and Insurance*, which I co-authored with Moshe A. Milevsky, Peng Chen, and Kevin X. Zhu. Subsequent to the 2007 book, Thomas M. Idzorek and Paul D. Kaplan had made substantial progress in modeling lifetime advice. Their current CFA Institute Research Foundation book has a similar title, *Lifetime Financial Advice: A Personalized Multi-Level Optimization Approach*. Idzorek and Kaplan refer to our original *Lifetime Financial Advice* (2007) as IMCZ.

Many of the concepts used in the current Idzorek and Kaplan book were in *Lifetime Financial Advice* (2007), including renditions of the first illustration in their chapter 1. The diagram shows how, for young people, human capital dominates their total wealth. As people embark on their careers, they have a whole lifetime of earning ahead of them, which causes their human capital to be near its maximum value. Meanwhile, most young people have little financial capital, and in many cases, it is negative because they may have borrowed to complete their education. As people age, they save and invest, converting some of their human capital into financial capital. By the time they reach retirement, they have little human capital left, but they are near their maximum in financial capital. Their total capital is the sum of their human capital and their financial capital. In retirement, they spend down the financial capital, potentially facing longevity risk, with their intended spending potentially exceeding their remaining wealth.

The earlier *Lifetime Financial Advice* (2007) modeled human capital, usually regarding it as bond-like. That book prescribed a constant implied equity/bond mix, summed across human and financial capital, over one's lifetime. Because human capital is dominant for young people, what little financial capital they may have should be invested entirely in equities, perhaps even levered up. Young people do need to protect their earning power, especially if they have dependents. Life insurance is prescribed to protect a portion of the present value of their earnings.

As people age, they save and invest, converting their human capital into financial capital. The less their bond-like human capital is a part of their total wealth, the more conservatively the financial capital should be managed. Thus, investors may wish to reduce the equities and increase the bonds or bond-like assets in their asset allocation mix as they age. As investors approach retirement, they need to continue to de-risk their financial capital. This can be done either by adding bonds to the portfolio or using various types of insurance protection products. At this stage of their life, they have little human capital to protect, and life insurance is no longer important, except perhaps as a way to pass on untaxed bequests.

In retirement, investors face a different problem: longevity risk. Of course, we all want to live a long time, but we might live so long that we run out of assets. Longevity risk is the danger that our diminished human capital plus our financial capital will not be sufficient to cover our entire life's spending. Payout annuities are a way to smooth one's consumption (or joint consumption with one's spouse) over an entire lifetime of uncertain length. With payout annuities, investors are able to spend more in their early retirement years, because they do not have to protect themselves against the contingency of a long life.

Lifetime Financial Advice, 2024

Idzorek and Kaplan would agree with this general direction about to how to manage our investments over our lifetimes. They also use similar lifetime concepts of human capital and financial capital. They also make use of life insurance and annuities as part of the solutions. Idzorek and Kaplan, however, provide much more detail and a full, holistic solution to lifetime investing.

One of the substantial contributions that Idzorek and Kaplan make is to solve the consumption problem—that is, how much to consume during each period of our lifetime. (In IMCZ, consumption is exogenous, that is, not capable of being varied as part of the solution.) In Idzorek and Kaplan, consumption is determined at each point in time, and even though consumption over one's entire lifetime is estimated, it is easily adapted to changing circumstances at the next point in time. Thus, the investor can determine how much to save, withdraw, or invest, reflecting not only on their current circumstances but also in the context of the life they have ahead, including their desire to make bequests.

In Part I of this book, Idzorek and Kaplan solve the complex problem of how much to consume each period, what their asset allocation mix is, how much life insurance is needed for bequests, and as retirement approaches, when and how much to annuitize. After human capital, financial capital, and liabilities are estimated, the models are parameterized with the individual's subjective discount rate, intertemporal elasticity, and risk tolerance. The authors provide sample questions to help practitioners estimate these and other preference parameters.

Separately, the strength of the bequest motive and its flexibility need to be provided. The maximization equation tells us what portion of total wealth to consume each period. Risk tolerance determines the asset mix in the context of both human capital and financial capital.

Taxes have a significant impact on investment results. Part II of this book addresses asset location. An investor has both a taxable account and tax-deferred accounts. An important contribution of this book is to connect the tax management of portfolios to the life-cycle framework. The techniques presented in this part use single-period optimization models. These models can be used because decisions are made at a point in time, even though they are in the context of an entire lifetime. Each asset class has different tax characteristics, so that the deferred accounts have different holdings than the taxable accounts. The linkage of the Part I life-cycle model across time with the Part II single-period "net-worth optimization" at each point in time is a true breakthrough in the field.

In Part III of this book, Idzorek and Kaplan introduce other preferences as well as potentially heterogeneous expectations. Here, they use the popularity asset pricing model (PAPM), about which I have previously co-authored several works with them, including the CFA Institute Research Foundation book *Popularity: A Bridge between Classical and Behavioral Finance* (2018) by Ibbotson, Idzorek, Kaplan, and Xiong. The PAPM can include numerous individual preferences, such as a desire for liquidity, recognizable brands, or particular protection insurance schemes. For those with ESG preferences, the model is an ideal framework for tilting toward characteristics that the investor likes and away from characteristics they dislike. This model is integrated with an approach that also allows for investors to seek alpha while taking account of tracking error (active risk) as a cost incurred in seeking that alpha.

In the penultimate chapter (chapter 11), the authors bring all these elements of lifetime financial decision making together holistically. This chapter, like much of the book, is full of equations. The authors determined that formal models are needed to address the complex problems of how to manage our investments and our human capital over our entire life. Many readers will find the equations challenging. It is fine to skip over these equations, focusing instead on the illustrations and the intuitions described in the text. The authors are to be commended, however, for their ability to integrate the numerous dimensions of investing into a holistic approach and to express them both conceptually and mathematically.

The final chapter (chapter 12) provides a very special kind of summary. Woven into most of the other chapters is the experience of a fictitious investor, Isabela. The authors follow her through her entire lifetime of planning, saving, investing, consuming, and bequeathing. She is guided by a fictitious planner, Paula, who uses the theories and methods in this book to provide lifetime financial advice. This final chapter thus ties together Isabela's lifetime of experiences using a series of economic balance sheets, asset mixes, and other elements of the advice Paula renders.

The general structure of the book labels the three parts of the analysis as a parent model, a child model, and a grandchild model. Each part then can be optimized using inputs from the previous level. This structure enables the whole approach to be modularized into parts, while keeping all the parts integrated so that they can all be used together. Because each part involves optimization, and we know from other work that optimization can intensify input errors, one might wonder whether serial optimizations worsen this error intensification. Actually, they do not. In this method, input errors do not grow exponentially, because each level of optimization tends to constrain the optimizations at the next level.

In closing, Thomas Idzorek and Paul Kaplan deserve a great deal of credit for providing a holistic set of solutions to some of the most important problems that we face over our lifetimes. They have integrated human capital, financial capital, life insurance, annuities, bequests, taxes, investor preferences, heterogeneous expectations, and alpha seeking into a formal lifetime consumption and investing model. This is a tremendous accomplishment. CFA Institute Research Foundation is extremely proud to present their work as part of an ongoing investigation into how to invest over the entire lifetime of a human being.

Roger G. Ibbotson
Professor in the practice emeritus of finance at Yale University
and chair of Zebra Capital Management LLC

AN INTEGRATED LIFE-CYCLE APPROACH

1. A NEW MULTILEVEL LIFE-CYCLE MODEL

Life-cycle finance is arguably the most important specialty in finance.

—Laurence B. Siegel[1]

Life-cycle finance is the specialty in finance that focuses on the financial issues faced by individuals over the course of their lifetimes. It is the economic approach to financial planning. This makes it arguably the most important specialty in financial economics because individuals have only one lifetime in which to get it right. Although much has been written on life-cycle finance, to our knowledge, a comprehensive and actionable treatment does not exist. Additionally and until now, life-cycle models have been disconnected from the much more prevalent single-period optimization models. In this book, we bring together various strands of research, some of which we originated, into a comprehensive and concrete framework for personalized optimal financial decision making over the course of one's life. Following advice based on this framework should lead to better financial outcomes and more gratifying, personalized results.

In 2007, CFA Institute Research Foundation published *Lifetime Financial Advice: Human Capital, Asset Allocation, and Insurance* by Roger Ibbotson, Moshe Milevsky, Peng Chen, and Kevin Zhu (hereafter IMCZ), which we think of as a precursor to this book.[2] **Exhibit 1.1** highlights many of the key takeaways from IMCZ as well as life-cycle finance.

Starting at the left vertical cross section, most young people do not have much in the way of financial assets, and their overall wealth is dominated by what is called human capital (the value of all future labor income). Human capital and financial capital are the primary components of the asset side of an individual balance sheet, in which the individual balance sheet provides a relatively holistic view of an investor's financial health. Life-cycle finance embraces the critical role of human capital as the most important asset for many investors.

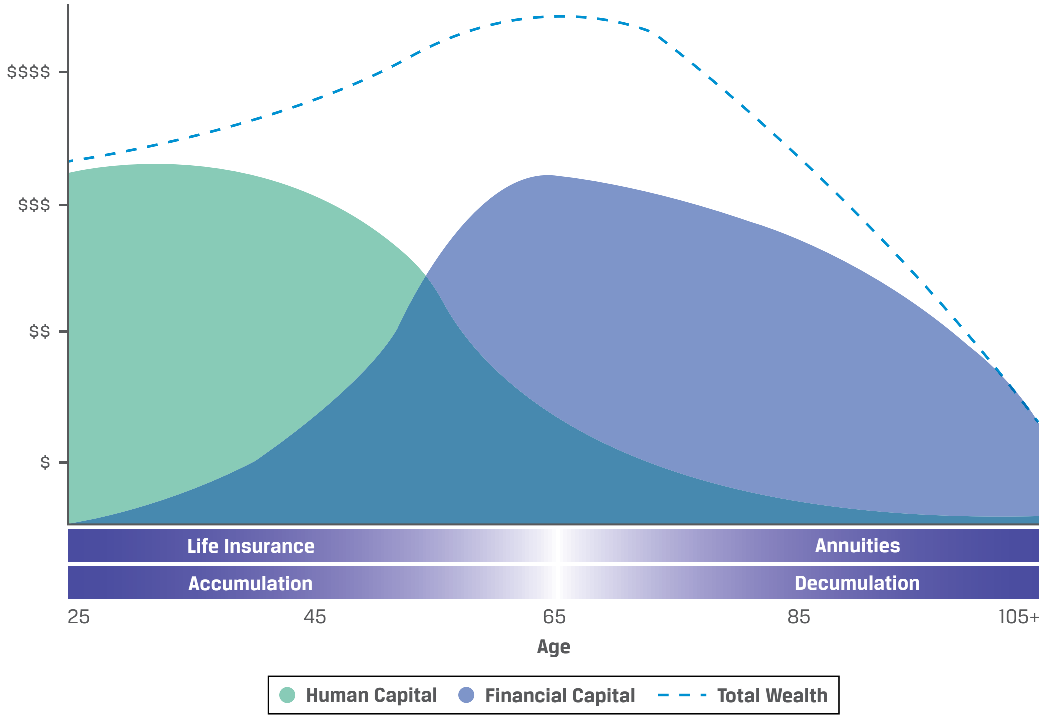
Moving from left to right, through time, most investors use the majority of their ongoing salary (labor income) to pay for ongoing expenses (consumption), but they also save and invest part of their earnings, building up financial capital. Starting in retirement, individuals begin to pay for ongoing expenses (consumption) with deferred labor income (e.g., social insurance or defined benefit pension income) *and* by drawing down accumulated financial capital. A primary objective is to accumulate enough assets (both deferred labor income and financial assets) to fund retirement—notice the dashed dark red line representing the desire for smooth real consumptions. Some investors risk living too frugally, whereas others risk not saving enough.

From a lifetime asset allocation perspective, if the cash flow characteristics of human capital are more bond-like than stock-like, young investors who have lots of human capital and little financial capital, are often overallocated to a fixed income–like asset that they cannot easily alter. To achieve a diversified holistic wealth portfolio that meets their risk tolerance, most younger investors should primarily invest their financial capital in equities. Moving through time, as human capital is saved and thus transformed into financial capital, the composition of total wealth evolves in such a way that financial capital is gradually invested more conservatively as it grows, and the value of human capital decreases. Turning to risks,

[1]From foreword to IMCZ (2007).

[2]IMCZ (2007) was largely written in 2005 and 2006, with Ibbotson, Chen, and Zhu working for Ibbotson Associates and Milevsky serving as a consultant to Ibbotson Associates. Both Idzorek and Kaplan also spent time at Ibbotson Associates, which was acquired by Morningstar in 2006.

Exhibit 1.1. The Investor's Life Cycle



during accumulation, life insurance can protect the value of human capital (mortality risk). During decumulation, annuities can protect against out living one's assets (longevity risk).

Building on the key insights and implied heuristics from IMCZ, in this book we move to a concrete three-stage model for providing optimal, personalized lifetime financial advice.

A major innovation in our approach is that it melds life-cycle models, which apply over an investor's entire lifetime, with single-period Markowitz optimization-based models to be run relatively frequently. As we will demonstrate, we base this new linkage in part on the Levy–Markowitz (1979) utility function and in part on the investor's economic balance sheet. In each period (such as a year), the life-cycle model gives the optimal amount that the investor should spend, save, or withdraw in aggregate and ascertains the optimal level of risk for financial assets (given the investor's human capital and liabilities taken from the investor's economic balance sheet). Taking information from the life-cycle model, the first level single-period optimization, what we call "net-worth optimization," is run each period to determine the optimal asset allocations in taxable and tax-advantaged accounts. Then, within each period and on an ongoing basis, a second optimization determines allocations to specific investment products (such as mutual funds) in each account. In this second optimization, the investor's nonpecuniary preferences for investment characteristics (one example being ESG) can be incorporated to tilt portfolios toward characteristics that the investor likes and away from those that they dislike.

Another feature that touches each level of our approach is its high degree of personalization. First, our approach personalizes the financial inputs at each level. These financial inputs include projected future labor income that forms the basis of human capital and projected future nondiscretionary spending that forms the basis of the investor's liabilities. We bring together human capital and liabilities in the investor's economic balance sheet. This personalized balance sheet serves as an intertemporal budget constraint on discretionary spending.

Second, our approach includes a wide set of preference parameters, both *pecuniary* and *nonpecuniary*. Pecuniary preferences include not only risk tolerance but also preferences regarding the timing and magnitude of discretionary consumption (spending) as well as bequests.

Third, we use a model for longevity that can be easily personalized. This model also allows us to price life insurance and annuities that the investor can use to manage mortality and longevity risk. This is an application of the argument in IMCZ that lifetime financial planning must treat insurance and annuities as well as securities (which traditionally receive most or all of the attention) as vital components of the planning process. The life-cycle model developed in this book gives the optimal amounts to deploy in each of these classes of products, in each period, and in a personalized way.

The elements of our three-stage model include the following:

1. Human Capital and Liabilities

   Human capital is the present value of all future earnings. During the working phase of a person's life, earnings are primarily wages. During the retirement phase, it includes social insurance (such as US Social Security) and income from defined benefit plans and annuities.

   The value of liabilities is the present value of future nondiscretionary spending (including debt repayment, such as mortgage payments). A person's financial wealth and human capital, taken together, must, at a minimum, be able to fund their liabilities.

2. The Investor Balance Sheet

   Just as a company's balance sheet provides a snapshot of its health, we introduce an investor balance sheet, which has financial wealth (in the form of securities and contracts) and human capital on the left or asset side, and liabilities and net worth on the right side. (We thus define net worth as financial wealth plus human capital minus liabilities.) In our model, the financial planning process centers on the investor balance sheet.

3. Financial or Pecuniary Preference Parameters

   We define a set of parameters that specify the investor's attitudes toward the magnitude and timing of discretionary consumption, risk, and bequests.

4. A Model of Longevity

   The probability of living to each year plays a central role in making rational decisions regarding consumption, bequests, life insurance, and annuities. We use a well-established parametric model of longevity, suggested by one of the authors of IMCZ (Milevsky 2012a), to calculate the probability of the investor living to each year.

5. Life-Cycle Models

   Life-cycle models are at the heart of our framework. A life-cycle model gives rational rules for annual spending, saving, withdrawing, and investing. During working years, spending is usually less than income, so the models provide rational rules for saving. During retirement, spending is usually more than income, so the model provides rational rules for drawing down accumulated financial wealth. In each year, the model also provides rational investment rules that set the asset allocation of net worth.

   At all times, a life-cycle model imposes the intertemporal budget constraint, which says that the present value of future discretionary spending must equal net worth.

6. Life Insurance and Annuities

In some of our life-cycle models, life insurance and annuities are available to the investor to manage mortality and longevity risk. Fixed payout annuities allow the investor to guarantee a pattern for lifetime income that does not depend on the path of wealth, regardless of how long they live. Variable payout annuities allow the investor a way to implement rational lifetime spending rules that do depend on the path of wealth. Life insurance allows the investor to guarantee a bequest of a given size, regardless of when they die. We also introduce a model to set the optimal size of the bequest.

7. An Asset Allocation and Location Optimizer Linked to the Investor Balance Sheet

We extend the Markowitz mean–variance asset allocation model in two ways. First, we link it to the investor balance sheet so that the model is providing asset allocation advice for financial assets, taking the asset allocation of human capital and of liabilities as a given. Second, the model determines the optimal location of assets in taxable and tax-advantaged accounts. We refer to this new simultaneous net-worth asset allocation and asset location that holds human capital long and liabilities short as *net-worth optimization*.

8. Nonfinancial or Nonpecuniary Investor Preferences

Based on work we presented in a previous CFA Institute Research Foundation publication (Ibbotson et al. 2018), we introduce nonpecuniary investor preferences into the portfolio construction process. By nonpecuniary preferences, we mean preferences for security characteristics other than risk and expected return. Although such characteristics could be almost anything, these are often preferences related to ESG issues, such as carbon emissions.

9. A Multi-Account Portfolio Optimizer with Taxes and Nonpecuniary Preferences

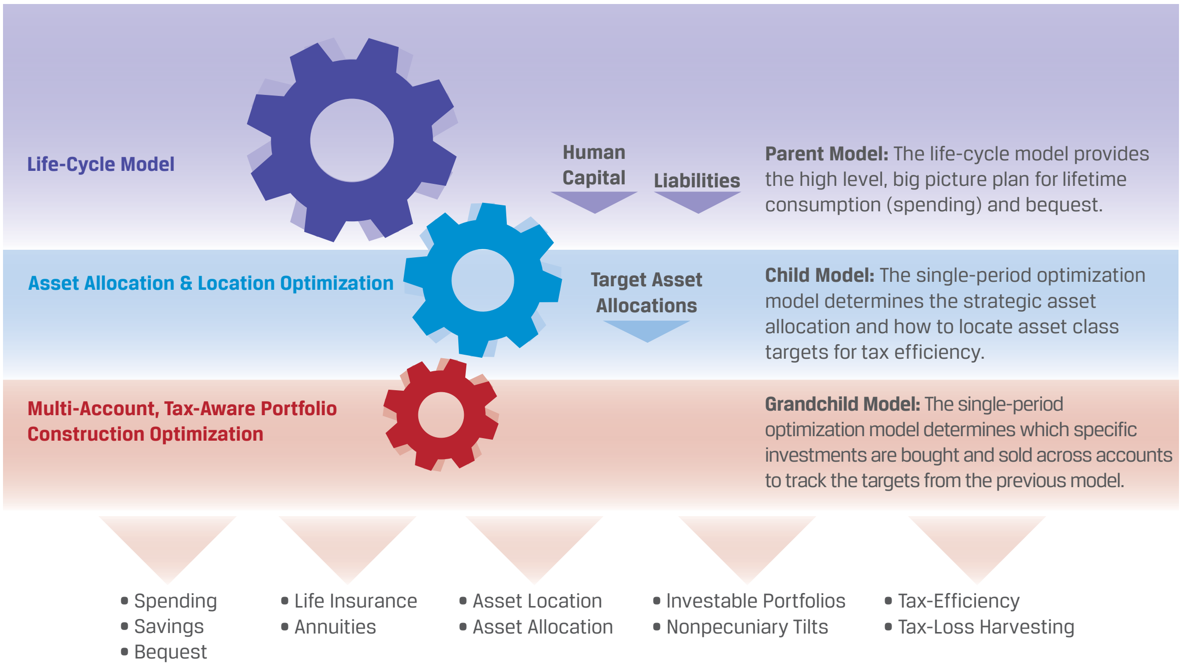
Because investor portfolios are typically implemented with managed products such as mutual funds and exchange-traded funds (ETFs), we include an optimizer that creates portfolios of funds. We do this by extending the manager structure optimizer of Waring et al. (2000) to include (1) multiple accounts, each with a different tax treatment; and (2) nonpecuniary preferences, so that the optimal portfolio tilts toward characteristics that the investor likes and away from the characteristics that the investor dislikes. In addition to the nonpecuniary characteristics, the optimizer favors expected active return (alpha) and disfavors active risk (sometimes called tracking error) with respect to target asset allocations that come for the asset allocation and location optimizer and with respect to manager-specific risk.

In this book, we assemble these elements into a new comprehensive multilevel model for providing optimal lifetime advice that brings together life-cycle and single-period optimization models. As **Exhibit 1.2** shows, our multilevel model starts with a "parent" life-cycle model that feeds into a "child" single-period net-worth optimization, which in turn feeds into a "grandchild" single-period, multi-account alpha-tracking error optimization. The bottom of Exhibit 1.2 identifies the numerous optimal outputs from the collective multilevel model. To the best of our knowledge, no other comprehensive model emanating from leading theories cohesively addresses these numerous practical decisions.

The parent life-cycle model builds on the work of Fisher (1930), Samuelson (1969), Merton (1969, 1971), Fama (1970), Lucas (1978), Kaplan (1986), and others. It considers a number of key financial or pecuniary investor preferences (beyond risk tolerance) and leads to optimal advice related to savings, spending, life insurance, leaving a bequest, and annuities. It forms the basis for a holistic view of the investor's economic balance sheet. In addition to the investor's risk tolerance, the estimated asset allocations associated with the investor's human capital and liabilities are outputs from the parent model that serve as inputs in a single-period, simultaneous, asset location and asset allocation optimization child model.

In the child asset location and asset allocation optimization model, we expand on Markowitz's (1952, 1959) mean–variance optimization (MVO) to jointly solve for separate target asset allocations based on the tax efficiency of the different asset classes. We further expand this model to incorporate the investor's

Exhibit 1.2. New Multilevel Model Linking Life-Cycle Models with Single-Period Optimization Models



balance sheet. We do this by extending the liability-relative optimization or surplus optimization framework of Leibowitz (1987) Sharpe (1990), Sharpe and Tint (1990), and Idzorek and Blanchett (2019) by including not only liabilities as an asset held short but also human capital as an asset held long. We call the resulting optimization framework "net-worth optimization." Importantly, we link the net-worth optimization problem to the utility maximization problem of the life-cycle model, essentially enabling them to talk the same language. The separate target asset allocation outputs (e.g., one for taxable accounts, one for tax-deferred accounts, and one for tax-exempt accounts) from the child model serve as inputs into a single-period, multi-account alpha-tracking error optimization grandchild model.

The grandchild single-period, multi-account alpha-tracking error optimization model expands the single account tax-free manager structure optimization framework of Waring et al. (2000) to optimize across all of an investor's accounts in a single optimization. This optimization determines which specific investments to buy and sell. Tax efficiency is driven by the asset location optimized targets as well as optimization parameters for investment options that differ in their tax treatment. The objective function includes not only a penalty for tracking error but also trading costs and realized taxes, enabling the model to serve as a tax-loss harvester and smart reoptimizer/rebalancer while minimizing trading costs and taxes. Based on the PAPM of Ibbotson et al. (2018) and Idzorek, Kaplan, and Ibbotson (2021, 2023), the objective function also includes a nonpecuniary preference utility term that allows the optimizer to further personalize the portfolio by tilting toward characteristics the investor likes and away from characteristics they dislike. The ability to look across accounts enables the optimizer to select the best possible investment option from across the various accounts while also considering their nonpecuniary characteristics.

Modularization

When it comes to uniting our parent, child, and grandchild models into a cohesive multilevel model, we believe the combined model is more powerful than the sum of the individual parts. With that said, the combined model can be modularized such that each of the models—the parent life-cycle model, the net-worth asset allocation and asset location optimization model, and the personalized multi-account alpha-tracking error optimization model—work as standalone models. Furthermore, each individual model can be coupled with other models or systems created by others. For example, as put forth in this book, the net-worth optimization creates the separate target asset allocations that serve as inputs into the multi-account alpha-tracking error optimization. The multi-account alpha-tracking error optimization is indifferent to how the separate target asset allocations are created.

Our comprehensive approach is an advancement over IMCZ who use different models to answer different questions, take consumption as a given rather than solve for it, and do not take liabilities into account.

The Rest of the Book

The rest of this book is organized around the three levels of our multilevel model. Part I covers the parent life-cycle model. Part II presents the single-period, simultaneous, asset location and asset allocation optimization child model. Part III puts forth the single-period, multi-account alpha-tracking error optimization grandchild model in which we include nonpecuniary preferences.

To illustrate how a practitioner, such as a wealth adviser or financial planner who may not understand the details of the models, could use the three-stage multilevel model for ongoing financial planning, we weave in an ongoing end-to-end applied example. In our applied example, we introduce a hypothetical investor Isabela, who is working with a financial planner Paula. We follow Isabela throughout her lifetime. We assume that Paula is using a state-of-the-art financial planning and investment management system based on the three-stage model and concepts presented in this book.

We conclude the book by pulling together the end-to-end example followed by a call for change and action.

The vast majority of the material in this book applies globally. Given the nature of financial planning and the critical role of taxation and social insurance, when presenting some of the methods, especially in conjunction with the specific example, we use the United States for our setting. Nevertheless, the methods presented in this book can be applied in any country in the world. Specific laws and institutions that vary from country to country will affect such application.

PART I: PARENT LIFE-CYCLE MODEL

> As far as I am aware, no one has challenged the view that, if people were capable of it, they ought to plan their consumption, saving, and retirement according to the principles enunciated by Modigliani and Brumberg in the 1950s.
>
> —Angus S. Deaton (2005)[3]

In homage to *Strategic Financial Planning over the Lifecycle* by Narat Charupat, Huaxiong Huang, and Moshe Milevsky (2012), which begins with this fantastic quote from Angus Deaton, we too are unaware of any such challenges.

In Part I of this book, we develop our parent life-cycle model and some variations, including models of lifetime consumption (spending), investing, and bequests. These models integrate human capital, asset allocation, life insurance, and annuities as constituents of a comprehensive life-cycle model.

Many practitioners are unfamiliar with life-cycle models and find them relatively foreign; thus, we pay particular attention to presenting them in a detailed manner with an emphasis on providing practical solutions and insights. We demonstrate how *life-cycle* models provide a holistic *lifetime* blueprint or financial plan. Later, in Parts II and III, we integrate the use of life-cycle models with single-period optimization models based on Markowitz (1952, 1959) MVO. In this book, the life-cycle model is the parent model, and the single-period optimization models act as child and grandchild models providing more detailed, time-specific financial advice.

The vast majority of humans are myopic, living in the here and now. Even when it comes to financial decision making, many of us make decisions on an as-needed basis. Financial planning is strategic planning for the financial aspects of one's life. Strategic planning of any kind requires discipline and, ideally, the ability to step back and see the broader and more holistic context. Like a master chess player, rather than looking one or two moves ahead, one should attempt to see all of the possible moves and anticipate a wider range of possible countermoves. Such a chess player is prepared with a new and complete game plan for every possible situation, assuring victory over a less-well-prepared player.

In the world of optimal financial planning, "life-cycle" models are the ultimate chess players. A good life-cycle model tries to understand all of the possible moves and countermoves. Although life-cycle models do not guarantee victory, one might think of them as Deep Blue, IBM's master computerized chess player, applied to financial planning. Like Deep Blue, life-cycle models are inherently sophisticated; most practitioners are unaware of them, making them unavailable and unused by financial planners for optimal financial planning. To put it bluntly, life-cycle models are the most powerful models for financial planning that we have, but they exist only in obscurity.

One of our key goals is not only to introduce practitioners to life-cycle models but also to present them in a way that demystifies them and develops enough intuition around them that practitioners will feel comfortable with their recommendations (even if the details of the model are not fully understood). Today, many practitioners are comfortable with the output and recommendations associated with things like Monte Carlo simulation, scenario analysis, returns-based style analysis, performance attribution analysis, and MVO; yet most practitioners do not understand the details of these techniques. The time has come to put life-cycle models in the hands of practitioners.

Antecedents

Our approach is based on the pioneering work of Nobel laureates Milton Friedman and Franco Modigliani, the originators of life-cycle modeling (Friedman 1957, Modigliani 1966). In their models, individuals base

[3]Deaton is the 2015 Nobel Laureate in Economic Sciences.

their consumption decisions not on current income alone, as postulated by Keynes (1936), but on expected lifetime wealth and income. During working years, the individual can save and invest a portion of their income to build up enough financial wealth to provide for consumption during retirement.

These models were further developed by Nobel laureates Paul Samuelson (1969) and Robert Merton (1969, 1971). Based on the insight of yet another Nobel laureate, Gary Becker (1993), life-cycle models came to view the present value of future income as "human capital," a form of wealth like any other. (As an abstract concept, human capital goes back centuries, arguably to Adam Smith.)

For many people, human capital is their most important and valuable asset. An understanding of its characteristics dramatically alters the way that financial wealth should be invested. Individuals can borrow against their human capital to finance large purchases, such as a car or home. For many, it provides a steady paycheck, which pays for ongoing living expenses so that financial capital can remain invested (and hopefully grow) over a long time horizon. In a life-cycle model, consumption decisions are based on total wealth, which includes both financial wealth and human capital, as well as on liabilities as we discuss in Chapter 4.

Siegel (2008, p. xiii) notes that mainstream finance has focused on how to build optimal portfolios, and little on life-cycle finance:

> But this body of work [mainstream finance] does not say (or least it does not say very clearly) how much to save, how quickly to spend down one's assets, or how to insure against untoward events. It does not answer the central question of life-cycle finance: How can I spread the income from my working life over my entire life? To address these questions, we need to look outside mainstream finance—in particular, at actuarial science and the theory of insurance.

Life-Cycle Finance and Economic Theory

In this part of the book, we use economic theory to illuminate life-cycle finance. Bodie, Treussard, and Willen (2008) present an outline of an economic theory of life-cycle finance. They discuss three economic principles that are essential for a comprehensive life-cycle model. In the following exhibit, we list these principles and note how they are manifested in the life-cycle models presented in this book.

Economic Principles of Life-Cycle Finance and Their Manifestations in Our Models

Economic Principle	Manifestation in Our Models
Focus not on the financial plan itself but on the consumption profile it implies.	In our models, investors maximize utility over the entire lifetime path of consumption.
Financial assets are vehicles for moving consumption from one location [time] in the life cycle to another.	Investors save and invest in financial assets during their working years and, in retirement, draw down their financial assets to fund consumption.
A dollar is more valuable to an investor in situations in which consumption is low than in situations in which consumption is high.	We make the standard economic assumption of diminishing marginal utility.

Source: Economic principles from Bodie, Treussard, and Willen (2008).

Bodie, Treussard, and Willen (2008) then identify five insights from the economic theory of life-cycle finance. In the exhibit below, we list these along with how they are manifested in our models.

Our models lead to some additional insights. In particular, the following:

- *The Optimal Use of Life Insurance and Annuities over the Life Cycle*

  In our models, if the investor wishes to leave a bequest of a given size, during their working years, they should accumulate funds in ordinary assets and fill the gap between what they have accumulated and the desired bequest with term life insurance, until they have accumulated enough assets to leave the bequest without insurance. At that point, they should generate income using annuities. They should not hold term life insurance and a lifetime payout annuity at the same time, because these contracts are opposites.

- *Immediate Variable Annuities (IVAs) Are the Optimal Instruments for Generating Income from Risky Assets*

  If annuities are available, the investor should generate income by buying IVAs (not to be confused with deferred variable annuities that are like mutual funds with insurance wrappers). An IVA is similar to an immediate fixed annuity (IFA) but with payouts that are linked to the value of a portfolio of risky assets. If, however, IVAs are not available, the investor should attempt to mimic the payout of IVAs, by selling ordinary securities or collecting dividends and interest on them, to generate income.

Insights from Economic Theories of Life-Cycle Finance and Their Manifestations in Our Models

Insight from Economic Theories	Manifestation in Our Models
The Lifetime Budget Constraint	In our models, an investor maximizes the utility of consumption over the entire lifetime, subject to a single lifetime intertemporal budget constraint.
The Importance of Constructing "Contingent Claims"	We assume a complete market for contracts that pay off contingent on whether or not the investor is alive at any given time. Hence there is a complete market for annuities and life insurance. We also assume a complete market for contracts with payments contingent on any possible state of the world at any time (these are what we ordinarily call risky assets).
The Prices of Securities Matter!	We assume that a complete market exists for risky assets that are priced with a stochastic discount factor.
Risky Assets in the Life-Cycle Model	In our models, the investor manages consumption across possible states with contingent claims contracts (risky assets).
Asset Allocation over the Life Cycle	In our models, the investor maintains a constant level of risk for their net worth by changing the level of risk in financial assets as the levels of human capital and liabilities evolve over time.

Source: Insight from economic theories from Bodie, Treussard, and Willen (2008).

Structure of Part I

Part I contains five chapters and is organized as follows: Chapter 2 sets the stage for the rest of Part I as well as later chapters by arguing for a truly holistic approach to financial planning that incorporates a variety of pecuniary and nonpecuniary factors. It goes well beyond risk tolerance to include pecuniary preferences related to consumption and bequest and introduces the idea that investors may also have nonpecuniary preferences they want incorporated into their portfolios.

Chapters 3 and 4 put forth the core ingredients and frameworks from which the life-cycle models in chapters 5 and 6 are built. More specifically, chapter 3 develops the framework of utility theory for making rational decisions. It focuses on the characteristics of the investor, including preferences, survival probabilities, and nondiscretionary spending.[4] We model survival probabilities using the two-parameter formula presented by Milevsky (2012a, chapter 2); it gives results very similar to those obtained from actuarial mortality tables. Based in part on the work of Larry Epstein and Stanley Zin (Epstein and Zin 1989), the life-cycle models we present in later chapters include five types of investor preferences that affect optimal personalized advice.

In chapter 4, we continue to focus on the investor from a pecuniary perspective and the development of the additional key ingredients and foundation for the life-cycle models presented in chapters 5 and 6. We elaborate on the investor's balance sheet, which shows the investor's assets, liabilities, and net worth holistically (i.e., including human capital and other potentially overlooked assets and liabilities). (This is called the "economic balance sheet" by Waring and Whitney 2009.) The investor's net worth is the difference between these.

As Wilcox, Horvitz, and DiBartolomeo (2006) note, "The key element in applying best-practice simulations is the time series of implied balance sheets … showing the relationships of discretionary wealth to assets" (p. 16). We follow this principle in all of our life-cycle models.

The most important nonfinancial asset included in this balance sheet is human capital. One of the insights made by a number of financial economists, notably Milevsky, is that human capital, like all assets, comes with risk and that changes in value of human capital can be correlated with the returns on risky assets, such as stocks and bonds. In his book, *Are You a Stock or a Bond?*, Milevsky (2012b) illustrates the potentially risky nature of human capital. For example, a tenured university professor has bond-like or safe human capital, whereas a stockbroker has stock-like or risky human capital. These are both polar cases and most people are somewhere in between; in any case, the risk of this asset must be understood and modeled. In chapter 6, we model human capital like a combination of stocks and bonds.

When making financial decisions, the investor needs to budget for current nondiscretionary expenses (consumption) as well as a future nondiscretionary spending stream. We treat the present value of this stream as the "liability" on the right side of the investor's balance sheet (financial assets and human capital being on the asset side). (It is not a legal liability as one would see on an ordinary balance sheet, but it behaves very much in the same way.) Like human capital, the value of the investor's liabilities can be risky. In chapter 4, we model the value of liabilities using the same approach as that which we use to model human capital. At any point in time, the value of financial assets, plus human capital, less liabilities is the investor's net worth. In the life-cycle models presented in this book, savings and spending decisions are based on net worth.

---

[4]By preferences, we mean how an investor ranks alternative combinations of consumption at difference times and under different market conditions, as well as alternative combinations of consumption and bequests. By survival probability, we mean the likelihood of the investor physically surviving to the end of a given period.

Over the course of a lifetime, an individual may face various forms of risk. These include (1) income risk (reflected in the risk of human capital), (2) investment risk, and (3) mortality and longevity risk. Of these, investment risk can be managed through investment decisions. Mortality risk (which we define as the risk of dying during one's income earning years) can be mitigated with life insurance, and longevity risk (the risk of outliving one's money, that is, the means to fund consumption) can be mitigated with annuities.

In chapters 5 and 6, we bring together the ingredients and methods from chapters 3 and 4 to form a series of life-cycle models. In chapter 5, we present models with a fixed market rate of return in which the investor uses life insurance and annuities to manage uncertainty about the time of their future death. In chapter 6, we introduce market uncertainty as well as uncertainty in income and nondiscretionary expenses into the models that we present in chapter 5.

As a brief warning, like most financial models, life-cycle models involve formulas, some of which are complicated. Where possible, we have written the content in such a way that practitioners can largely skip the formulas while still developing a general understanding of the model, its inputs, and most important, the practical advice of the model output. The key outputs from the parent life-cycle models are as follows:

- A holistic estimate of the investor's economic balance sheet (through time)
- A lifetime spending and saving schedule, including
    - nondiscretionary consumption; and
    - discretionary consumption.

2. HOLISTIC INVESTOR PROFILING

Context

Focusing on the investor, we emphasize that investors have (1) pecuniary or financial preferences that go beyond risk tolerance and that they have an impact on the output (e.g., consumption and bequest) of the life-cycle model; and (2) nonpecuniary or nonfinancial preferences that have an impact on investment selection. This chapter sets the stage for the rest of Part I and the ensuing life-cycle models; note, however, that the nonpecuniary investor preferences introduced here as part of a holistic investor profile are not incorporated until Part III.

Key Insights

- Life-cycle finance takes a more holistic view of the investor and requires a more holistic investor profile.
- Many life-cycle models, including the ones developed in this book, have three to five key financial (pecuniary) investor preferences.
- The industry focuses on one preference (i.e., risk tolerance) and largely ignores the other key investor preferences.
- A complete investor profile and assessment system should attempt to measure all of the investor's preferences, including all pecuniary preferences as well as the investor's nonpecuniary preferences. We call on the industry to develop and adopt such systems.
- Risk tolerance is ubiquitous in financial planning. This emphasis leads to a high degree of comfort (and lack of scrutiny) that may not be justified. Different definitions and assumptions related to risk tolerance and the investor's risk profile, the assumptions of the tools used to measure them, and the portion of the investor's portfolio to which they are ultimately applied can lead to material errors relative to what was intended.
- Among practitioners, risk tolerance is typically treated myopically and directly applied only to financial investments. This is in contrast to life-cycle finance, where it is applied holistically to total wealth as described by Merton (1969) and Samuelson (1969). We recommend broadening the application of risk tolerance beyond specific financial asset accounts to include all the components of economic net worth, which is given by the investor's full economic balance sheet.
- Financial planners and advisers need a *holistic-balance-sheet-estimator* that enables them to apply risk tolerance to the individual's net worth and to "infer" the appropriate risk level for the part of the investor's total wealth under advisement.

In this chapter, we note the difference between what we call a holistic *investor profile* and the much more common, myopic, and investment-centric investor *risk profile*. We provide a blueprint for what should be included in the more holistic investor profile and how one might ascertain the investor's key pecuniary (financial) and nonpecuniary (nonfinancial) preferences. Just as an investor profile is often confused or conflated with a risk profile, a risk profile is often incorrectly thought to be the same as risk tolerance. Whether one is discussing the investor profile, the risk profile, or risk tolerance, it is important to have clear definitions of each, to understand the differences, and to apply them in a coherent manner.

Investor Profiling

What is known as "investor profiling" or "client profiling" is often viewed and treated as an independent step within an overall financial planning process. Far too often, the financial planning process focuses on just one element of the investor profile (i.e., risk tolerance) and is reduced to the following sequence:

1. have the client take a risk tolerance questionnaire (RTQ) or assessment;
2. file the RTQ in a safe place (where it is unlikely ever to be seen again) to serve as documentation of suitability; and
3. for the specific account in question, slot the client into a risk-based model or portfolio.

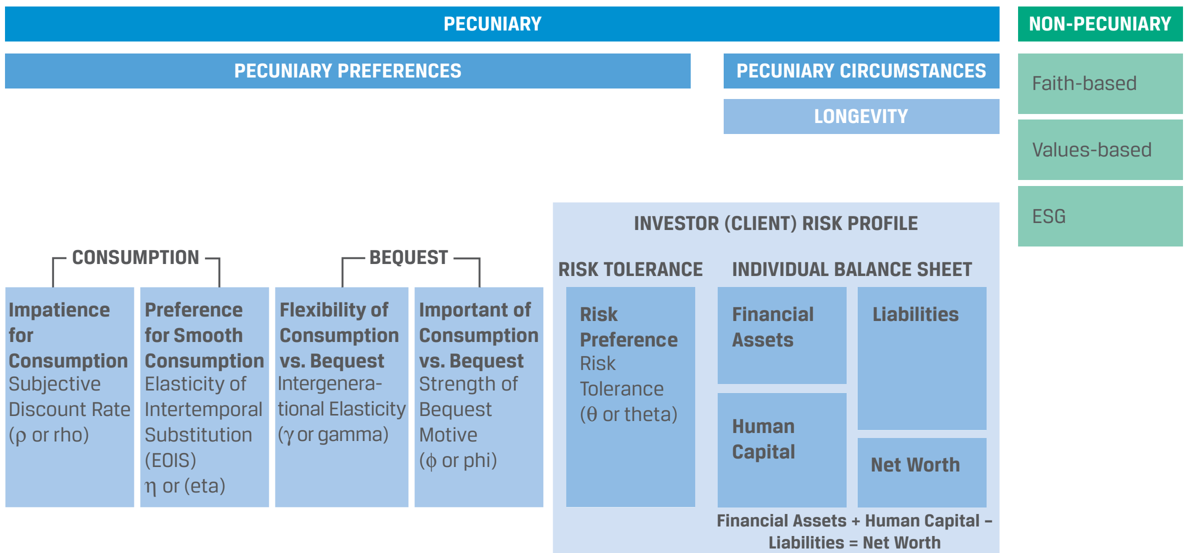
Ironically, depending on the jurisdiction in question, such suitability requirements can appear to require a process that, when viewed from the holistic perspective of what is truly best for an investor, would preclude the adviser from actually acting in the investor's best interest.

We move from a myopic investment-centric view that we might call "investor *risk* profiling" to a holistic, life-cycle finance view that we call an "investor profile" (hence the name of this chapter, "Holistic Investor Profiling"). Investor profiling and risk profiling are often thought of as the same thing, with the result that an investor profile frequently contains no more information than simply the client's risk profile. In other words, major parts of the holistic investor profile needed for life-cycle planning are missing.

In **Exhibit 2.1**, we begin with a high-level vision of the major pieces of a holistic investor profile. Notice that a holistic investor profile should include multiple types of pecuniary investor preferences as well as nonfinancial or nonpecuniary preferences. We elaborate on these different elements shortly.

The industry is largely focused on (investor) risk profiling and investor risk tolerance, which we see as two smaller segments of an overall investor profile as indicated in Exhibit 2.1.[5] Although we include risk tolerance and all of the elements of an individual balance sheet as part of a risk profile, all too often, practitioners and regulators think of risk profiling more narrowly. The box labeled "Risk Preference: Risk Tolerance (θ or theta)" might be thought of as "willingness" to take on risk, whereas "risk capacity" relates to the

Exhibit 2.1. A Holistic Investor Profile



[5]As an example, see Brayman et al. (2015), which was commissioned by the Ontario Securities Commission to study global best practices in risk profiling. The CFA Institute Research Foundation book *Risk Profiling and Tolerance: Insights for the Private Wealth Manager*, edited by Joachim Klement (2018), is dedicated to risk profiling within a wealth management setting, but does not mention life-cycle finance, other investor preferences, or a broader investor profile.

individual balance sheet and the investor's ability to withstand adverse outcomes if the risk happens. Collectively, these two factors form the essence of the investor's risk profile.

At times, a myopic approach may be desirable, but those times should be the exception rather than the rule. Good financial planning and life-cycle finance are holistic and go well beyond finding the right risk level for a single account among an investor's financial assets. In the rest of Part I, we develop life-cycle finance models. These models help us answer important financial planning questions, such as how much to save, how much to spend, how much to consume now versus later, how to invest, and how to properly plan for a bequest. To answer these questions in a way that is best for the investor given their personal characteristics, we need to know their pecuniary preferences. Risk tolerance is just one of these key investor preferences, but many of the life-cycle models that we present include up to four other key investor preferences:

1. Impatience for Consumption: Subjective Discount Rate (ρ or rho)
2. Preference for Smooth Consumption: Elasticity of Intertemporal Substitution (EOIS, η or eta)
3. Risk Tolerance (θ or theta)
4. Flexibility of Consumption versus Bequest: Intergenerational Elasticity (γ or gamma)
5. Importance of Consumption versus Bequest: Strength of Bequest Motive (ϕ or phi)

These financial preferences are included in Exhibit 2.1, and a detailed discussion is picked up in the next chapter. Nevertheless, assessing these preferences is what makes optimal planning possible.

Having emphasized that risk tolerance is only one of the five key pecuniary investor preferences, we ask: why are these other preferences not being measured and incorporated in financial planning? Determining how best to measure these other investor preferences is beyond the scope of this book, but we would think that many of the techniques used to estimate risk tolerance could be adapted to help measure these other preferences.

Sample Questions for Understanding Investor Pecuniary Preferences

In the hope of inspiring others, we suggest some potential questions that could be used to measure these other key investor pecuniary preferences. Admittedly, these are ripe for improvement and expansion.

Question 1 Example. Impatience for Consumption: Subjective Discount Rate (ρ or rho)

Imagine that you expect to retire in 20 years and that you will live for 20 years after retirement. Your total budget in "real" (today's) dollars, including labor income (generated by your human capital) and investment income (generated by your financial capital) is $100,000 per year for 40 years—that is, $4 million. (Note that this is different from having $4 million in assets now.) As shown in the following table, with Option A, you can choose to spend the same real amount before retirement and after retirement; with Option B you spend less prior to retirement and more after retirement, and with Option C you spend more before retirement and less after retirement.

	Annual Real PreRetirement Spending	Annual Real Spending in Retirement
Option A	$100,000	$100,000
Option B	$80,000	$120,000
Option C	$120,000	$80,000

Which option would you prefer?

- Option A
- Option B
- Option C

Question 2 Example. Preference for Smooth Consumption: EOIS (η or eta)

The following table contains three possible real (inflation-adjusted) consumption paths.

Options	Year 1	Year 2	Year 3	Year 4	Year 5	Average	Standard Deviation	Excess Income/ Standard Deviation
A	$100,000	$100,000	$100,000	$100,000	$100,000	$100,000	$0	N/A
B	$100,000	$110,000	$96,000	$110,000	$96,000	$102,400	$6,375	0.38
C	$100,000	$120,000	$90,000	$120,000	$90,000	$104,000	$13,565	0.12

Path A is stable. Relative to Path A, Path B is more volatile but on average is $2,400 higher. Path C is significantly more volatile than Path A and Path B. On average, Path C is $4,000 higher than Path A and $1,600 higher than Path B. Although the average consumption amount of Paths B and C are higher than Path A, notice that in some years, one must significantly reduce consumption. Which option would you prefer?

- Option A
- Option B
- Option C

Question 3 Example. Flexibility of Consumption versus Bequest: Intergenerational Elasticity (γ or gamma)

The following three possible scenarios relate to your standard of living in retirement and ability to leave a bequest to loved ones, charities, or causes:

Scenario 1: You prefer to maximize your standard of living knowing that you will not be able to leave a bequest.

Scenario 2: You prefer a frugal standard of living to maximize the size of bequest.

Scenario 3: You prefer a moderate standard of living and would like to plan to leave a moderate bequest.

As we move toward your preference, please note this is not about which specific scenario you prefer. Rather, which of the three options (A, B, or C) most aligns with your feelings regarding the three scenarios?

- Option A: I am indifferent to the three scenarios.
- Option B: I am more or less okay with two of the scenarios.
- Option C: I strongly prefer one of the scenarios over the other two.

Question 4 Example. Importance for Consumption versus Bequest: Strength of Bequest Motive (ϕ or phi)

Please select the scenario that most aligns with your preference related to your standard of living in retirement and ability to leave a bequest to loved ones, charities, or causes:

Scenario 1: You prefer to maximize your standard of living knowing that you will not be able to leave a bequest.

Scenario 2: You prefer a frugal standard of living in order to maximize the size of bequest.

Scenario 3: You prefer a moderate standard of living and would like to plan to leave a moderate bequest.

At the end of chapter 1, we introduced Isabela, a 25-year-old investor, working with Paula the planner. We assume that Paula's state-of-the-art financial planning system included a system for evaluating Isabela's financial preferences. For our purposes, we assume Isabela had the following responses: Question 1 – Option B, Question 2 – Option B, Question 3 – Option C, and Question 4 – Scenario 3. We also assume that the investor profiling system also helped assess Isabela's risk tolerance. **Exhibit 2.2** summarizes Isabela's *financial* preferences.

Exhibit 2.2. Financial Preferences for Life-Cycle Model: Isabela, Age 25

Financial Preferences	Qualitative Assessment	Numeric Input
Impatience for Consumption: Subjective Discount Rate (ρ or rho)	**Patient:** Isabela is patient, wanting to live somewhat frugally now in hopes of a higher living standard later in life.	2%
Preference for Smooth Consumption: EOIS (η or eta)	**Moderate:** Because she plans ahead, Isabela is willing to have moderate interruptions to her consumption.	50%
Risk Tolerance (θ or theta)	**Low:** Isabela has a somewhat low tolerance for risk.	35%
Flexibility of Consumption versus Bequest: Intergenerational Elasticity (γ or gamma)	**Low:** Isabela has low flexibility when it comes to her desire to have both a moderate standard of living and her ability to leave a bequest.	25%
Importance for Consumption versus Bequest: Strength of Bequest Motive (ϕ or phi)	**Moderate:** Isabela prefers a moderate standard of living and would like to plan to leave a moderate bequest.	1.5%

Later, we demonstrate how these parameters feed directly into various life-cycle finance models to provide personalized, optimal advice.

Moving from these five *pecuniary* preferences, a growing body of research indicates that many investors also have *nonpecuniary* or nonfinancial preferences. For example, some investors have values-based preferences and, as a result, attempt to avoid or minimize certain types of investments or exposures. From an industry perspective, this exclusionary-based approach is largely classified as socially responsible investing (SRI).

Some investors prefer to invest in firms or industries that they believe are making the world a better place, such as green energy firms, firms that are curing diseases, firms that promote equality (however defined), or firms with diverse independent boards. From an industry perspective, this is largely bucketed into what is called ESG investing. From this nonpecuniary perspective, ESG is about avoiding or minimizing exposure to disliked characteristics and seeking exposures to liked characteristics. We should mention that a group of *pecuniary* ESG investors believe, correctly or not, that the market does not properly price all relevant ESG information so they can earn higher returns through ESG-oriented active management; and others believe (this is a different matter) that, because ESG is good for the future of the world, it must also produce superior returns over any foreseeable period.

Long ago, the Nobel Prize–winning economist Milton Friedman expressed a view on what would later become known as SRI, but it also applies to ESG. In a famous *New York Times* op-ed titled, "A Friedman Doctrine: The Social Responsibility of Business Is to Increase Its Profits" (Friedman 1970), he said that expressing your values through your portfolio is inefficient and investors and companies should base their decisions purely on pecuniary considerations. Then, having received the best possible returns from a portfolio that is unencumbered by nonpecuniary pursuits, the investor can use those returns (if they so choose) to directly support charities, initiatives, and causes that are the most important to them.

Measuring Nonpecuniary Preferences

Just as a complete investor profile should incorporate all of the investor's pecuniary preferences, if an investor wants their nonpecuniary preferences incorporated into their portfolio, one must attempt to measure them too. In the spirit of the Friedman doctrine, the adviser should determine whether or not the investor wants their nonpecuniary preferences reflected in their portfolio at all. This involves a trade-off: all else being equal, is the investor willing to give up some amount of return to have their portfolio reflect their nonpecuniary preferences?

For some investors, the answer is "no" and we don't have to dig any deeper. For other investors, the answer is "yes" and we must determine (1) how strong their nonpecuniary preferences are (how much return they are willing to give up), (2) what their nonpecuniary preferences are, and (3) what are the relative strengths of each nonpecuniary preference relative to each of the other nonpecuniary preferences. For example, one might like diversity, love green energy, dislike tobacco, disdain handguns, and have no view on other issues.

Continuing to try to inspire our readers, we present some potential methods for measuring an investor's nonpecuniary preferences. Again, these are ripe for improvement and expansion.

The following is a sample question to determine the strength of an investor's nonpecuniary preferences:

All else being equal, how much of an annual return reduction would you be willing to accept to have a portfolio that fully aligns with your nonfinancial preferences:

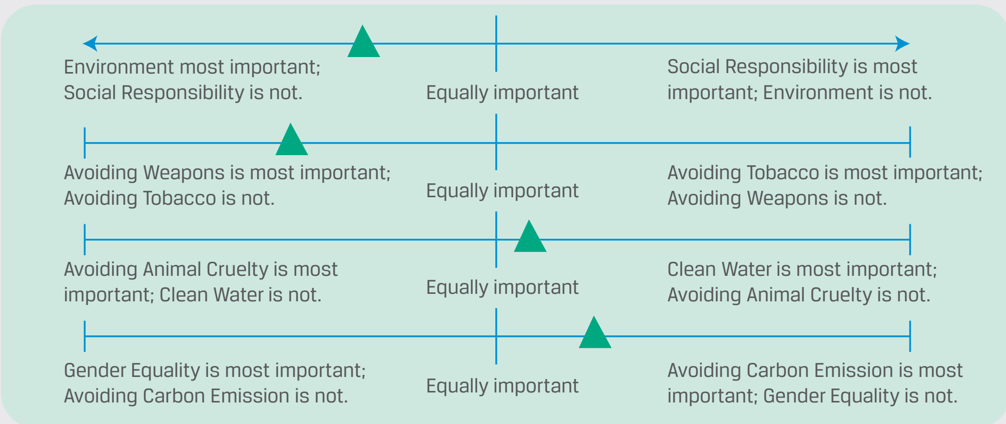
- None: I am not willing to sacrifice return
- 0.5% (e.g., a 7.5% return would be reduced to 7.0%)
- 1.0% (e.g., a 7.5% return would be reduced to 6.5%)

- 1.5% (e.g., a 7.5% return would be reduced to 6.0%)
- 2.0% (e.g., a 7.5% return would be reduced to 5.0%)

In our experience, some ESG advocates object to any sort of framing that says ESG can reduce returns. They are misinformed. As such, it may be necessary to recast this first question as a portfolio with a "customization fee" rather than a direct reduction in the return of the hypothetical portfolio's expected return.

Clearly, efficiently measuring an investor's nonpecuniary preferences requires a lot of work. One potential technique uses a series of sliders that start in a *neutral* position in which two items that an investor likes (or dislikes) are contrasted with one another. **Exhibit 2.3** provides an illustration of how one might do this.

Exhibit 2.3. Measuring Nonpecuniary Preferences



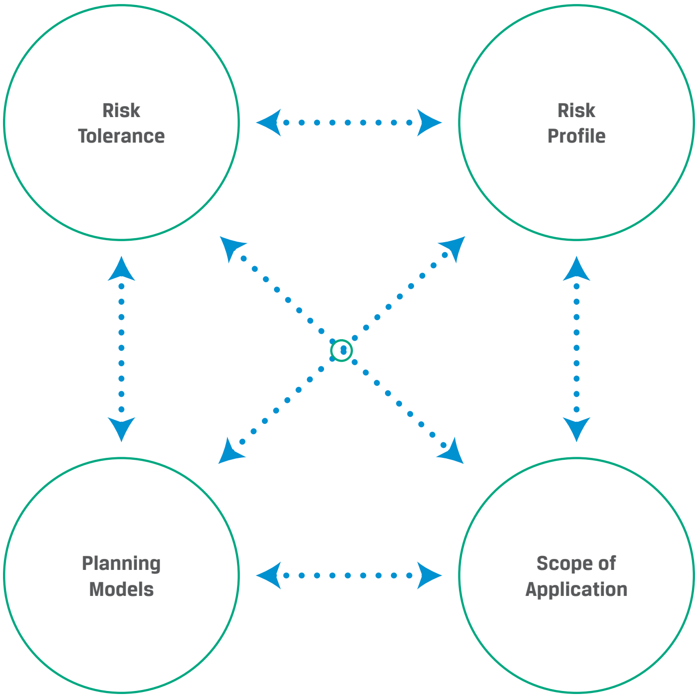
In chapters 9, 10, and 11, we demonstrate how nonpecuniary preferences can be directly incorporated into portfolio construction (expected portfolio utility problems).

Despite the strong logic of this view, a number of investors nevertheless want their nonpecuniary preferences reflected in their portfolios. In chapters 9, 10, and 11, we explain how to do this in an optimization framework. Here, just as we think it important to measure and understand the investor's pecuniary preferences, part of a complete investor profile involves an understanding of whether the investor has nonpecuniary preferences that need to be considered.

A Coherent System

Drilling down into the sub-elements of the overall investor profile, the notions of a risk profile and correspondingly of risk tolerance are central to financial planning as well as the single-period optimization

Exhibit 2.4. A Coherent Financial Planning System

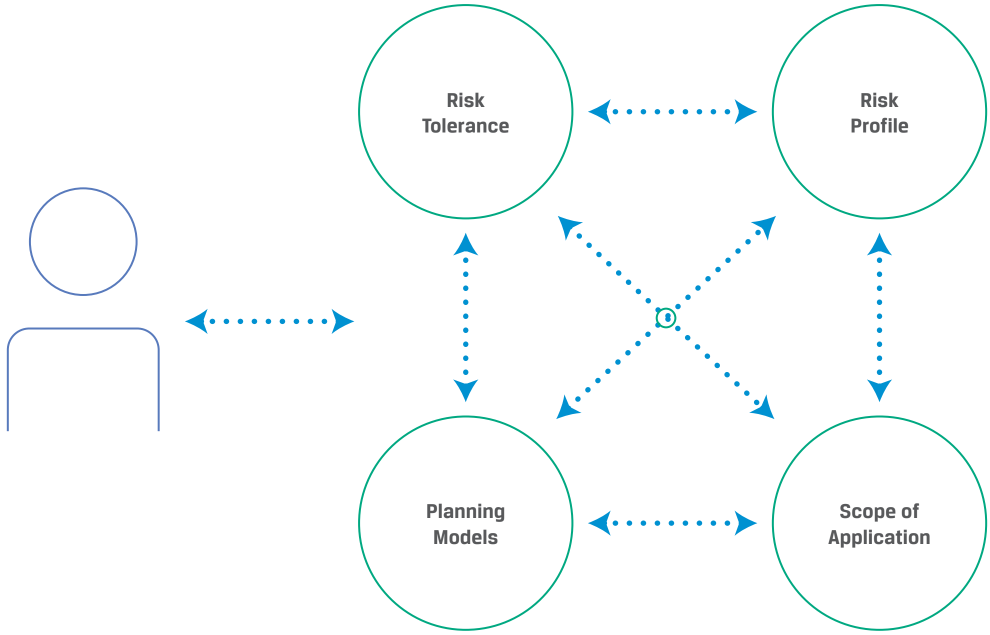


models, such as Markowitz's MVO, and the life-cycle models developed in chapters 5 and 6. A risk profile and risk tolerance assessment can and should go hand-in-hand, but depending on the context, they may or may not be the same thing. Even when they are different, only the most disciplined practitioners clearly distinguish between the two. One of our key goals in this chapter is to highlight the importance of explicitly defining these terms and the scope with which they are applied and to ensure that the definition and scope of application are consistent with each other as well as with the fundamental financial planning model in question.

Regardless of whether an adviser moves to a holistic investor profile, as illustrated in **Exhibit 2.4**, it is critical that (1) risk tolerance; (2) investor risk profiling; (3) the scope of application—for example, specific investment accounts versus total assets (financial assets plus human capital) versus net worth (total assets minus liabilities); and (4) the financial model in question work together in a coherent manner. We believe that financial planning is fraught with material inconsistencies, related to these four items, that are largely unrecognized by the industry. We might think of these as four critical trees that form a forest. Although each tree is important, one needs to be able to step back and see the entire forest and how the different elements of financial planning fit together to form a single system.

Individuals are complicated. Proper financial planning requires a detailed understanding of a wide variety of factors as depicted in **Exhibit 2.5**. Other factors may influence risk tolerance. A coherent financial planning process needs to disentangle these many factors into distinct parts or trees.

Exhibit 2.5. From Investor to Coherent Planning



Risk Tolerance

The term "risk tolerance" is so widely used that few of us bother to really think about what it really means, how it is measured, and how it is applied. These first-order questions are seldom pondered, and the various disconnects can lead to unrecognized mistakes in the creation and implementation of a financial plan.

Given that almost all financial professionals agree to some extent on the meaning of the term, it is hard to imagine that seemingly benign nuances of risk tolerance can lead to malignant results. But they can.

Following are just some of the issues that arise regarding risk tolerance measurement:

- When investors respond to questions from an RTQ, are those responses a pure reflection of their attitudes toward risk in isolation, or do some of those investors allow knowledge about their circumstances, time horizon, or financial goals, to influence their attitudes and responses?
- Conversely, does the scoring of responses (and the actions taken) from an RTQ assume the former or the latter?
- What did the investor assume? What did the RTQ provider assume? What did the planner assume? And, how is this information used in financial planning models?

Moving from the world of practitioners to that of economists, the answers to these questions can be equally unclear. When economists think about and model investors' attitudes toward risk, do they assume that risk tolerance is a pure reflection of the investors' attitudes toward risk in isolation, or are they assuming that investors' attitudes toward risk are fully or partially informed by the investors' circumstances, time horizon, or financial goals?

The idea of risk tolerance is prevalent in the worlds of both the practitioner and the economist, and it is foundational to financial planning and the models underpinning it. Nevertheless, the way that risk tolerance is defined, measured, and applied is inconsistent.

Two sources of confusion are as follows: (1) the term "risk tolerance" is used to describe two different things, and (2) the "scope" of the investor's wealth to which the risk tolerance is applied is often unclear.

Two Interpretations of "Risk Tolerance"

Let us begin with the dual meaning of risk tolerance. Starting with the first issue, in the narrow sense, risk tolerance is used to describe an *element* or *part* of the broader client risk-profiling process. This narrow definition of risk tolerance (Interpretation A in the following example) often feeds into a broader process that we call the "client risk profile." The broader client risk profile typically considers additional factors, such as time horizon, investment experience, goals, and risk capacity. Then, the output or bottom line of the broader client risk profile (Interpretation B in the following example) is often referred to as the client's "risk tolerance." Depending on the situation, the narrow and broad meanings of risk tolerance (Interpretations A and B) can be *opposite*.

An example illustrates. **Exhibit 2.6** summarizes two different interpretations of risk tolerance. Interpretation A, the independent isolationist view, narrowly defines risk tolerance as the investor's attitude toward risk and asserts that that attitude is independent of other seemingly meaningful factors, such as the investor's time horizon, risk capacity, and other assets. Interpretation B, the informed attitudes view, asserts that the investor's risk tolerance is informed by the investor's circumstances, such as time horizon, risk capacity, and other assets.

Exhibit 2.6. Two Interpretations of Risk Tolerance

	Interpretation A: Independent Isolationism (Independent Risk Attitude)	**Interpretation B: Informed Attitudes (Client Risk Profile)**
Interpretation	Investors' attitudes toward risk are independent of their circumstances, time horizon, risk capacity, market expectations, and goals.	Investors' attitudes toward risk are informed based on their circumstances, time horizon, risk capacity, market expectations, and goals.
Implication	Risk tolerance is an ingredient and must be jointly considered with other factors (e.g., circumstances, time horizon, risk capacity, market expectations, goals) when making financial decisions.	Risk tolerance is a summation informed by other factors (e.g., circumstances, time horizon, risk capacity, market expectations, goals) when making financial decisions.
Potential Implementation Challenge	In this path, if one fails to incorporate the other factors at some point into financial decisions, critical information has not been considered.	In this path, if one mixes risk tolerance with other factors when making a financial decision, a form of double counting occurs.

Our point is not that one interpretation is necessarily better than the other. Rather, one must decide which interpretation to use and then integrate that decision into the rest of the planning process. Many commercial risk tolerance assessment tools are unclear on this process and do not provide accurate descriptions of what they are really measuring, making the challenge for the financial adviser even bigger. If advisers are not clear, relevant information could be ignored or double counted.

Scope of Wealth to Which Risk Tolerance Is Applied

Moving to the second issue, the "scope" or breadth of the investor's wealth to which the risk tolerance is applied is often unclear. Should it be applied to each investment account? To all investment accounts combined? Does it apply to other financial assets, such as bank accounts, one's home, investment real estate, and other financial assets? Does an adviser consider human capital to be part of the investor's assets? Does risk tolerance apply to net worth, which is composed of all financial assets plus human capital minus liabilities?

In chapter 4, we explain how to calculate the various elements of the investor's balance sheet. For now, we skip over those details but move forward with key elements of that balance sheet. As illustrated in Exhibit 2.6, the left-hand side of the balance sheet contains the investor's assets, in which the two major groupings are financial assets and human capital. We elaborate on human capital in chapter 4, but it is basically the net present value of current and future labor income. For many investors, human capital behaves somewhat like a bond and its net present value often makes it the single largest asset of an investor.

The right-hand side of the balance sheet consists of the investor's liabilities or net present value of nondiscretionary consumption (expenditures for food, clothing, shelter, and healthcare); and their net worth, that which could be spent discretionarily or bequeathed.[6] The various red callout boxes shown in **Exhibit 2.7** highlight the different potential "scopes" to which one could apply a risk tolerance or risk profile measure.

When thinking about the scope of the application of risk aversion, both Merton (1969, 1971) and Samuelson (1969) have demonstrated that, based on an assumption of *constant relative risk aversion* (CRRA), the same fraction of total assets (both financial assets and human capital) should be allocated in combination to risky assets regardless of age.[7] The Samuelson (1969) and Merton (1969, 1971) models, which are similar, have become standard in the literature on rational lifetime financial planning.

This is not to suggest that time horizon is unimportant nor that the allocation to risky assets should never change. In life-cycle models, time horizon is part of the model affecting both human capital and liabilities and, if one takes a holistic view, changes in the amount of human capital drive changes in the way financial assets should be invested, even if the overall amount of risk at the net-worth level is constant. This perspective was made clear in Bodie, Merton, and Samuelson (1992) and in Ibbotson et al. (2007). As a practical matter, in many practitioner-oriented approaches and in life-cycle models, because of the time horizon, the amount allocated to risky assets tends to decrease with age, even if the details and perspective around why that occurs differ.

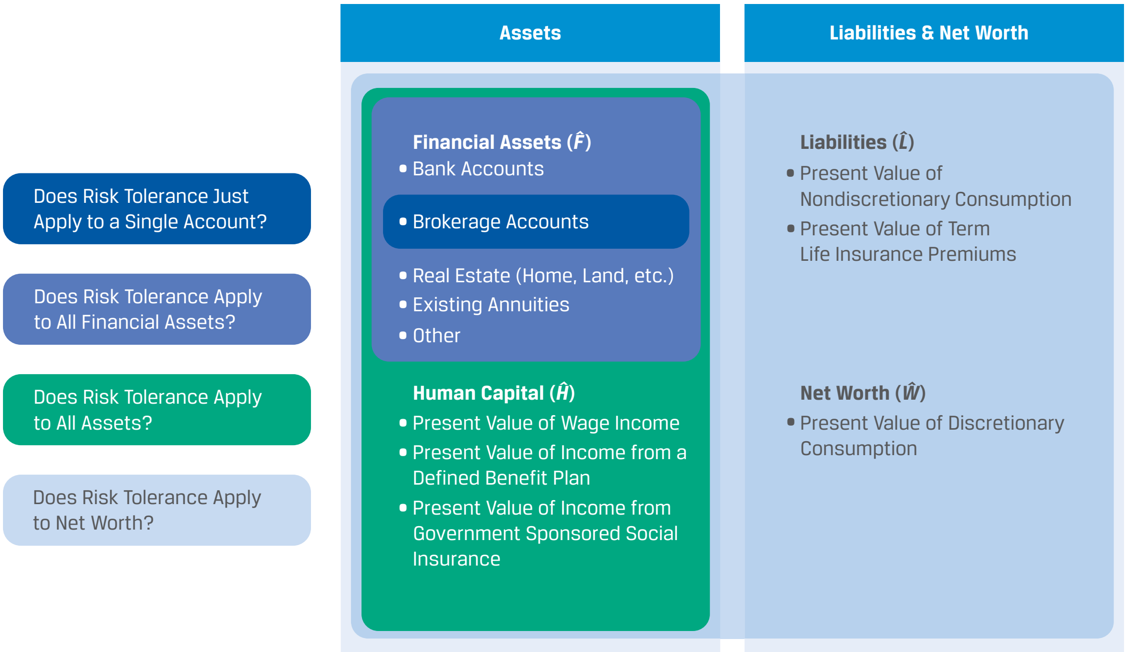
At a 2006 conference,[8] Paul Samuelson, the 1970 Nobel laureate in economic science, emphasized this key takeaway from his 1969 paper. In the life-cycle models advanced in this book, we go beyond the *total*

[6]Distinguishing between nondiscretionary and discretionary expenses is highly subjective. Where one precisely draws that distinction is beyond the scope of what we cover, although we encourage advisers to do this in a thoughtful and consistent manner.

[7]Financial economists often assume that investors have CRRA because it is a unitless measure that can be applied in various settings. Most importantly, in an investment setting, it leads to optimal portfolios that when are expressed in percentage terms, are independent of wealth, and thus serves as a justification for the common practice of expressing portfolios in percentage terms.

[8]Proceedings of the conference, which was on the future of life-cycle finance, are in a CFA Institute Research Foundation book (Bodie, McLeavey, and Siegel 2008).

Exhibit 2.7. Scope of the Application of Risk Tolerance



asset approach of the Merton and Samuelson models to focus on the investor's *entire balance sheet* and *net worth*.

We move to the work of another Nobel laureate in economic science, Daniel Kahneman, who has a similarly broad perspective. In a seminal piece with Mark Riepe (Kahneman and Riepe 1998) on matching investors to appropriate portfolios, they recommend the following:

> Encourage clients to adopt as broad a frame as possible when making investment decisions.
>
> When developing a client's investment policy, follow a top-down process, which accounts for all of the investor's objectives simultaneously. Avoid the common bottom-up approach in which a separate policy is established for each objective.

We interpret this statement as meaning that risk tolerance should be holistic and applied to the entire balance sheet and hence net worth. Then, following a holistic top-down perspective, the appropriate risk level for *financial accounts* depends on the investor's liabilities, the nature of their human capital, and any other existing assets that are deemed to be nontradeable. According to one perspective, the risk level of a portfolio of tradable financial assets becomes a "dial" of sorts that should be adjusted so that the investor's entire balance sheet reflects the investor's risk tolerance. As we shall see in chapters 8 and 11, once an appropriate risk level is determined for the financial assets in question, from a tax efficiency perspective, the policy asset allocation should be created and then implemented appropriately across accounts based on their tax treatment (this is called "asset location").

To demonstrate how "the scope of application" matters, we examine how different scopes would affect Isabela, our hypothetical investor.

Isabela, the Investor

At this point in time, we assume Isabela is 25 years old. A year earlier, Isabela received her master's degree in marine biology and started working at a large scientific-research-oriented aquarium earning $75,000 per year (after taxes). To encourage employees to save, Isabela's employer provides a 50% match up to a total *employer* contribution of $6,833 (this is an arbitrary number established by the employer). To take advantage of the maximum possible match of $6,833 Isabela chooses to contribute $13,667, even though this is more money than she needs to currently save.[9] After one year of working and contributing to her employer-sponsored defined contribution, tax-deferred retirement plan (coupled with matching employer funds), she had a tax-deferred account balance of $20,500. Contrary to best practices, Isabela's defined contribution plan defaults to a 100% allocation to a money market fund. Isabela also has $250,000 in a taxable brokerage account resulting from the sale of her grandmother's home that she (along with her two siblings) had inherited a year earlier when her 97-year-old grandmother passed away. Based on some online research and talking with friends, Isabela purchased a 60/40 balanced fund in her taxable brokerage account. Inheriting the $250,000, combined with uncertainty about how to invest across her accounts, was part of what motivated Isabela to seek out a financial planner, Paula.

In the upcoming chapters, we will get into the details, but Paula captures Isabela's financial information and preferences, all of which seamlessly feed into Paula's financial planning and ongoing investment management software system. Using equations from chapter 4, Paula's system estimated Isabela's human capital (the net present value of labor income, social insurance, and other labor-generated cash flows) at $2,767,869, her consumption-related liabilities (the net present value of nondiscretionary consumption) at $1,392,064, and the (net present) value of a series of annual term life insurance purchases at $220,087.

We elaborate more on this in chapter 4, but like many investors Isabela's risky human capital is deemed to be more bond-like than stock-like and is thus "modeled" as 20% stocks and 80% bonds. We also treat a component of her human capital as riskless. This brings it to a stock/bond/cash mix of 18.7/74.9/6.4. This asset mix provides the basis for discounting the future cash flows in the net present value calculation. Additionally, for holistic asset-allocation-balance-sheet purposes, Isabela's human capital of $2,767,689 is treated as $518,373 in stocks, $2,073,492 in bonds, and $175,824 in cash.

Moving to Isabela's consumption-related liabilities, they are deemed to be much more bond-like than stock-like and thus are "modeled" as 15% stocks, and 85% bonds. This 15/85/0 is the basis for discounting the future cash flows in the net present value calculation. Additionally, for holistic asset-allocation-balance-sheet purposes, Isabela's consumption-related liabilities are treated as $175,824 in equities and $996,180 in fixed income.

(The relevant equations are in chapter 4. "Mortality weighting" means adjusting future cash flows by the probability that the person will be alive to collect them and whether or not the person can annuitize.)

Asset Allocation Results Using Different "Scope" Interpretations

In **Exhibit 2.8**, we demonstrate how the assumed scope or application of a target 35% stock/65% bond asset allocation inferred from Isabela's risk tolerance of 35% can lead to different results.

---

[9]Without getting into too much detail, as of 2022, for individuals under 50 the US tax code allows an individual to contribute up to $20,500 with a total employee plus employer contribution of up to $61,000. Isabela could have contributed more but chose to stop once the match was maximized.

Exhibit 2.8. Impact on Overall Asset Allocation Based on the Scope of Application of Risk Tolerance

Financial Assets	Stocks	Bonds	Cash	Total
Panel A. Applied to Brokerage Account Only				
Brokerage Account	$87,500	$162,500	$0	$250,000
Retirement Account	$0	$0	$20,500	$20,500
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689
Total Assets	$605,873	$2,235,992	$196,324	$3,038,189
A. Brokerage Account Allocation (%)	35.0%	65.0%	0.0%	
B. All Financial Assets Allocation (%)	32.3%	60.1%	7.6%	
C. All Assets Allocation (%)	19.9%	73.6%	6.5%	
D. Net-Worth Allocation (%)	26.1%	75.3%	−1.4%	
Panel B. Applied to All Financial Assets				
Brokerage Account	$87,500	$162,500	$0	$250,000
Retirement Account	$7,175	$13,325	$0	$20,500
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689
Total Assets	$613,048	$2,249,317	$175,824	$3,038,189
A. Brokerage Account Allocation (%)	35.0%	65.0%	0.0%	
B. All Financial Assets Allocation (%)	35.0%	65.0%	0.0%	
C. All Assets Allocation (%)	20.2%	74.0%	5.8%	
D. Net-Worth Allocation (%)	26.6%	76.1%	−2.7%	
Panel C. Applied to All Assets				
Brokerage Account	$250,000	$0	$0	$250,000
Retirement Account	$20,500	$0	$0	$20,500
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689
Total Assets	$788,873	$2,073,492	$175,824	$3,038,189
A. Brokerage Account Allocation (%)	100.0%	0.0%	0.0%	
B. All Financial Assets Allocation (%)	100.0%	0.0%	0.0%	
C. All Assets Allocation (%)	26.0%	68.2%	5.8%	
D. Net-Worth Allocation (%)	37.2%	65.4%	−2.7%	

(*continued*)

Exhibit 2.8. Impact on Overall Asset Allocation Based on the Scope of Application of Risk Tolerance (*continued*)

Financial Assets	Stocks	Bonds	Cash	Total
Panel D. Applied to Net Worth				
Brokerage Account	$216,250	$33,750	$0	$250,000
Retirement Account	$17,733	$2,768	$0	$20,500
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689
Total Assets	$752,356	$2,110,010	$175,824	$3,038,189
A. Brokerage Account Allocation (%)	86.5%	13.5%	0.0%	
B. All Financial Assets Allocation (%)	86.5%	13.5%	0.0%	
C. All Assets Allocation (%)	24.8%	69.4%	5.8%	
D. Net-Worth Allocation (%)	35.0%	67.7%	−2.7%	

In the top panel, Panel A, the 35/65/0 risk tolerance-based allocation target is applied only to the *brokerage* account resulting in $87,500 or 35% allocated to stocks, and $162,500 or 65% allocated to bonds. In the bottom section of Panel A, we identify the implied or inferred asset allocations corresponding to the scope of risk tolerance application in the given panel.

Moving to Panel B, the 35/65/0 risk tolerance-based target is applied to both the *brokerage account* and the *retirement account* (i.e., total financial assets). Summing the brokerage and retirement account allocations, we have $94,675 or 35% allocated to stocks, and $175,825 or 65% allocated to bonds. Each of the two financial asset accounts has this allocation. As we did in the bottom section of Panel A, in the bottom section of Panel B, we identify the implied or inferred asset allocations corresponding to the scope of risk.

Moving to Panel C, we attempt to apply the 35/65/0 risk tolerance-based target to all asset (i.e., financial assets plus human capital). In this case, because of a large amount of bond-centric human capital, the best one can do is to allocate 100% of the financial assets to stocks. It is not possible to achieve a total assets allocation of 35% stocks/65% bonds.

The most theoretically sound approach is to apply risk tolerance to the entire balance sheet, in which the investor's net worth corresponds to the 35/65/0 target. This is shown in Panel D. As one can see at the bottom of Panel D, it is possible to allocate Isabela's financial assets in such a way that her net-worth allocation more or less matches the 35/65/0 target.

Adding the Liability (Right-Hand Side of the Balance Sheet)

We now add the liability. The liability (net present value of nondiscretionary consumption and life insurance) is $1,392,064, resulting in a net worth of $1,646,126.

In Exhibit 2.8 we expand on Exhibit 2.7, showing the right-hand side of the investor's balance sheet. Notice that stock/bond split percentages associated with the left-hand side of the balance sheet can vary significantly from the implied stock/bond/cash split percentages associated with net worth. The stock/bond/cash split associated with net worth, which is calculated by subtracting the present value of liabilities from total assets, differs from the split for total assets in Exhibit 2.7 because of the effect of the liabilities.

It is not until we reach Panel D of Exhibit 2.8 that *financial* wealth is organized in such a way that it nearly achieves a split of 35/65/0 in net worth, the *holistic* definition of wealth.[10] Unfortunately, most advisers and investors do not use this holistic view and application of risk tolerance. We recommend that they do so.

Critically, it is only by observing the holistic balance sheet view that we begin to have a complete understanding of the investor's overall economic well-being, and thereby, know-the-client! Through this balance sheet lens, we can begin to understand the magnitude and asset class characteristics of the investor's human capital and nondiscretionary consumption and how they relate to the overall picture.

The degree to which the systematic asset classlike cash flow characteristics of human capital are significantly helping to offset or defease the somewhat-similar systematic asset class-like cash flow characteristics of the liabilities, is extremely important. Human capital is clearly funding or offsetting discretionary consumption; however, the power of human capital is somewhat masked within the net-worth rows of **Exhibit 2.9**. Note also that there is more to the story than systematic characteristics, especially as it pertains to the liabilities. Although liabilities do have systematic characteristics, they also have large amounts of nonsystematic or idiosyncratic risk. One can attempt to manage some nonsystematic liability risk through health insurance, car insurance, or home insurance.

An important takeaway from both Exhibit 2.8 and Exhibit 2.9 is that the application of risk tolerance to a subset of the balance sheet can dramatically alter the investor's asset allocation at the holistic net-worth level. In other words, knowing an investor's risk tolerance is inadequate without the ability to estimate the composition of the investor's balance sheet.

All of this analysis leads us to call on the industry to develop and adopt an expanded tool kit. Financial planners and advisers need a "holistic balance sheet estimator" that enables them to apply risk tolerance to the individual's net worth and thereby calculate the appropriate risk level for the part of the investor's portfolio that is under advisement.

[10]Coincidentally, the investor's asset allocation of financial assets is almost identical in Panel A and Panel D.

Exhibit 2.9. Holistic Balance Sheet View of Net Worth

Financial Assets	Stocks	Bonds	Cash	Total	Liabilities	Stocks	Bonds	Cash	Total
Panel A. Applied to Brokerage Account Only									
Brokerage Account	$87,500	$162,500	$0	$250,000	PV of Nondiscretionary Consumption	$175,797	$996,180	$0	$1,171,977
Retirement Account	$0	$0	$20,500	$20,500	PV of Term Life Insurance for Bequest	$0	$0	$220,087	$220,087
					Net Worth				
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689	PV of Discretionary Spending	$430,077	$1,239,812	($23,763)	$1,646,126
					Total (in Percent of Net Worth)	26.1%	75.3%	−1.4%	100.0%
Total Assets	$605,873	$2,235,992	$196,324	$3,038,189					
A. Brokerage Account Allocation (%)	35.0%	65.0%	0.0%						
B. All Financial Assets Allocation (%)	32.3%	60.1%	7.6%						
C. All Assets Allocation (%)	19.9%	73.6%	6.5%						
D. Net-Worth Allocation (%)	26.1%	75.3%	−1.4%						

(*continued*)

Financial Assets	Stocks	Bonds	Cash	Total	Liabilities	Stocks	Bonds	Cash	Total
Panel B. Applied to All Financial Assets									
Brokerage Account	$87,500	$162,500	$0	$250,000	PV of Nondiscretionary Consumption	$175,797	$996,180	$0	$1,171,977
Retirement Account	$7,175	$13,325	$0	$20,500	PV of Term Life Insurance for Bequest	$0	$0	$220,087	$220,087
					Net Worth				
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689	PV of Discretionary Spending	$437,252	$1,253,137	($44,263)	$1,646,126
					Total (in Percent of Net Worth)	26.6%	76.1%	–2.7%	100.0%
Total Assets	$613,048	$2,249,317	$175,824	$3,038,189					
A. Brokerage Account Allocation (%)	35.0%	65.0%	0.0%						
B. All Financial Assets Allocation (%)	35.0%	65.0%	0.0%						
C. All Assets Allocation (%)	20.2%	74.0%	5.8%						
D. Net-Worth Allocation (%)	26.6%	76.1%	–2.7%						

(*continued*)

Exhibit 2.9. Holistic Balance Sheet View of Net Worth (*continued*)

Financial Assets	Stocks	Bonds	Cash	Total	Liabilities	Stocks	Bonds	Cash	Total
Panel C. Applied to All Assets									
Brokerage Account	$250,000	$0	$0	$250,000	PV of Nondiscretionary Consumption	$175,797	$996,180	$0	$1,171,977
Retirement Account	$20,500	$0	$0	$20,500	PV of Term Life Insurance for Bequest	$0	$0	$220,087	$220,087
					Net Worth				
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689	PV of Discretionary Spending	$613,077	$1,077,312	($44,263)	$1,646,126
					Total (in Percent of Net Worth)	37.2%	65.4%	–2.7%	100.0%
Total Assets	$788,873	$2,073,492	$175,824	$3,038,189					
A. Brokerage Account Allocation (%)	100.0%	0.0%	0.0%						
B. All Financial Assets Allocation (%)	100.0%	0.0%	0.0%						
C. All Assets Allocation (%)	26.0%	68.2%	5.8%						
D. Net-Worth Allocation (%)	37.2%	65.4%	–2.7%						

(*continued*)

Financial Assets	Stocks	Bonds	Cash	Total	Liabilities	Stocks	Bonds	Cash	Total
Panel D. Applied to Net Worth									
Brokerage Account	$216,250	$33,750	$0	$250,000	PV of Nondiscretionary Consumption	$175,797	$996,180	$0	$1,171,977
Retirement Account	$17,733	$2,768	$0	$20,500	PV of Term Life Insurance for Bequest	$0	$0	$220,087	$220,087
					Net Worth				
Human Capital	$518,373	$2,073,492	$175,824	$2,767,689	PV of Discretionary Spending	$576,559	$1,113,829	($44,263)	$1,646,126
					Total (in Percent of Net Worth)	35.0%	67.7%	−2.7%	100.0%
Total Assets	$752,356	$2,110,010	$175,824	$3,038,189					
A. Brokerage Account Allocation (%)	86.5%	13.5%	0.0%						
B. All Financial Assets Allocation (%)	86.5%	13.5%	0.0%						
C. All Assets Allocation (%)	24.8%	69.4%	5.8%						
D. Net-Worth Allocation (%)	35.0%	67.7%	−2.7%						

A Vision for an Individual Economic Balance Sheet Estimator

For financial modelers, estimating an economic balance sheet for an investor is a relatively straightforward exercise. It requires a handful of inputs and several assumptions. The most difficult part is to calculate the net present value of human capital, which we explain in chapter 5.

Exhibit 2.10 contains an image of a spreadsheet workbook demonstrating a basic implementation. The various input cells are in orange. After completing the inputs, the green output cells on the right populate. The spreadsheet file is available as part of the supplementary materials associated with this book. Our hope is that the creators of a financial planning software solution would create more sophisticated versions of this example, applying the holistic financial planning concept.

Exhibit 2.10. Workbook Template for Estimating an Individual's Balance Sheet

Inputs:	
Target Investor Risk Level for Net Worth	35%
Balance in Taxable Account	$250,000
Balance in Tax-Advantaged Account	$20,500
Current Age	25
Retirement Age	65
Average Annual Income in $	$75,000
Expected Income from SS + DB in $	$25,000
Gender	F
Life Expectancy	94
Nondiscretionary Annual Spending	$ 40,000
Bequest Amount	$ 1,000,000

Outputs: Stocks	Bonds	Cash	
36%	64%	0%	<-Target Allocation for Net Worth
93%	0%	0%	<-Target Allocation in Taxable Account
0%	4%	3%	<-Target Allocation in Tax-Advantaged Account
93%	4%	3%	<-Overall Target Allocation for Financial Assets

Individual Balance Sheet

Asset	Stocks	Bonds	Cash	**Liabilities**	Stocks	Bonds	Cash
Taxable Account	$251,565	$0	$0	PV of Nondisc. Consump.	$175,400	$1,031,416	$0
Tax-Advantaged Account	$0	$10,820	$8,115	PV of Life Insurance Bequest	$0	$0	$185,248
Human Capital in $	$517,558	$2,072,999	$177,132	**Net Worth**	$593,723	$1,052,403	$0
Total:	$769,123	$2,083,819	$185,247	Total:	$769,123	$2,083,819	$185,247

In this example, notice that the first orange input is the target risk level for the individual's net worth. Then, using a bit of algebra, the spreadsheet estimates the target the target equity percentage for the advisable portfolio, that is, that part of the investor's net worth that is under the control of the adviser.

Conclusion and Key Takeaways

- The most holistic view of an investor is an investor profile. A risk profile is one part of an investor profile. Risk tolerance is one part of a risk profile.
- The industry is largely focused on risk profiling and risk tolerance.
- Life-cycle finance models go well beyond risk tolerance, to include important investor preferences, which are both pecuniary and, for some, nonpecuniary. The time has come to measure these additional investor preferences and to incorporate them into holistic financial planning. We provide examples of how one might begin to access these additional preferences.
- Despite the considerable focus on risk profiling and risk tolerance, confusion related to the two and the scope with which they are applied can lead to dramatically different asset allocations. Regardless of the scope and interpretation of risk tolerance, the implications for the investor can only truly be seen in the context of the investor's total balance sheet and net worth.
- Although we are advocates of a holistic approach, regardless of the scope of application, practitioners need to be cognizant of the investor's holistic situation to ensure a coherent financial planning process. We fear that this holistic and important view is largely missing from common financial planning practice.
- Practitioners need tools that enable them to generate an investor's economic balance sheet. That individual-specific economic balance sheet should be at the center of the financial planning process.

3. THE INVESTOR'S FINANCIAL PREFERENCES, SURVIVAL PROBABILITIES, AND NEEDS

Context

This chapter, as well as chapter 4, develops the background and foundation for the life-cycle models that are forthcoming in chapters 5 and 6. This and the next three chapters (chapters 4 through 6) focus on the investor from a *pecuniary* or *financial* perspective. In chapter 9, we expand our models to incorporate the investor's *nonpecuniary* or *nonfinancial* preferences.

Key Insights

- This chapter and chapter 4 introduce core ingredients and frameworks that will culminate with life-cycle models in chapters 5 and 6, all of which stem from the core economic theories that are the basis for optimal lifetime financial decision making.
- The theory of rational decision making takes the *form* of utility theory and the *idea* of utility maximization.
- Utility theory is broad and can be applied to lifetime financial planning in a variety of ways, including deciding how much and when to consume and save, how large a bequest to leave, how much and when to purchase life insurance and annuities, and how to invest taxable and tax-advantaged assets.
- In the life-cycle models that we are working toward, five key *pecuniary* investor preferences enable life-cycle models to tailor or personalize the advice by finding the utility-maximizing solution for a given investor.
- In addition to the five key pecuniary preferences, a key factor that goes into optimal lifetime financial planning is the *probability that the investor will survive* to each possible age in the future.

In this chapter, we introduce the idea of optimal decision making and other key concepts that lead to life-cycle models in later chapters. We introduce the five key pecuniary investor preferences that enable life-cycle models to tailor or personalize the advice by finding the utility maximizing solution for a given investor. In this chapter and the next, we lay the groundwork for the life-cycle models that we present in chapters 5 and 6.

The Foundation of Optimal Decision Making

People make numerous economic decisions over the course of their lifetimes. These decisions always involve trade-offs. The decisions that one makes today influence what may or may not be feasible in the future, including one's ability to leave money to family, friends, or causes. Although many of these decisions may seem to lack cohesion, economists have developed a model in which all of a person's economic decisions *should* flow from a single logical calculation—that is, *utility maximization*. Ultimately, the class of models that we develop in this book are *life-cycle utility maximization models*. These are models of pure rational behavior for which mainstream economics is sometimes criticized; this critique led to the development of behavioral economics. Nevertheless, we believe that utility-based models are an excellent starting point for financial planning because they can be used to help investors improve their economic decisions.

We are not alone in our view. Angus S. Deaton (2005, p. 18), who later was named Nobel laureate in 2015, directly addressed this behavioral-based concern, writing,

> Even if behavioral economics manages to replace the lifecycle theory in providing a successful empirical description of the way that people actually behave—and it is still someway from having achieved that aim—the life-cycle model will still be the baseline to which people aspire. The role of behavioral perspectives is to help make people better-off by making life-cycle behavior a better description of behavior. Perhaps we are witnessing the movement of Modigliani's life-cycle hypothesis from a positive to a normative theory, away from description and towards prescription.

At the heart of the *theory of rational decision making* (under uncertainty) is the assumption that all of the person's preferences related to a complicated set of interdependent trade-offs can be captured by a utility function. The term "utility" refers to the happiness or satisfaction that a person enjoys as a result of consuming goods or services. A *utility function* is a mathematical formula (or set of formulas) that provides an *internally consistent* ranking on all possible bundles of the goods and services that a person consumes or uses or gives away. The multidimensional baseball player ranking system developed by Billy Beane and Bill James (the statistician) chronicled in the book (Lewis 2004) and movie *Moneyball* is an example of an applied utility function. In economic theory, a ranking is internally consistent if the following conditions hold:

1. *Completeness*.[11] Given two bundles of goods and services, Consumption Bundle A and Consumption Bundle B, one and only one of the following must be true:
    - **a.** The person strictly prefers A to B.
    - **b.** The person strictly prefers B to A.
    - **c.** The person is indifferent between A and B.
2. *Transitivity*. In the case of three consumption bundles A, B, and C, if the person prefers A to B and B to C, then the person prefers to A to C.
3. *Nonsatiation*. Holding all else equal, the person prefers more of a single good or service to less of it.

Importantly, if these conditions hold, there is *always* a utility function that results in the same ranking as an investor's preferences. Furthermore, we usually assume that the rate of increase in utility decreases as the amount of a good or service increases. This is called *diminishing marginal utility*.

Given behavioral finance's focus on what is often suboptimal decision making, where appropriate, we address the behavioral perspective on our approach. Note that some behavioral economists believe that one or more of these conditions are contrary to observed human behavior and therefore models that assume these conditions may be of questionable value. Some of them have developed utility functions that while leading to rankings that are internally inconsistent, are more in line with how people actually make decisions. We will discuss some of the implications of behavioral economics and finance for financial planning in chapters 9, 10, and 11. In this and most other chapters, however, we assume fully rational behavior because we believe that this leads to the best financial advice.

An Illustration of Utility Theory and Utility Maximization

To help understand utility theory, imagine that you are at a party where only two items are being served: beer and pizza. When you arrive, you first have a glass of beer, as that gives you the highest possible immediate satisfaction. Then, you have a slice of pizza, bringing you to the next level of satisfaction. Then

[11]Later, when we move from general utility theory to expected utility theory, Completeness and Transitivity become Axiom 1 and Axiom 2.

another slice of pizza, then another glass of beer, and then another slice of pizza, one more glass of beer, and then one more slice of pizza before you go. From these selections, we can infer the ranking of bundles of beer and pizza shown in **Exhibit 3.1**.

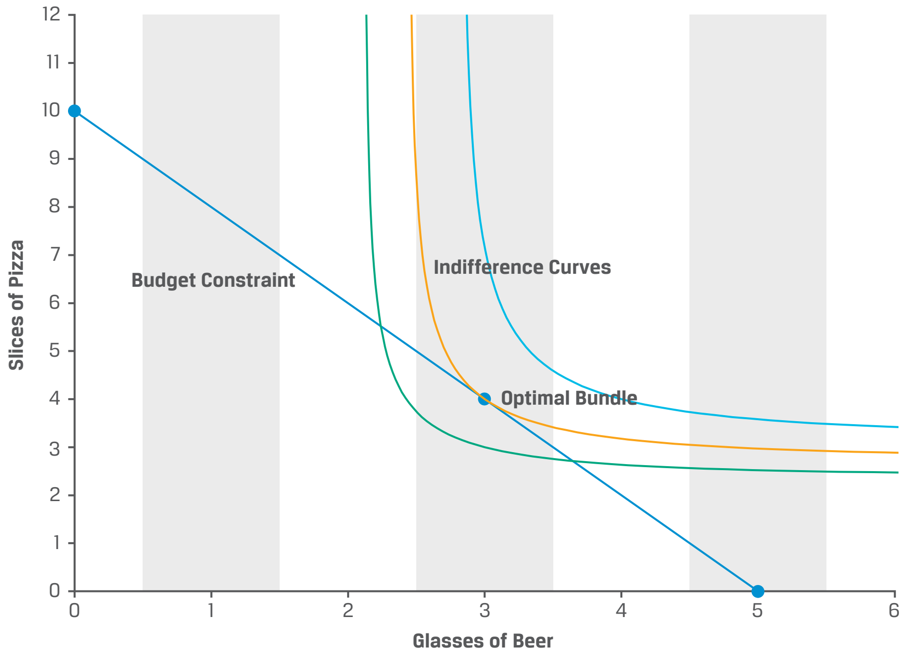
From a diminishing marginal utility perspective, presumably the first glass of beer resulted in the largest absolute increase in utility from a starting utility of zero. Next, the second consumable, the first slice of pizza in this example, increased absolute utility by an amount that is less than the first consumable. Then, the third consumable, the second slice of pizza, increased absolute utility by an amount that is less than the second consumable, and so forth. That is, although each additional consumable increased utility from an absolute perspective, there is diminishing marginal utility.

Exhibit 3.1. Ranked Preferences for Party Attendee 1

Rank	Glasses of Beer	Slices of Pizza
1	1	0
2	1	1
3	1	2
4	2	2
5	2	3
6	3	3
7	3	4

To illustrate how a utility function provides a ranking of bundles of goods and services (consumption bundles), economists have developed a graphic device called *indifference curves*. An indifference curve is made up of all bundles of two goods or services that have the same level of utility for a given investor based on the investor's preferences. Continuing our example of beer and pizza, **Exhibit 3.2** shows three indifference curves for combinations of these two goods for Party Attendee 1. For any given indifference curve, all points above the curve represent bundles that the person prefers to the bundles on the curve, and all points below the curve represent bundles that are less preferred.

Exhibit 3.2. Finding the Optimal Bundle of Goods for Party Attendee 1



To illustrate the economic application of utility functions, suppose that to defray costs, the host of the party charges $4 for a glass of beer and $2 for a slice of pizza. You come to the party with $20. How many glasses or beer and how many slices of pizza do you buy? Importantly, the optimal answer may be different for different party attendees (even if they all have $20). One party attendee may be super hungry, another party attendee may be thirsty, another may not drink alcohol, and yet another may be allergic to cheese. The utility maximizing solution for one attendee may not be the same as others—utility is personalized!

To answer this question for Party Attendee 1, whose ranks are depicted in Exhibit 3.1, we introduce a *budget constraint* in Exhibit 3.2. In the case of only two goods or services, a budget constraint is a line that shows all combinations of the two goods or services that can be purchased with the money that is available. The idea of a budget constraint is a critical element of life-cycle finance. Economic theory concludes that each person uses all available money because not to do so would be suboptimal.

In our beer and pizza example, on the horizonal axis of Exhibit 3.2, we see that $20 buys five glasses of beer. On the vertical axis, we see that the same amount buys 10 slices of pizza. Notice that the blue indifference curve represents bundles of pizza and beer (consumption bundles) that are *infeasible* given Party Attendee 1's budget constraint of $20. Conversely, the green indifference curve is not optimal because Party Attendee 1 could consume more. The actual bundle that will be purchased is that for which the budget constraint is tangent to an indifference curve (the orange one) because this is the bundle that maximizes utility. We label this point the Optimal Bundle for Party Attendee 1.

To reinforce the notion that different people derive utility in different ways, **Exhibit 3.3** shows the ranked preference for Party Attendee 2, who is extremely hungry and does not want to drink too much.

Using the same $20 budget constraint, **Exhibit 3.4** shows the indifference curves and optimal bundle for Party Attendee 2. As before, the blue indifference curve is not feasible given the budget constraints and the green indifference curve is not optimal because Party Attendee 2 could consume more given their $20 budget. The optimal bundle is that for which the orange indifference curve is tangent to the budget constraint. Notice that the budget constraint line is the same for both party attendees given that they both have $20.

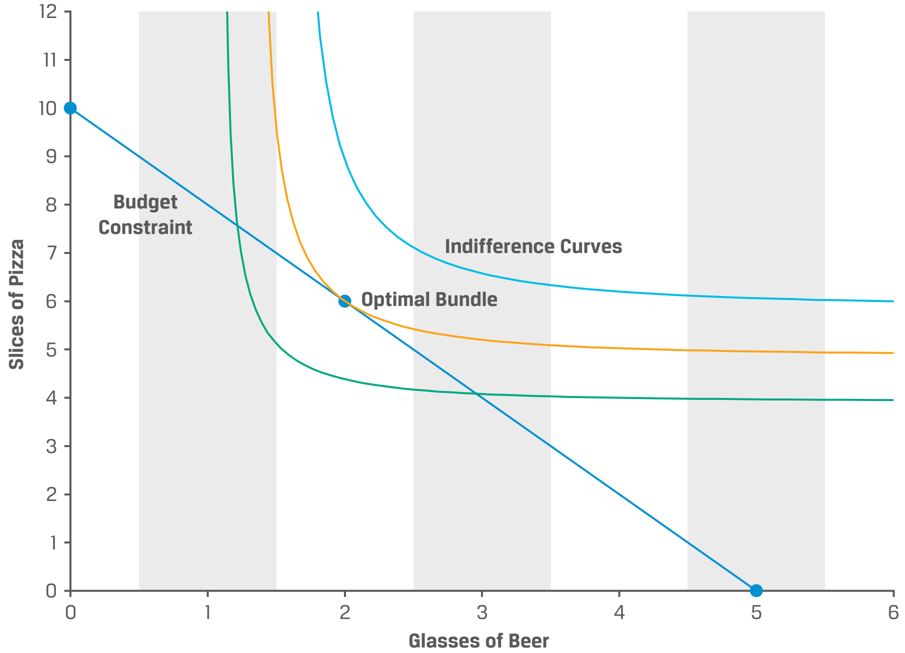
Both party attendees maximize utility; however, the optimal, utility-maximizing consumption bundle is different for each attendee. The optimal bundle for Party Attendee 1 consists of four slices of pizza and three glasses of beer. The optimal bundle for Party Attendee 2 consists of six slices of pizza and two glasses of beer.

We could certainly expand the complexity of this party problem. What if there were actually *two* parties, one Friday night and one Saturday night, both of which were selling pizza and beer. This complicates the problem and adds a *timing* or *temporal* element to the problem. A given person may prefer to evenly spread their consumption across the two nights, consume more Friday night, or delay consumption consuming more on Saturday night. What if the person wanted to bring home a combination of pizza and beer for a

Exhibit 3.3. Ranked Preferences for Party Attendee 2

Rank	Glasses of Beer	Slices of Pizza
1	0	1
2	0	2
3	0	3
4	1	3
5	1	4
6	1	5
7	2	5
8	2	6

Exhibit 3.4. Finding the Optimal Bundle of Goods for Party Attendee 2



friend that was unable to attend the parties? This adds something analogous to a *bequest* element to the problem.

Of course, financial planning deals with much larger economic issues than finding the right combination of beer and pizza to consume across two different parties. But the principles of rational decision making and the desire to make the optimal decisions to maximize a given investor's utility, as illustrated in Exhibits 3.1 and 3.2, can be used to make many key lifetime financial decisions, including the following:

- How much to save each year before retirement?
- How much to spend each year during retirement?
- How to invest before and during retirement?
- How much life insurance to have before retirement?
- How much to have in annuities during retirement?

Based on the foundation presented in this chapter and the next chapter, the life-cycle models of chapters 5 and 6 will answer these questions.

Importantly, instead of applying utility theory to specific goods and services, in life-cycle finance, we apply it to real (inflation-adjusted) dollar amounts at specific times and under specific market conditions. Hence, a bundle could consist of spending *x* dollars this year, *y* dollars next year if the stock market is up, and *z* dollars next year if the market is down. We refer to each real dollar spending amount as consumption at a specific time (e.g., year 1, year 2, and so on) and specific condition (e.g., up market, flat market, down market).

By casting the utility optimization this way (and making some additional assumptions that we shall discuss), we can boil down the investor's key pecuniary preferences to the five key preference parameters that we briefly introduced in chapter 2. We have attempted to adopt intuitive, descriptive names for these preference parameters, but we also include their common name from the economics literature:

1. Impatience for Consumption: Subjective Discount Rate (ρ or rho)
2. Preference for Smooth Consumption: EOIS (η or eta)
3. Risk Tolerance (θ or theta)
4. Flexibility of Consumption versus Bequest: Intergenerational Elasticity (γ or gamma)
5. Importance for Consumption versus Bequest: Strength of Bequest Motive (ϕ or phi)

Given the importance of these parameters and their heavy use throughout this book, **Exhibit 3.5** provides a glossary of sorts that identifies the common Greek symbol that we will use in various equations, the most common name from the economics literature, a brief description, the more intuitive name we have adopted for each parameter, and the primary practical impact on advice of a ceteris paribus increase in the parameter. We will discuss each of these parameters in more detail, although various callouts are necessary to weave together the various concepts into a cohesive whole.

As we discuss later in this chapter, the first two parameters, *impatience for consumption* and *preference for smooth real consumption*, together largely determine the pattern of consumption over time. We demonstrate how to incorporate these two parameters into a basic life-cycle model that considers the timing of

Exhibit 3.5. Glossary of Five Key Life-Cycle Model Parameters

Symbol	Preference Parameter	Description	Intuitive Name	Ceteris Paribus, Impact of Increasing Parameter
ρ rho	Subjective Discount Rate	Preference to consume now versus later.	Impatience for Consumption	Higher values indicate a stronger preference for consumption today versus in the future.
η eta	EOIS	Preference for smooth consumption from one period to the next.	Preference for Smooth Consumption	Higher values indicate more flexibility and a lower preference for smooth consumption.
θ theta	Risk Tolerance	Investor's attitude towards risk.[12]	Risk Tolerance	Higher values indicate a greater wiliness to take risk.
γ gamma	Intergenerational Elasticity	Flexibility in choosing between consumption and bequest.	Flexibility of Consumption versus Bequest	Higher values lead to higher sensitivity of size of bequest to strength of bequest motive.
ϕ phi	Strength of Bequest Motive	Importance of consumption versus importance of bequest.	Importance of Consumption versus Bequest	Higher values indicate a greater preference for bequests.

[12]Risk here is in reference to net worth as we discuss in chapters 4, 6, and 8. In chapters 9, 10, and 11, risk is in reference to financial assets.

consumption. Likewise, the final two parameters, *flexibility of consumption versus bequest* and *importance of consumption versus bequest*, together largely determine the size of the investor's bequest. We demonstrate how to include them in a life-cycle model. In chapter 2, in the spirit of an RTQ, we introduced possible questions to assess an investor's preference. As we introduce these various preference parameters that go into the life-cycle models in chapters 5 and 6, along the way, we will present several necessary callouts.

Impatience for Consumption: The Subjective Discount Rate (ρ or rho)

The "subjective discount rate," as it is called in economics, measures the degree to which the investor is impatient, preferring immediate or near-term consumption relative to future consumption. Practically speaking, we think of this as patience versus impatience, in which a patient investor is relatively indifferent to consuming now or later, while an impatient investor strongly prefers the immediate gratification of current consumption at the expense of lower future consumption. The parameter does not indicate whether or not delaying consumption is a better financial decision; its purpose is just to quantify how this decision affects the utility for a given investor. Within a life-cycle model, this parameter is similar to a discount rate in a time value of money calculation in which future cash flows are converted into current values; however, in this case, it represents how an investor *personally* discounts future consumption rather than a market-based discount rate. As we will discuss, the value of the impatience for consumption (subjective discount rate) relative to a certain market rate of return determines whether the investor will increase, decrease, or keep consumption constant over time. We denote the subjective discount factor by the Greek letter ρ.

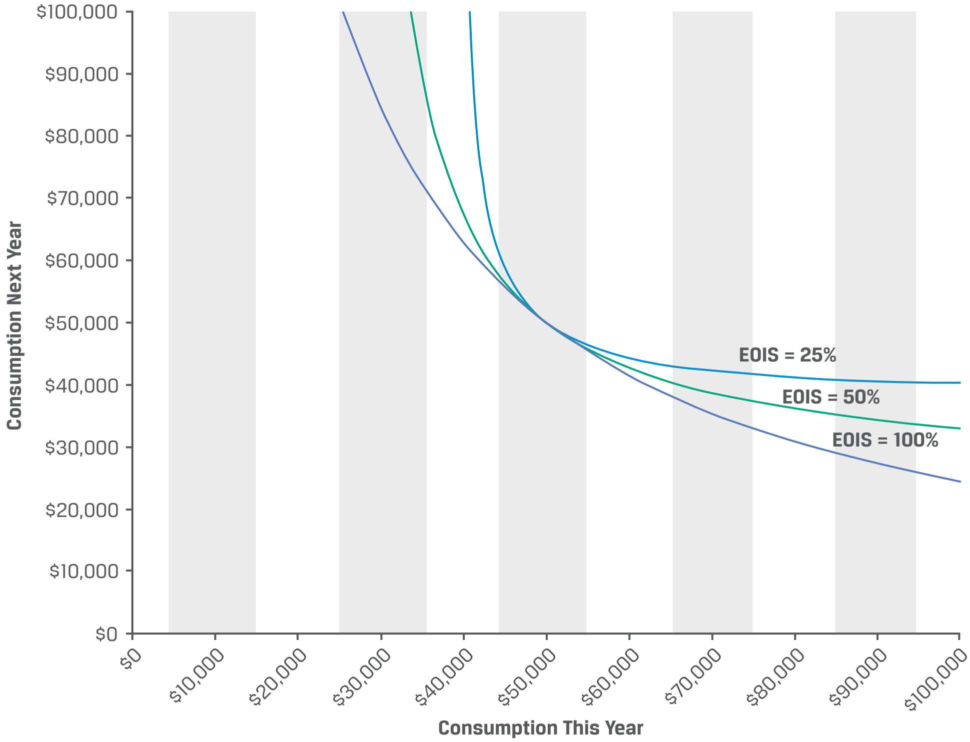
Preference for Smooth Real Consumption: Elasticity of Intertemporal Substitution (EOIS, η or eta)

The EOIS measures how flexible a person is in substituting consumption in one period for another, regardless of impatience. It typically takes on a value between 0% and 100%. We denote it by the Greek letter η.

As we demonstrate in **Exhibit 3.6**, this parameter is about the investor's preference for smooth consumption. Exhibit 3.6 shows three indifference curves for *real* or inflation-adjusted consumption this year and consumption next year. Each curve represents a utility function with a different value for η that an investor might have. Of these three curves, the one with the lowest value for η (25%) is the most L-shaped, indicating that the investor is fairly inflexible when it comes to substituting consumption in one period for another. The one with $\eta = 50\%$ has wider curvature and the one with $\eta = 100\%$ has even wide curvature. This shows that the greater the value of η, the greater the flexibility.

First, notice that for all three values of η, the investors are indifferent to (1) $50,000 real this year and $50,000 real next year and (2) all other combinations of consumption this year and next on their respective indifference curves. This indifference is due to the fact that all investors like smooth consumption. To depart from smooth consumption of $50,000 in each year and obtain the same level of utility, each investor requires some level of additional consumption in one year and less consumption in the other year, in which the total is greater than the total of $100,000 of the smooth consumption pattern. The investor with $\eta = 25\%$ really likes smooth consumption and requires the largest increase in total consumption to remain indifferent. The investor with $\eta = 100\%$ likes smooth consumption but, of the three investors in this example, requires the smallest increase in total consumption to remain indifferent. In other words, lower values of η mean less flexibility around departures from smooth consumption, and larger value of η mean more flexibility around departures from smooth consumption.

Exhibit 3.6. Indifference Curves for Different Preferences for Smooth Real Consumption (EOIS)



Intertemporal Decision Making

Now that we have defined the impatience for consumption (subjective discount rate, ρ) and the preference for smooth consumption (EOIS, η), we can create a simple two-period life-cycle model (with a 100% probability of surviving through the second period) of how a rational investor decides between consumption this year and next year.

Note that starting here, and throughout this book, we use equations to give precise expression to our models. However, for the nontechnical reader, we do our best to explain the various formulas in words and believe one can skip the more complicated formulas while still advancing one's intuition and understanding.

To keep the model simple, we assume that there is a real risk-free market return, which we denote *r*, and that there are no risky assets. (We introduce risky assets after we have introduced the risk preference parameter.)

To complete the simple two-period life-cycle model, let:

c_0 = consumption this year,

c_1 = consumption next year,

W_0 = wealth at the beginning of this year, and

r, = real risk-free rate of market.

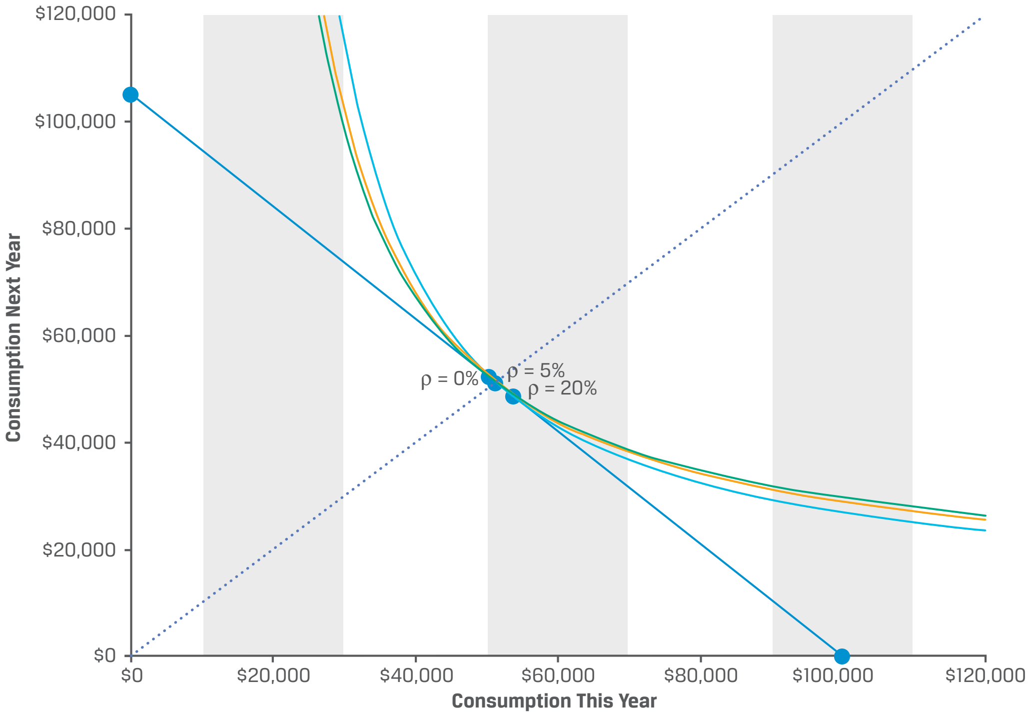
We have the following *budget constraint*:

$$c_0 + \frac{c_1}{1+r} = W_0. \quad (3.1)$$

That is, the consumption this year (c_0), plus the consumption next year (c_1) discounted by 1 plus r, must equal the wealth at the beginning of this year (W_0).

In **Exhibit 3.7**, we have drawn this budget constraint based on the assumption that the real risk-free market return (r) is 5%. The downward-sloping blue line is the budget constraint. It shows all possible combinations of consumption this year and next that satisfy Equation 3.1 (adhere to the budget constraint). The three points (blue dots) at the middle of the budget constraint are the optimal bundles or combinations of consumption this year *and* consumption next year, for three different values of the impatience for

Exhibit 3.7. Optimal Consumption Bundles for Different Values of the Subjective Discount Rate



	ρ = 0%	ρ = 5%	ρ = 20%
Consumption This Year (c_0)	$50,305	$51,220	$53,717
Consumption Next Year (c_1)	$52,180	$51,220	$48,598
Percentage Difference $\left(\frac{c_1 - c_0}{c_0}\right)$	3.73%	0.00%	−9.53%

consumption (subjective discount rate), assuming that η is 75% in all three cases.[13] The values of ρ for the blue dots are 0%, 5%, and 20%. For each value of the three impatience-for-consumption subjective discount rates, a corresponding indifference curve is tangent to the budget constraint line. We have also introduced a 45-degree line as the dividing line between bundles in which consumption increases over time (above the line) and bundles in which it decreases (below the line). For bundles on the line, consumption is the same in both years.

Equation 3.2 is a simple formula for the percentage change in consumption over the two years when surviving to next year is certain. (We will later expand the models to incorporate the probability of survival.):

$$\frac{c_1 - c_0}{c_0} = \left(\frac{1+r}{1+\rho}\right)^{\eta} - 1. \tag{3.2}$$

The right side of Equation 3.2 shows that consumption will increase in the next period if the risk-free market rate of return (numerator) is higher than the impatience-for-consumption subjective discount rate (denominator), will decrease if the opposite is true, and will stay the same if the two rates are the same. The extent of this effect depends on the investor's preference for smooth real consumption (EOIS, η). In other words, the optimal consumption bundle (amount to consume *this* year and amount to consume *next* year) is a function of (1) the investor's impatience for consumption (subjective discount rate, ρ), (2) the risk-free rate of return (r), and (3) preference for smooth real consumption (EOIS, η). Exhibit 3.7 illustrates how optimal consumption changes based on the three impatience-for-consumption subjective discount rates ($\rho = 0\%$, 5%, and 20%). Given the scaling in the graph, the blue dots appear to be in relatively similar positions, but the table highlights that annual real consumption between the two years differs by nearly 10% for $\rho = 20\%$.

The key takeaway from a life-cycle model perspective is that the investor's preference around impatience for consumption (subjective discount rate, ρ) and the preference for smooth consumption (EOIS, η) interact with one another to impact the timing and magnitude of utility maximizing consumption. Recall that in chapter 2, we provided sample questions designed to gain insight on the investor's preferences in which the different responses might correspond to different values of ρ and η. Although we have not yet talked about optimal savings rates, these two preferences will clearly affect saving decisions as one chooses to either spend more and save less now or to spend less and save more now.

Risk Tolerance (θ or theta)

Returning to the third of our five pecuniary investor preference parameters, as we mentioned in chapter 2, risk preference (risk tolerance, θ or theta) measures the degree to which a person is *willing* to take on risk to obtain potential higher levels of future consumption.

Before we can really describe the role of risk tolerance, one needs a foundational understanding of expected utility theory and how it is used to make decisions under uncertainty. After all, the vast majority of financial planning decisions, including investment choice, involve uncertainty.

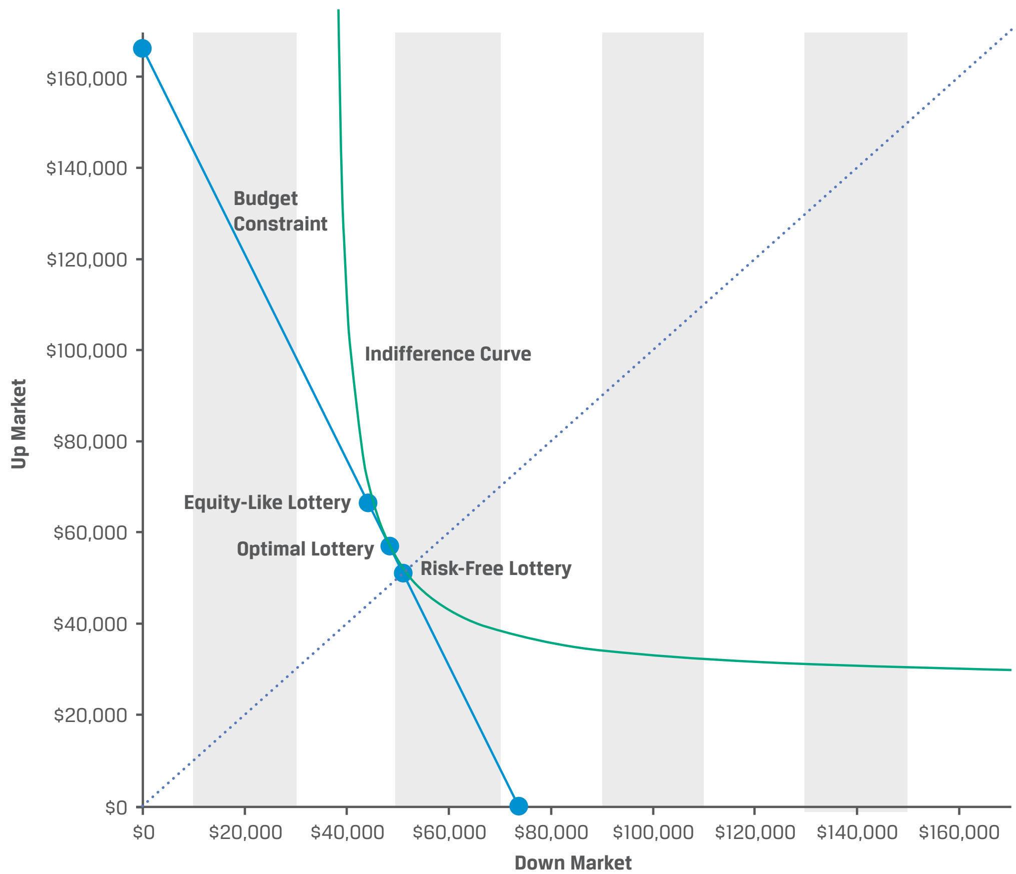
In its most elementary form, the application of *utility theory to decision making under uncertainty* is similar to the *theory of intertemporal decision making*, the difference being that, instead of deciding between bundles of consumption at different times, the investor decides between bundles of consumption under alternative randomly determined conditions called *states*. We assume that the probability of each state occurring is known. Examples of states include up markets and down markets.

[13]We chose these values to illustrate the impact of the impatience-for-consumption parameter (subjective discount rate) on intertemporal consumption decisions. They are not meant to be realistic.

In the language of the *economics of uncertainty*, bundles of consumption under alternative states are referred to as *lotteries*, which builds on the idea that the "state" that occurs is random. An important assumption that we make is that of *complete markets*. Complete markets have a price *today* for delivery of one (real) dollar in each state in the future and that price is different depending on the state.[14] Furthermore, it is possible to transact at that price in any volume desired. Because of *diminishing marginal utility*, in an up-market "state," the price *today* for delivery of one (real) dollar will be *lower*, and conversely, in a down-market "state," the price *today* for delivery of one (real) dollar will be *higher*. In other words, additional money is most valuable when you need it most and least valuable when you need it the least.

Exhibit 3.8 illustrates how this works. For simplicity we assume two possible states: down market and up market and a budget of $50,000 that one can allocate between these two states. We have assigned a price and a probability to these two states: a price of $0.30 today for $1 in the future if the up-market state with a 60% probability is realized and a price of $0.68 today for $1 in the future if the down-market state with a 40% probability is realized. This allows us to define and draw a budget constraint that reflects all of the possible states and outcomes. The bottom end of the budget constraint corresponds to using all $50,000 to purchase contracts that only pay in the down-market state, with the payout being about $50,000/0.68,

Exhibit 3.8. Finding the Optimal Lottery



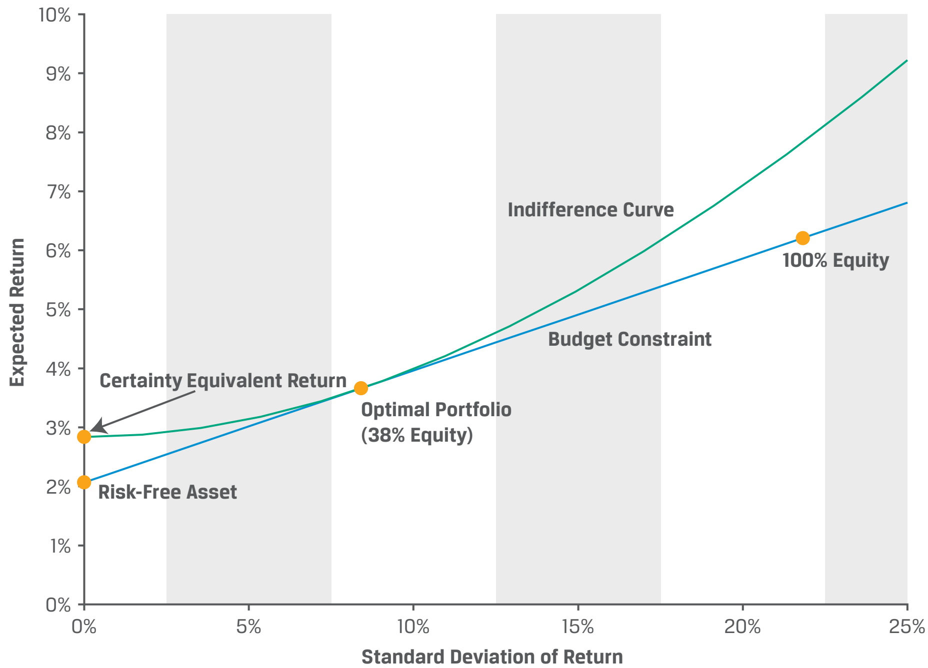
[14]This assumption is also called the assumption of *complete contingent claims markets*.

which is about $73,500. The upper end of the budget constraint corresponds to using all $50,000 to purchase contracts that only pay in the up-market state, the payout being $50,000/0.30, which is about $166,700. The danger of allocating *all* $50,000 to either of the two states is the other state could be realized and the individual would end up with nothing, which is why diversifying between these two states by creating a consumption bundle is optimal. We have also drawn a 45-degree line that shows which lotteries (consumption bundles) are risk-free (because they pay the same amount in both states). We have drawn an indifference curve that is tangent to the budget constraint, showing the optimal lottery.

In the real world, we cannot purchase lotteries that pay only in a specific state. Instead, we form portfolios of assets, each of which pay various amounts in different states, forming the distribution of possible future portfolio values. Also, performance is typically measured as a rate of return rather than a dollar figure; thus, the different states of a portfolio are represented by the distribution of possible returns as defined by the expected return and standard deviation of return. In the portfolio theory developed by Markowitz (1952, 1959), portfolios are ranked by combinations of expected return and standard deviation of return, the latter being a measure of risk. Similarly, the *utility* of a portfolio for a given investor is a function of expected return and standard deviation and the "ranking" of different portfolios are specific to a given investor's risk tolerance. We elaborate on this shortly.

In **Exhibit 3.9**, we recast the optimal lottery problem from the economics of uncertainty, depicted in Exhibit 3.8, as an optimal expected return/standard deviation problem. This likely will be more familiar to most readers. Having defined the vertical axis as expected return and the horizontal axis as the standard deviation of return, the budget constraint now appears as an upward-sloping line. This form of the budget constraint is similar to the Markowitz efficient frontier.

Exhibit 3.9. Finding the Optimal Lottery in Terms of Expected Return and Standard Deviation



In risk and expected return space, the indifference curve is also upward sloping because the investor prefers to have a higher rather than lower expected return, and a lower rather than higher standard deviation. The indifference curve touches the vertical axis at the *certain-equivalent rate of return*; in this case, at 2.8% on the vertical axis. This is the rate of return that if it were riskless (standard deviation equals 0%), gives the same utility as all of the combinations of expected return and standard deviation along the indifference curve, including the optimal combination.

In Exhibit 3.9, we identified a combination of expected return and standard deviation that has characteristics similar to those of a diversified 100% equity portfolio. (In Exhibit 3.8, we indicated the position of this portfolio as a "lottery" and labeled it "Equity-Like Lottery.") Treating this portfolio as a 100% equity allocation, based on the indifference curve of the investor, the optimal portfolio can be looked at as a portfolio that is 38% equity and 62% in the risk-free asset.

Expected Utility Theory[15]

The *theory of rational behavior under uncertainty* is more specific than the *general utility theory* that we have discussed so far. This specific form of utility theory is called *expected utility theory*. Its main result is that, under certain assumptions, the *utility of a lottery* can be written as a probability-weighted average of the values of a function of each of the lottery's payoffs. This probability-based decomposition of utility greatly simplifies the problem of finding optimal portfolios of risky assets and other problems in finance that deal with making decisions under uncertainty: it reduces the problem of comparing lotteries to calculating the utility of consumption under different states and multiplying by the probabilities of the possible states.

Expected utility theory was first laid out in the 1940s by the mathematicians John von Neumann and Oskar Morgenstern ([1944] 1967). Their theory, that rational investors maximize expected utility, is the foundation of classical finance. As we discuss next, *expected utility theory* is the basis for Markowitz's theory of MVO. Here, we show how von Neumann and Morgenstern define rationality and describe the conclusions that they reach.

The Framework

Von Neumann and Morgenstern modeled how investors rank different investments that have *uncertain* payoffs. They considered investments as a more general form of lotteries than we have so far. For Von Neumann and Morgenstern, a lottery is generalized as any set of possible outcomes, each with a probability. The only requirement is that the possible outcomes are mutually exclusive and that the probabilities add up to 100%. For example, a lottery L could consist of three possible outcomes, payoffs of $25, $50, and $65, with probabilities of 20%, 30%, and 50%, respectively.

Given two lotteries L_1 and L_2, we use the following notations to describe how the investor ranks them:

- $L_1 \sim L_2$ means that the investor is indifferent between L_1 and L_2.
- $L_1 < L_2$ means that the investor strictly prefers L_2 over L_1.
- $L_1 \leq L_2$ means that the investor either prefers L_2 over L_1 or is indifferent between the two.

The Assumptions (Axioms)

Von Neumann and Morgenstern made four assumptions about how a rational investor would rank lotteries known as *axioms*. The first two are completeness and transitivity, which we mentioned in our general discussion on utility. The third and fourth are as follows:

[15]This section is adapted in part from Kaplan (2015b).

Axiom 3: Continuity

If $L_1 \leq L_2 \leq L_3$, the probability p exists such that $pL_1 + (1 - p)L_3 \sim L_2$. This says that, for a given lottery, a less desirable one, and a more desirable one, it is possible to blend the less and more desirable ones into a lottery that is as desirable as the one they surround.

Axiom 4: Independence

Given $L_1 < L_2$, and third lottery L_3, the investor's preference for L_2 over L_1 is not affected by the possibility of L_3. Hence, given a probability p, $pL_1 + (1 - p)L_3 < pL_2 + (1 - p)L_3$.

The Theorem

If we assume that an investor's preferences obey these four seemingly innocuous assumptions, von Neumann and Morgenstern proved that the investor's preferences can be expressed with a von Neumann–Morgenstern utility function that we denote $u(.)$. (When the context makes clear that we are referring to a von Neumann–Morgenstern utility function, we will simply call it a utility function.) Given a lottery L_1 with payouts $x_1, x_2, \ldots, x_m$ with corresponding probabilities $p_1, p_2, \ldots, p_m$, and a second lottery L_2 with payouts $y_1, y_2, \ldots, y_n$ with corresponding probabilities $q_1, q_2, \ldots, q_n$, the *expected utility* of each of these lotteries is given by the following:

$$EU(L_1) = \sum_{i=1}^{m} p_i u(x_i), \tag{3.3}$$

$$EU(L_2) = \sum_{i=1}^{n} q_i u(y_i). \tag{3.4}$$

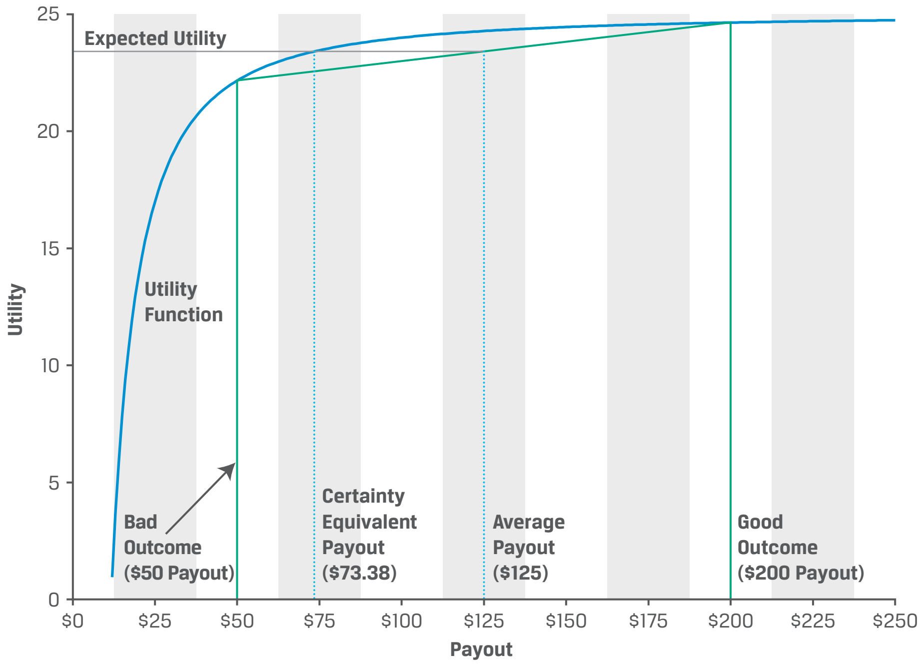
In Equation 3.3, the expected utility for lottery L_1 is simply the weighted sum of the utility of each of its possible payouts $[u(x_1), u(x_2), \ldots, u(x_m)]$ weighted by the corresponding probabilities $p_1, p_2, \ldots, p_m$. Likewise, in Equation 3.4, the expected utility for lottery L_2 is simply the weighted sum of the utility of each of its possible payouts $[u(y_1), u(y_2), \ldots, u(y_m)]$ weighted by the corresponding probabilities $q_1, q_2, \ldots, q_m$.

The von Neumann–Morgenstern theorem is that $L_1 < L_2$ if and only if $EU(L_1) < EU(L_2)$. The power and beauty of the von Neumann–Morgenstern theorem is that it reduces the abstract problem of comparing choices with uncertain outcomes to a straightforward numerical problem. This is foundational to optimal, probabilistic decision making.

The form of the utility function has two restrictions. First, it must be increasing so that more is always better than less, i.e. nonsatiation. Second, it must be increasing at a decreasing rate. Thus, the plot of $u(.)$, has a concave shape. Concavity is necessary for the investor to be risk averse. In other words, given the choice between (1) a certain outcome and (2) a lottery with an uncertain outcome and the same expected value, the investor prefers the certain outcome.

Exhibit 3.10 shows how a von Neumann–Morgenstern utility function works. Utility is on the vertical axis and the payout is on the horizonal axis. In this exhibit, we assume that the investor is evaluating a lottery in which two equally likely possible outcomes have payouts of $50 and $200, respectively. The utility curve

Exhibit 3.10. Example of Finding the Expected Utility of a Lottery Using a von Neumann–Morgenstern Utility Function



shows the amount of utility associated with various payouts. We represented each of the two possible payouts with vertical lines from the horizonal axis to the utility function. We read off the utility of each outcome from the utility function. Because the probability of each outcome is 50%, the *expected* utility of the lottery is the average of these two utility values.

Because of the uncertainty associated with the lottery that has an expected payoff of $125, its expected utility is less than the expected utility of receiving $125 with certainty. Exhibit 3.10 also shows how to calculate the certainty-equivalent payout. This is the payout level that if the investor could receive with certainty would yield the same level of expected utility as a lottery with an expected payoff of $125.

Incorporating the Risk Tolerance Parameter

The von Neumann–Morgenstern *expected utility theory* and the corresponding *utility function*, which was built upon the *theory of rational decision making*, creates the theoretical framework for both the single-period Markowitz optimization models and life-cycle models. As explained earlier, the more an investor dislikes risk, the more concave the von Neumann–Morgenstern utility function. If the investor has CRRA, the shape of the utility function is described by a single *risk tolerance parameter*, which we denote θ.[16] With CRRA, the utility function is as follows:[17]

[16]The concepts of relative risk aversion and CRRA were introduced by Arrow (1965) and Pratt (1964).

[17]The reciprocal of the risk tolerance parameter is sometimes called the risk aversion parameter. However, we use the term somewhat differently in Part III.

$$u_\theta(x) = \begin{cases} ln(x), & \theta = 1 \\ \dfrac{\theta}{\theta - 1}\left(x^{\frac{\theta-1}{\theta}} - 1\right), & \theta \neq 1. \end{cases} \tag{3.5}$$

For those unfamiliar with the layout of Equation 3.5, let us explain. The utility function $u_\theta(x)$ on the left has two different functional forms depending on the value associated with the risk tolerance parameter, θ. When $\theta = 1$ the form is $ln(x)$. When $\theta \neq 1$, it takes the following power form:

$$\frac{\theta}{\theta - 1}\left(x^{\frac{\theta-1}{\theta}} - 1\right).$$

Here, x generically represents some value that is of concern to the investor such as discretionary consumption in a given period, or a bequest. Note that in Equation 3.5, we place the risk tolerance parameter as a subscript in the name of the function. Next, we introduce utility functions for other preferences that take the same mathematical form as the function defined in Equation 3.5, but with a different parameter. We will denote each of these utility functions $u_p(.)$, where p is replaced by the preference parameter in question.

As we will see in chapters 5 and 6, the power form of the utility function (as well as the logarithmic form when $\theta = 1$) leads to a simple and elegant optimal consumption recommendation in which discretionary consumption is proportional to the investor's net worth.

A useful property of CRRA utility functions is that the amount being invested does not affect how the investor ranks alternative portfolios.[18] This is why CRRA is the most common assumption about investor preferences in finance.

Values of θ are typically between 0% and 100%. A value of 0% means that the investor has no willingness to take risk and, therefore, if possible, it holds only riskless assets. A value of 100% indicates a high level of risk tolerance, leading to highly risky portfolios (assuming that the expected returns of risky assets are high enough to make then attractive). Although values greater than 100% are possible, they are not typical.

Note that the most common alternative to CRRA from behavioral finance is prospect theory (Kahneman and Tversky 1979). In prospect theory, individuals tend to anchor on a specific reference point and evaluate outcomes relative to that point. Initially, for gains or losses of the same magnitude, losses are more painful than corresponding gains. In contrast to CRRA, however, as losses grow large relative to the reference point, with prospect theory, investors become risk seeking, fatalistically swinging for the fences in an attempt to recoup previous losses.

Importantly, expected utility theory and the assumption of CRRA are normative (how people *should* behave). That is, if investors can in fact be coached into behaving rationally, they should follow advice based on expected utility theory and the assumption of CRRA. Prospect theory is useful in explaining the nonrational decision making that is often observed in real-world decision making when a person does not have the assistance of an expert. The adviser is the expert.

If we assume that people cannot be coached and cannot be saved from their own bad decision making, one should arguably start with a utility function that accepts the investor's bad decision making. We believe the reason that investors seek advice from professional advisers and wealth managers is not to receive suboptimal recommendations rooted in the notion that the investor's decision making is flawed,

[18]See Kaplan (2015b).

but rather, investors recognize they are not perfect decision makers and thus want advice from an expert. One of the most important goals of behavioral finance is to understand bad decision making and to find ways to avoid it.

Mean–Variance Optimization[19]

In the preface to Markowitz and Blay (2014, p. xxi), Harry Markowitz, the father of modern portfolio theory, states, "This field [risk-return analysis] is plagued by a Great Confusion." What he means by this "Great Confusion" is the wide acceptance of the assertion that MVO is valid only if either (1) returns follow a normal distribution or (2) investors have quadratic utility functions.[20] It is true that if either of these assumptions holds, investors should seek mean–variance efficient portfolios, but neither is necessary to justify MVO. Beginning with his 1959 book, Markowitz has always justified MVO with expected utility theory. In this section, we explain how Markowitz uses it to justify MVO.[21]

According to Markowitz (1999), what was lacking before Markowitz (1952) was a *theory of portfolio construction*. He specifically refers to the year 1952 rather than the paper that he published in 1952 because in that same year, A. D. Roy (1952) also published a paper on selecting a portfolio on the basis of mean and variance. Markowitz considers Roy the cofounder of modern portfolio theory and laments that he never received the credit that he deserved.[22]

James Tobin (1958) published a paper in which he used MVO to model how investors allocate between cash and bonds. He justified using MVO by assuming that either the investor has quadratic utility or the distribution of returns is from a two-parameter family, such as the normal distribution. Hence, it was Tobin, not Markowitz, who made the assumptions that perhaps gave rise to the Great Confusion.

The great statistician Leonard J. Savage was a proponent of expected utility theory, and as Markowitz recalls, he had been "indoctrinated at point-blank range in expected utility theory" by Savage (Markowitz, Savage, and Kaplan 2010).

[19]This section is based in part on Kaplan (2014).

[20]A quadratic utility function takes the form $u(x) = \frac{1}{2}(x - \overline{x})^2$, where $\overline{x}$ is the point of maximum utility, beyond which utility decreases rather than increases, thus violating the axiom of nonsatiation we introduced earlier.

[21]Markowitz (1959).

[22]In 2006, *The Journal of Investment Management* published an English translation of a paper on mean-variance analysis that was originally published in 1940 in Italian, 12 years before Markowitz published his famous paper. The author was the Italian mathematician and actuary Bruno de Finetti (2006). There are several reasons why de Finetti's paper was unknown before 2006. First of all, it was in Italian and thus unknown among English-speaking researchers. Second, de Finetti was trying to solve a problem in reinsurance, not in investments. Markowitz has summed up the situation this way: "It was dead end, not because it deserved to be a dead end, but that was, in fact, its historical destiny." (Markowitz, Savage, and Kaplan 2010).

Markowitz (1959) wrote an entire book on MVO. In it, he reviews expected utility theory and demonstrates that, under a broad set of assumptions, MVO leads to a solution that is a good approximation of maximizing expected utility (recall that Markowitz sees expected utility theory as the justification for MVO). Twenty years later, Markowitz and Haim Levy (1979) published a paper in which they formally show how expected utility can be approximated by a function of expected return and variance. Of often overlooked importance, they demonstrated the accuracy of their approach with a variety of both *utility functions* and *return distributions*. As explained, the efficient frontier from MVO traces out the risk and expected return of the efficient asset mixes, in which case each point on the efficient frontier identifies the maximum possible expected return for that particular level of risk. Hence, the asset mix that maximizes Levy and Markowitz's approximation should be close to the asset mix that maximizes expected utility, regardless of the *utility function* and the *distribution of returns*.

Levy and Markowitz (1979) also showed that expected utility can be approximated by a function of mean and variance (standard deviation squared). They did this by first approximating the utility function with a second order Taylor expansion around the expected return (μ):

$$u(1+\tilde{R}) \approx u(1+\mu) + u'(1+\mu)(\tilde{R}-\mu) + \frac{1}{2}u''(1+\mu)(\tilde{R}-\mu)^2, \qquad (3.6)$$

where $\tilde{R}$ denotes the random rate of return on a portfolio.

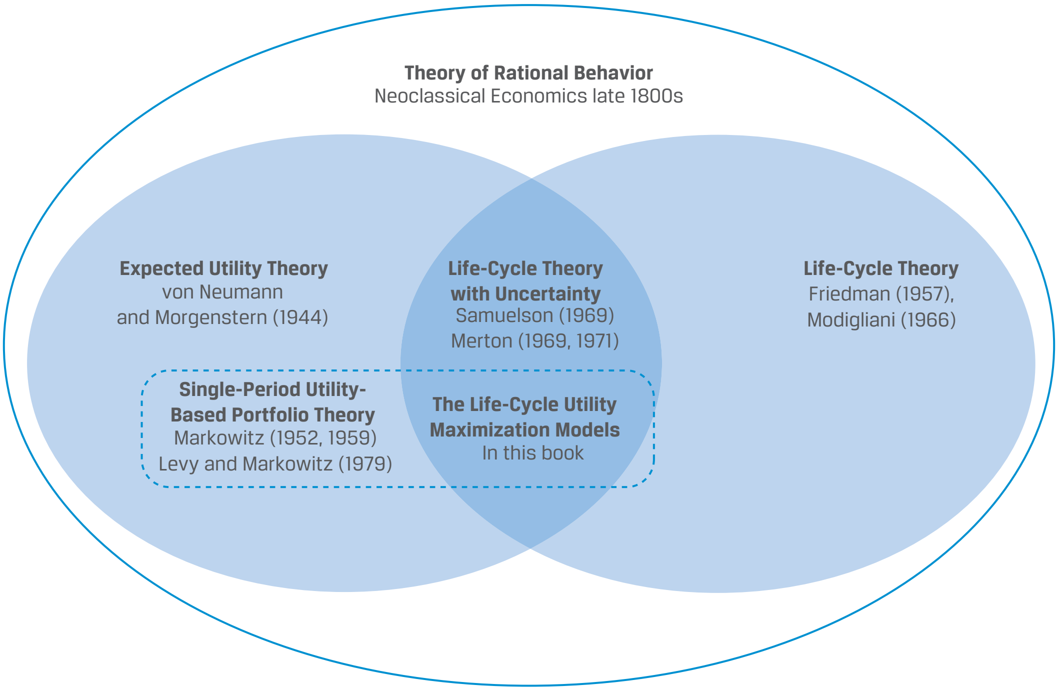
Let σ denote the standard deviation of return. Because $\mu = E[\tilde{R}]$ and $\sigma^2 = E[(\tilde{R}-\mu)^2]$, taking the expected value of the approximated utility functions, we have the following:[23]

$$E[u(1+\tilde{R})] \approx u(1+\mu) + \frac{1}{2}u''(1+\mu)\sigma^2, \qquad (3.7)$$

where $u''(.)$ denotes the second derivative of the utility function. We discussed the nonsatiation principle, which states that more is always better than less. This implies that the utility function is always increasing, so that its first derivative is always positive. We also discussed the principle of diminishing marginal utility, which tells us that as we increase consumption, the *rate of change* in utility is decreasing. This implies that the second derivative is always negative. So, in Equation 3.7, because variance is multiplied by the second derivative of the utility function, which is always negative, the investor dislikes variance. Thus, we can approximate expected utility as a function of mean (μ) and variance (σ^2). Furthermore, this approximating function is increasing in expected return and decreasing in the standard deviation of return. Hence, if the approximation is good, an investor who seeks to maximize expected utility will do well by choosing the portfolio along the mean–variance efficient frontier that maximizes the approximating function.

[23] $E[\tilde{X}]$ denotes the mathematical expectation of $\tilde{X}$; i.e., the mean of the distribution of $\tilde{X}$.

Exhibit 3.11. The Pedigree of Life-Cycle Theory



As depicted in **Exhibit 3.11**, *the theory of rational behavior* was primarily developed in the late 1800s as part of neoclassical economics and encompasses the *expected utility theory* of von Neumann and Morgenstern ([1944] 1967) and life-cycle theories of Friedman (1957) and Modigliani (1966). Life-cycle theory with uncertainty, as put forth by Samuelson (1969) and Merton (1969, 1971), is at the intersection of *expected utility theory* and *life-cycle theory*. The *life-cycle utility maximization models* developed in this book primarily live within this intersection as life-cycle models with uncertainty. Finally, the *single-period utility-based portfolio theory* of Markowitz (1952, 1959) and Levy and Markowitz (1979) is another embodiment of *expected utility theory*. One of our primary innovations in chapter 8 is to show how single-period portfolio optimization can be used periodically to implement a *life-cycle utility maximization model* in real time. We have drawn a dashed line around these two distinct model types to reflect that by simultaneously using both model types in conjunction with one another, we will attempt to offer optimal financial advice.

Survival Probabilities[24]

When discussing the first two preference parameters, we demonstrated how *impatience for consumption* (subjective discount rate, ρ) and the *preference for smooth consumption* (EOIS, η) can be incorporated into a simple two-period life-cycle model for determining the optimal amount of consumption in each of the two periods. We made the simplifying assumption that the lifetime was exactly two periods and thus we did not need to consider the uncertainty of the planning horizon.

Before discussing the final two preference parameters related to bequests, we need to cover how we model longevity. Modeling longevity is critical in lifetime financial planning. While planning based on life

[24]This material is adopted in part from Kaplan (2015a).

expectancy is common practice, we do not think that it is advisable because about half of all investors will outlive their median life expectancies. Instead, we take a probabilistic approach in which we consider the probability of surviving until each year in the future. From the survival probabilities, we calculate the probability of dying each year.

Human mortality was first studied, quantified, and modeled by Benjamin Gompertz (1779–1865), a British actuary.[25] (More than a century earlier, John Graunt and Edmond Halley—the latter of Halley's Comet—constructed the first "life tables," containing life expectancy estimates, but they did not study the statistical properties of the estimates.) Gompertz discovered that the probability of a person surviving to age a_2 could be well approximated by a formula that only has three parameters:

1. the person's current age (a_1);
2. the mode of the distribution of the age of death (m); and
3. the dispersion of the age of death around the mode (b). The b parameter is similar to the standard deviation parameter of a normal distribution.

We present the details of the Gompertz formula and explain how to estimate the parameters m and b in Appendix 3A. The following two critical terms are based on the Gompertz formula and are an important part of our life-cycle models:

q_v^t = the probability of the person surviving to at least year v, given that the person was alive in year t; and

q_v^{tJ} = using the additional subscript J, this is the joint probability that one member of a couple survives to at least year v, given that both people were alive in year t.

A key advantage of the Gompertz model is that, by setting just two parameters, the mode (m) and the dispersion around the mode (b), we can easily calculate survival probabilities for a given country or population group, or even a household or individual investor. This allows us to personalize recommendations.

In **Exhibit 3.12**, based on a given set or parameters, we plot the probability of survival of 65-year-old men and women. We also plot the joint survivor probability for a heterosexual couple, both age 65, surviving for a given number of years.

As we mentioned earlier, the m parameter is the mode of the distribution of the age of death. The value of this parameter can be set using a relevant table of mortality rates. We say relevant because mortality rates differ by sex and geography. The value can be refined based on a person's health and genetics. This flexibility makes the Gompertz formula useful in the personalization of financial planning and related applications.

When modeling the bequest motive, we need to calculate the probability distribution of age of death. We can calculate the probability of dying in year v, given that the investor is alive in year t, p_v^t, from survival probabilities as follows:

$$p_v^t = q_v^{t-1} - q_v^t. \tag{3.8}$$

In **Exhibit 3.13**, we plot the probability of death in each year for a 65-year-old man and a 65-year-old woman as of year 0 using the parameters we presented earlier. For each year, we also plot the probability that it will be the year in which the last member of the couple dies. This is the curve labeled "Joint."

[25]For a detailed discussion of the Gompertz formula and some interesting facts about Benjamin Gompertz, see Milevsky (2012a, chapter 2).

Exhibit 3.12. Survival Probabilities, Age 65

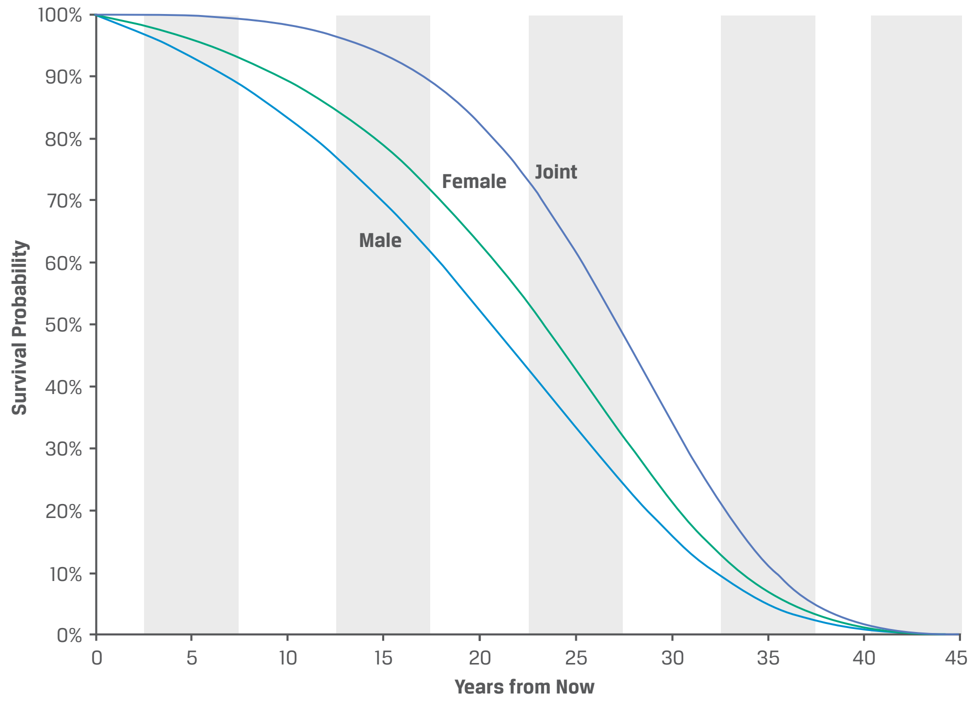


Exhibit 3.13. Death Probabilities, Age 65

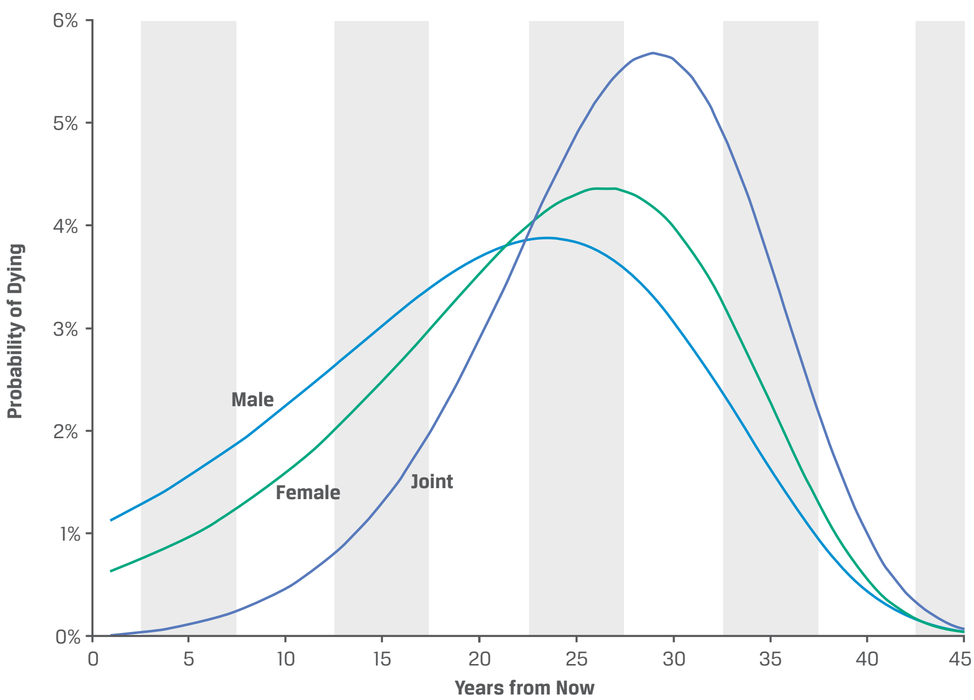
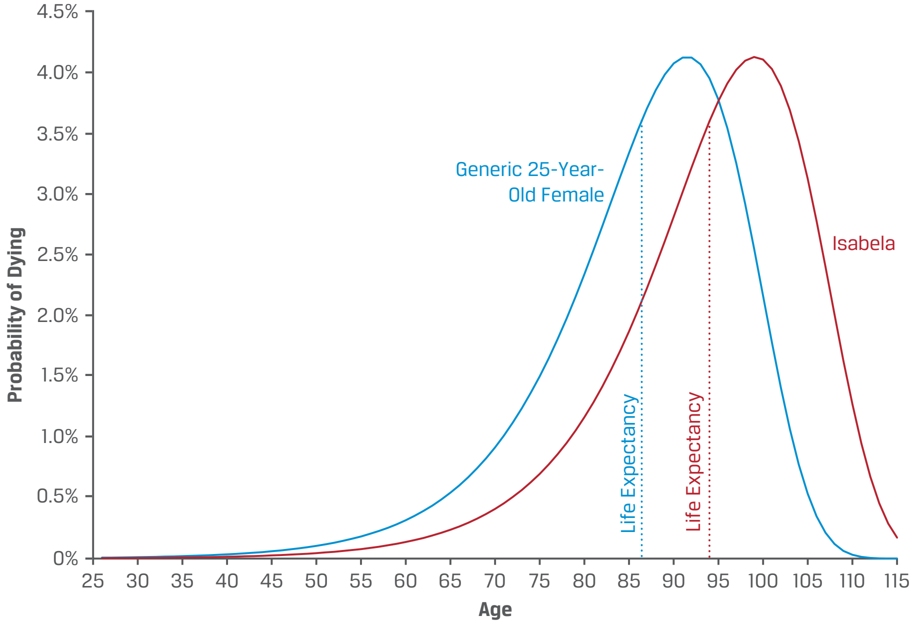


Exhibit 3.14. Probability Distribution of Age of Death: Generic versus Personalized (Isabela)



To make this more concrete, we demonstrate how one can use the Gompertz model to create personalized survival probabilities for our hypothetical investor Isabela, who is currently 25 years old.

Recall that Isabela recently started working with Paula the planner. In talking with Isabela, Paula discovered that Isabela's paternal grandfather had died two years earlier at 99, both of her maternal grandparents were still alive at ages 92 and 94, and two of her paternal great aunts were still alive at ages 100 and 102! Based on this information, within Paula's financial planning software, she indicated that Isabela was likely to live longer than the default life expectancy for a 25-year-old woman. Based on our standard set of parameters for a 25-year-old woman, life expectancy is 86.4. Paula overrides the default life expectancy with age 94 to reflect the high longevity in Isabela's family. **Exhibit 3.14** shows the impact of raising the life expectancy on the probability distribution of age of death.

As we move forward with various calculations for Isabela, we use her personalized survival probabilities.

The Groundwork for Our Base Case Life-Cycle Models

We still have our two bequest-related preference parameters to discuss, but importantly, we now have all of the elements of *utility theory* needed to write the formula for the *utility of a lifetime stream of consumption* without bequests. In the following discussion, we begin with the utility of a lifetime stream of consumption and then introduce bequests.

Utility of Lifetime Consumption

Equation 3.9 does not specifically tell us how to maximize the utility of lifetime consumption by varying consumption across the investor's lifetime. Instead, it simply tells us how to calculate the utility of a consumption stream based on two of the investor's preference parameters. At this point, we treat all future consumption as *known* in advance. The utility of lifetime consumption starting from year t is as follows:

$$U_t^L = \underbrace{\sum_{v=t}^{T} \underbrace{q_v^t}_{\substack{\text{Probability}\\ \text{of}\\ \text{surviving}\\ \text{to year } v}} \underbrace{\frac{1}{(1+\rho)^{v-t}}}_{\substack{\text{Time Value}\\ \text{Discount}}} \underbrace{u_\eta(c_v)}_{\substack{\text{Utility of}\\ \text{Consumption}\\ \text{in year } v}}}_{\substack{\text{Utility of mortality weighted}\\ \text{lifetime consumption}}}, \tag{3.9}$$

where $u_\eta(.)$ is the single-period utility function. In words, the total mortality weighted lifetime utility of consumption is the sum of the probability of surviving to year v multiplied by the time value discount factor multiplied by the utility of consumption in each year. The utility in a given year v has the same mathematical form as the CRRA utility function in Equation 3.5, but with the EOIS, η, rather than the risk tolerance parameter, θ. That is:

$$u_\eta(c_v) = \begin{cases} ln(c_v), & \eta = 1 \\ \dfrac{\eta}{\eta-1}\left(c_v^{\frac{\eta-1}{\eta}} - 1\right), & \eta \neq 1. \end{cases} \tag{3.10}$$

Consistent with our earlier callout discussion on intertemporal decision making, Equation 3.9 includes two of the investor preference parameters: *impatience for consumption* (subjective discount rate, ρ) and the *preference for smooth consumption* (EOIS, η), which together affect the timing and magnitude of utility-maximizing consumption. Importantly, this utility of lifetime consumption does not use a generic or market-based discount rate to discount future consumption but instead is personalized, being the investor's subjective discount rate, ρ.

Equation 3.9 also highlights the importance of having a probabilistic longevity model. The further the investor looks out into the future, the less likely it is that they will be alive to enjoy whatever consumption is planned. Therefore, many investors will place more weight on consumption in the near future than in the far future.

As we shall see in chapter 5 when we fully develop a life-cycle model, it is useful to calculate the constant level of consumption that would result in the same lifetime utility as a given series of future consumption. We denote the constant level of consumption $\hat{c}_t$. We find it by solving the following equation for $\hat{c}_t$:[26]

$$\sum_{v=t}^{T} q_v^t \frac{1}{(1+\rho)^{v-t}} u_\eta(c_v) = \left[\sum_{v=t}^{T} q_v^t \frac{1}{(1+\rho)^{v-t}}\right] u_\eta(\hat{c}_t). \tag{3.11}$$

In models that have uncertainty, we often use a Monte Carlo simulation in which we run n trials. For each trial i, we calculate the constant level of consumption that would result in the same lifetime utility as the series of future consumption in the trial. Denote this $\hat{c}_{ti}$. We now define certainty equivalent consumption, $\hat{c}_t$ as the solution to:[27]

$$\frac{1}{n}\sum_{i=1}^{n} u_\theta(\hat{c}_{ti}) = u_\theta(\hat{c}_t). \tag{3.12}$$

Note that the risk tolerance parameter, θ, is now incorporated into the framework.

[26]Appendix 3B explains how to solve this equation. It and Equations 3.12 and 3.14 are based on Blanchett and Kaplan (2013).

[27]Appendix 3B explains how to solve this equation.

Bequest Preferences

The two bequest preference parameters are *flexibility of consumption versus bequest* (intergenerational elasticity, γ) and *importance of consumption versus bequest* (strength of the bequest motive, ϕ).

We now show how to incorporate the two bequest preference parameters into a life-cycle model. In chapters 5 and 6, we present models in which the investor selects the size of the bequest, B, which can be guaranteed using life insurance. Given the focus on incorporating bequest, we refer to this as an intergenerational model. To incorporate bequest into a life-cycle model, we use an intergenerational utility function that includes the certainty equivalent consumption, $\hat{c}_t$, developed in the previous section and the size of the desired bequest, B. Using the ^{IG} superscript, the *intergenerational* utility function, U_t^{IG}, is as follows:

$$U_t^{IG} = \underbrace{(1-\phi)u_\gamma(\hat{c}_t)}_{\text{Utility from Consumption while Alive}} + \underbrace{\phi u_\gamma\left(\frac{B}{D_t}\right)}_{\text{Utility from Bequest when Dead}}, \quad (3.13)$$

where D_t is a divisor that we use to express B as an annual amount comparable to $\hat{c}_t$:

$$D_t = \sum_{v=t}^{T} q_v^t (1+\rho)^{t-v}. \quad (3.14)$$

In Equation 3.13, the *total intergenerational utility* is the weighted sum of two different utilities: (1) the utility when the person is alive $u_\gamma(\hat{c}_t)$, and (2) the utility of the bequest when the person has died, $u_\gamma\left(\frac{B}{D_t}\right)$.

Noting that the *importance of consumption versus bequest* (strength of the bequest motive, ϕ) is between 0% and 100%, it becomes clear that ϕ controls the weight on the utilities when alive and dead.

Exhibit 3.15 illustrates how the intergenerational-elasticity parameter and the strength-of-bequest-motive parameter determine the size of the bequest. For a given investor, each curve shows the relationship between the strength of the bequest motive and the size of the bequest, for a given value of the intergenerational elasticity parameter. As expected, each curve is positively sloped, indicating that, as expected, the size of bequest increases with increasing values of the strength-of-bequest motive. The *slope* of the curve, however, increases with the value of the intergenerational elasticity parameter, showing that the higher the value of this parameter, the more sensitive the size of the bequest is with respect to the strength of the bequest motive.

The Investor's Needs and the Level of Consumption

Each year, investors have to spend money on basic necessities, such as food, shelter, and clothing. We call this *nondiscretionary consumption*.[28] We call any additional spending *discretionary consumption*.[29]

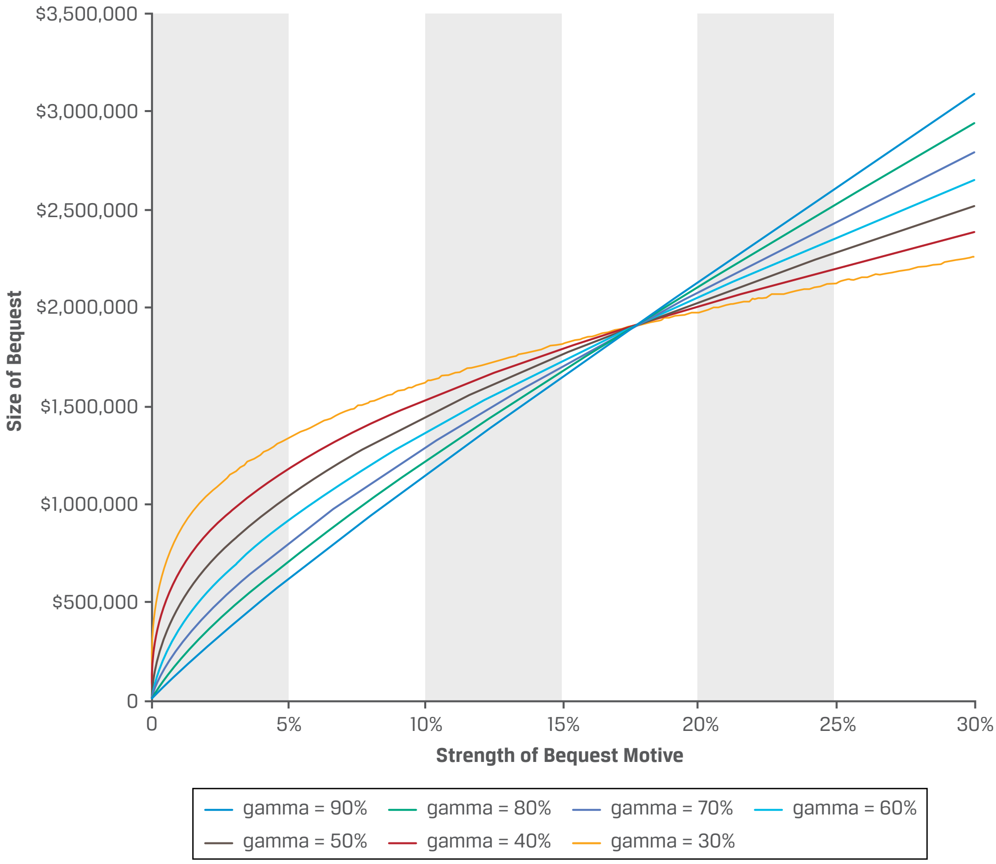
Let $\bar{c}_v$ denote nondiscretionary consumption in year v. To account for nondiscretionary consumption, we modify the intemporal utility function in Equation 3.9 as follows:

$$U_t^L = \sum_{v=t}^{T} q_v^t \frac{1}{(1+\rho)^{v-t}} u_\eta(c_v - \bar{c}_v). \quad (3.15)$$

[28]We also include interest and principal repayment on loans, such as mortgages, in nondiscretionary consumption.

[29]See Wilcox et al. (2006), p. 28, for a detailed discussion of discretionary and nondiscretionary consumption.

Exhibit 3.15. How the Bequest Parameters Affect the Size of the Bequest



Nondiscretionary consumption is how we model the needs and goals of investors. As we discuss in the next chapter, these needs and necessary goals are thought of as soft liabilities of the investor.

Conclusion and Key Takeaways

Like many economic-oriented decisions, *optimal* financial advice and decision making involves a web of complicated trade-offs. Although investors may not always behave rationally, it is arguably very useful to know the optimal answer to a problem when approached rationally. Utility theory and utility maximization provide the foundation for optimal decision making, whether that is solving life-cycle models or single-period mean–variance problems.

General utility theory hails from neoclassical economics, providing a theoretically rich framework for decision making. *General utility theory*, the *theory of rational decision making under uncertainty*, and more specifically the *expected utility theory* of von Neumann and Morgenstern ([1944] 1967), provide the foundation for optimal financial planning decisions.

In addition to introducing decision making under uncertainty based on expected utility theory, we introduce key investor preferences that in part determine how an investor ideally should make financial decisions.

Thus, an investor's preferences play a key role in decisions, such as how much to consume now versus later or how much one values their own consumption versus leaving a bequest. Ultimately, these preferences will be included in the life-cycle models presented in chapter 5 and 6. In particular, the first two investor preferences, *impatience for consumption* and *preference for smooth real consumption*, influence the timing of consumption and thus will influence saving rate and spending rate decisions. The final two investor preferences, *flexibility of consumption versus bequest* and *importance of consumption versus bequest*, influence the degree to which an investor derives utility from consumption versus leaving a bequest and thus will influence decisions involving life insurance and immediate annuities.

Finally, given the uncertainty of being alive in the future, we introduce the Gompertz formula for survival probability, which is particularly flexible and easy to use for estimating the probability of being alive, and as such, works well with a lifetime utility function and our life-cycle models.

Appendix 3A. Measuring the Probability of Survival with the Gompertz Formula

As we discuss in chapter 3, Gompertz discovered that the probability of a person surviving to age a_2 can be well approximated by a formula that has three additional parameters:

1. the person's current age (a_1);
2. the mode of the distribution of the age of death (m); and
3. the dispersion of the age of death around the mode (b). The b parameter is similar to the standard deviation parameter of a normal distribution.

We present the Gompertz formula in this appendix.

Given these parameters, the Gompertz formula for the probability of surviving to age a_2 is as follows:

$$g(a_2,a_1;m,b) = exp\left\{exp\left(\frac{a_1-m}{b}\right)-exp\left(\frac{a_2-m}{b}\right)\right\}. \tag{3A.1}$$

To make the Gompertz formula operational, we need parameter values for m and b. Blanchett and Kaplan (2013) estimated values of these parameters using mortality data on American men and women, which we have updated. The results are shown in **Exhibit 3A.1**.

We assume that the investor will not live past year T (typically between 40 and 50 for a 65-year-old), so that we can truncate the Gompertz function.[30] Let

a_0 = the age of the person in year 0, and

a_D = age at which death is certain = $a_0 + T + 1$.

We define the truncated Gompertz function as follows:

$$g_T(a_2,a_1;a_D,m,b) = \frac{g(a_2,a_1;m,b)-g(a_D,a_1;m,b)}{1-g(a_D,a_1;m,b)}. \tag{3A.2}$$

Exhibit 3A.1. Estimated Parameters for the Gompertz Function Based on US Mortality Rates

Parameter	Men	Women
M	88	91
B	10.65	8.88

[30]While that value of T is somewhat arbitrary, it should be chosen so that $g(a_D,a_1;m,b)$ is very close to zero.

If the investor is a single person, we need to use the truncated Gompertz function to calculate survival probabilities based on the sex of the person. Let:

q_v^t = the probability of the person surviving to at least year v, given that the person was alive in year t.

We then have:

$$q_v^t = g_T(a_0 + v, a_0 + t; a_D, m_G, b_G), \tag{3A.3}$$

where the $_G$ subscript denotes that the m and b parameters are gender-specific.

For couples, we calculate the probability that at least one member of the couple will be alive at in year v, given that they were both alive in year t. We call this the *joint survivor* probability. The general formula for the probability of at least one person of a couple composed of two people, A and B, each with a given age, m parameter, and b parameter, surviving at least to year v is as follows:

$$q_v^{tJ} = q_v^{tA} + q_v^{tB} - q_v^{tA} q_v^{tB}, \tag{3A.4}$$

where the superscript ^{J} denotes the joint survivor probability and the superscripts ^{A} and ^{B} denote the two members of the couple.

Appendix 3B. Solving Equations 3.11 and 3.12

Inverse Utility Functions

Our general utility function takes the following form:

$$u_p(x) = \begin{cases} ln(x), & p = 1 \\ \dfrac{p}{p-1}\left(x^{\frac{p-1}{p}} - 1\right), & p \neq 1, \end{cases} \tag{3B.1}$$

where p is the parameter of the utility function in question. Suppose that:

$$y = u_p(x). \tag{3B.2}$$

The inverse of the utility function, $u_p^{-1}(.)$ takes y as the input and gives x as the output:

$$x = u_p^{-1}(y). \tag{3B.3}$$

The inverse of the utility function given in Equation 3B.1 is as follows:

$$u_p^{-1}(y) = \begin{cases} exp(y), & p = 1 \\ \left(\dfrac{1-p}{p} y + 1\right)^{\frac{p}{1-p}}, & p \neq 1. \end{cases} \tag{3B.4}$$

Solving Equation 3.11

From Equation 3.11, we have the following:

$$u_\eta(\hat{c}_t) = \frac{\sum_{v=t}^{T} q_v^t \frac{1}{(1+\rho)^{v-t}} u_\eta(c_v)}{\sum_{v=t}^{T} q_v^t \frac{1}{(1+\rho)^{v-t}}}. \tag{3B.5}$$

We solve for $\hat{c}_t$ using the inverse utility function:

$$\hat{c}_t = u_\eta^{-1}\left(\frac{\sum_{v=t}^{T} q_v^t \frac{1}{(1+\rho)^{v-t}} u_\eta(c_v)}{\sum_{v=t}^{T} q_v^t \frac{1}{(1+\rho)^{v-t}}}\right). \tag{3B.6}$$

Solving Equation 3.12

Equation 3.12 states the following:

$$\frac{1}{n}\sum_{i=1}^{n} u_\theta(\hat{c}_{ti}) = u_\theta(\hat{c}_t). \tag{3B.7}$$

We then solve for $\hat{c}_t$ using the inverse utility function:

$$\hat{c}_t = u_\theta^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} u_\theta(\hat{c}_{ti})\right). \tag{3B.8}$$

4. THE INVESTOR'S BALANCE SHEET WITH HUMAN CAPITAL AND LIABILITIES

Context

Like the previous chapter, in this chapter, we continue to focus on the investor from a pecuniary or financial perspective. This chapter presents further details that will be included in the life-cycle models presented in chapters 5 and 6 by introducing human capital and liabilities and thus creating a more holistic view of the investor and their situation. Our goal in later chapters is to present detailed life-cycle models that capture all of the investor's pertinent pecuniary and nonpecuniary information and preferences, to guide optimal lifetime advice.

Key Insights

- The left-hand side of an investor's balance sheet consists of two primary assets: financial assets and human capital. The right-hand side of the investor's balance sheet consists of the value of liabilities and net worth. Liabilities consist of (1) the economic value of present and future nondiscretionary consumption and (2) the economic value of present and future term life insurance premiums for a bequest (if any). Net worth is the economic value of present and future discretionary consumption.
- The intertemporal budget constraint says that the total value of the left-hand side on an investor's balance sheet must equal the total value of the right-hand side.
- Human capital produces a somewhat uncertain income stream, in which the uncertain income stream is often relatively stable or more bond-like than stock-like. The relative stability of income from human capital and its ability to pay for ongoing consumption provides a form of risk-taking capacity that allows financial capital to be invested in riskier assets.
- For many people, human capital is their single most valuable and important asset that pays for consumption during the accumulation phase, contributing to financial capital through ongoing savings, and continuing to pay for living expenses in retirement as deferred labor income in the form of social insurance (social security) and defined benefit pensions.
- Risk capacity is the extent to which investors can take on market risk in their investments, based on their individual balance sheets. The more bond-like their human capital—and the greater their human capital relative to their financial wealth—the greater their risk capacity. Differences in risk capacity among investors with the same risk tolerance lead to differences in optimal equity allocations.
- Income, its net present value (human capital), nondiscretionary consumption, and its net present value (part of the value of liabilities), evolve through time. In our models with uncertainty, we apply random unexpected shocks (innovations) to each of these to model the wide range of possible future values/levels of each and the probability of each value/level. In other words, all possible realizations can be modeled and can thus be incorporated into probabilistic financial planning decision rules that are based in part on each of these components of the investor's balance sheet.

An Investor's Balance Sheet

When it comes to understanding the health of a business, one of the two most important tools is the balance sheet (the other being the income statement). The same is true when it comes to investors. Businesses have revenues and expenses that relate directly to assets, liabilities, and equity value.

Likewise, investors have income and consumption that relate directly to assets, liabilities, and net worth. In both cases, the left-hand side of the balance sheet accounts for all assets and the right-hand side of the balance sheet accounts for all liabilities and net worth.

We generally divide an investor's assets into two primary buckets—financial assets and human capital. It is relatively straightforward exercise to list all of an investor's financial assets. Bank account balances, brokerage statements, and defined contribution retirement account statements are often a click or two away. The value of real estate, such as a home, rental property, or land, can likely be approximated by Zillow or comparable web sites/apps. The value of private businesses can be estimated. Given the ease with which financial capital can be estimated, in this chapter, we focus on the human capital part of the balance sheet as well as an investor's liabilities.

Similar to the balance sheet presented in chapter 2, **Exhibit 4.1** presents the primary entries one would expect to find on an investor's balance sheet. Any undefined notation will be formally defined later in this chapter.

As explained in chapter 3, we define nondiscretionary consumption as the money that the investor must spend on basic necessities, such as food, shelter, and clothing, and additional spending constitutes discretionary consumption. From Exhibit 4.1, we can see the importance of this distinction. For our purposes, nondiscretionary consumption constitutes the investor's liabilities and, although they are typically not legal liabilities, it is critical that one can pay for these expenses. To fund nondiscretionary consumption, the value of the assets must be at least that of the value of the liabilities. Should the value of the assets exceed that of the liabilities, the investor has positive net worth, which they can use to fund discretionary consumption and possibly a planned bequest. In light of our life-cycle models, which we present in chapters 5 and 6, we have chosen not to label discretionary consumption as a liability given that it is more akin to a want than a need; nevertheless, we recognize that it too can be treated as a type of soft liability.

Drawing up an investor's balance sheet can be a useful step in forming a lifetime financial plan. Starting with the left side of the balance sheet, the first step is to make a list of financial assets and assess their value to come up with $\hat{F}$. The second step is to project lifetime income from its various sources, select an appropriate discount rate (as we will discuss), and estimate the present value of lifetime income to come up with an estimate of human capital. Taken together, financial assets and human capital constitute a person's total assets. A person can use these assets to finance both their nondiscretionary and discretionary spending.

Exhibit 4.1. The Investor's Balance Sheet

Assets	Liabilities & Net Worth
Financial Assets ($\hat{F}$) • Bank Accounts • Brokerage Accounts • Real Estate (Home, Land, etc.) • Existing Annuities • Other	**Liabilities ($\hat{L}$)** • Present Value of Nondiscretionary Consumption • Present Value of Term Life Insurance Premiums
Human Capital ($\hat{H}$) • Present Value of Wage Income • Present Value of Income from a Defined Benefit Plan • Present Value of Income from Government Sponsored Social Insurance	**Net Worth ($\hat{W}$)** • Present Value of Discretionary Consumption

Moving to the right side of the balance sheet, the next step is to project lifetime *nondiscretionary* consumption. In financial planning, the distinction between nondiscretionary and discretionary consumption is essential. Nondiscretionary consumption includes any spending that must be funded in all circumstances. It includes all of the essentials such as food, clothing, and shelter. A financial plan should include creating a budget to make sure that nondiscretionary consumption can be covered using available income and assets.

The present value of nondiscretionary consumption is the part of the liabilities that the assets, including human capital, *must* cover. The present value of the assets, net or minus the value of the liabilities, is the net worth of the investor. Net worth is the amount of wealth available to the investor for discretionary consumption. Discretionary consumption includes all spending beyond nondiscretionary consumption. It includes items such as vacation trips, dining out, the purchase of luxury items, and paying for a grandchild's tuition. As long as the investor has sufficient income and net worth to pay for these expenses, they can be incorporated into the investor's lifetime financial plan. In practice, distinguishing between nondiscretionary consumption and discretionary consumption can be somewhat arbitrary.

Human Capital

For most people, the single largest source of income over the course of their lives are their careers. Based on the work of Nobel laureate Gary Becker (1993), financial economists view the present value of all of a person's future labor income as the value of an asset—namely, human capital. For many people, human capital is their single most valuable asset. From their labor income, people not only pay for consumption during their working years, but also accumulate the financial capital that they will need to fund retirement. Savings during their working years, out of the income from human capital, enables financial capital to be accumulated and invested for decades—weathering turbulent markets and growing over time. In addition to paying for nondiscretionary consumption during retirement, many people also save for other goals, such as paying for a grandchild's education or going on a dream vacation. Beyond discretionary consumption, many people want to leave a bequest.

For people who are still working, an important part of financial planning is projecting future wage income. At Morningstar, we developed a model of annual wage income in which wages not only vary by age but also by education and gender. The model has four levels of education:

1. generic (for when the education level is unknown),
2. high school,
3. college (four-year undergraduate degree), and
4. post college (graduate degree).

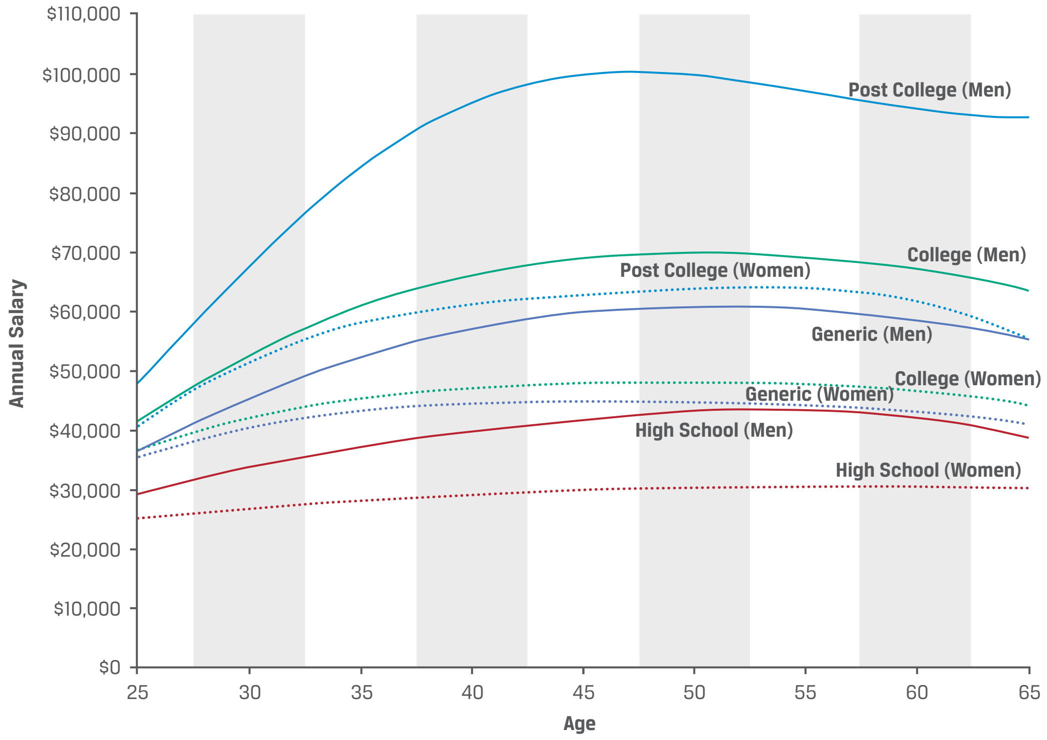
Exhibit 4.2 plots the model's projection of annual wage income for men and women at these education levels between the ages of 25 and 65.

If the salary of the investor is known, we rescale the projections of this model by multiplying the projected salary by the ratio of the known salary to the model's value for the investor's current age.

In chapter 2, we noted that Paula's financial planning system had estimated that the value of Isabela's human capital was $2,767,689.[31] Recall that Isabela, currently age 25, had recently completed her master's degree in marine biology and was earning $75,000 per year. Notice that the starting point for the salary

[31]We will discuss the distinction between human capital with and without mortality weighting, the former being for when the investor will be using annuities at retirement. Since we will be assuming that Isabela will be using annuities at retirement, the figure we show here is with mortality weighting.

Exhibit 4.2. Projected Annual Wages for Ages 25–65, by Gender and Education Level



curve for a post-college woman at age 25 is approximately $40,000 for a ratio of $75,000/$40,000 or 1.875. Thus, when estimating Isabela's lifetime human capital of $2,767,689, the planning software applied the post-college salary curve for women shown in Exhibit 4.2 and multiplied it by 1.875. The estimate of human capital also captures the net present value of assumed social insurance (Social Security) payments for life.

Nontradeable Sources of Retirement Income

After a person retires, they may have some income that at least in part, replaces their wages. These sources of retirement income can include the following:

1. Income from a defined benefit plan. This is a lifetime source of income that functions like an annuity.
2. Government-sponsored social insurance. Many countries provide lifetime retirement income to their residents on the basis of salary history, contributions, years of residence, starting age, and other factors.[32]
3. Income from a preexisting annuity. If the investor purchased an annuity before seeking advice, we count the income from that annuity as part of exogenous (external) retirement income.

[32]In the United States, this is Social Security. In Canada, this is the Québec Pension Plan (for those who have worked in Québec) and the Canada Pension Plan (for all other workers). Canada also has Old Age Security and the Guaranteed Income Supplement, which are not tied to salary history or contributions. Other countries have similar social insurance systems.

4. Inheritances and life settlements from someone else's life insurance policy or as a beneficiary of a period certain annuity. We also count income from nonpassive investments, such as real estate, and the cash value of any insurance policies one might own.

All four of these nontradeable sources of retirement income can be thought of assets on the left-hand side of an investor's balance sheet. Conceptually, the first two are forms of *deferred* labor income and thus a form of human capital that one earns or accrues from working and is paid out in retirement contingent on the person being alive. From a purist perspective, given that annuities are purchased, we typically think of them as financial assets. From both a contingent and observable cash flow perspective during retirement, the payout characteristics of IFAs are extremely similar to those of defined benefit pensions and social insurance payments. As such, it is often computationally convenient to group them with human capital and thus include annuities in that category.

Lifetime Income and Human Capital

As we will discuss, the risk of exogenous income can be modeled using a portfolio of risky assets. Let:

k_y = the expected return on the portfolio used to model risk (risk-free rate if no market risk);[33]

c_v = consumption in year v;

y_v = exogenous income in year v, either from wages prior to retirement or from the exogenous retirement income sources;

t = the current year;

F_t = the level of financial wealth as of year t;

H_t = the value of human capital in year t, which we will define; and

T = the final year to consider.

Letting k denote the discount rate on consumption, the investor's lifetime consumption, when there are no annuities, is subject to an intertemporal budget constraint, which in principle is like the following:

$$\underbrace{\sum_{v=t}^{T} \underbrace{\frac{1}{(1+k)^{v-t}}}_{\text{Time Value Discount}} \underbrace{c_v}_{\text{Consumption in year } v}}_{\text{Lifetime Consumption}} = \underbrace{\underbrace{F_t}_{\text{Financial Wealth at time } t} + \underbrace{H_t}_{\text{Human Capital Value at time } t}}_{\text{Total Assets}}. \tag{4.1}$$

Although Equation 4.1 may look complicated, it simply says that the summed present value of lifetime consumption (the left-side of the equation) must always equal the combined value of financial assets and human capital (the right-side of the equation). This type of intertemporal budget constraint, when incorporated into a life-cycle model, grounds the solution in reality: Beyond some modest level of debt, one cannot dramatically consume more than their total assets.

[33]In chapter 5, we assume that there is no market risk, so k_y is the risk-free rate. In chapter 6, we introduce market risk, so k_y is the expected return on the portfolio of risk assets used to model the risk of income. When there is both risky income and riskless income, we calculate risky human capital using k_y and riskless human capital using the risk-free rate, the sum of the two being total human capital. In the case of Isabela, we model her employer's matching contribution to their retirement plan to be the maximum amount of $6,833 allowed by Isabela's plan, treat it as riskless income, and include its present value as part of human capital in the exhibits in this chapter and in the remainder of this book. We assume that each year, Isabela contributes $13,667 to her retirement plan to receive the maximum match, and either withdraws from or adds to her taxable account to get to her optimal level of consumption.

Although the value of financial wealth is mostly observable through various bank statements, account statements, and estimated asset values, the capitalized or present value of human capital is not readily available and must be estimated. Human capital (without mortality weighting) is the present value of all labor-based income:

$$H_t = \underbrace{\sum_{v=t}^{T} \underbrace{\frac{1}{(1+k_y)^{v-t}}}_{\substack{\text{Time Value}\\ \text{Discount}}} \underbrace{y_v}_{\substack{\text{Labor Income}\\ \text{in year } v}}}_{\text{Human Capital}}. \tag{4.2}$$

Notice that the *right-hand* side of Equation 4.2 is nearly identical to the *left-hand* side of Equation 4.1. The only difference is the final term in each expression, where y_v is annual labor-related income (e.g., salary, defined benefit payments, Social Security payments) being summed and c_v is the annual consumption being summed. Intuitively, if annual consumption c_v is *lower* than annual labor-related income y_v, then financial wealth and thus net worth *increase*. Conversely, if annual consumption c_v is *greater* than annual labor-related income y_v, then financial wealth and thus net worth decrease.

Equations 4.1 and 4.2 assume that the investor has no access to annuities (beyond what they might have purchased before year *t*). Suppose that once the investor retires, there is a complete annuities market so that at time *t*, for each year *v*, the investor can buy fairly priced annuities that pay one real dollar in year *v*, if the investor is still alive (or if at least one person in a couple is still alive). Letting *a* denote the year of retirement, when calculating the present value of a cash flow in year *v* back to year *t*, following are the three possibilities regarding mortality weighting when annuities are available during retirement:

1. *t* and *v* are before retirement ($t \le v < a$). In this case, there is no mortality weighting.
2. *t* is before retirement and *v* is during retirement ($t \le a \le v$). In this case, there is mortality weighting from year *a* to year *v*.
3. *t* and *v* are during retirement ($a < t \le v$). In this case, there is mortality weighting from year *t* to year *v*.

As shown in chapter 3, we denote the probability of surviving from year *t* to at least year *v* as q_v^t. To account for the three possibilities just listed, we define a set of mortality weights, based on the values of q_v^t, for calculating present values—namely, the following:

$$\hat{q}_v^t = \begin{cases} 1, & \text{if } t \le v < a \\ q_v^a, & \text{if } t \le a \le v \\ q_v^t & \text{if } a < t \le v. \end{cases} \tag{4.3}$$

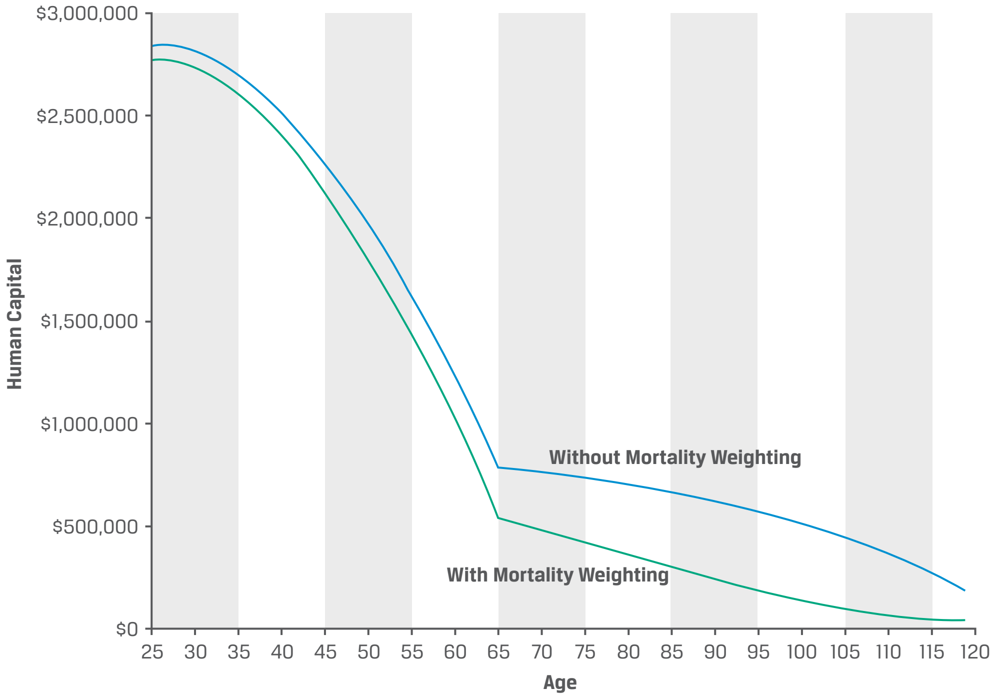
Hence, the intertemporal budget constraint is as follows:

$$\underbrace{\sum_{v=t}^{T} \underbrace{\hat{q}_v^t}_{\substack{\text{Mortality}\\ \text{Weight}}} \underbrace{\frac{1}{(1+k)^{v-t}}}_{\substack{\text{Time Value}\\ \text{Discount}}} \underbrace{c_v}_{\substack{\text{Consumption}\\ \text{in year } v}}}_{\text{Lifetime Consumption}} = \underbrace{\underbrace{\hat{F}_t}_{\substack{\text{Financial}\\ \text{Wealth at}\\ \text{time } t}} + \underbrace{\hat{H}_t}_{\substack{\text{Human Capital}\\ \text{Value at time } t}}}_{\text{Total Assets}}, \tag{4.4}$$

where $\hat{H}_t$ is human capital with mortality weighting:

$$\hat{H}_t = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+k_y)^{v-t}} y_v. \tag{4.5}$$

Exhibit 4.3. The Evolution of Expected Human Capital with and without Mortality Weighting for Isabela (25-Year Old Woman with Post-College Education)



As we will discuss in chapter 5, when annuities are available, financial assets take the form of annuities; so, we denote it with a hat $\hat{F}_t$ to distinguish it from financial wealth when annuities are not available, F_t.

Exhibit 4.3 shows the evolution of Isabela's expected human capital with and without mortality weighting. It demonstrates how expected human capital with mortality weighting can be quite a bit less than without mortality weighting because the weighting scheme lowers the value of each term of the sum.

Interestingly and somewhat unintuitively for many, the value of expected human capital depicted in Exhibit 4.3 temporarily increases and then begins a long steady decrease. This initial increase occurs because wage growth during the early years of employment outpaces the impact of the discount rate and increasing age.

Risky Human Capital

Thus far, we have only taken the riskiness of human capital into account when setting the discount rate, k. Human capital is subject to unpredictable changes. As Milevsky (2012b) discusses, and as we indicated when discussing Isabela's human capital, the degree of uncertainty depends on the nature of a person's income. For example, a stockbroker's income is tied to the stock market, which consists of a fixed base income and a variable component that is highly linked to the stock market. One might choose to model or represent a stockbroker's human capital asset as a mix of 45% stocks and 55% bonds. In contrast, a tenured university professor's income is subject to little risk. One might choose to model or represent a

Exhibit 4.4. Capital Market Assumptions for Modeling Human Capital

			Correlation with:		
	Expected Return	Standard Deviation	Domestic Stocks	International Stocks	Bonds
Domestic Stocks	4.72%	15.88%	1.00		
International Stocks	5.04%	17.18%	0.87	1.00	
Bonds	2.75%	5.62%	0.21	0.37	1.00

Exhibit 4.5. Risk and Expected Return of Stock/Bond Mixes Representing Different Investors

Investor	Domestic Stocks	International Stocks	Bonds	Expected Return	Standard Deviation
Isabela	15%	5%	80%	3.19%	6.13%
Tenured Professor	10%	0%	90%	2.98%	5.61%
Stockbroker	30%	15%	55%	3.70%	8.51%

tenured university professor's human capital asset as a mix of 10% stocks and 90% bonds. The discount rate should be personalized based on the nature of the income, as Blanchett and Straehl (2015) discuss. We model Isabela's to be that of an asset mix consisting of 20% stocks and 80% bonds.[34]

Exhibit 4.4 shows the capital market assumptions for the three asset classes (domestic stocks, international stocks, and bonds) that we use to model the riskiness of human capital (and as we discuss later, the riskiness of some liabilities).[35] Based on the capital market assumptions shown in Exhibit 4.4, **Exhibit 4.5** and **Exhibit 4.6** show the expected returns and standard deviations of different combinations of the three asset classes for modeling the riskiness of human capital.

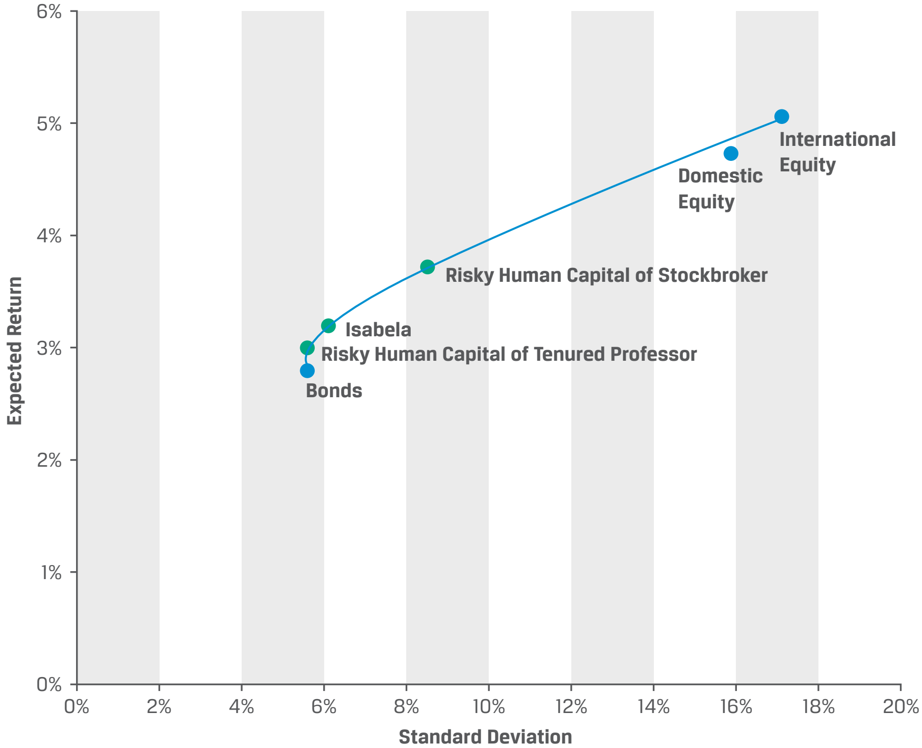
Under these assumptions, the 20/80 asset mix representing the riskiness of Isabela's income has an expected (real) return of 3.19% and a standard deviation of 6.13%. We denote the expected return and standard deviation of the representative portfolio by k_y and s_y, respectively.

To model income as being risky, we take the model of income we described as expected income and introduce statistical noise. Let y_v^t denote expected income level of income in year v as of year t. To introduce

[34]In our models, for simplicity, we assume that the degree of riskiness of human capital (and of consumption-related liabilities as we will discuss) remains constant over time, when in fact, it changes. Straehl, ten Brincke, and Gutierrez Mangas (2023) explore modeling company-specific human capital.

[35]In the Monte Carlo model discussed in chapter 6, there are three random variables; thus, we create a three-asset class model using the capital market assumptions shown in Exhibit 4.4. This three-asset class model is derived from the 10-asset class model that we discuss in chapters 7 and 8.

Exhibit 4.6. Risk and Expected Return Plot of Stock/Bond Mixes Representing Different Investors' Human Capital



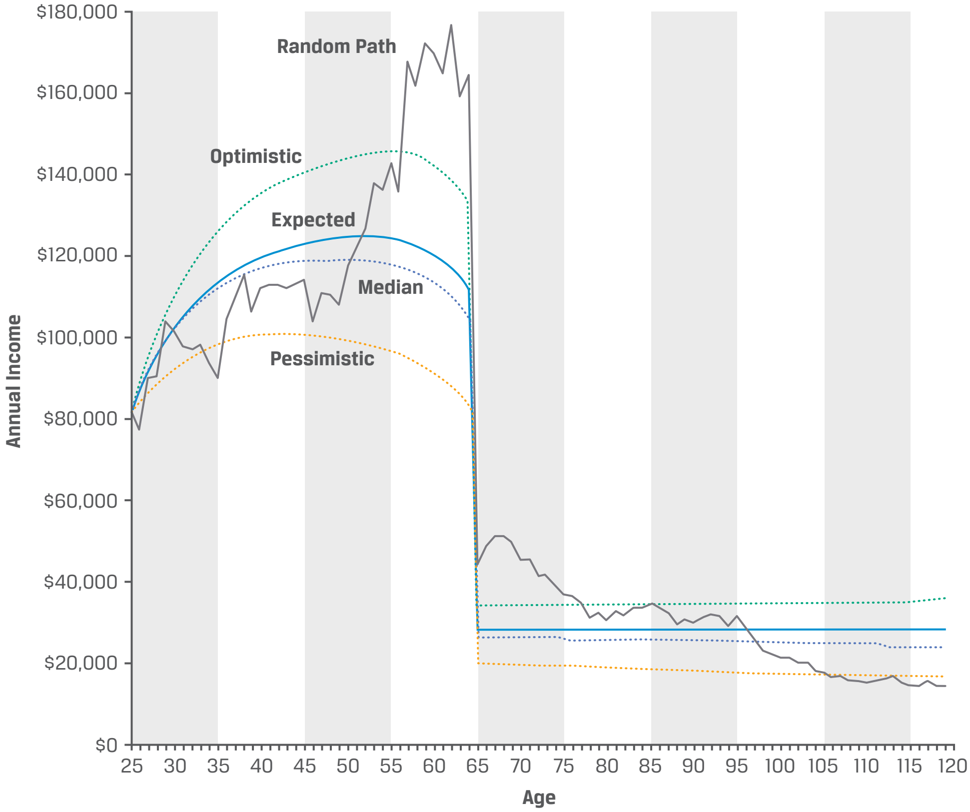
statistic noise, we assume that each year $t + 1$, the expected values as of year t of all future years are multiplied by the same statistical innovation (i.e., independent shock), $\tilde{I}_{t+1}^{y}$ so that:

$$y_v^{t+1} = \tilde{I}_{t+1}^{y} y_v^{t}. \quad (4.6)$$

The expected value of $\tilde{I}_{t+1}^{y}$ is 1. To be consistent with an expected return of k_y and standard deviation of s_y, the standard deviation is $\frac{s_y}{1+k_y}$. Importantly, these random shocks allow us to model a wide variety of potential income paths representing the distribution of possible outcomes.

The impacts of the annual innovations are cumulative, so that deviations from the expected values potentially can grow over time. To illustrate this, **Exhibit 4.7** shows the expected income of a 25-year-old woman with a post-college education (Isabela) as well as the 25th, 50th, and 75th percentiles of possible future income levels (the smooth lines). We label these three percentile curves "Pessimistic," "Median," and "Optimistic," respectively. This exhibit also shows a randomly generated possible future income path (the black line with significant variability). Importantly, the random unexpected shocks (innovations) enable us to quantify the wide range of possible future income levels and the probability of each income level.

Exhibit 4.7. Percentiles of Possible Future Income of Isabela (25-Year-Old Woman with a Post-College Education)



In other words, a wide range of possible realizations are modeled and thus can be incorporated into probabilistic financial planning decisions that are influenced by different income levels.

When income is risky, so too is human capital (because it is just a net present value of future income). The basic principle of valuation, that the net present value of an asset is the sum of all future cash flows discounted at an appropriate discount rate that reflects their uncertainty, applies. The discount rate for risky human capital is the expected return on the representative stock/bond asset mix, k_y. Thus, letting H_v^t denote expected human capital without mortality weighting in year v as of year t, we have the following:

$$H_t^t = \sum_{v=t}^{T} \frac{1}{(1+k_y)^{v-t}} y_v^t. \tag{4.7}$$

Note that Equation 4.6 is similar to Equation 4.2, except that the discount rate in Equation 4.7 is specific to income and human capital. Letting $\hat{H}_v^t$ denote expected human capital with mortality weighting, we have the following:

$$\hat{H}_t^t = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+k_y)^{v-t}} y_v^t. \tag{4.8}$$

Again, note that Equation 4.8 is similar to Equation 4.4, except that the discount rate in Equation 4.8 is specific to income and human capital. The innovations affect human capital the same way that they affect income levels:

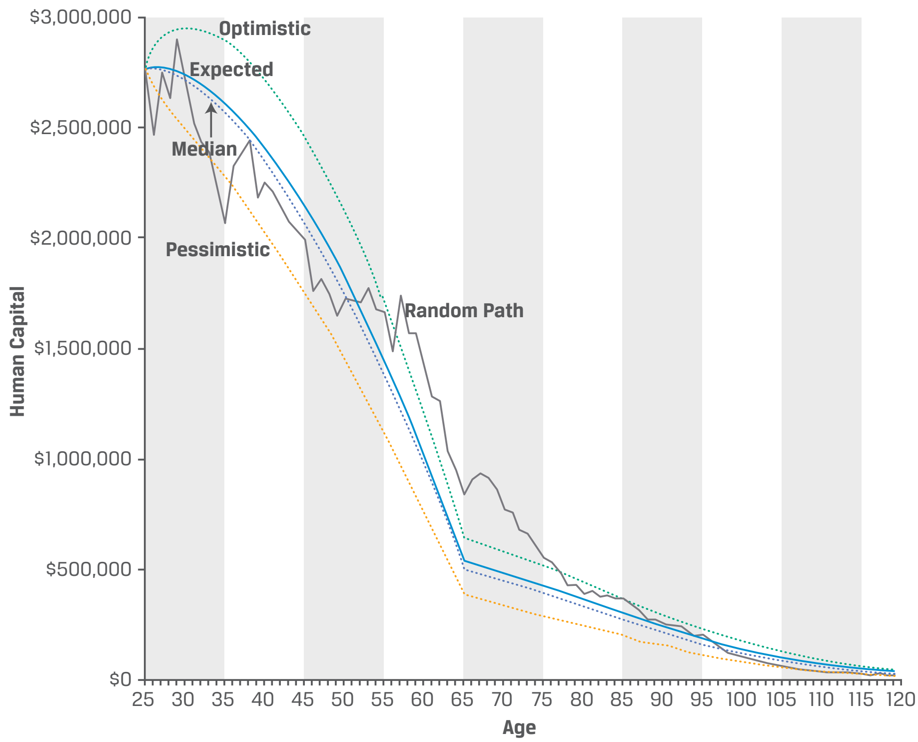
$$H_v^{t+1} = \tilde{I}_{t+1}^{y} H_v^{t}, \tag{4.9}$$

$$\hat{H}_v^{t+1} = \tilde{I}_{t+1}^{y} \hat{H}_v^{t}. \tag{4.10}$$

As with income, the impacts of the annual innovations on human capital are cumulative, and therefore, deviations from the expected values can potentially grow over time. We illustrate this in **Exhibit 4.8**, which shows the expected human capital (with mortality weighting) of a 25-year-old woman with a post-college education (Isabela) along with percentiles of possible paths of human capital. This exhibit also shows a randomly generated possible future path of human capital.

As we saw with income levels in Exhibit 4.6, moving from possible human capital (with mortality weighting) values in Exhibit 4.8, the random unexpected shocks (innovations) enable us to model a wide range of possible future human capital (with mortality weighting) values and the probability of each value. In other words, a wide range of possible realizations have already been modeled and can thus be incorporated into probabilistic financial planning decisions that are influenced by different human capital (with mortality weighting) values.

Exhibit 4.8. Percentiles of Possible Human Capital (with Mortality Weighting) of Isabela (25-Year-Old Woman with a Post-College Education)



Incorporating the Uncertainty and Variability of Nondiscretionary Consumption

As with human capital, we can model nondiscretionary consumption and consumption-related liabilities as being risky. In the same way that we introduced random unexpected shocks (innovations) to income, we introduce them for nondiscretionary consumption. In this example, we assume that nondiscretionary consumption and consumption-related liabilities behave like a mix that is 15% stocks and 85% bonds. Under our assumptions for stock and bond returns, the 15/85 asset mix has an expected real return of 3.07% and a standard deviation of 5.77%. We denote the expected return and standard deviation of the representative asset mix for nondiscretionary consumption and consumption-related liabilities by k_c and s_c, respectively.

Let $\overline{c}_v^t$ denote expected *nondiscretionary* consumption in year v as of year t. We assume that each year $t+1$, the expected values as of year t of all future years are multiplied by the same innovation (independent shock), $\tilde{I}_{t+1}^c$, so that:

$$\overline{c}_v^{t+1} = \tilde{I}_{t+1}^c \overline{c}_v^t. \tag{4.11}$$

As is the case with $\tilde{I}_{t+1}^y$, the expected value of $\tilde{I}_{t+1}^c$ is 1. To be consistent with an expected return of k_c and standard deviation of s_c, its standard deviation is $\frac{s_c}{1+k_c}$.

Multiplying the evolving nondiscretionary consumption by random unexpected shocks (innovations) enables us to model a wide range of possible future levels of nondiscretionary consumption and the probability of each of the levels of nondiscretionary consumption. A wide range of possible realizations has already been modeled and thus can be incorporated into probabilistic financial planning decisions that are influenced by different income levels.

As is the case with income, the impacts of the annual innovations are cumulative, so that deviations from the expected values can potentially grow over time. To illustrate this, **Exhibit 4.9** shows the 25th, 50th, and 75th percentiles of possible future levels of nondiscretionary consumption. Treating nondiscretionary consumption as a "bad," we label these percentiles as "Optimistic," "Median," and "Pessimistic," respectively. This exhibit also shows a randomly generated possible future path. Note the similarity between the cumulative innovations in this exhibit and those in Exhibit 4.7 because of the high correlation between the innovations in income and the innovations in nondiscretionary consumption.

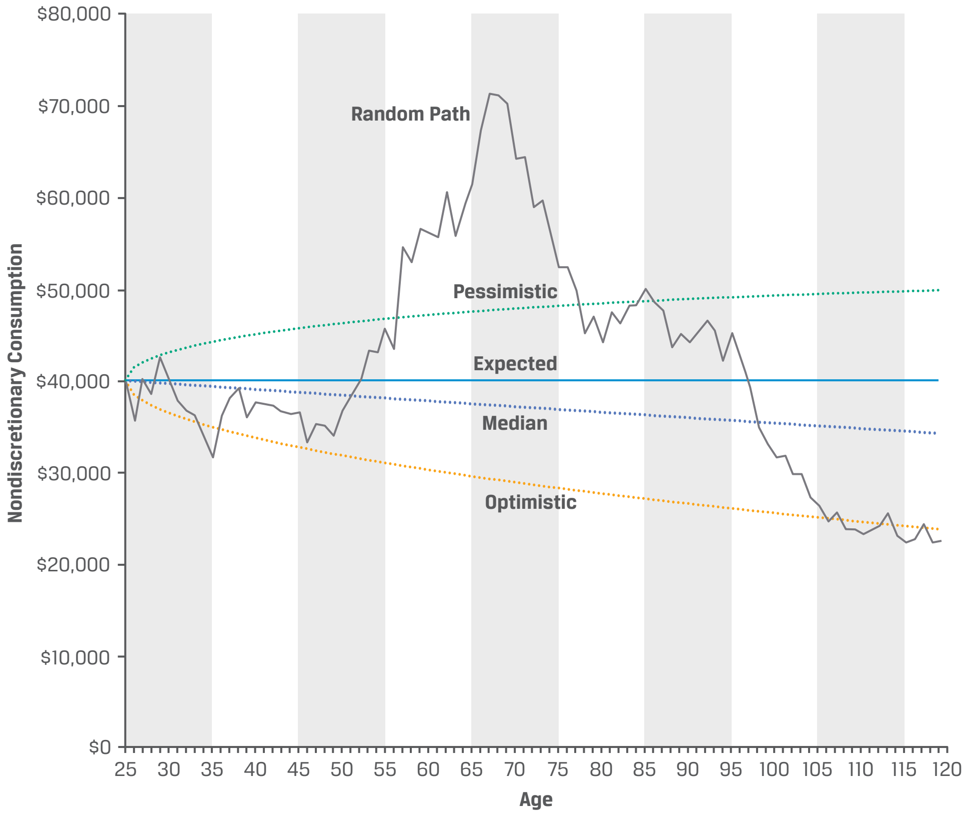
Liabilities

We can now move from the left side of an investor's balance sheet representing the investor's assets—financial capital and human capital—to the right side of the balance sheet that represents the investor's liabilities. As Exhibit 4.1, shows, we model two types of liabilities: those arising from nondiscretionary consumption, and those arising from series of annual term life insurance purchases to guarantee a bequest. We call the first type consumption-related liabilities and the second type life insurance–related liabilities. We next consider each in turn.

Consumption-Related Liabilities

We treat the present value of current and future expected nondiscretionary consumption as the value of consumption-related liabilities. In some contexts, discretionary consumption can be thought of as a liability, albeit a very soft type of liability; however, in the context of life-cycle models, it is essential to keep it distinct from liabilities.

Exhibit 4.9. Percentiles of Possible Future Nondiscretionary Consumption



When it comes to funding an investor's liabilities, if there is an interruption to income during the investor's working years, the investor will need an alternative source of income to pay for nondiscretionary consumption. Similarly, in retirement, to the degree that the four nontradeable sources of retirement income (income from a defined benefit plan, government-sponsored social insurance, income from preexisting annuities, and inheritances and life settlements) mentioned earlier do not fully pay for retirement consumption, an alternative source of income to pay for consumption is needed. The alternative sources of income include a drawdown of accumulated financial assets, and the possibility of borrowing (presumably against future income from either financial capital or human capital).

Recall from chapter 3, that we denote *nondiscretionary* consumption in year v by $\overline{c}_v$, which for now we assume is deterministic. Hence, we can write *discretionary* consumption as $c_v - \overline{c}_v$. That is, discretionary income is the difference between actual spending, c_v, in year v and the nondiscretionary part, $\overline{c}_v$, in year v. Denoting L_t^C as the value of consumption-related liabilities (present value of nondiscretionary consumption) without mortality weighting, in the absence of life insurance, we can rewrite the pro forma intertemporal budget constraint (without annuities available) as follows:

$$\sum_{v=t}^{T} \frac{1}{(1+k)^{v-t}} (c_v - \overline{c}_v) = F_t + H_t - L_t^C, \tag{4.12}$$

where L_t^C is the value of consumption-related liabilities without mortality weighting:

$$L_t^C = \sum_{v=t}^{T} \frac{1}{(1+k_c)^{v-t}} \overline{c}_v. \tag{4.13}$$

If annuities are available at retirement, the pro forma intertemporal budget constraint is as follows:

$$\sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+k)^{v-t}} (c_v - \overline{c}_v) = \hat{F}_t + \hat{H}_t - \hat{L}_t^C, \tag{4.14}$$

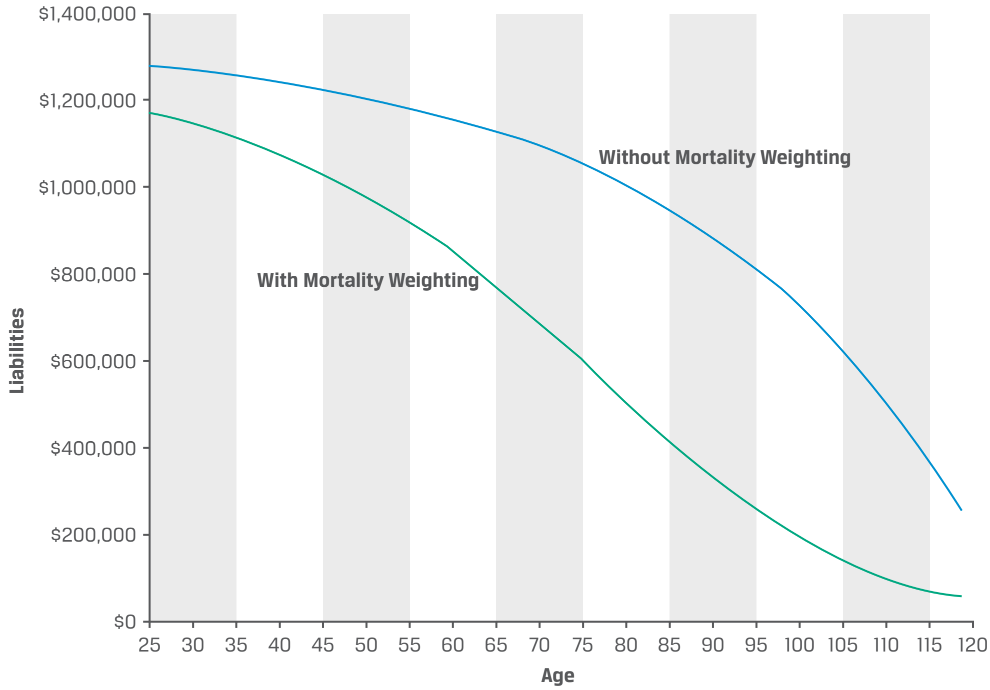
where $\hat{L}_t^C$ is the value of consumption-related liabilities with mortality weighting:

$$\hat{L}_t^C = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+k_c)^{v-t}} \overline{c}_v. \tag{4.15}$$

Notice that Equation 4.13 is nearly identical to Equation 4.11 and Equation 4.14 is nearly identical to Equation 4.12, with the addition of q_v^t incorporated into the equations to reflect the uncertainty of being alive.

In the case of Isabela, we assume that Isabela's expected nondiscretionary consumption is a constant real $40,000 per year. **Exhibit 4.10** shows the evolution of expected consumption-related liabilities (the net present value of nondiscretionary consumption) with and without mortality weighting under

Exhibit 4.10. The Evolution of Expected Consumption-Related Liabilities with and without Mortality Weighting



this assumption. It also shows how expected liabilities with mortality weighting can be quite a bit less than without mortality weighting, just as we saw with human capital.

When nondiscretionary consumption is risky, so too is the value of consumption-related liabilities. The discount rate for consumption-related liabilities is the expected return on the representative portfolio for nondiscretionary consumption, k_c. So, letting L_v^{Ct} denote the expected value of consumption-related liabilities without mortality weighting in year v as of year t, we have the following:

$$L_t^{Ct} = \sum_{v=t}^{T} \frac{1}{(1+k_c)^t} \overline{c}_v^t. \tag{4.16}$$

Letting $\hat{L}_v^{Ct}$ denote the value of consumption-related liabilities with mortality weighting in year v as of year t, we have the following:

$$\hat{L}_t^{Ct} = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+k_c)^t} \overline{c}_v^t. \tag{4.17}$$

The innovations affect consumption-related liabilities the same way that they affect income:

$$L_v^{Ct+1} = \tilde{I}_{t+1}^{y} L_v^{Ct}, \tag{4.18}$$

$$\hat{L}_v^{Ct+1} = \tilde{I}_{t+1}^{y} \hat{L}_v^{Ct}. \tag{4.19}$$

The impacts of the annual innovations on human capital are cumulative, so that deviations from the expected values potentially can grow over time. We illustrate this in **Exhibit 4.11**. This exhibit shows the expected value of consumption-related liabilities (with mortality weighting) with 25th, 50th, and 75th percentiles of possible paths of consumption-related liabilities, labeled "Optimistic," "Median," and "Pessimistic," respectively. All paths start when Isabela is 25 years old and the value of her consumption-related liabilities is $1,171,977. In contrast with the previous charts showing income, human capital, and consumption in which higher amounts were associated with the optimistic scenario, when we move to liabilities, the optimistic scenario corresponds to lower liabilities. This exhibit also shows a single randomly generated possible future path of consumption-related liabilities.

The riskiness of liabilities has an important implication for how to invest. Namely, it means that the investor needs to dedicate a portion of investments to a portfolio that matches the changing values of the liabilities.

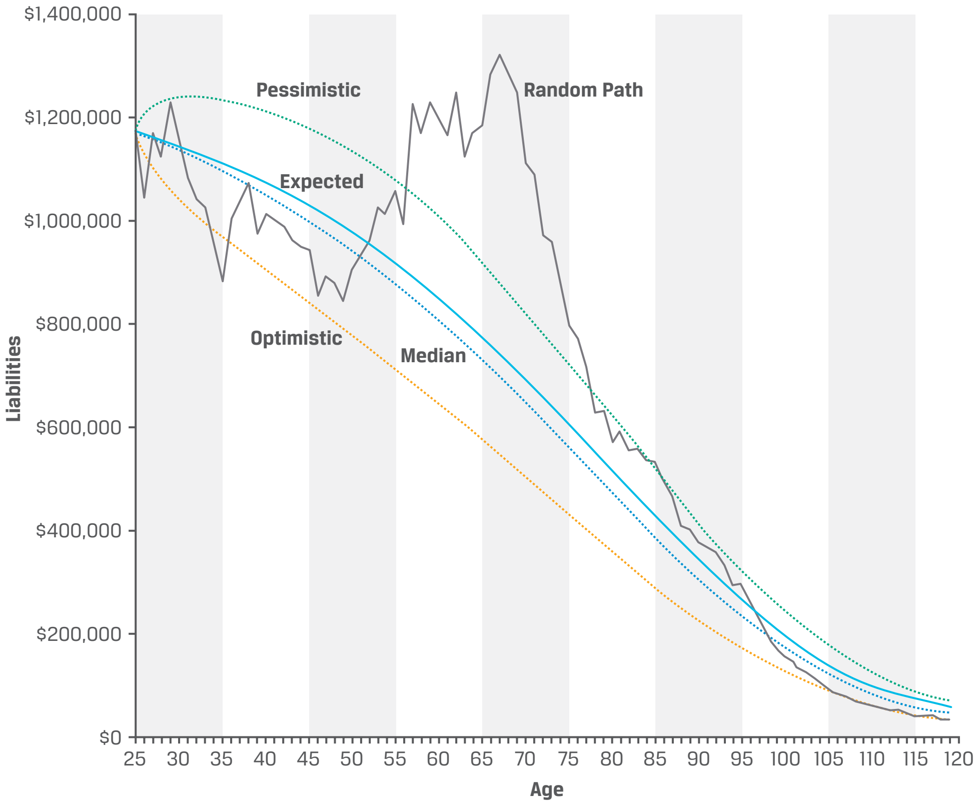
Life Insurance–Related Liabilities

As we will see in chapters 5 and 6, Isabela decides to leave a real bequest just over $1,000,000. To do this, she plans on buying a term life insurance policy and renewing it every year. We assume that she can buy this policy at a fair price. Let $LI1_{t-1}$ denote the price $1 of life insurance paid in year $t-1$ for year t:

$$LI1_{t-1} = \frac{1-q_t^{t-1}}{1+r}, \tag{4.20}$$

where q_t^{t-1} is the probability of surviving from year $t-1$ to year t, and r is the risk-free rate. Letting B denote the size of the bequest, the premium for term life insurance in year $t-1$ is $LI1_{t-1}B$. This premium has a similar impact on cash flows as nondiscretionary spending, so we treat it the same way.

Exhibit 4.11. Percentiles of Possible Values of Consumption-Related Liabilities (with Mortality Weighting)



The life insurance–related liability is the present value of the term premiums. As we discuss in chapter 5, this is equal to the amount of money needed to guarantee the bequest in a single lump-sum payment. To calculate this, we first calculate the probability of the investor dying in year v (given that the investor is alive in year t). We denote this p_v^t. We have the following:

$$p_v^t = q_{v-1}^t - q_v^t. \tag{4.21}$$

The lump-sum price for guaranteeing a bequest of $1 is as follows:

$$LI_t = \sum_{v=t}^{T} \frac{1}{(1+r)^{v-t}} p_v^t. \tag{4.22}$$

Thus, the life insurance–related liability is LI_tB.

Risk Capacity

In chapter 6, we discuss how the investor's risk tolerance determines the optimal risk level for that investor's net worth. From the investor's balance sheet, we see that the risk to net worth depends on the risk of financial assets, human capital, and liabilities. The risk of human capital and of liabilities may be beyond the investor's control. As discussed in chapter 2, this leaves the risk of financial assets as the only lever or dial available for setting the risk of net worth.

This gives rise to the concept of risk capacity. Consider the three investors in Exhibit 4.6: the stockbroker, the typical investor (Isabela), and the tenured university professor. Suppose that they differ only in the risk of human capital but have the same risk tolerance. Because the stockbroker has the riskiest human capital, they are the most constrained of the three in terms of how much risk they can take in their financial assets to avoid having a net worth that is too risky. We say that they have the lowest risk capacity of the three. The tenured university professor is at the opposite extreme. To get their net worth to the right level of risk, they need to take on the most risk of the three in their financial assets. We say that they have the greatest risk capacity of the three. The typical investor falls somewhere between these extremes and has moderate risk capacity. Importantly, by modeling the riskiness of human capital, we capture an important element of risk capacity. It is then by applying risk tolerance at a broad, holistic level that we can and should arrive at an appropriate and personalized risk level.

Conclusion and Key Takeaways

In chapter 3, we covered the inherent characteristics, assumed by economists, of an investor that go into the formation of an optimal financial plan. In this chapter, we have done the same thing for the financial characteristics of the investor that go into the formation of a financial plan. We did this using the investor's balance sheet, which shows that the value of an investor's assets must equal the value of the investor's liabilities plus net worth. The investor's assets consist of financial assets, such as securities and real estate, and human capital, which is the present value of present and future exogenous income. These assets together must be able to finance the investor's liabilities (i.e., the present value of present and future nondiscretionary consumption and the cost of life insurance).

The difference between the value of assets and the value of liabilities is net worth. The investor finances discretionary consumption out of net worth. As we will see in the next two chapters, the main task of lifetime financial planning is making optimal consumption and investment decisions based on the investor's preferences, survival probabilities, needs, and the intertemporal budget constraint as shown in the investor's balance sheet.

5. LIFE-CYCLE MODELS WITH A DETERMINISTIC MARKET

Context

While continuing to focus on the investor from a pecuniary or financial perspective, in this chapter, we bring together the investor's pecuniary preferences and other intrinsic characteristics of the investor that we discussed in chapter 3 as well as the investor's balance sheet and intertemporal budget constraint that we discussed in chapter 4 to determine the investor's optimal *consumption path*. In this chapter, we assume a constant rate of investment return. In chapter 6, we introduce risky investment returns.

Key Insights

- We show how to use a life-cycle model to solve for the investor's optimal path or schedule of consumption by maximizing the lifetime utility function that we introduced in chapter 3, subject to a lifetime intertemporal budget constraint.
- If annuities are *not* available, the investor should schedule or plan lifetime consumption that is somewhat greater in the upcoming years and somewhat less in their later years because the likelihood of being alive decreases over time.
- If annuities are available at retirement, the investor does not need to consider the likelihood of being alive when making consumption decisions once retired. Discretionary consumption can grow or shrink at a steady rate no matter how long the investor lives from that point forward.
- If the investor would like to leave a bequest, they can do so using life insurance.
- During their working years, the investor can accumulate wealth in a regular account (without annuities) and buy term life insurance to make up the difference between the bequest target and the amount in this account.
- Once the amount in the regular account reaches the target bequest, and once the investor is retired, the investor can stop accumulating wealth in the account, stop purchasing life insurance, and switch to buying annuities.
- A trade-off exists between the size of the bequest and the level of consumption.
- The bequest-related preference parameters that we discussed in chapter 3 define the intergenerational utility function. By optimizing the intergenerational utility function, subject to the trade-off between the size of the bequest and the level of consumption, we find the optimal level for the bequest.

Optimal financial planning involves solving the main life-cycle finance problem of deciding how much to save or withdraw from savings each year, how much to spend each year, and how to invest. In this chapter, we explain how to solve the main life-cycle finance problem, albeit under the simplifying assumption of a constant rate-of-investment return. In the next chapter, we expand the model to include risky investment returns.

We consider three cases:

1. *The base case, without annuities and without life insurance*. In this case, the investor does *not* have access to either annuities or life insurance. Additionally, the investor funds all consumption with

exogenous income (i.e., income that the model takes as a given, such as labor income), plus income that can be generated using financial wealth.

2. *The annuity case*. In this case, we assume that once retired, the investor has access to fairly priced single-premium (fixed) immediate annuities (SPIAs). This allows the investor to have guaranteed lifetime income and thus avoid facing longevity risk.[36] In the next chapter, we incorporate return variability into the models and then will discuss single-premium *variable* immediate annuities.
3. *The life insurance case*. In this case, we continue to assume that once retired, the investor has access to SPIAs but also would like to leave a bequest. Life insurance allows the investor to guarantee a bequest of a given size. In this case, we show how the investor can jointly select the level of consumption and the size of the bequest based on the intergenerational utility function that we introduced in chapter 3.

Intertemporal Utility: The Optimal Lifetime Consumption Schedule

In all three cases, the investor seeks to maximize the intertemporal utility function that we introduced in chapter 3. That is, the investor seeks to maximize intertemporal utility by selecting the sequence of consumption:

$$\max_{c_t, c_{t+1}, \cdots, c_T} \underbrace{\sum_{v=t}^{T} \underbrace{q_v^t}_{\substack{\text{Probability}\\ \text{of}\\ \text{surviving}\\ \text{in year } v}} \underbrace{\frac{1}{(1+\rho)^{v-t}}}_{\substack{\text{Time Value}\\ \text{Discount}}} \underbrace{u_\eta(c_v - \bar{c}_v)}_{\substack{\text{Utility of}\\ \text{Discretionary}\\ \text{Consumption}\\ \text{in year } v}}}_{\substack{\text{Utility of Lifetime Discretionary}\\ \text{Consumption}}}, \tag{5.1}$$

where:

c_v = consumption in year v;

q_v^t = probability of the investor surviving to at least year v, given that the investor is alive in year t;

ρ = the investor's *impatience for consumption* (subjective discount rate);

$u_\eta(.)$ = the investor's utility function for *preference for smooth consumption* (EOIS, η); and

$\bar{c}_v$ = nondiscretionary consumption in year v.

This utility-maximization problem is at the heart of the life-cycle models in this book. Although Equation 5.1 may look complicated, its application is straightforward. Maximizing the expression in Equation 5.1 provides the optimal *discretionary* consumption schedule. Adding the *discretionary* consumption schedule to the *nondiscretionary* consumption schedule provides the *overall lifetime consumption schedule*.

The Base Case

In the base case, the intertemporal budget constraint is as follows:

$$\sum_{v=t}^{T} \frac{1}{(1+r)^{v-t}} (c_v - \bar{c}_v) = F_t + H_t - L_t^C, \tag{5.2}$$

[36]In theory, annuity payments should be real (inflation-adjusted) amounts. In some countries such as the United States, most annuity payments are nominal. In the models in this book, we assume that annuity payments are real.

where:

r = the constant market rate of return;

H_t = human capital without mortality weighting in year t; and

L_t^C = the value of consumption-related liabilities without mortality weighting in year t.

According to Equation 5.2, the total value of discretionary spending is equal to net worth (financial capital plus human capital minus liabilities) and this must hold through time.

As we discussed in chapter 4, human capital and the value of consumption-related liabilities without mortality weighting are given by the following:

$$H_t = \sum_{v=t}^{T} \frac{y_v}{(1+r)^{v-t}}, \tag{5.3}$$

$$L_t^C = \sum_{v=t}^{T} \frac{\bar{c}_v}{(1+r)^{v-t}}, \tag{5.4}$$

where y_v is exogenous income in year t.

We arrive at the solution to maximizing intertemporal utility, subject to the intertemporal budget constraint, in steps. First we define a discretionary consumption growth rate g:

$$g = \left(\frac{1+r}{1+\rho}\right)^{\eta} - 1. \tag{5.5}$$

If the market rate of return (r) is greater than the subjective discount rate (ρ), the market return overcomes the investor's impatience, so that the growth rate is positive. If the market rate of return is not high enough to overcome the subjective discount rate, the growth rate will be negative. Either way, the magnitude of the growth rate also depends on the investor's preference for smooth consumption (η). In practice, for common and realistic values of r, ρ, and η, the range of growth rates (g) are between −3% and +3%. In this case, this is what the growth rate of discretionary consumption would be if survival were always certain.

The solution is the sum of nondiscretionary consumption and optimal discretionary consumption. It turns out that optimal discretionary consumption is proportional the ratio of net worth to a term that we will call the *divisor*. The divisor is given by the following:

$$\Delta_t = \sum_{v=t}^{T} (q_v^t)^{\eta} \left(\frac{1+g}{1+r}\right)^{v-t}. \tag{5.6}$$

We can write the solution in terms of nondiscretionary consumption, the probability of survival, the growth rate, current net worth ($F_t + H_t - L_t^C$), and the current divisor (Δ_t) as follows:

$$c_v = \bar{c}_v + (q_v^t)^{\eta}(1+g)^{v-t} \frac{F_t + H_t - L_t^C}{\Delta_t}. \tag{5.7}$$

Note that consumption in the current and all future periods is proportional to current net worth. In other words, if net worth were, say, doubled, discretionary consumption in all periods would be doubled.

An interesting mathematical exercise would be to substitute the right-hand side of Equation 5.7 for each c_v in the intertemporal budget constraint, Equation 5.2. Working through the math, we would see how the divisor makes the solution in Equation 5.7 satisfy the intertemporal budget constraint.

The optimal level of consumption in each period *v* can also be expressed as the sum of nondiscretionary consumption in period *v*, and the ratio of net worth in period *v* to the divisor in period *v*:

$$c_v = \bar{c}_v + \frac{F_v + H_v - L_v^C}{\Delta_v}. \tag{5.8}$$

This allows us to see how consumption varies over time in terms of how net worth and the divisor evolve over time. The reciprocal of the divisor, $1/\Delta_v$, is the fraction of net worth spent on discretionary consumption. Using assumptions that we made about Isabela in previous chapters, the curve labeled "Without Mortality Weighting" in **Exhibit 5.1** shows how this starts low and rises at an increasing rate in the base case. Hence, as Isabela (or anyone) ages, she will spend more of her net worth on discretionary consumption.

Exhibit 5.2 displays annual consumption for the three cases:

- *With* Annuities Available at Retirement and *without* Life Insurance (green line): Notice that Isabela's annual consumption is the greatest under this scenario because she can spend money that might have gone to purchase life insurance (which is deemed unavailable in this scenario).
- *With* Annuities Available at Retirement and *with* Life Insurance (blue line): In this scenario, consumption is a relatively steady and around $5,350–$6,050 lower than in the previous scenario. This lower level of consumption is due to the reduction in Isabela's net worth because of her plan for a bequest.

Exhibit 5.1. Fraction Isabela Spends of Her Net Worth for Discretionary Consumption at Each Age

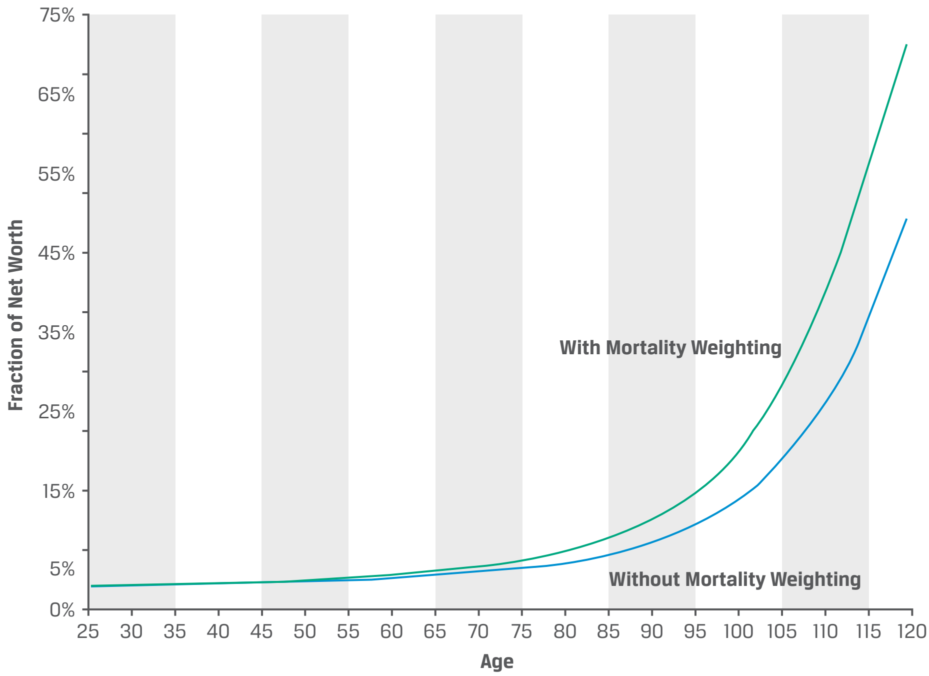
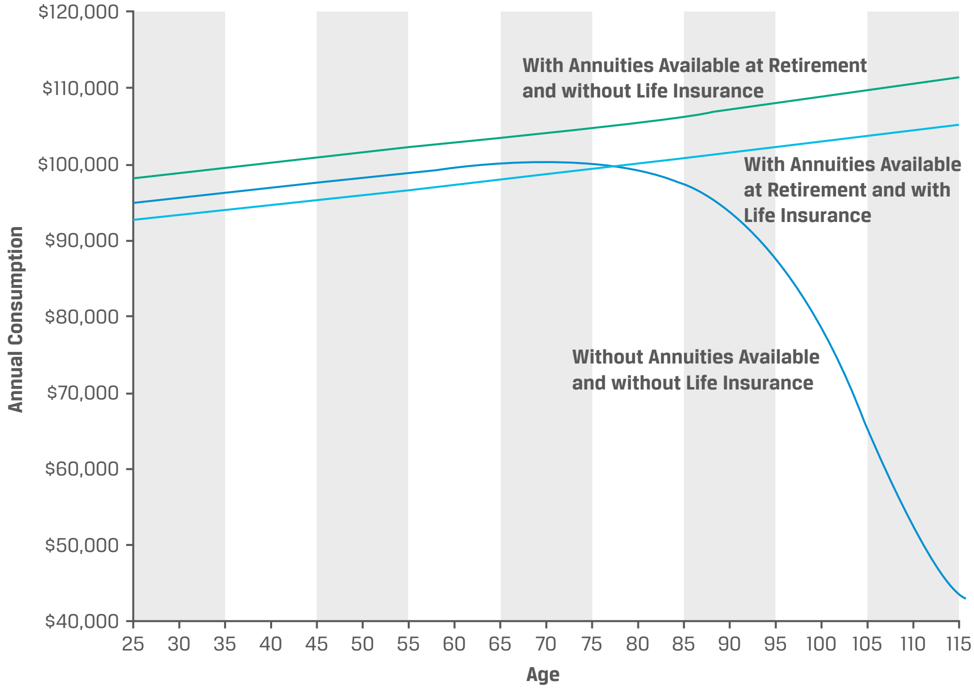


Exhibit 5.2. Evolution of Consumption for Isabela in Different Cases



- *Without* Annuities Available at Retirement and *without* Life Insurance (dark blue curve): In this scenario, Isabela's consumption is between the other two scenarios during accumulation, but then it drops relatively dramatically because of unhedged longevity risk. When no annuities are available to guarantee income, an investor should focus on consumption in the near to intermediate future and plan on reducing consumption in later years. The investor plans on consuming the most when the probability of being alive to enjoy it is high and consuming the least when the probability of being alive to enjoy it is low. Equation 5.7 shows that the degree to which the investor does this depends on their intertemporal flexibility as given by the investor's *preference for smooth consumption* (EOIS, η).[37]

Financial wealth evolves as follows:

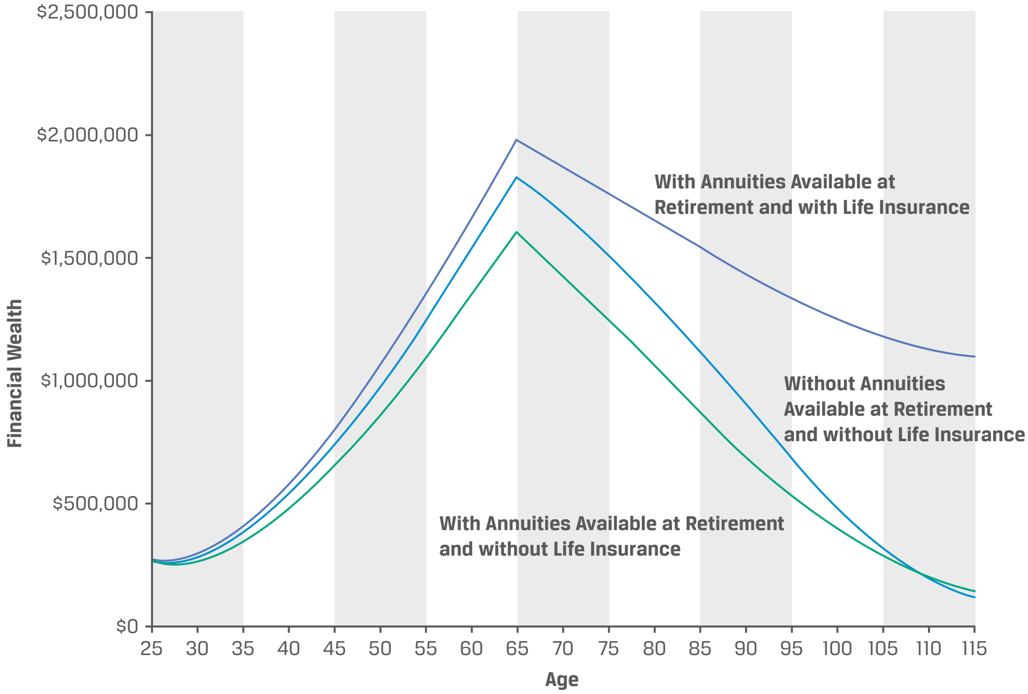
$$F_t = (1 + r)(F_{t-1} + y_{t-1} - c_{t-1}). \tag{5.9}$$

In general, before retirement, $y_t > c_t$, the investor is saving $y_t - c_t$. During retirement, $c_t > y_t$, the investor is withdrawing $c_t - y_t$.

Exhibit 5.3 shows how Isabela's financial wealth will evolve over time. Except for a small dip at the beginning, her financial wealth will increase over the entire time before retirement, when she will be saving. Once she retires and is withdrawing, financial wealth declines.

[37]Milevsky and Huang (2011) and Habib, Huang, and Milevsky (2017) also derive the consumption rule given by Equation 5.7.

Exhibit 5.3. Evolution of Isabela's Financial Wealth in Different Cases



Annuity Case

We now assume that a *complete market* for annuities is available to Isabela when she retires. The term "complete market" is important. A complete market for annuities means that in year t, the investor can effortlessly and instantaneously purchase a fairly priced contract that makes a one-time payment of $1 in year v, contingent on being alive then. The fair price of this contract is $q_v^t \frac{1}{(1+r)^{v-t}}$. Because annuities will not be available to Isabela until she retires, we need to use the adjusted mortality weights, $\hat{q}_v^t$, that we defined in chapter 4. Hence, modifying the intertemporal budget constraint of Equation 5.2 for when annuities will be available at retirement, the intertemporal budget constraint becomes as follows:

$$\sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+r)^{v-t}} (c_v - \bar{c}_v) = \hat{F}_t + \hat{H}_t - \hat{L}_t^C, \tag{5.10}$$

where $\hat{H}_t$ and $\hat{L}_t^C$ are human capital and the value of consumption-related liabilities modified versions of Equations 5.3 and 5.4 with mortality weighting in year t, respectively, as follows:

$$\hat{H}_t = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+r)^{v-t}} y_v, \tag{5.11}$$

$$\hat{L}_t^C = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+r)^{v-t}} \bar{c}_v. \tag{5.12}$$

The critical step is to solve for the optimal lifetime consumption plan or schedule, the investor maximizes utility given in Equation 5.1 subject to the intertemporal budget constraint given in Equations 5.10–5.12. As with the base case, we solve this maximization problem by defining a divisor for net worth:

$$\hat{\Delta}_t = \sum_{v=t}^{T} (q_v^t)^{\eta} (\hat{q}_v^t)^{1-\eta} \left(\frac{1+g}{1+r} \right)^{v-t}. \tag{5.13}$$

Returning to the definition of $\hat{q}_v^t$ in chapter 4, if v (and therefore t) is before retirement, $\hat{q}_v^t$ is 1, so the term being summed in Equation 5.13 is the same as the corresponding term in Equation 5.6. For terms in which v is past the year of retirement, however, a different mortality weighting, as Exhibit 5.1 shows, leads to discretionary consumption being a higher fraction of net worth with annuities than without.

Writing the solution in terms of initial net worth, we have the following:

$$c_v = \overline{c}_v + \left(\frac{q_v^t}{\hat{q}_v^t} \right)^{\eta} (1+g)^{v-t} \frac{\hat{F}_t + \hat{H}_t - \hat{L}_t^C}{\hat{\Delta}_t}. \tag{5.14}$$

Again, returning to the definition of $\hat{q}_v^t$ in chapter 4, if v (and therefore t) is before retirement, $\hat{q}_v^t$ is 1, so the mortality term being summed is the same as the corresponding term in Equation 5.7. Once the investor is retired so that t and v are past the year of retirement, $\hat{q}_v^t = q_v^t$, so that the mortality term becomes 1. Hence, starting at retirement, discretionary consumption grows at the constant rate g, which for Isabela is 0.24%.[38] This is shown in Exhibit 5.2 where, for the two cases in which annuities are available at retirement, consumption grows at a nearly steady rate until age 65, and then at a steady rate no matter how long Isabela shall live.

Annuitized financial wealth evolves not only with market returns but also with adjusted survival probabilities, as follows:

$$\hat{F}_t = \frac{1+r}{\hat{q}_t^{t-1}} (\hat{F}_{t-1} + y_{t-1} - c_{t-1}). \tag{5.15}$$

For example, if the market rate is 2.5% (the risk-free rate that we assume throughout this book) and the probability of surviving from year $t-1$ to year t is 96%, (as is nearly the case for Isabela at age 65), the combined effect would be an effective rate of return on an annuitized investment of $1.025/0.96 - 1 = 6.8\%$. The additional 4.3% is a *mortality credit*, described in Milevsky (2006). Mortality credits arise because, in any given cohort of annuitants, those that die forfeit their shares of the underlying investment portfolio to the survivors. Mortality credits are part of what makes annuities valuable to investors who have wealth that they do not plan on leaving as part of a bequest.

For convenience, we define a SPIA as an annuity contract paying $1 per year until the annuitant (investor) dies. The fair price of such a SPIA in year t that pays $1 every year starting in year t until the investor dies is as follows:[39]

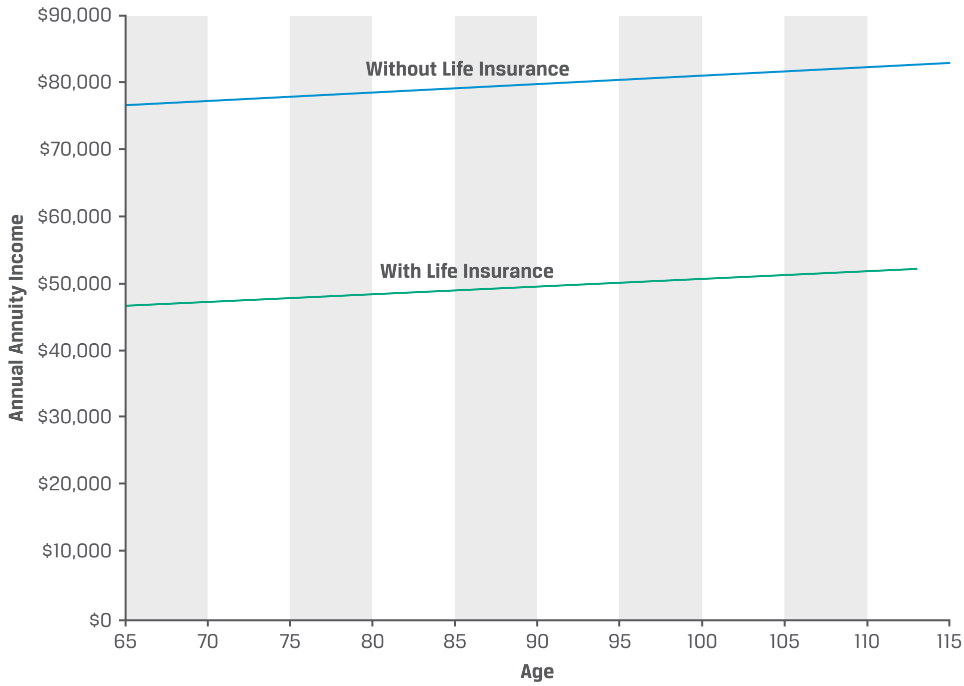
$$A_t = \sum_{v=t}^{T} q_v^t \frac{1}{(1+r)^{v-t}}. \tag{5.16}$$

---

[38]If g were negative, discretionary consumption would shrink at a constant rate.

[39]This definition means that the investor must simply buy as many SPIAs as they need in dollars of real annual income from the annuity. Typically, SPIAs involve a large primary purchase potentially followed by incremental purchases.

Exhibit 5.4. Evolution of Annuity Income



Annuity income is equal to the number of SPIAs that make up financial wealth, as follows:

$$AI_t = \frac{\hat{F}_t}{A_t}. \tag{5.17}$$

As **Exhibit 5.4** shows, after Isabela makes her initial purchases of annuities at age 65, her annuity income will gradually increase over time.

Note that, in the case of annuities becoming available at retirement and in the absence of life insurance, at retirement, the investor should move all financial wealth solely to annuities, and stay 100% in annuities.[40] When there is no bequest, it is not necessary to accumulate assets that will pass on to one's heirs. Instead, once retired, annuities are the only asset the investor needs: They deliver the underlying market return (which in this chapter is the riskless rate) *plus* the mortality credits. In other words, by holding annuities, the investor can earn mortality credits and thus be able to fund more consumption with annuities than without. The practical implication of this is that, if an investor has no plans for leaving a bequest (or has already funded a planned bequest as we shall see from the life insurance case), annuities should be the primary investment.

The Life Insurance Case[41]

Life insurance allows an investor to leave a bequest of a given size should the investor die before accumulating enough financial wealth to fulfill the desired bequest. Although life insurance policies take a variety

[40]Yaari (1965) showed that it is optimal to be 100% in annuities when no bequest is planned.

[41]This section is based on Kaplan (2022a).

of different forms, we assume that the investor purchases term life insurance each year. In chapter 4 (Equation 4.20), we presented the formula for the price per dollar of term life insurance, which we restate here as Equation 5.18:

$$LI1_{t-1} = \frac{1-q_t^{t-1}}{1+r}. \tag{5.18}$$

If the level of bequest is B, the premium in year $t-1$ is $LI1_{t-1}B$.

Another way to view the cost of life insurance is to assume that the investor is willing to pay a lump sum in year t to guarantee the bequest, no matter when the investor dies. To price the lump-sum policy, we first need to know the probability of dying in each year. As we discussed in chapter 4, we can calculate this from the survival probabilities. Letting p_v^t denote the probability of the investor dying in year v (given that the investor is alive in year t) and restating Equation 4.21 as Equation 5.19, we have the following:

$$p_v^t = q_{v-1}^t - q_v^t. \tag{5.19}$$

The lump-sum price for guaranteeing a bequest of $1 is as follows:

$$LI_t = \sum_{v=t}^{T} \frac{1}{(1+r)^{v-t}} p_v^t. \tag{5.20}$$

Hence, the lump-sum price for guaranteeing a bequest of size B is LI_tB. This is equivalent to the present value of the term premiums, and therefore, it is equal to the life insurance–related liability on the balance sheet. Hence, it enters the intertemporal budget constraint as follows:

$$\sum_{v=t}^{T} q_v^t \frac{1}{(1+r)^{v-t}}(c_v - \bar{c}_v) = F_t + \hat{F}_t + \hat{H}_t - \hat{L}_t^C - LI_tB. \tag{5.21}$$

Note that financial wealth takes two forms: regular financial wealth in the form of conventional investments (F_t) and annuity wealth ($\hat{F}_t$). The optimal level of consumption in year v is thus:

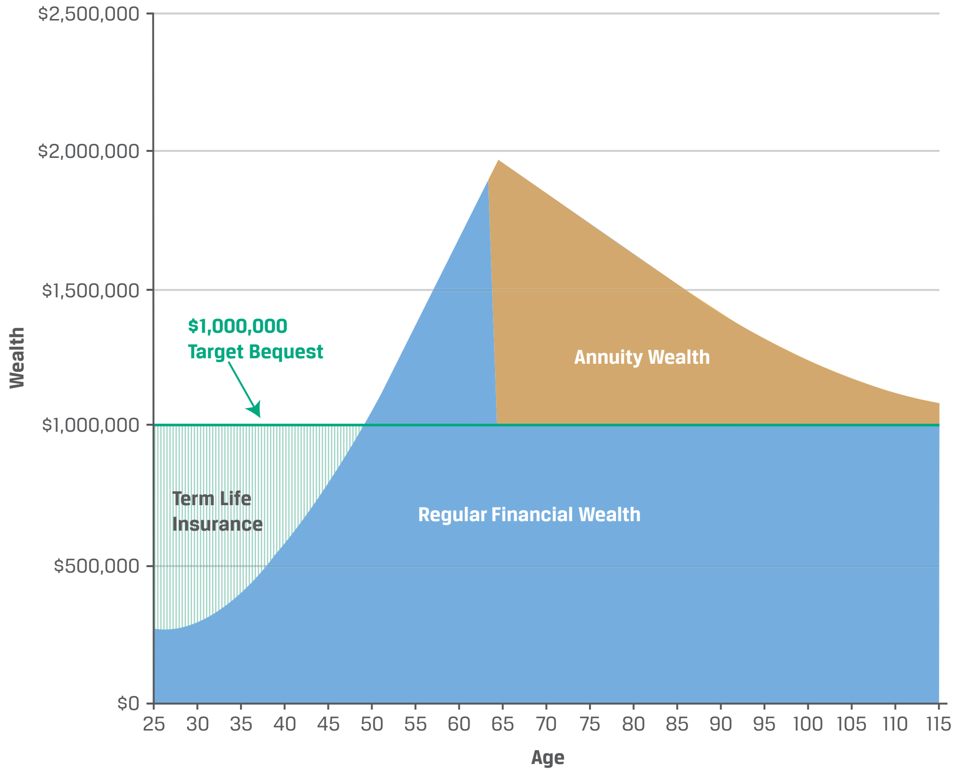
$$c_v = \bar{c}_v + \left(\frac{q_v^t}{\hat{q}_v^t}\right)^{\eta} (1+g)^{v-t} \frac{F_t + \hat{F}_t + \hat{H}_t - \hat{L}_t^C - LI_tB}{\hat{\Delta}_t}. \tag{5.22}$$

This is basically the same as Equation 5.14 but with wealth in conventional assets and the cost of life insurance taken in account. As in the annuity case, discretionary consumption grows (or decreases) at rate g. But, as Exhibit 5.2 shows, the level of consumption is lower than in the annuity case. As Equation 5.21 shows, the cost of life insurance—that is, the cost of guaranteeing a bequest of known size—reduces the amount of net worth available for consumption.

As we will see, the investor first accumulates wealth in conventional investments, until they reach the desired bequest (B) or retirement. While accumulating conventional assets, the investor purchases term life insurance to fill the gap between B and F_t. If this happens before retirement, they continue to accumulate conventional assets until they retire. Then, they invest any accumulated conventional assets in excess of B in annuities so that their financial wealth consists of both F_t (which is now equal to B) and $\hat{F}$.

Exhibit 5.5 shows how for Isabela, regular financial wealth, annuity wealth, and term life insurance will evolve over time, assuming that she plans on leaving a bequest of $1,000,000. Starting from age 25, she will accumulate conventional assets, reaching $1,047,424 at age 50. Over this period, she will buy enough

Exhibit 5.5. Evolution of Isabela's Regular Financial Wealth, Annuity Wealth, and Term Life Insurance



term life insurance to fill the gap between $1,000,000 and the amount accumulated.[42] At age 51, because she now has more than she needs to meet her bequest goal, she stops buying life insurance. Because she does not yet have access to annuities, she will continue to accumulate conventional assets, reaching $1,976,912 at age 65. At that point, she puts $976,912 into annuities, keeping $1,000,000 in conventional assets to fulfill her bequest goal. From this point forward, all additional financial wealth is in annuities.

Note that Isabela first buys life insurance and then, after a hiatus while waiting for annuities to become available, moves money that is in excess of the amount needed for the bequest into payout annuities. This illustrates the principle that term life insurance and SPIAs should not be held at the same time. We would argue that this principle should be put into practice so that young investors should purchase term life insurance to protect those who depend on their income, but once they have accumulated enough wealth to meet that need, they should discontinue the term insurance and begin to accumulate annuities to fund their consumption because of the mortality credits that they offer.

[42]Insurance companies offer a bundled product, called whole life, that has both a savings component and a life insurance component. Each year, the investor pays a constant premium, some of which goes into the savings component and the rest pays for term life. Over time, as the value of the savings component grows, the amount of term life decreases, much as it does in Exhibit 5.5. There is some debate about whole life, with some financial planners advising their clients to avoid the costs imposed by insurance companies by investing their savings in less expensive options, while buying just the amount of inexpensive term life insurance needed to guarantee the desired bequest. Other financial planners think that whole life is worth the costs because it imposes the discipline of saving each year and maintaining the right level of life insurance.

Exhibit 5.6. Sources and Uses of Isabela's Retirement Income at Age 65

Sources of Income		Uses of Income	
US Social Security	$28,356	Nondiscretionary Consumption	$40,000
Interest on Conventional Assets	$24,390[43]	Discretionary Consumption	$57,870
Annuity Income	$47,069	Additional Annuity Purchases	$1,946
Total	$99,816	Total	$99,816

Sources of Income during Retirement

Isabela will have three sources of retirement income: (1) US Social Security, (2) interest income from her conventional financial assets, and (3) annuity payouts. She should not sell assets, because to do so would undo the strategy that Paula has devised and she has so carefully followed to provide both for her own income and for leaving a bequest of $1,000,000. Each year, Isabela uses this income for two purposes: (1) to fund current consumption and (2) to purchase additional annuities to fund increases in consumption in subsequent years. **Exhibit 5.6** shows the sources and uses of Isabela's income at age 65.

Selecting the Bequest Level

Now that we have gone through the mechanisms by which an investor can fund consumption while guaranteeing a bequest, we can examine the trade-off between these two and develop an approach to selecting the optimal combination of them.[44]

Recall from chapter 3 that we measure lifetime utility by determining the constant level of consumption that would result in the same lifetime utility as a given (nonconstant) series of future consumption. We denote this constant level of consumption $\hat{c}_t$. Here we use $\hat{c}_t$ as the measure of consumption in the consumption-bequest trade-off. In chapter 3, we also introduced an intergeneration utility function in $\hat{c}_t$ and B to represent the investor's preferences regarding their consumption and leaving a bequest. To select the level of the bequest, we maximize this utility function subject to the constraint that the combination of $\hat{c}_t$ and B be feasible.

Feasible combinations are on the trade-off line between consumption ($\hat{c}_t$) and the size of the bequest (B) as shown in **Exhibit 5.7**. The trade-off line runs from the point on the vertical axis where $B = 0$, and $\hat{c}_t$ is what it is in the annuity case, to the point on the horizonal axis where $\hat{c}_t = 0$ and B is at its maximum possible value, as follows:

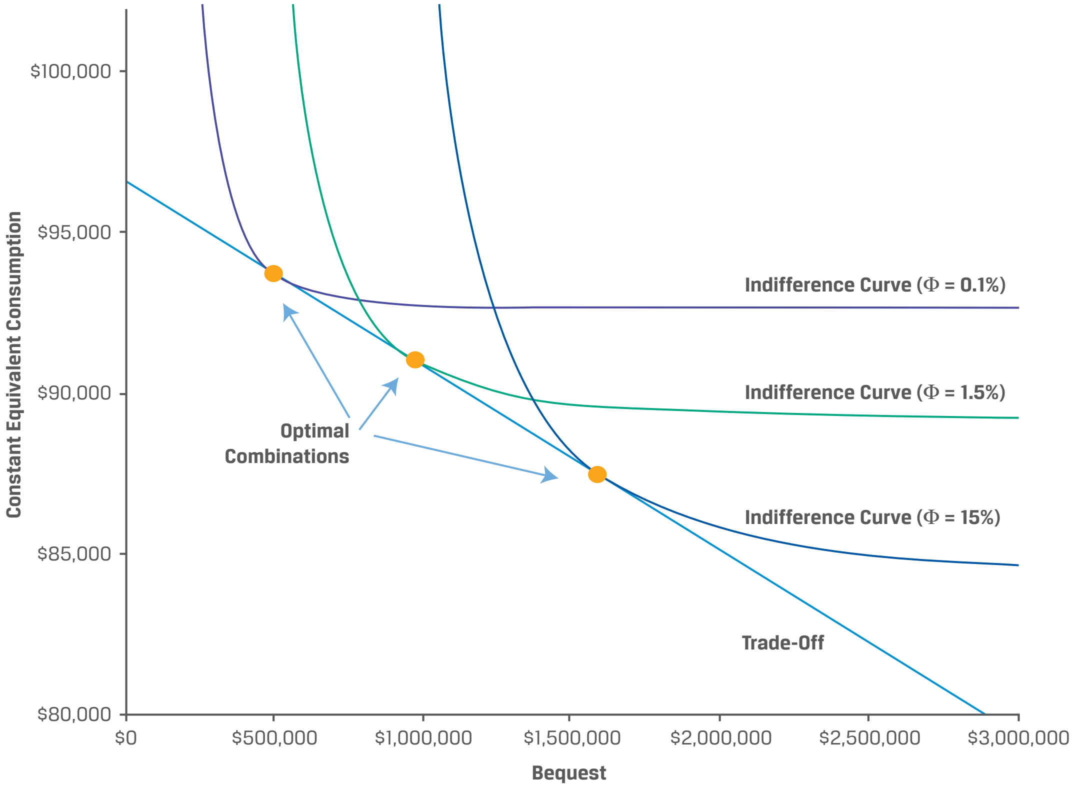
$$\bar{B} = \frac{FW_0 + \hat{H}_0 - \hat{L}_0^C}{LI_0}. \tag{5.23}$$

One of the parameters of the intergenerational utility function in chapter 3 is γ, the elasticity of intergenerational substitution. As we stated in chapter 2, the value of this parameter for Isabela is 25%. The other

[43]The reason that interest on conventional assets is $24,390 and not $25,000 is that we assume that money is withdrawn from the account just before interest is paid. So withdrawing $24,390 from a balance of $1,000,000 leaves $975,610. Then, 2.5% interest on this brings the balance back to $1,000,000.

[44]Our approach to modeling bequests differs from the approach in the academic literature. See chapter 3.

Exhibit 5.7. The Consumption-Bequest Trade-Off and Optimal Combinations



parameter is ϕ, which is the strength of the bequest motive. It can be between 0% (no bequest motive) and 100% (all resources fund the bequest). In Exhibit 5.7, we set ϕ to 0.1%, 1.5% (Isabela), and 15% to create three cases.

To illustrate maximizing intergenerational utility, in Exhibit 5.7, we included an indifference curve for each of the three values for ϕ. Recall from chapter 3 that an indifference curve shows all combinations of either (1) two desirable quantities, leading to an indifference curve that is downward sloping or (2) one desirable and one undesirable quantity, such as expected return and standard deviation in the Markowitz model (Kaplan 2020d). Here, two desirable quantities (consumption and the bequest) lead to indifference curves that are downward sloping.

We could draw indifference curves that are above the trade-off line. We could also draw indifference curves that intersect the trade-off line. However, for each value of ϕ, a unique indifference curve is tangent to the trade-off line—that is, the point of tangency identifying the optimal combination of consumption and the bequest. This is the point at which intergenerational utility is highest on the trade-off line. From Exhibit 5.7, we see that the stronger the bequest motive (the higher the value of ϕ), the lower the level of consumption and the larger the bequest. In practice, this means that an adviser needs to discern the investor's preferences regarding consumption and bequests, so that they can quantify the economic trade-off between them.

Conclusion and Key Takeaways

Building on the foundations of chapter 3 and 4, we introduced our first life-cycle models in this chapter. We will continue to add to the base life-cycle model and demonstrate how such models can be used to provide optimal financial advice.

In our initial or base life-cycle model, we can solve for a lifetime discretionary consumption schedule that can be paired with nondiscretionary consumption to create an optimal lifetime consumption schedule. By optimal, we mean that it maximizes lifetime utility in which the utility function captures the investor's two key preferences related to the timing of consumption: (1) the investor's *impatience for consumption* (subjective discount rate, ρ) and (2) the *preference for smooth consumption* (EOIS, η).

We then expanded on the base life-cycle model to demonstrate how the optimal lifetime discretionary consumption schedule changes based on the presence of annuities, and then annuities and life insurance. Life insurance and annuities are powerful tools that investors can use to manage the uncertainty around the time of death—dying early (mortality risk) or late (longevity risk).

Through life insurance and wealth accumulation, an investor can guarantee a bequest. Through annuities, an investor can guarantee income over their lifetime, no matter how long that period is. The investor needs to choose, however, between consumption and the size of the bequest. An intergenerational utility model provides a way to make this choice.

As an introductory chapter to life-cycle models, we have assumed a constant market rate of return (which is the risk-free rate). In such models presented here, all of the investor preferences that we discuss in chapter 3, except risk tolerance, are relevant. In the next chapter, we will carry over all of the elements of the models that we present in this chapter, but with risky market returns so that risk tolerance comes into play.

6. LIFE-CYCLE MODELS WITH UNCERTAINTY

Context

In chapter 5, we present a set of life-cycle models that provide formulas for optimal consumption and bequests and integrate longevity uncertainty with annuities and life insurance. The models in chapter 5 assume that the return on assets is a constant. In this chapter, we carry over all of the elements of the models presented in chapter 5, but we incorporate asset return uncertainty. By incorporating return uncertainty into these models, the investor's risk tolerance becomes central in selecting asset mixes or portfolios that have the optimal level of risk and expected return. The models that we present in this chapter are the most complete of those that we call "the life-cycle utility maximization models in this book" in Exhibit 3.11. As the exhibit shows, the models are built on foundations laid by Friedman (1957), Modigliani (1966), Samuelson (1969), Merton (1969, 1971), and others.

Key Insights

- Market uncertainty can be modeled using a random variable called the stochastic discount factor (SDF).
- The market price of an asset is the probability-weighted average of the future payoff of the asset times the SDF.
- An investor's optimal portfolio is an inverse power function of the SDF, with the investor's risk tolerance parameter being the power. Hence, the higher the risk tolerance parameter, the more risk the investor is willing to tolerate in pursuit of expected return.
- A number of spending rules have been developed. Some set a fixed rate of spending amount, whereas others dynamically vary spending with wealth. However, none of these rules take into account the investor's preferences, needs, and circumstances that we have laid out in chapters 3 through 5.
- Adding the SDF to the models that we present in chapter 5, we form a parallel set of models that take market risk and return into account as well as the investor's preferences, needs, and circumstances.
- Just as immediate fixed payout annuities (often called fixed SPIAs) are the relevant annuity instruments if the market return is constant, IVAs are the relevant annuity instruments when market returns are uncertain.
- If actual IVAs are not available, the investor should mimic them to create the optimal level of discretionary consumption each year.

In this chapter, we extend the life-cycle model to account for the uncertainty associated with investing in assets that are not risk free. We begin with a discussion of the SDF, which enables us to incorporate investment uncertainty into the models. This is what enables us to incorporate risk tolerance into the life-cycle models. We then develop a lifetime spending rule.

The Stochastic Discount Factor[45]

Although many practitioners may be unfamiliar with the subject matter of this section, what we cover in this section is the foundation of all that follows in this chapter. It is how we handle the central fact about financial assets, namely, the uncertainty of their future values. In principle, that uncertainty should be fully

[45]This section is based on Kaplan (2020a).

reflected in the prices at which they trade in the financial markets. How those prices are determined is one of the main questions that financial economics seeks to answer.

In formal financial economics, in models going back at least half a century when Merton (1973) introduced the intertemporal capital asset pricing model (ICAPM), the price of risky assets is determined with a construct known as the SDF or *pricing kernel*. The SDF is the discount factor for each possible future state of nature. In pricing an asset, it is applied to what the asset's cash flow is in each state. This is in contrast to older or less formal approaches, which may be more familiar to readers, that involve the discounting of the expected values of uncertain cash flows at rates that reflect the risk of the cash flows, rather than considering the full probability distributions of the cash flows. This is the DCF model. In other words, the DCF approach is based on point estimates, whereas the SDF approach is based on the distributions of a series of cash flows.

The key difference between the single-period CAPM, which is taught in business schools, and the ICAPM, which is taught in more advanced finance courses such as in PhD-level classes, is in the meaning of beta. In the single-period CAPM, beta is the sensitivity of an asset's returns to the returns on the market portfolio. In the ICAPM, beta is based on the sensitivity of the asset's returns to the SDF.

In this section of this chapter, we explain the SDF concept and what it means for asset prices, expected returns, and portfolio selection. We also show how the ICAPM can be derived in the SDF asset pricing framework.

Financial economics models uncertainty using probability theory; thus, unsurprisingly, concepts from probability theory are at the heart of the SDF construct. A key concept of probability theory is that of a *random variable*. A random variable takes on a potentially different value under each possible scenario. The likelihood or the probability of each scenario occurring is known before the actual scenario occurs. For each scenario, the SDF gives a price today to a $1 payoff should the given scenario occur. Both the payoff from holding an asset and the SDF can be modeled as random variables. The price of the asset today is the probability-weighted average of the SDF times the asset's future payoff. This probability-weighted average is called the *expected value*.

Let us express the asset pricing formula mathematically with two points in time, today (time t) and one period from today (time $t+1$). Let:[46]

$\tilde{Q}^t_{t+1}$ = the SDF (a random variable);

$\tilde{X}_{t+1}$ = the payoff of the asset in question[47] (a random variable); and

$P_t[\tilde{X}_{t+1}]$ = the price of the asset at time t (a value).

We then have the following:

$$P_t[\tilde{X}_{t+1}] = E_t[\tilde{Q}^t_{t+1}\tilde{X}_{t+1}], \tag{6.1}$$

where $E_t[.]$ means the expected value taking all information available at time t into account.

A risk-free asset is one that pays the same in all scenarios. The price of a risk-free asset that pays $1 in all scenarios follows:

[46]We place a ~ over a variable to indicate that its value is unknown as of the present year (t) and will not be known until a future year.

[47]In a multiperiod setting $\tilde{X}_{t+1}$ could include the price of at asset at time $t+1$, which is also a random variable.

$$E_t[\tilde{Q}^t_{t+1}] = \frac{1}{1+R_{ft}}, \tag{6.2}$$

where R_{ft} is the one-period risk-free rate of return at time t.

It is very important to understand the SDF concept, but estimation of the SDF is beyond the scope of the book. We take it as given.

Exhibit 6.1 illustrates how the SDF prices an asset with an uncertain payoff. This illustration shows two possible future states of the world: a Down Market, which has a 30% chance of occurring, and an Up Market, which has a 70% chance of occurring. If the Down Market occurs, the SDF is 1.3011 and if the Up Market occurs, it is 0.8429. In the last column of Exhibit 6.1, we show the expected value of each variable listed in the first column. This is 0.3 times the value in the Down Market plus 0.7 times the value in the Up Market. So, for the SDF, this is price of a risk-free asset, as given by the right-hand side of Equation 6.2. In this example, we have assumed a risk-free rate of 2%, so this 1/1.02 = 0.9804 as shown in the final column of Exhibit 6.1.

Next, we introduce an asset to price. As Exhibit 6.1 shows, in a Down Market, it pays its owner $81.01, and in an Up Market, it pays its owner $114.65. In the next row of Exhibit 6.1, we apply the formula in Equation 6.1 by multiplying the payout by the SDF in each state and then calculating the expected value of that product. As Exhibit 6.1 shows, this gives us a price of $99.27.

Note that the SDF applies to all securities. That is, every possible security (e.g., stock, bond) has a payoff in each possible future state, to which the value of the SDF in that state, and the probability of the state occurring, apply. Also, there are no arbitrage opportunities because any two assets with identical payoffs across scenarios have the same price. This is sometimes called *the law of one price* or *no-arbitrage condition*. Put slightly differently, ignoring the idea that nonpecuniary preferences may affect price, two identical cash flow series that have the same payouts in each future state must also have the same price.[48]

We can restate the asset pricing formula in Equation 6.1 in terms of returns (Equation 6.3). The return on the asset over the period t to $t+1$ is as follows:

$$\tilde{R}_{t+1} = \frac{\tilde{X}_{t+1}}{P_t[\tilde{X}_{t+1}]} - 1. \tag{6.3}$$

Exhibit 6.1. Example of Pricing an Asset with the SDF

	Down Market (Probability = 30%)	Up Market (Probability = 70%)	Expected Value (0.3 × Down + 0.7 × Up)
SDF	1.3011	0.8429	1/1.02 = 0.9804
Asset Payoff	$81.01	$114.65	$104.56
SDF × Payoff	$105.41	$96.64	Asset Price = $99.27
Asset Return	$81.01/$99.27 – 1 = –18.39%	$114.65/$99.27 – 1 = 15.49%	Expected Return = 5.33%

[48]In chapter 9, building on the popularity asset pricing model of Idzorek, Kaplan, and Ibbotson (2021, 2023), we expand on the idea that as a result of investor nonpecuniary preferences or tastes, two assets with the same cash flows may in fact be priced differently.

We illustrate this in the last row of Exhibit 6.1.

Therefore,

$$E_t[\tilde{Q}^t_{t+1}(1+\tilde{R}_{t+1})]=1. \tag{6.4}$$

We can rewrite this as follows:

$$E_t[\tilde{Q}^t_{t+1}\tilde{R}_{t+1}]=\frac{R_{ft}}{1+R_{ft}}. \tag{6.5}$$

The expected value of the product of two random variables, say, $\tilde{X}$ and $\tilde{Y}$, is related to their covariance (a measure of how much they move together) as follows:

$$E[\tilde{X}\tilde{Y}]=Cov[\tilde{X},\tilde{Y}]+E[\tilde{X}]E[\tilde{Y}]. \tag{6.6}$$

Applying this to our earlier equation showing $E_t[\tilde{Q}^t_{t+1}\tilde{R}_{t+1}]$, that is, Equation 6.5 with some rearranging of terms, we have the following equation for the expected excess return on the asset:

$$E_t[\tilde{R}_{t+1}-R_{ft}]=-(1+R_{ft})Cov_t[\tilde{Q}^t_{t+1},\tilde{R}_{t+1}]. \tag{6.7}$$

From the data in Exhibit 6.1, we find that $Cov_t[\tilde{Q}^t_{t+1},\tilde{R}_{t+1}]=0.0326$ and $E_t[\tilde{R}_{t+1}-R_{ft}]=3.325\%$.

This equation can be applied to any asset or portfolio. We can apply it to a benchmark portfolio, such as a broad market index and then combine the result with Equation 6.7 to get the following:

$$E_t[\tilde{R}_{t+1}-R_{ft}]=\beta_t E_t[\tilde{R}_{Bt+1}-R_{ft}], \tag{6.8}$$

where $\tilde{R}_{Bt+1}$ is the return on the benchmark and:

$$\beta_t=\frac{Cov_t[\tilde{Q}^t_{t+1},\tilde{R}_{t+1}]}{Cov_t[\tilde{Q}^t_{t+1},\tilde{R}_{Bt+1}]}. \tag{6.9}$$

Equations 6.8 and 6.9 present a generalized form of the CAPM, the ICAPM, derived solely from the SDF-based asset pricing formula. In particular, β_t is a generalized measure of systematic risk.

In **Exhibit 6.2**, we apply the SDF to our benchmark asset. From the data in this table and the data in Exhibit 6.1, we find that $\beta_t=0.8111$ and $[\tilde{R}_{Bt+1}-R_{ft}]=4.10\%$. From these values, we can see that Equation 6.8 holds:

$$3.325\%=0.8111\times 4.10\%.$$

How exactly the SDF is determined depends on the specifics of the asset pricing model that gives rise to it. In many models, the SDF varies inversely with the growth rate of the economy because of the *principle of diminishing marginal utility*. According to this principle, the benefit (utility) of an additional dollar to a person's wealth or income varies inversely with the starting level wealth or income. So, if economic growth is strong, incremental increases in wealth or income will be less beneficial to investors than when economic growth is weak or negative. This leads to an inverse relationship between economic growth and

Exhibit 6.2. Example of Pricing a Benchmark Asset with the SDF

	Down Market (Probability = 30%)	Up Market (Probability = 70%)	Expected Value (0.3 × Down + 0.7 × Up)
SDF	1.3011	0.8429	1/1.02 = 0.9804
Asset Payoff	$76.86	$118.63	$106.10
SDF × Payoff	$100.00	$100.00	Asset Price = $100.00
Asset Return	$76.86/$100.00 – 1 = –23.14%	$118.63/$100.00 – 1 = 18.63%	Expected Return = 6.10%

the SDF. Because the risk-free rate is inversely related to the expected value of the SDF, increases in the expected rate of economic growth lead to increases in the risk-free rate, and decreases in the expected rate of economic growth lead to decreases in the risk-free rate. At least, that is the relationship in theory.

Again, estimation of the SDF is beyond the scope of this book. For our purposes, we take the SDF as a given random variable.

The SDF plays a central role in multiperiod models of spending and investing. For any plan or strategy for investing and spending over multiple periods to be feasible without leaving anything on the table, the market value of all spending, current and future, must equal the current level of net worth. Recall that we introduced the concept of an *intertemporal budget constraint* in chapter 3 when we contemplated stretching our $20 budget over two different parties. In the context of the SDF, to write this *intertemporal budget constraint*, we need to first define SDFs over multiple periods as the product single-period SDFs. Hence, the SDF over the period t to v is as follows:

$$\tilde{Q}_v^t = \tilde{Q}_{t+1}^t \tilde{Q}_{t+2}^{t+1} \cdots \tilde{Q}_v^{v-1}. \tag{6.10}$$

Thus, in the absence of annuities, the intertemporal budget constraint is as follows:

$$W_t = c_t + E_t[\tilde{Q}_{t+1}^t \tilde{c}_{t+1}] + E_t[\tilde{Q}_{t+2}^t \tilde{c}_{t+2}] + \cdots, \tag{6.11}$$

where:

W_t = net worth at time t;

c_t = consumption (spending) at time t; and

$\tilde{c}_v$ = consumption at time v (a random variable).

We discussed expected utility theory in chapter 3 as well as the investor's preference regarding *impatience for consumption* (subjective discount rate, ρ). In single-period expected utility theory, the investor seeks to maximize the expected utility of ending-period wealth. Expected utility theory can be expanded to a multiperiod setting by incorporating an intertemporal preference parameter, the investor's *impatience for consumption* (subjective discount rate, ρ). Thus, letting $u(.)$ denote the single-period utility function, multiperiod expected utility is given as follows:[49]

[49]Multiperiod expected utility, as formulated in Equation 6.12, assumes that the investor's life span is fixed and known. In the next section, we introduce uncertain lifespans, in the same manner as in chapter 3.

$$U_t = u(c_t) + E_t\left[\frac{u(\tilde{c}_{t+1})}{1+\rho} + \frac{u(\tilde{c}_{t+2})}{(1+\rho)^2} + \cdots\right]. \tag{6.12}$$

Focusing just on time t and time $t + 1$, if the investor is to maximize U_t subject to the intertemporal budget constraint, the following condition must hold:

$$\frac{1}{1+\rho}\frac{u'(\tilde{c}_{t+1})}{u'(c_t)} = \tilde{Q}^t_{t+1}. \tag{6.13}$$

Note that this equation does not have any expectations. It simply says that the investor needs to set things up so that consumption at time $t + 1$ in some fashion tracks the realized value of the SDF.

To develop a specific formula for consumption needing to track the SDF, we assume that the single-period utility function is of the CRRA form, as we discussed in chapter 3; namely:

$$u(c) = \begin{cases} ln(c), & \theta = 1 \\ \dfrac{\left(c^{1-\frac{1}{\theta}}\right) - 1}{1-\dfrac{1}{\theta}}, & \theta \neq 1 \end{cases} \tag{6.14}$$

where θ is the risk tolerance parameter.[50]

With this CRRA utility function, the conditions for maximizing expected utility subject to the intertemporal budget constraint imply the following:

$$\frac{1}{1+\rho}\left(\frac{\tilde{c}_{t+1}}{c_t}\right)^{-\frac{1}{\theta}} = \tilde{Q}^t_{t+1}. \tag{6.15}$$

Solving for the ratio of consumption at time $t + 1$ to consumption at time t, we have the following:

$$\frac{\tilde{c}_{t+1}}{c_t} = \frac{1}{(1+\rho)^{\theta}}(\tilde{Q}^t_{t+1})^{-\theta}. \tag{6.16}$$

To meet this condition, the investor needs to construct a portfolio with return as follows:

$$\tilde{R}_{\theta t+1} = \frac{(\tilde{Q}^t_{t+1})^{-\theta}}{E_t[(\tilde{Q}^t_{t+1})^{1-\theta}]} - 1. \tag{6.17}$$

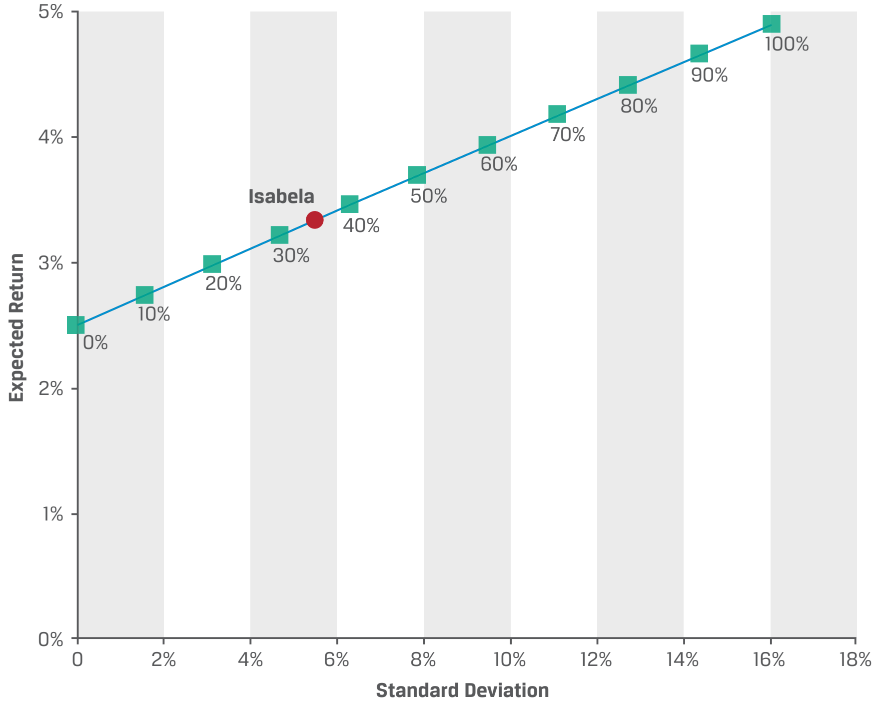
Hence, the change in consumption is directly related to the return on the portfolio:

$$\frac{\tilde{c}_{t+1}}{c_t} = \frac{E_t\left[(\tilde{Q}^t_{t+1})^{1-\theta}\right]}{(1+\rho)^{\theta}}(1+\tilde{R}_{\theta t+1}). \tag{6.18}$$

---

[50]The reciprocal of the risk tolerance parameter is called the risk aversion parameter.

Exhibit 6.3. Risk and Expected Return of Optimal Portfolios for Different Levels of Risk Tolerance



If investors have different degrees of risk tolerance, each one will manage their own portfolio such that the returns on the portfolio track the SDF according to Equation 6.18. Because there are no inefficiencies, the relationship between risk and expected return should be positive across the portfolios. To demonstrate this, we assume that $\tilde{Q}_{t+1}^{t}$ follows a lognormal distribution. We now assume that the risk-free rate is 2.5% so that, as Equation 6.2 shows us, the expected value of the SDF is 1/1.025. According to the three-asset class model that we presented in chapter 4, we assume that the standard deviation of $ln(\tilde{Q}_{t+1}^{t})$ is 15.18%. Based on these assumptions, we calculate the expected return and standard deviation of $\tilde{R}_{\theta t+1}$ for values of θ between 0% and 100%. **Exhibit 6.3** plots standard deviation versus expected return for these portfolios. It also shows where Isabela lands, given that, as we stated in chapter 2, θ for her is 35%. The resulting curve is similar to a Markowitz efficient frontier, albeit much more linear.

Solving Multiperiod Utility Maximization Problems

Multiperiod utility maximization problems are often solved using a numerical method called dynamic programming or the Bellman equation (Bellman 1957). In this approach, one first solves the problem for the last period (which is in the future and often corresponds to the assumed death date) and then uses backward recursion, period by period, until the solution for the present period emerges. But, as the number of decision variables and constraints increase, the problem can become unwieldy. This is known as "the curse of dimensionality." As a result, many dynamic programming-based life-cycle models make lots of simplifying assumptions to limit the number of decision variables and constraints. Although some people find this method somewhat intuitive, we think that this (or any numerical method) can quickly become a "black box" that obscures the intuition and understanding of how the inputs to the model result in the outputs.

So, as Equations 6.12–6.18 show, we take a different approach. We represent risky assets with what is known as a complete contingent claims market and impose no other constraint other than the single intertemporal budget constraint. A complete contingent claims market is one in which the investor can buy a contract with a payout at any time in the future that is based on any contingency. Under these assumptions, we can use the approach of Kaplan (1986), which allows us to write the solution to the utility maximization problem with a set of equations that are easy to interpret and helpful in understanding the economics of life-cycle finance.

Goals-Based Investing

Just as we commented on our decision to avoid dynamic programming (given its popularity among researchers), we feel compelled to comment on goals-based or goals-centric approaches (given their popularity among practitioners).[51] One reason that goals-centric investing is popular may be the loosely defined nature of the solutions it prescribes; yet, it is often presented to investors as an actionable plan for obtaining one's goals. From a life-cycle finance perspective, many such approaches put the cart before the horse, with a myopic view focused on specifics while failing to see the larger, more important, big picture.

From one behavioral finance perspective, goals-based investing embraces separate mental accounting, a way of compartmentalizing financial problems in one's head to avoid seeing the interactions between the different problems (Thaler 1985). From a more positive behavioral finance perspective, the key advantages of goals-centric approaches seem to be investor engagement, helping the investor to understand and to trust in the plan. Conversely, the creation of mental accounts leads to artificial constraints that result in one's total assets divided into separate buckets of money with corresponding policy portfolios. This is the antithesis of holistic life-cycle planning and should be avoided by financial planners.[52] In contrast, our models are relatively consistent with mainstream academic life-cycle models, especially that of Epstein and Zin (1989) with separate parameters for (1) flexibility in planning consumption across periods and (2) risk tolerance.[53]

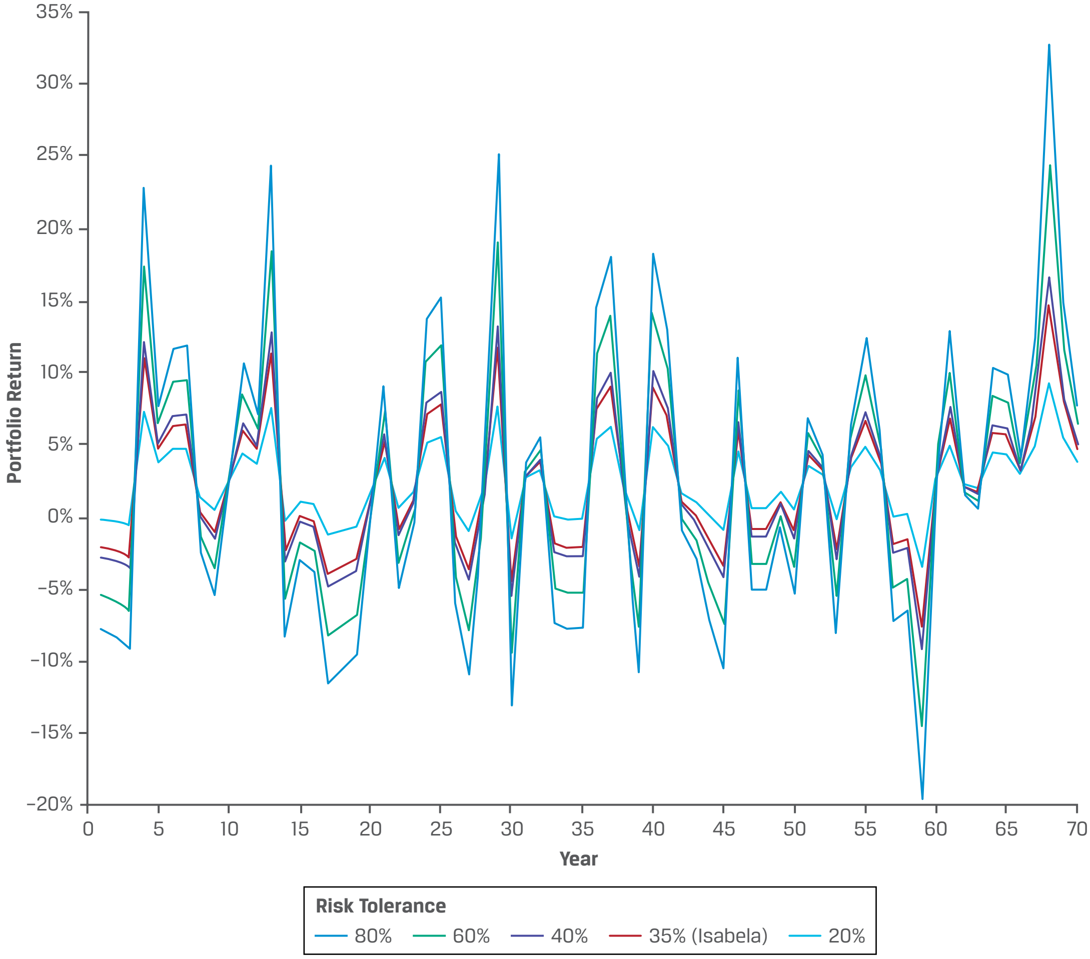
To illustrate how differences in risk tolerance can lead to differences in returns and spending, we generated 30 random values for the SDF, each corresponding to a year, and calculated the portfolio return for risk tolerance levels 20%, 35% (Isabela), 40%, 60%, and 80%. **Exhibit 6.4** plots these returns over 70 years, which is about the expected remaining lifespan for Isabela. Because the returns on all of the portfolios are calculated from the values of the SDF, they are in perfect synch. We then calculated the optimal level

[51]See Shefrin and Statman (2000), Nevin (2004), and Parker (2021) for overviews of goal-centric investing.

[52]An area for future research is finding a way to bridge the gap between the standard life-cycle model goal of smooth lifetime consumption and the specific consumption linked to goals. A recent example of this direction is Daga, Smart, and Pakula (2023).

[53]In the earlier literature on life-cycle models, such as Samuelson (1969), Merton (1969, 1971), and Kaplan (1986), these were the same parameter.

Exhibit 6.4. Simulated Portfolio Returns for Different Levels of Risk Tolerance

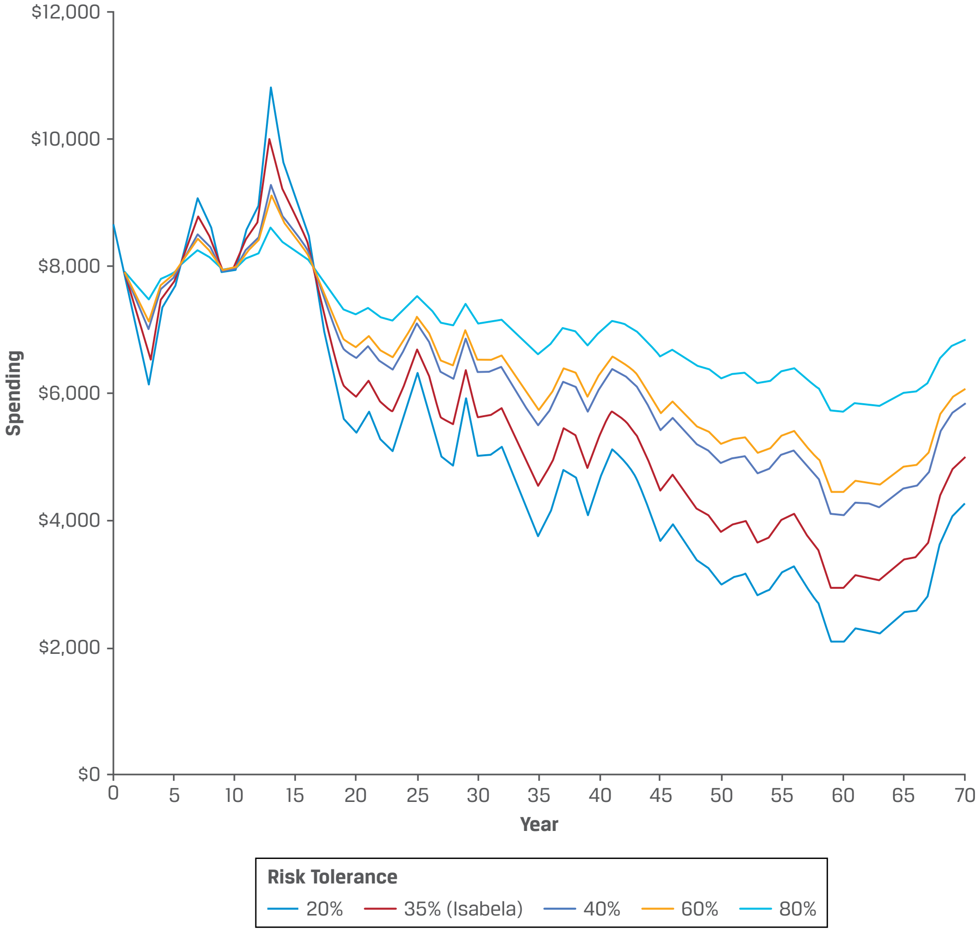


of spending (consumption) each year given these returns, assuming an initial portfolio value equal to Isabela's financial wealth at age 25, $270,500 and that Isabela lives to age 95.

Exhibit 6.5 shows the results. The spending levels move up and down in accordance with the asset mix/portfolio returns, with the largest changes corresponding to the riskiest asset mix/portfolio ($\theta = 80\%$). Note that the consumption levels are quite low because we have not yet introduced human capital into the spending model. Next, we introduce human capital as part of net worth. Because Isabela has a high level of human capital, as we shall see, she will be able to sustain much higher levels of consumption than shown in Exhibit 6.5.

The concept of an SDF or a pricing kernel, especially in conjunction with expected utility theory, is a powerful theoretical tool in understanding how assets are priced and how investors should manage their portfolios and spend over time. Key themes that emerge from the models are as follows:

Exhibit 6.5. Simulated Spending Levels for Different Levels of Risk Tolerance



1. The market value of an asset depends on the covariance between the future value of the asset and the SDF.
2. Returns on optimal portfolios are linked to the SDF through the level of risk tolerance.
3. Changes in optimal spending go hand in hand with portfolio returns.

Although the SDF is a theoretical construct, not visible in the real world, the insights it provides are useful in practice as we will demonstrate.

In the remainder of this chapter, we present stochastic versions of the three cases of deterministic models that we presented in chapter 5, namely:

1. the base case (without annuities and life insurance),
2. the annuities case, and
3. the life insurance case.

For the base case, we start with a discussion of the literature on spending rules and what is absent from that literature. We then add those missing elements one by one to arrive at our rule.

After discussing these three cases, we show how each spending rule can be restated in terms of the payout of an IVA.

Spending Rules[54]

The models that we present in this chapter are part of a broader literature on spending rules, especially for retirement.[55] The literature on retirement spending follows two general approaches: static and dynamic. In static approaches, the retiree selects an amount to spend in the first year of retirement and then grows that amount at the rate of inflation. The problem with a static approach is that the retiree runs the risk of running out of money before dying.

The best-known static rule is the 4% rule of Bengen (1994) in which the retiree spends 4% of the initial value of retirement funds in the first year then increases spending by the (realized) rate of inflation after that. Note that this rule does not take into account any specific knowledge about the retiree or what is in their retirement portfolio. It is a once-size-fits-all approach.

A more refined static approach is a success probability model. In this approach, for each spending level being considered, the model calculates the probability of not running out of money before dying. (The better models incorporate survival probabilities into the calculations.) This yields a trade-off between spending level and success probability for any given investment strategy, which is usually expressed as an asset allocation. The retiree can pick a desired success probability and select the asset allocation that maximizes spending. Kaplan (2006) presents a model that does this using Monte Carlo simulation. Milevsky and Robinson (2005), in contrast, developed a version that uses formulas rather than Monte Carlo simulation.

Dynamic spending rules avoid running out of money by varying spending with portfolio value. A very simple (we would argue too simple) approach was proposed by Waring and Siegel (2015). In their model, the retiree picks a date far enough into the future to be nearly certain of not surviving until then. Call this date T. At each year $t \leq T$, the retiree calculates (and next year recalculates) the market price of a sequence of $1 payments from year t through year T, creating a series of prices. The amount of spending each year is wealth divided by the corresponding price.

The only parameter in the Waring–Siegel model that is specific to the retiree is T. In every other way, the Waring–Siegel model is a one-size-fits-all approach. This is why we consider their model to be too simple and incomplete. In our view, a retirement spending model should take into account the preferences, needs, and circumstances of each retiree. Kaplan and Blanchett (2020) present such a model. In this chapter, we present models based on the Kaplan–Blanchett approach with some simplifications to ease exposition, but also to accommodate generalizations for including the preretirement period. Because, in the preretirement period, spending is usually less than income, there is saving (the difference between spending and income). Hence, the models also guide preretirement saving advice.

[54]This and the following four sections are based on Kaplan and Blanchett (2020) and Kaplan (2020b).

[55]We use the term "consumption" in the context of life-cycle models. In this section, we discuss a broader set of models in which the term "spending" is used. The two terms basically denote the same thing, so we use them interchangeably. Which word we use is mainly determined by which type of model or which literature we are discussing.

Preferences, Needs, and Circumstances

As we discuss in chapters 2, 3, and 4, investors differ from one another in their preferences, needs, and circumstances. Here, we focus on three of those specific preferences:

1. Risk Tolerance (θ or theta) is the investor's attitude toward risk. Higher values indicate greater willingness to take on risk.
2. Preference for Smooth Consumption: EOIS (η or eta) is preference for smooth consumption from one period to the next. Higher values indicate more flexibility and a lower preference for smooth consumption.
3. Impatience for Consumption: Subjective Discount Rate (ρ or rho) is the investor's preference to consume now versus later. Higher values indicate a stronger preference for consumption today versus in the future.

By needs, we mean nondiscretionary consumption for necessities like food and housing. Any spending strategy must cover this.

The circumstances of an investor are as follows:

1. Longevity is the probability of surviving to each possible age. This depends on age, gender, and health. Also, when planning for a couple, the spending strategy needs to consider the probability of each person surviving the other in each possible year.
2. Financial wealth is the amount of financial wealth that the investor has available at the time of retirement, which can impose a constraint on how much the retiree can spend each year.
3. Income, as discussed in chapter 4, typically is the salary an investor receives before retirement and then the annuity-like income stream they receive from a social insurance program, such as US Social Security. Some retirees might have additional sources of income, such as from a defined benefit plan. A spending strategy should take all such income into account.
4. Market return distributions are not necessarily specific to the retiree, but the return distributions on the assets that are available to the investor form part of the circumstances in which retirees make their spending decisions. In the model that we present in this chapter, market returns are determined by the SDF, as we discussed.

Creating a Spending Rule for Everyone

We start with the Waring–Siegel spending rule and introduce the elements that it is missing, one by one in stepwise fashion, to get to our complete model.

The Waring–Siegel Rule

Before creating an investor-specific spending rule, let us first review the Waring–Siegel rule. In their model, at each year t, the retiree calculates the price of a stream of $1 annual payments until the last possible year (30 years in their example). This is:

$$\Delta_t^{WS} = \sum_{v=t}^{T} \frac{1}{(1+r)^{v-t}}, \tag{6.19}$$

where r is the real rate of return, which Waring and Siegel estimate by taking an average of rates on Treasury inflation-protected securities (TIPS) with different maturities over the assumed 30-year period. In year t, spending is as follows:

$$c_t = \frac{W_t}{\Delta_t^{WS}}, \tag{6.20}$$

where, again, W_t is net worth in year t. If wealth is at least partially invested in risky assets, it is subject to market-generated fluctuations over time. So, although the fraction of wealth being spent follows a predetermined path, the amount of spending fluctuates over time as market returns vary.

Using Certainty Equivalent Return

Because spending varies with market fluctuations, the discount rate for the future stream of spending should not be a risk-free rate. Rather, it should reflect the riskiness of spending. In chapter 4, we discussed how the discount rates for human capital and the value of consumption-related liabilities are their corresponding expected returns. But for the spending rule, it is the *certainty equivalent return* on net worth. The certainty equivalent return is the constant rate of return that would make the investor indifferent between that constant return rate and the risky return. To see how it works, suppose that returns are lognormally distributed with expected return ER and logarithmic standard deviation σ_{log}. Letting h denote the certainty equivalent return, we have the following:

$$h = (1+ER)exp\left(-\frac{\sigma_{log}^2}{2\theta}\right) - 1. \tag{6.21}$$

For Isabela, $ER = 3.33\%$ and $h = 2.91\%$.

The certainty equivalent return takes into account (1) the investor's preferences (risk tolerance) and (2) an important component of the circumstances (market return volatility). We now have:

$$\Delta_t^h = \sum_{v=t}^{T} \frac{1}{(1+h)^{v-t}}. \tag{6.22}$$

Adding a Growth Rate

As Waring and Siegel (2018) and others discuss, the pattern of spending can be reshaped by introducing a growth rate, g. If g is positive, relative to the original Waring and Siegel (2015) model (Equations 6.19 and 6.20), spending shifts from the early years to the later years. If negative, spending shifts from the later years to the early years. Incorporating the growth rate into the spending model, we have the following:

$$\Delta_t^{h,g} = \sum_{v=t}^{T} \left(\frac{1+g}{1+h}\right)^{v-t}. \tag{6.23}$$

As we discuss in chapter 5, however, the growth rate is not an arbitrary constant. Rather it is a specific function of the investor's preferences and circumstances. From Equation 5.5 in chapter 5, the formula for g is as follows:

$$g = \left(\frac{1+h}{1+\rho}\right)^{\eta} - 1. \tag{6.24}$$

The *preference for smooth consumption* (EOIS, η) is usually between 0% (no flexibility) and 100% (high level of flexibility).

For Isabela, $g = 0.45\%$.

Whether consumption increases, decreases, or remains constant depends on whether the certainty equivalent return is above, below, or the same as the investor's *impatience for consumption* (subjective discount rate, ρ). If $h > \rho$, it is worthwhile for the investor to forgo some consumption in the earlier years to potentially earn a higher market return. If $h < \rho$, the preference for earlier consumption over later consumption outweighs the potential benefit of market returns so consumption is higher in the earlier years than in the later years. If $h = \rho$, the potential benefit of market rates of return offsets the preference for earlier consumption over later consumption, resulting in a flat pattern of consumption.

When consumption changes over time, the rate of change depends on the *preference for smooth consumption* (EOIS, η). This parameter captures how responsive an investor is to changes in intertemporal trade-offs. We can see exactly how this works. Consider the case in which $h < \rho$ so that consumption is declining. The higher the value of η, the greater the rate of decline.

Taking Longevity into Account

In chapter 3, we presented the Gompertz function, which is a formula for the probability of surviving for any given number of years. Recall that the Gompertz function has three parameters:

1. current age,
2. the mode of the distribution of the age of death, and
3. the dispersion around the mode of the age of death.

Taking these parameters as part of the retiree's circumstances, the Kaplan–Blanchett model takes longevity into account by calculating the survival probabilities for the investor using the Gompertz formula. Let

$q_v^t =$ the probability of the retiree surviving to at least year v, given that the retiree is alive in year t.

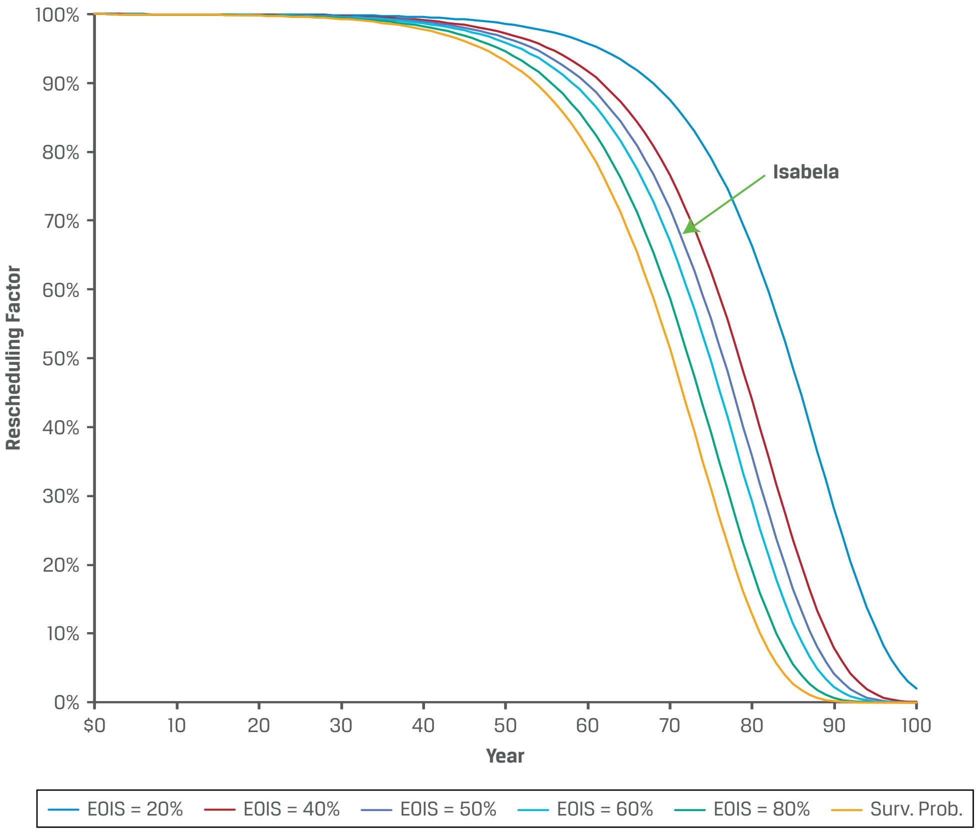
Exhibit 6.6 includes a plot of the survival probabilities for a 62-year-old woman. The other curves are what we call *rescheduling factors*. A rescheduling factor indicates how much spending to move from later years to nearer years because of the lower survival probabilities in later years. The rescheduling factor for a given investor is as follows:[56]

$$RF_v^t = (q_v^t)^\eta. \tag{6.25}$$

In the Waring–Siegel model, the retiree is effectively planning for the same level of spending until the final year. But this may not be appropriate for many investors given that the survival probability declines over the planning horizon. Hence, planned spending out into the future should decline as well. However, the investor may not be fully flexible when it comes to scheduling spending. This is why the survival probability is adjusted by the investor's *preference for smooth consumption* (EOIS, η) in the formula for the rescheduling factor. The closer the η is to zero, the less the rescheduling is, as shown by the shapes of the curves in Exhibit 6.6 for η values of 20%, 40%, 50% (Isabela), 60%, and 80%.

[56]The rescheduling factor is part of the solution to a model with a deterministic market in chapter 5. See Equations 5.6 and 5.7. We reintroduce it here as an element missing from the Waring–Siegel model.

Exhibit 6.6. Survival Probabilities and Rescheduling Factors



We incorporate the rescheduling factor into the spending rule by including it in the present value calculation. We have the following:

$$\Delta_t = \sum_{v=t}^{T} RF_v^t \left(\frac{1+g}{1+h} \right)^{v-t}. \qquad (6.26)$$

Taking Needs and Exogenous Income into Account

Investors typically have some minimum amount that they must spend to fulfill their basic needs of food, housing, and other necessities. This is nondiscretionary spending. As in chapter 4, we denote the expected level of nondiscretionary spending in year v as of t, $\bar{c}_v^t$. Also following chapter 4, we denote the expected return of the representative portfolio for nondiscretionary consumption and consumption-related liabilities as k_c.

At a minimum, the retiree must have enough wealth to cover nondiscretionary expenses in present and all future years. The present discounted value of present and future nondiscretionary spending is the value of the investor's consumption-related liabilities. The value of consumption-related liabilities (without mortality weighting) in year t is given by Equation 4.16, which we restate here as Equation 6.27:

$$L_t^{Ct} = \sum_{v=t}^{T} \frac{1}{(1+k_c)^t} \overline{c}_v^t, \tag{6.27}$$

where, following chapter 4, we denote the expected return of the representative portfolio for nondiscretionary spending and consumption-related liabilities as k_c.

Also, as in chapter 4, we denote the expected amount of exogenous income in year *v* as of year *t* with y_v^t. As we discuss in chapter 4, before retirement, this is wages, and once retirement starts, it includes guaranteed lifetime income, such as income from a defined benefit plan, income from a preexisting annuity, and any government-sponsored social insurance payments (such as US Social Security).

Also following chapter 4, we denote the expected return of the representative portfolio for exogenous income and human capital k_y. Human capital (without mortality weighting) is given by Equation 4.7 and restated here as Equation 6.28:

$$H_t^t = \sum_{v=t}^{T} \frac{1}{(1+k_y)^{v-t}} y_v^t. \tag{6.28}$$

Recall from the investor's balance sheet that we presented in chapter 4 that the net worth is financial assets plus human capital minus liabilities. Hence, when there is no life insurance, we can spell out net worth in terms of these components, as follows:

$$W_t = F_t + H_t^t - L_t^{Ct}. \tag{6.29}$$

We can now write the spending rule taking needs and exogeneous income into account through their corresponding components on the investor balance sheet:

$$c_t = \overline{c}_t^t + \frac{F_t + H_t^t - L_t^{Ct}}{\Delta_t}. \tag{6.30}$$

In words, for each year the spending rules are as follows: (1) estimate your net worth, (2) estimate the divisor given in Equation 6.26, (3) divide your net worth by the divisor to get your discretionary consumption, and (4) add your nondiscretionary consumption to get total spending for the year.

The Annuities Case

We now move to the case in which annuities become available at retirement. Human capital *with* mortality weighting is given by Equation 4.8 and restated here as Equation 6.31. As before, letting $\hat{H}_v^t$ denote the expected human capital (with mortality weighting) in year *v* as of year *t*, we have the following:

$$\hat{H}_t^t = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+k_y)^{v-t}} y_v^t. \tag{6.31}$$

Letting $\hat{L}_v^{Ct}$ denote the expected value of consumption-related liabilities (with mortality weighting) in year *v* as of year *t*, we have the following:

$$\hat{L}_t^{Ct} = \sum_{v=t}^{T} \hat{q}_v^t \frac{1}{(1+k_c)^t} \overline{c}_v^t. \tag{6.32}$$

The divisor becomes:

$$\hat{\Delta}_t = \sum_{v=t}^{T} (q_v^t)^{\eta} (\hat{q}_v^t)^{1-\eta} \left(\frac{1+g}{1+h} \right)^{v-t}. \tag{6.33}$$

When annuities are available, the spending rule is as follows:

$$c_t = \bar{c}_t^t + \frac{F_t + \hat{H}_t^t - \hat{L}_t^{Ct}}{\hat{\Delta}_t}. \tag{6.34}$$

In words, for each year, this version of the spending rules says to (1) estimate your net worth using mortality weighting on human capital and liabilities; (2) estimate the divisor given in Equation 6.33; (3) divide your net worth by the mortality-adjusted divisor to get your discretionary consumption; and then, (4) add your nondiscretionary consumption to get total spending for the year.

The models we presented in chapter 5 do not include market risk, and the investors fund discretionary consumption with SPIAs. This type of annuity is also called an IFA, to emphasize that payments are all the same fixed amount, which we assume to be real (inflation-adjusted). With risky markets, however, payments need to vary with market performance. This is what an IVA does. Next, after discussing asset allocation and the life insurance case, we discuss IVAs and how to use them to fund discretionary consumption.

The Life Insurance Case

As in chapter 5, we assume that the investor can purchase life insurance to guarantee a bequest. To determine the size of the bequest, we use an intergenerational utility function to pit the investor's discretionary consumption against the size of the bequest. In chapter 5, with no market uncertainty, we use constant equivalent consumption, as defined in chapter 3, as the measure of consumption (this is the constant level of consumption that results in the same utility as the optimal path of consumption). Here, with market uncertainty, we use certainty equivalent discretionary consumption, which is also defined in chapter 3. To measure constant equivalent discretionary consumption, we run a Monte Carlo simulation that includes a set of trials, each one containing a time series of discretionary consumption. **Exhibit 6.7** shows the resulting distribution of constant equivalent discretionary consumption with and without annuities available (and without life insurance) using the assumptions that we have made for Isabela. From these, we calculate the certainty equivalent discretionary consumption, with and without annuities available, using the investor's risk tolerance parameter, θ (see Equation 3.12).

Note in Exhibit 6.7 that the distribution of constant equivalent discretionary consumption with annuities available is entirely to the right of the distribution without annuities available. This distribution reflects how annuitization can result in more consumption regardless of market behavior.

Certainty equivalent discretionary consumption with no life insurance represents one end of the intergenerational trade-off between the investor's consumption and the bequest for the next generation. When annuities are available, this means putting all wealth in the life payout annuities, which we describe in the next section.

At the other end of the consumption/bequest trade-off is zero certainty equivalent discretionary consumption and the maximum possible bequest (see Equation 5.23). **Exhibit 6.8** shows the consumption/bequest trade-off curve for Isabela with annuities available at retirement. The trade-off curve is a straight line connecting the two extremes. The upper left end is at the maximum certainty-equivalent discretionary level of consumption and no bequest; and the lower right end (not shown) is at zero certainty-equivalent discretionary consumption and the maximum bequest. Exhibit 6.8 also shows the indifference curve for Isabela's maximum intergenerational utility; the curve is tangent to the trade-off line at her optimal bequest level, $1,157,671.

Exhibit 6.7. Distribution of Constant Equivalent Discretionary Consumption for Isabela, with and without Annuities Available at Retirement

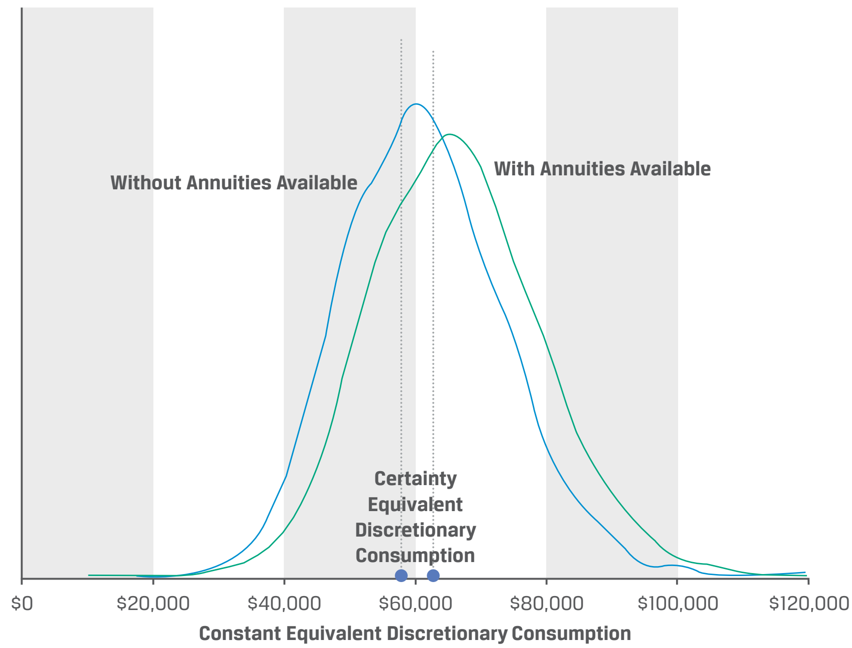
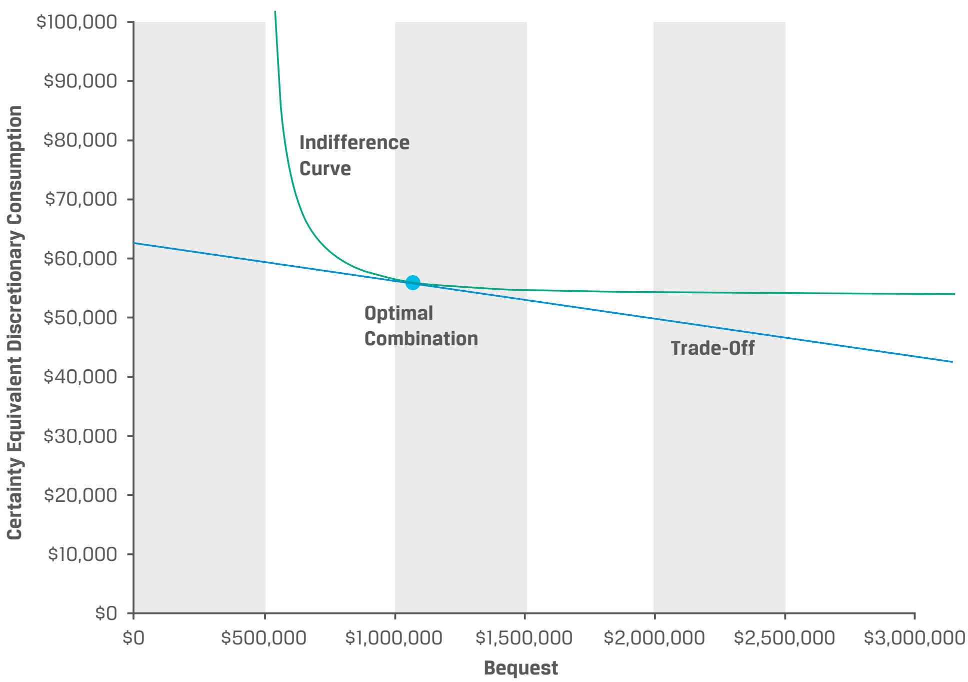


Exhibit 6.8. Maximizing Isabela's Intergenerational Utility with Annuities Available at Retirement



Immediate Variable Annuities: Real and Imagined[57]

Many investors face the prospect of outliving their money. Fortunately, this risk, known as *longevity risk*, can be mitigated with annuities. There are various types of annuities, however, and picking the right one is essential in dealing with longevity risk.

The most basic type of annuity is an IFA. Upon purchase (hence "immediate") an IFA pays a fixed amount (hence "fixed") to its holder at regular intervals for the remaining life of the holder. If the IFA is held by a couple, the IFA can have a provision that should one member of the couple die first, the surviving member will continue to receive a payment until death, possibly in an amount different than the original payments (e.g., the surviving spouse receives 75% of the original annuity income). In chapter 5, we discuss how in the context of a life-cycle model, investors can use IFAs and life insurance to guarantee lifetime income and guarantee a bequest in a world in which market returns are constant.

How IVAs Work

Of course, market returns vary. Earlier in this chapter, we discuss life-cycle models with risky market returns. When market returns are risky, the relevant type of annuity is an IVA.[58] Rather than making fixed payments, an IVA pays the investor the ratio of the value of one unit of a portfolio of risky assets (similar to a share of a mutual fund) to the value of one dollar growing at a fixed rate, called the *assumed interest rate* (AIR).

To see how an IVA works, recall that $\tilde{R}_{\theta t+1}$ denotes the realized return from year t to year $t+1$ on a portfolio of risky assets formed for an investor with risk tolerance θ. These returns can be linked over time to form an evolving cumulative index, S_v, from year t to year v, as follows:

$$\tilde{S}_v = S_t(1+\tilde{R}_{\theta t+1})(1+\tilde{R}_{\theta t+2})\cdots(1+\tilde{R}_{\theta v}). \tag{6.35}$$

The optimal value for the AIR is the certainty equivalent return, which we denote as h. Hence, the payoff in year v of an IVA bought in year t is as follows:

$$\tilde{P}_v^t = \frac{\tilde{S}_v/S_t}{(1+h)^{v-t}}. \tag{6.36}$$

Exhibit 6.9 shows selected percentiles of the payouts, over time, of an IVA that is particularly suited for Isabela because its AIR is equal to the certainty equivalent return of her net worth.

The Optimal Spending Rule when Annuities Are Not Available

Earlier in this chapter, we presented the optimal spending rule based on the investor's preferences, needs, and circumstances when annuities are not available. Here, we show how it is related to IVAs.

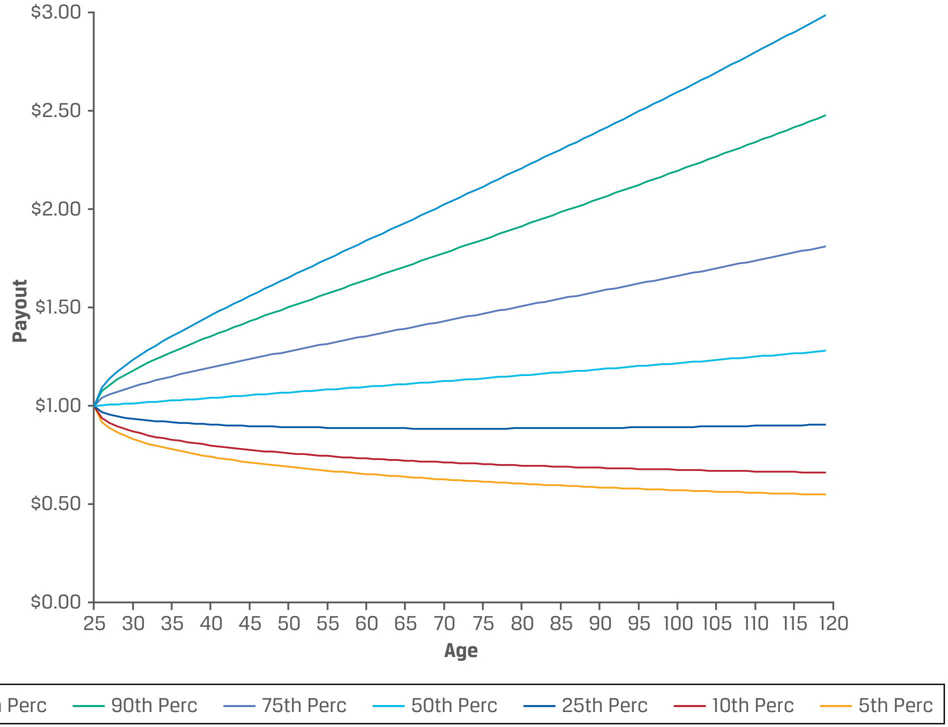
We continue to distinguish between nondiscretionary and discretionary consumption. Recall that for year t, nondiscretionary consumption is $\bar{c}_t$, total consumption c_t, and discretionary consumption is $c_t - \bar{c}_t$. As we

---

[57]This section is based on Kaplan (2022b). It builds off an insight in chapter 6 of Milevsky (2006).

[58]See section 6.12 in Milevsky (2006) for a more detailed mathematical discussion of IVAs. Also, "[n]ote that immediate variable annuities are distinct from and should not be confused with deferred variable annuities [VAs], which are tax-deferred accumulation polices that allow the investor to allocate funds to risky or variable investment funds." (Milevsky 2006, p. 131).

Exhibit 6.9. Percentiles of the Payout of an IVA for Isabela over Time



discussed in chapter 4, nondiscretionary consumption, and hence the value of consumption-related liabilities, can vary over time in ways that are correlated to asset class returns.

Human capital is also an element of the spending rule. As we discussed in chapter 4, exogenous income, and hence human capital, can vary over time in ways that are correlated to asset class returns.

The return that we introduced in Equation 6.17, $\tilde{R}_{\theta t+1}$, is actually the return on net worth. Hence, net worth evolves as follows:

$$\tilde{W}_{t+1} = (1+\tilde{R}_{\theta t+1})(W_t - c_t + \bar{c}_t). \tag{6.37}$$

As discussed earlier in this chapter, the decisions of the investor are constrained by an intertemporal budget constraint that says that net worth must be able to pay for present and all future discretionary consumption. When annuities are not available, imposing the intertemporal budget constraint leads to the following spending rule:[59]

$$c_t = \bar{c}_t + \frac{W_t}{\Delta_t}, \tag{6.38}$$

where the divisor Δ_t is given by Equations 6.25 and 6.26 taken together.

[59] Note Equations 6.29 and 6.30 taken together are equivalent to Equation 6.38.

Another form of this spending rule is as follows:

$$c_t = \bar{c}_t + N_v^t \tilde{P}_v^t, \tag{6.39}$$

where N_v^t is the number of IVAs "held" in year v, given that $\frac{W_t}{\Delta_t}$ were "held" in year t. This is given as follows:

$$N_v^t = RF_v^t (1+g)^{v-t} \frac{W_t}{\Delta_t}. \tag{6.40}$$

We put "held" in quotation markets because we are assuming that the investor does not hold actual IVAs. This spending rule says that the investor can achieve the optimal pattern of consumption by mimicking IVAs as a form of self-annuitization. They can do this by making withdrawals on a portfolio of conventional assets. In this case, IVAs are imagined.

As discussed earlier in this chapter, the rescheduling factor, RF_v^t in Equation 6.40, captures the effect of time on discretionary consumption in the absence of annuities (see Equation 6.25 for the definition of RF_v^t, and Exhibit 6.6 to see how it varies over time). Because the investor cannot guarantee future income, in the absence of any other effects, the investor will plan on reducing discretionary consumption in future years because the likelihood of being alive to enjoy it diminishes. This size of the reduction, however, depends how flexible the investor is in shifting consumption between periods.

How the number of IVAs evolves over time depends on the combined effect of the rescheduling factor and the growth factor. **Exhibit 6.10** presents an example of this that is based on our assumptions regarding Isabela. For the case with no annuities, we first have the number of IVAs "held" by a 25-year-old increasing until age 65. In this case, the growth term prevails. The number of IVAs declines gradually until age 75 and then declines rapidly as the likelihood of being alive does.

Exhibit 6.11 shows percentiles of discretionary consumption. We calculated these by multiplying the number (without annuities available) of IVAs shown in Exhibit 6.10 by the percentiles of IVA payoffs in Exhibit 6.9. Reflecting the number of IVAs "held" in Exhibit 6.10, in the higher percentiles, discretionary consumption rises from ages 25 to about 75, and then it declines.

The Optimal Spending Rule when Annuities Are Available at Retirement

Now, suppose that actual annuities are available. In this case, human capital with mortality weighting in year t is $\hat{H}_t^t$ as defined in Equation 4.8 from chapter 4. Similarly, the value of consumption-related liabilities with mortality weighting in year t is $\hat{L}_t^{Ct}$ as defined in Equation 4.17. Hence, when annuities are available, net worth (without life insurance) is as follows:

$$\hat{W}_t = F_t + \hat{H}_t^t - \hat{L}_t^{Ct}. \tag{6.41}$$

With annuities available, net worth evolves not only with returns on assets but also with survival probabilities as follows:

$$\tilde{\hat{W}}_{t+1} = \frac{(1+\tilde{R}_{\theta t+1})}{\hat{q}_{t+1}^t}(\hat{W}_t - c_t + \bar{c}_t). \tag{6.42}$$

Exhibit 6.10. Number of IVAs with and without Actual Annuities Available

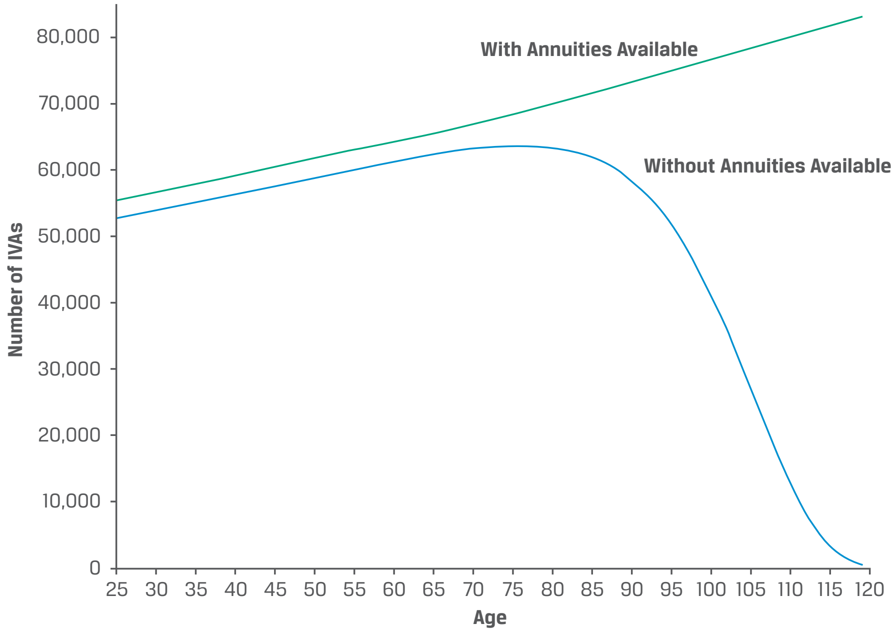
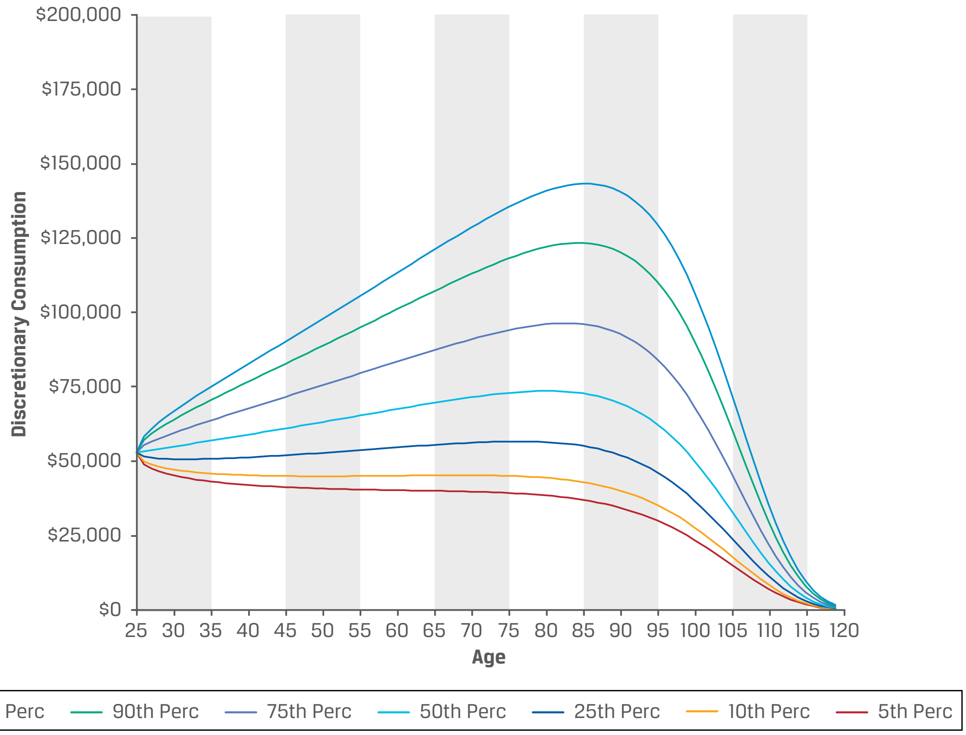


Exhibit 6.11. Percentiles of Discretionary Consumption when Annuities Are Not Available



For example, if the return from year t to year $t+1$ was 10% and the probability of surviving from year t to year $t+1$ was 95%, the combined effect would be $1.10/0.95 - 1 = 15.79\%$. The additional 5.79% is a mortality credit, as described in Milevsky (2006). Mortality credits arise because in any given cohort of annuitants, those that die forfeit their shares of the underlying investment portfolio to the survivors. Mortality credits are part of what makes annuities valuable to investors who have wealth that they do not plan on leaving as part of a bequest.

With annuities available, the divisor in the formula for discretionary consumption is $\hat{\Delta}_t$ as defined in Equation 6.33. Again, note that the formula for $\hat{\Delta}_t$ does *not* include survival probabilities.

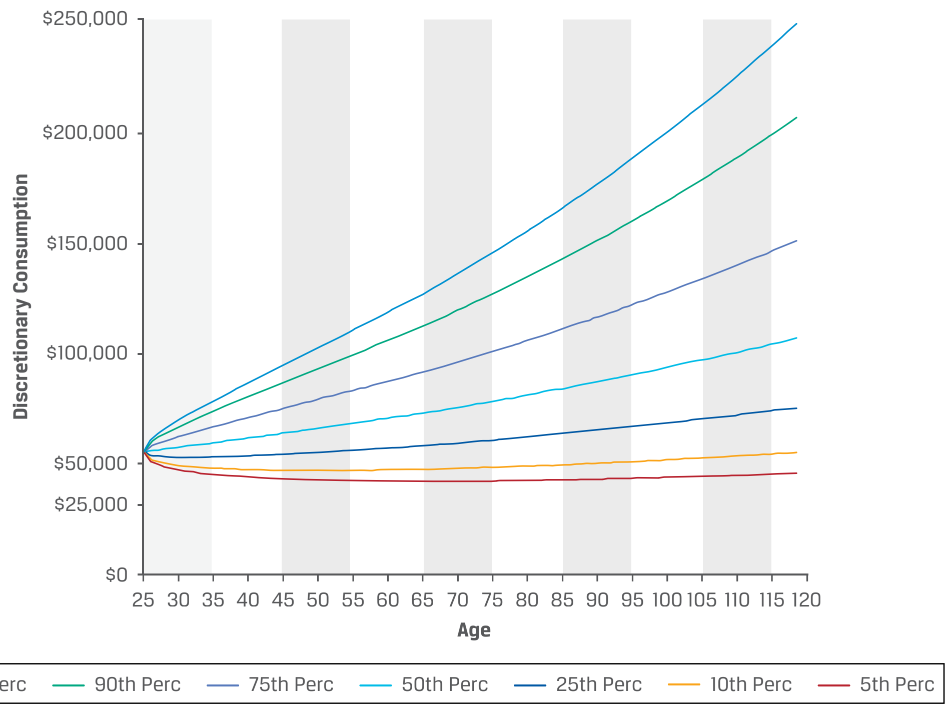
When annuities are available at retirement, the spending rule is as follows:

$$c_t = \bar{c}_t + \frac{\hat{W}_t}{\hat{\Delta}_t}. \qquad (6.43)$$

This spending rule can also be written in terms of IVAs:

$$c_t = \bar{c}_t + \hat{N}_v^t \tilde{P}_v^t, \qquad (6.44)$$

Exhibit 6.12. Percentiles of Isabela's Discretionary Consumption when Annuities Are Available at Retirement



where $\hat{N}_v^t$ is the number of IVAs held in year v, given that $\frac{\hat{W}_t}{\hat{\Delta}_t}$ were held in year t. This is given as follows:

$$\hat{N}_v^t = \left(\frac{q_v^t}{\hat{q}_v^t}\right)^{\eta} (1+g)^{v-t} \frac{\hat{W}_t}{\hat{\Delta}_t}. \tag{6.45}$$

As Exhibit 6.10 shows, when using actual IVAs, their number grows at the constant rate g.

Exhibit 6.12 shows percentiles of discretionary consumption for Isabela, assuming that she follows the strategies that we present in this chapter for the remainder of her life. We calculated these by multiplying the number of IVAs (with annuities available) shown in Exhibit 6.9 by the percentiles of IVA payoffs.

The percentiles of discretionary consumption in Exhibit 6.12 are quite different from those in Exhibit 6.11. Rather than rising and then falling, the higher percentiles only rise. Thus, at retirement, the annuities provide the income needed to fund discretionary consumption.

Conclusion and Key Takeaways

Investors can take advantage of risky capital markets that provide expected returns greater than the risk-free rate. To do so, however, they must take on risk. The combination of risk and expected return that the investor ends up with depends on both the trade-off between risk and expected return that the capital markets offer and the risk tolerance of the investor.

The expected returns that markets offer largely depend on the market prices of assets. Financial economists have developed an asset pricing framework in which the price of an asset that pays an uncertain cash flow on a known future date is the expected value of the SDF times the cash flow. This framework applies to the investor's consumption starting in the present and extending into the future and thus gives us the market value of the entire consumption stream.

The investor's consumption and investment decisions are subject to the intertemporal budget constraint, which states that the market value of the consumption stream must equal the investor's wealth. From the intertemporal budget constraint and the investors' preferences (including risk tolerance), needs (nondiscretionary consumption), and circumstances, in this chapter, we derive the investor's optimal investment and consumption decisions.

We derive two forms of the optimal consumption rule: one in which annuities are not available, and one in which they are available at retirement. When annuities are not available, investors schedule their discretionary consumption to be highest during the near future and to be lowest in the far future when the probability of being alive to enjoy consumption is lowest. In contrast, when annuities are available at retirement, investors are free to plan for higher discretionary consumption during retirement, because the annuities will provide income at every possible age during retirement.

The type of annuity that investors should use in risky markets is the IVA. IVAs combine market returns with mortality credits and provide investors with lifetime variable income. IVAs can also serve as models for spending rules in which the investor mimics IVA payments by making withdrawals on a portfolio of conventional assets. This is the application of imaginary IVAs. Real or imagined, IVAs can be useful tools for managing longevity risk.

In chapters 3 through 6, we have presented a theoretical framework of life-cycle finance and have illustrated how it works for a hypothetical investor, Isabela. In chapters 7 through 11, we present some practical investment tools designed to help investors form portfolios that are consistent with the theory that we have presented here.

PART II: CHILD ASSET LOCATION AND ALLOCATION MODEL

> Taxes are one of the most persistent frictional costs investors face. Unfortunately, asset allocation methods used by investment practitioners have generally been implemented with either complete or varying degrees of indifference towards the impact of taxes on long-term wealth.
> —Kenneth A. Blay and Harry M. Markowitz (2016, p. 26)

Kenneth Blay and Harry Markowitz open their 2016 article with this quote outlining their approach to tax-cognizant portfolio construction in the pursuit of after-tax wealth creation.

In part II of this book, using outputs from the parent life-cycle model developed in part I, we develop an extension to Markowitz single-period optimization that simultaneously solves for both asset location and asset allocation. The result is an implementation of the investment strategy at the asset allocation level for the current period given by the parent model. For simplicity, we call the child model "net-worth optimization," although it also an asset allocation and location model.

The model we propose has a number of similarities to the one presented by Wilcox, Horvitz, and DiBartolomeo (2006) as well as to the one presented by Blay and Markowitz (2016). These similarities include keeping money with different tax treatments separate and taking into account the different tax rates on income and realized capital gains. There are important differences, however, especially with the Blay and Markowitz (2016) model. Most notably, although our model leads to an efficient frontier in after-tax returns, their model leads to an efficient frontier in the weighted average of the present values of the future cash flows of the asset classes separately. Thus, their model not only segregates assets by tax treatment, but once allocations are made, their model also segregates assets by asset class within each account, as if no rebalancing occurs across asset classes over time. In contrast, in our model (and in that of Wilcox et al. 2006), the optimal weights are for a single period and are meant to be updated each period, so that portfolios are regularly reoptimized in response to new conditions (e.g., changes to the individual balance sheet, capital market assumptions, tax rates). We believe that such regular reviewing and reoptimizing of asset allocation targets are good investment practices and that investment models should be consistent with this practice.

Although we embraced life-cycle modeling in part I as part of our parent model, in part II, our child model for simultaneously solving for asset allocation and asset location is based on a direct extension of single-period MVO. As such, our approach is distinct from life-cycle-based approaches that simultaneous explore optimal asset location and asset allocation across taxable and tax-deferred accounts, such as Dammon, Spatt, and Zhang (2004) and Huang (2008).

In chapter 7, we explain how to estimate the effective tax rate of each asset class, based on the tax rates of the investor and the tax properties of the asset class.

In chapter 8, using the effective tax rates from chapter 7, we extend the single-period, asset-only Markowitz optimization model to simultaneously solve for both asset location and asset allocation, taking the investor's balance sheet into account. In each period, the parent life-cycle model from part I calculates the values of the three distinct components of the investor's balance sheet (i.e., financial assets, human capital, and liabilities) and passes them on to the single-period child model. Chapter 8 describes the net-worth optimization child model, which creates asset allocations to represent the portfolio of financial assets. It does this using an extension of the surplus optimization framework of Leibowitz and Henriksson (1988) and Sharpe and Tint (1990). To apply this approach to net worth, from the parent model, the parent model passes values for financial assets, human capital, and liabilities to the child model. Then based on the principals we discuss in chapter 4, we assign an asset allocation to human capital and to liabilities. Within the extended surplus optimization framework, we treat human capital as a portfolio held long and liabilities as a portfolio held short. We then run the optimizer to maximize the utility of the asset allocation of net worth, using a utility function based on the risk tolerance parameter of the parent model.

In the child model stage, we assume that there are two types of accounts: taxable and tax advantaged. The outputs from the net-worth optimizer are separate, tax-efficient asset allocation targets for each account type.

Looking ahead to part III, we extend the after-tax parameter adjustment process for asset classes to funds. We also expand our account types to include taxable, tax-deferred, and tax-exempt accounts.

7. EFFECTIVE TAX RATES FOR ASSET CLASSES[60]

Context

In this chapter, we examine how taxes affect asset class returns to develop a set of asset-class-specific effective tax rates. We use these effective tax rates in chapter 8 to simultaneously find the optimal asset allocation and location of assets between taxable and tax-advantaged accounts, while taking the investor's balance sheet from part I into account. The separate asset location and asset allocation targets, which come out of that analysis, serve as inputs into the multi-account portfolio construction optimization in chapter 11.

Key Insights

- When thinking about the impact of taxes on asset class returns, it is important to start with the pretax return generation process.
- Reverse optimization is a technique for forming a set of pretax expected returns such that a given asset mix is on the mean–variance efficient frontier.
- The expected return on tax-exempt (municipal) bonds can be modeled as the expected return that comes from reverse optimization, with the tax rate of the marginal tax-exempt bond investor applied.
- To develop effective rates, we take the ratio of each asset class's after-tax expected return to its pretax expected return.
- To develop preliquidation after-tax expected returns, for each asset class, we apply the tax rate on income to the income portion of the expected return and the capital gains rate to the capital gains portion of the expected return.
- To calculate the capital gains portion of the expected return on an asset class, we need an estimate of turnover and of the cost basis.
- Because the cost basis for an asset class is generally not known, we estimate it using a model that assumes that all of the parameters for the asset class have remained unchanged over a long period.
- We assume that the securities representing each asset class are sold off after a long period so that a long-term capital gains tax must be paid. We apply the long-term capital gains tax to come up with a final postliquidation after-tax expected return.

As Benjamin Franklin famously wrote, "In this world nothing can be said to be certain, except death and taxes."[61] Yet the standard asset allocation paradigm, the MVO model of Markowitz (1952, 1959), which we touched on in chapter 3, does not consider taxes. The life-cycle models in chapters 5 and 6 accept and model death. In this chapter, we begin to accept and model the inevitability of taxes.

For many individual investors, different accounts have different tax consequences. Hence, an individual investor not only faces the problem of asset allocation but also *asset location*—that is, which asset classes and which investments to locate in which account.

---

[60]The chapter is based on Kaplan (2020c).

[61]Although the quote is attributed to Franklin, the "death and taxes" saying did not originate with him. See https://en.wikipedia.org/wiki/Death_and_taxes_(idiom).

In this chapter, we discuss how to form effective tax rates as an additional set of inputs for tax-aware MVO. In the next chapter (chapter 8), we discuss how to use this type of MVO to jointly perform asset allocation and location in the context of a life-cycle model with risky assets as discussed in chapter 6. In chapter 11, we discuss how to simultaneously solve for optimal asset allocation and location across accounts with different tax treatments in a single optimization.

Tax-Efficiency and Asset Location

For individual investors, multiple accounts with different tax treatments and available investments result in a complicated web of choices with real-world tax implications. Successfully navigating this web leads to "tax alpha." The relative quality and costs of available investments may vary across different accounts. Moreover, the ability to look across accounts while considering tax efficiency allows one to select the best funds. Similarly, for investors who want additional personalization based on their nonpecuniary preferences, the ability to contemplate a wider range of available investments from across their different accounts may allow for greater personalization while concurrently considering investment quality, costs, and tax efficiency.

In the United States, various evolving tax rules related to different account types (tax-exempt, tax-deferred, and taxable), differences in the taxation of short-term versus long-term capital gains, different tax rates for qualified versus nonqualified dividends, different tax rates for income versus capital appreciation, and the ability to offset taxable gains with taxable losses collectively create opportunities for tax-aware, multi-account portfolio management to add significant tax alpha. Additionally, for many investors, the prospect of a lower federal income tax rate in retirement creates the opportunity to add value through tax-aware investment decisions.

To provide a real-world example of the potential impact on returns when funds are held in different account types, in **Exhibit 7.1**, we identify two well-known equity funds and two well-known fixed-income funds. For each of these two asset classes—equities and fixed income—we have purposely selected one passive (index) fund and one active fund. Using information reported in each fund's prospectus, we identify the 10-year annual pretax return, the 10-year annual after-tax (including distributions) return, and the effective tax rate.[62]

In general, equity funds are more tax efficient than fixed-income funds because the majority of their returns come from capital appreciation rather than interest, which is often taxed at higher rates. In this simple example, the average effective tax rate is 7.39% for equity funds and 36.02% for fixed-income funds—a difference of 28.63%. Moving to the active–passive dimension, the average effective tax rate is 24.59% for the two active funds and 18.82% for the two passive funds—a difference of 5.77%. Although this example contains only four funds, this pattern is indicative of what we would expect to see across a larger sample: Equity funds are more tax-efficient than fixed-income funds, and passive funds are more tax-efficient than active funds. These prospectus-based figures often assume the highest possible tax rates; as we shall see, another way in which the proposed framework personalizes the recommendation is by using the investor's specific tax rates.

[62]As we discuss later in this chapter, an effective tax rate is the combined effect of ongoing tax on income, ongoing tax on capital gains, and capital gains tax on an assumed date of liquidation. Letting R_B denote pre-tax return, R_A denote after-tax return, and τ the effective tax rate, we have $R_A = (1 - \tau)R_B$. See Kaplan (2020c).

Exhibit 7.1. Potential Impact of Account Type on Returns

Fund/Manager	10-Year Annual Pretax Return	10-Year Annual After-Tax Return	Effective Annual Tax Rate (Authors' Calculations)
Fidelity Magellan Fund	12.82%[a]	11.36%[a]	11.39%
Vanguard 500 Index Fund	12.97%[a]	12.53%[a]	3.39%
Average Effective Tax Rate: *Equity* Funds			7.39%
Pimco Total Return Fund	4.71%[a]	2.93%[a]	37.79%
iShares Core US Agg. Bond ETF	3.30%[a]	2.17%[a]	34.24%
Average Effective Tax Rate: *Bond* Funds			36.02%
Average Effective Tax Rate: *Active* Funds			24.59%
Average Effective Tax Rate: *Passive* Funds			18.82%

[a]Data sources are Fidelity Magellan Fund Prospectus (2019), Vanguard 500 Index Fund Prospectus (2019), Pimco Total Return Fund Summary Prospectus (2019), and iShares Core US Aggregate Bond ETF Summary Prospectus (2019).

Notes: Assumed tax rates in US fund prospectuses usually assume the highest possible federal and state tax rate on each type of income. Most investors pay lower rates. We calculated the effective annual tax rate as 1 – (10-year annual *after*-tax return/10-year annual *pre*tax return) to highlight the potential drag from taxes.

Pretax Reverse Optimization

Before getting into the calculation of effective tax rates, we need a method for developing pretax capital market assumptions. We do this through the method of *reverse optimization*.

The standard MVO model of Harry Markowitz (1952, 1959) includes three sets of pecuniary inputs, which we denote as follows:

μ_i = the expected returns of asset class *i*;

σ_i = the standard deviation of returns of asset class *i*; and

ρ_{ij} = the correlation between the returns of asset classes *i* and *j*.

In principle, all of these parameters should be forward-looking. In practice, the standard deviations and correlations are often estimated from long-term historical return data under the assumption that these parameters are stable over time. Given how poor past performance is as a predictor of future performance, it is poor practice to use *past* returns as optimization inputs. We need an alternative forecasting method.

One method of forming pretax expected returns is *reverse optimization*. First proposed by William Sharpe (1974), reverse optimization takes standard deviations and correlations as given and assumes that a particular portfolio or asset allocation (usually a very well-diversified one or even a "world wealth" portfolio that attempts to include all asset classes in proportion to their market capitalization) is mean–variance efficient. From these assumptions, it infers the set of expected returns that would in fact make the portfolio efficient.

We call the asset mix assumed to be efficient the *reference portfolio*. Typically, the reference portfolio is constructed using the aggregate market values of the asset classes as noted previously. In reverse optimization, the covariance matrix is used to calculate the sensitivity of each asset class to the reference portfolio. The sensitivity to each asset class to the reference portfolio is called its *beta*. It is similar to the beta in the CAPM. If the reference portfolio is the market portfolio, the betas in reverse optimization are the same as those in the CAPM.

In addition to the beta of each asset class, reverse optimization requires an assumption regarding the expected return of two asset classes or asset mixes. The two assumed expected returns are typically for cash and for the reference portfolio. If the reference portfolio is the market portfolio, the difference between the expected return on the market portfolio is the market premium of the CAPM.

In chapter 6, Exhibit 6.3, we showed that the trade-off between expected return and the standard deviation for the investor is our life-cycle model with uncertainty. The capital market assumptions that we derive in this chapter are on a set of asset classes and are meant to be used as inputs to a mean–variance optimizer that generates an efficient frontier. Ideally, the trade-off between risk and expected return that this efficient frontier depicts *should* be similar to the trade-off depicted in Exhibit 6.3, thus linking the parent life-cycle model with the child single-period optimization model. We create this linkage by selecting the expected return of the reference portfolio to bring the two trade-off curves close to each other, as **Exhibit 7.2** demonstrates. Kaplan and Idzorek (2023), written immediate following this manuscript, focuses explicitly on the linkage of the life-cycle model with the mean–variance optimizer and provides further details on the process.

Exhibit 7.3 shows the reverse optimized expected returns for our child model. To create this example, we estimated a covariance matrix from historical returns on indexes that represent the 10 asset classes listed

Exhibit 7.2. Trade-Offs between Risk and Expected Return in Life-Cycle and Single-Period Optimization Models

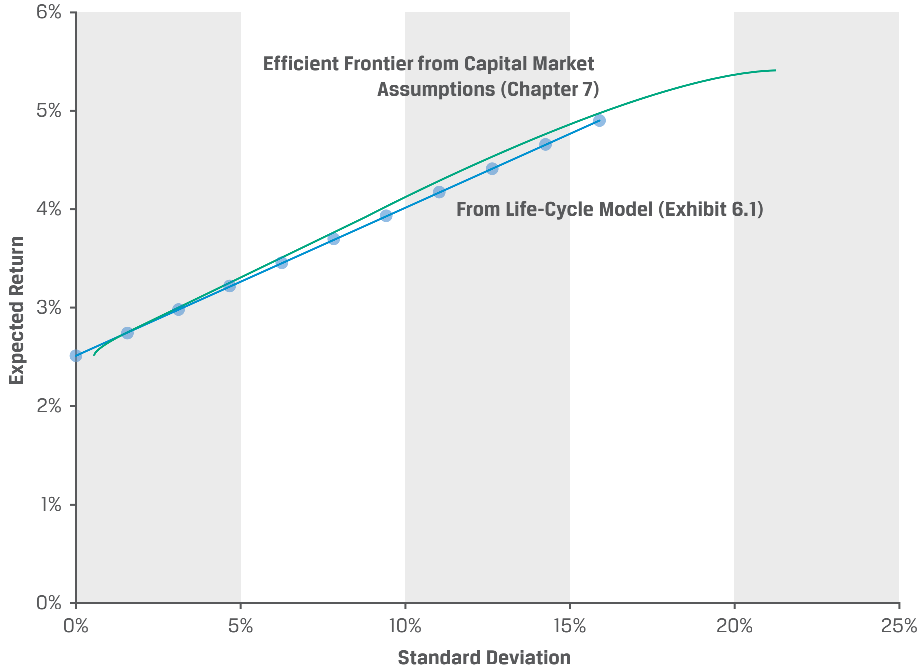


Exhibit 7.3. Reverse Optimization without Taxes in the Child Model

Asset Class	Reference Portfolio Weights	Beta	Expected Return	Standard Deviation
US Large-Cap Stocks	17.31%	1.43	4.68%	15.42%
US Mid/Small-Cap Stocks	7.42%	1.65	5.01%	17.95%
Global DM × US Stocks	21.89%	1.67	5.05%	16.71%
Emerging Market Stocks	5.68%	1.91	5.40%	21.42%
US Bonds	18.66%	0.12	2.69%	3.79%
Inflation-Linked Bonds	6.22%	0.24	2.88%	5.81%
Muni Bonds	0.00%	0.14	N/A	N/A
Global Bonds × US	22.82%	0.51	3.29%	8.33%
Cash	0.00%	0.00	2.50%	0.55%
Reference Portfolio	100.00%	1.00	4.02%	9.55%

in Exhibit 7.3 (see Appendix 7A).[63] Exhibit 7.3 shows the standard deviations on the 10 asset classes that come from the covariance matrix. It also shows the reference portfolio, betas, and expected returns for the 10 asset classes.

The weights of the reference portfolio are somewhat based on the market values of the asset classes.[64] Note that although this model includes municipal bonds and cash, their allocations in the reference portfolio are both zero. For municipal bonds, we did this because while municipal bonds would not be held in a nontaxable account, we include them as an asset class in tax-advantaged accounts. We do not show the results for municipal bonds because they are not relevant.

To model municipal bonds, we come up with a pretax expected return through reverse optimization, and then we apply a tax rate to it to come up the after-tax expected return. The tax rate that we apply need not be the same as the tax rate of the investor in question. Rather, it should be the tax rate of the *marginal investor*. The market for municipal bonds includes investors who have different tax rates. Some investors have low tax rates, so they find that taxable bonds offer higher after-tax returns than nontaxable municipal bonds. Other investors find that, after taxes, municipal bonds offer better returns than taxable bonds. The marginal investor is the investor who is on the fence between which scenario is better, and this investor's tax rate is what we should use when modeling municipal bonds. In this case, we use the pretax expected return from reverse optimization, 2.72% (not shown in Exhibit 7.3) and apply a tax rate of 30% as the tax rate of the marginal investor, for an expected return of 1.90%. Because the after-tax return is what is realized whether municipal bonds are held in the taxable or tax-advantaged account, we set the pretax

[63]One could certainly use a forward-looking estimate of the variance–covariance matrix.

[64]For further applied examples of reverse optimization, including the challenges of estimating asset class market values, see Idzorek (2007).

expected return to be the after-tax expected return (1.90%) and set the effective tax rate on municipal bonds to 0%.

We included cash with a zero allocation because, although cash is not part of the reference portfolio, we use it to set the assumptions that allow us to use the model to come up with expected returns on all the other asset classes. We assume that the expected return on cash is 2.5% (matching the assumed risk-free rate in previous chapters) and that the expected return on the reference portfolio is about 4.07% (which we determined by aligning the trade-offs between risk and expected returns shown in Exhibit 7.2).

Preliquidation After-Tax Expected Return

To form after-tax expected returns, in addition to the before-tax expected return, four additional pieces of information are needed for each asset class:

1. The division of before-tax expected return into expected income and capital gains returns.
2. The division of expected income return into qualified and nonqualified dividend (income) returns.
3. The division of the expected capital gains return into expected short-term and long-term capital gains.
4. Turnover.

From these data, we can estimate preliquidation and postliquidation expected after-tax returns.[65] Preliquidation returns are estimated under the assumption that unrealized capital gains are not taxed. Postliquidation returns are estimated under the assumption that, at a specific day in the future, the cumulative value of the assets invested in an asset class are sold and previously unrealized capital gains are realized and taxed. In the remainder of this section, we discuss how to estimate preliquidation expected after-tax returns. In the next section, we discuss how to use preliquidation returns to estimate postliquidation returns.

For the division of before-tax expected return and expected income on each asset class *i*, we have the following:

$$\mu_{Bi} = IR_i + CG_i, \tag{7.1}$$

where:

μ_{Bi} = the before-tax expected total return on asset class *i*;

IR_i = the expected income return on asset class *i*; and

CG_i = the expected capital gain on asset class *i*.

We assume two tax rates:

τ_{OI} = tax rate on ordinary income; and

τ_{LTCG} = tax rate on long-term capital gains.

[65]Wilcox et al. (2006) also estimate effective tax rates for use in MVO to jointly solve for asset allocation and location. Their approach, however, only takes into account the tax rate on income (dividends) and the long-term capital gains tax at liquidation. See their Appendix A.

Under the US tax code, there are two kinds of dividends: qualified and nonqualified. Qualified dividends are taxed at the long-term capital gains rate and nonqualified dividends are taxed at the ordinary income rate. (Interest income is taxed at the ordinary income rate.) Let:

q_i = the fraction of income that is qualified; and

τ_{Ii} = the blended income tax rate for income for asset class *i*.

The tax rate for income for asset class *i* is the blended rate, as follows:

$$\tau_{Ii} = q_i \tau_{LTCG} + (1 - q_i)\tau_{OI}. \quad (7.2)$$

Also, under the US code, long-term capital gains are taxed at the long-terms capital gains rate and short-term capital gains are taxed at the same rate as ordinary income. Let:

LT_i = the fraction of capital gains that is long-term; and

τ_{CGi} = the blended capital gains for asset class *i*.

The tax rate for capital gains for asset class *i* is the blended rate:

$$\tau_{CGi} = LT_i \tau_{LTCG} + (1 - LT_i)\tau_{OI}. \quad (7.3)$$

To model realized capital gains, we assume that each asset class is held as a portfolio of shares of a single fund that represents the asset class. We assume that this portfolio is subject to the same turnover that investors in the asset class typical experience. Let:

TO_i = the turnover rate for asset class *i*; and

$COST_i$ = the average historical cost of all shares as a fraction of beginning-of-year market value.

The preliquidation after-tax expected return follows:

$$\mu_{Ai} = (1 - \tau_{Ii})IR_i + CG_i - TO_i(1 + CG_i - COST_i)\tau_{CGi}. \quad (7.4)$$

Every variable on the right-hand side of this equation can be estimated from asset class data except for $COST_i$. If the asset allocation model is being applied to an actual investor's portfolio, $COST_i$ should be calculated using actual cost basis data. If it is being applied without reference to any actual investor, $COST_i$ needs to be estimated using a model.

In this case, we use a model. To create the model, we assume that all of the variables remain constant over a long period time. This assumption leads to a second relationship between μ_{Ai} and $COST_i$. We find μ_{Ai} and $COST_i$ by solving two equations in two unknowns. See Appendix 7B for details.

Exhibit 7.4 shows data on the 10 asset classes in Exhibit 7.3 as well as the results of the model for the cost basis and after-tax expected return. We use the before-tax asset expected total returns from the Exhibit 7.2. We assume that our hypothetical investor, Isabela, pays 20% on long-term capital gains and 40% on ordinary income. For the additional data on asset classes, we use data for asset class indexes. Hence, these results reflect what an index fund investor should experience.

Exhibit 7.4. Preliquidation After-Tax Expected Returns

Asset Class	Pre-Tax Expected Return			Fraction of Income Qualified	Fraction of Capital Gains Long Term	Blended Tax Rate		Turnover	Cost Basis	After-Tax Expected Return
	Total	Income	Capital Gains			On Income	On Capital Gains			
US Large-Cap Stocks	4.68%	1.19%	3.49%	97.62%	95.02%	20.48%	21.00%	33.00%	93.64%	3.75%
US Mid/Small-Cap Stocks	5.01%	1.14%	3.87%	89.47%	90.32%	22.11%	21.94%	36.52%	93.93%	3.97%
Global DM × US Stocks	5.05%	1.69%	3.36%	79.98%	89.51%	24.00%	22.10%	18.00%	87.50%	4.01%
Emerging Market Stocks	5.40%	1.52%	3.88%	73.87%	90.23%	25.23%	21.95%	33.00%	93.01%	4.23%
US Bonds	2.69%	2.69%	0.00%	0.00%	0.00%	40.00%	40.00%	100.00%	100.00%	1.61%
Inflation-Linked Bonds	2.88%	2.88%	0.00%	0.00%	0.00%	40.00%	40.00%	100.00%	100.00%	1.73%
Muni Bonds	1.90%	1.90%	0.00%	0.00%	0.00%	0.00%	40.00%	100.00%	100.00%	1.90%
Global Bonds × US	3.29%	3.29%	0.00%	0.00%	0.00%	40.00%	40.00%	100.00%	100.00%	1.97%
Cash	2.50%	2.50%	0.00%	0.00%	0.00%	40.00%	40.00%	100.00%	100.00%	1.50%

Postliquidation Capital Market Assumptions

The after-tax expected returns in Exhibit 7.4 are preliquidation returns with unrealized capital gains not taxed. It is likely that, at some point, the assets will be liquidated, and the previously unrealized capital gains will be realized and taxed. Let T be the number of years from now until the assets are liquidated.

Let:

V_{0i} = the initial value of assets in asset class i; and

V_{Ti} = the value of assets in asset class i at time T just before liquidation.

We have:

$$V_{Ti} = V_{0i}(1 + \mu_{Ai})^T. \tag{7.5}$$

The fraction of assets that are taxed at liquidation is as follows:

$$\Theta_i = \frac{CG_i}{\mu_{Bi}}(1 - TO_i). \tag{7.6}$$

The taxes on the realized capital gains at liquidation are as follows:

$$TL_i = (V_{Ti} - V_{0i})\Theta_i \tau_{LTCG}. \tag{7.7}$$

Therefore, at liquidation, the value of assets is as follows:

$$V_{Ti}^L = V_{Ti} - TL_i. \tag{7.8}$$

The postliquidation after-after tax expected return is as follows:

$$\mu_{Ai}^L = \left(\frac{V_{Ti}^L}{V_{0i}}\right)^{\frac{1}{T}} - 1. \tag{7.9}$$

Exhibit 7.5 shows the calculation of postliquidation capital market assumptions on the 10 asset classes in Exhibits 7.3 and 7.4, assuming a \$1,000 initial investment in each asset class and a 20-year holding period. **Exhibit 7.6** demonstrates the logic of the calculation of postliquidation after-tax returns. It shows the relationship between the preliquidation after-tax expected return, the taxes paid at liquidation, and the postliquidation after-tax expected return on US Large-Cap Stocks. A \$1,000 investment in this asset class growing at its preliquidation after-tax expected return of 3.75% grows to \$2,089.66 in 20 years, as the top curve shows. At liquidation, \$108.84 is paid in taxes, leaving a postliquidation value of \$1,980.82 The rate of return at which \$1,000 grows to \$1,980.82 is 3.48%, as the bottom curve shows.

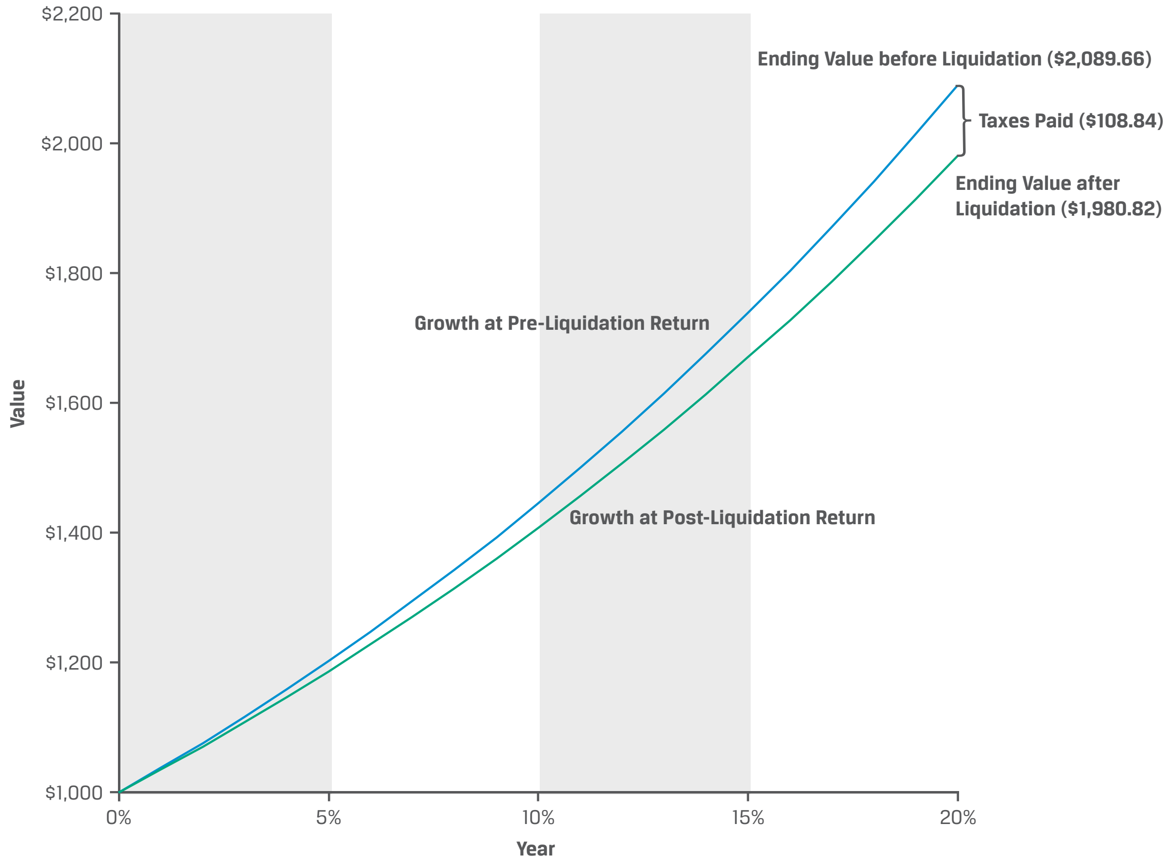
To calculate the after-tax standard deviation of return on each asset class i, we first need to calculate the asset class's *effective tax rate*. The effective tax rate is the single tax rate that, if applied to asset class's pretax return, yields the after-tax return. It is given by the following:

$$\tau_{Ei} = 1 - \frac{\mu_{Ai}^L}{\mu_{Bi}}. \tag{7.10}$$

Exhibit 7.5. Postliquidation After-Tax Expected Returns and Standard Deviations

Asset Class	Pre-Liquidation		Postliquidation					
	Expected After-Tax Return	End-of-Period Value	Fraction of Gain Taxed	Tax Paid	Final Value	After-Tax Expected Return	Effective Tax Rate	After-Tax Standard Deviation
US Large-Cap Stocks	3.75%	$2,089.66	49.94%	$108.84	$1,980.82	3.48%	25.71%	11.45%
US Mid/Small-Cap Stocks	3.97%	$2,176.72	49.01%	$115.33	$2,061.39	3.68%	26.56%	13.18%
Global DM × US Stocks	4.01%	$2,195.17	54.53%	$130.34	$2,064.83	3.69%	26.84%	12.23%
Emerging Market Stocks	4.23%	$2,291.34	48.15%	$124.35	$2,167.00	3.94%	27.05%	15.63%
US Bonds	1.61%	$1,377.24	0.00%	$0.00	$1,377.24	1.61%	40.00%	2.28%
Inflation-Linked Bonds	1.73%	$1,408.12	0.00%	$0.00	$1,408.12	1.73%	40.00%	3.49%
Muni Bonds	1.90%	$1,457.38	0.00%	$0.00	$1,457.38	1.90%	0.00%	3.14%
Global Bonds × US	1.97%	$1,477.76	0.00%	$0.00	$1,477.76	1.97%	40.00%	5.00%
Cash	1.50%	$1,346.86	0.00%	$0.00	$1,346.86	1.50%	40.00%	0.33%

Exhibit 7.6. Logic of Calculation of Postliquidation After-Tax Expected Return, US Large-Cap Stocks



The after-tax standard deviation is as follows:

$$\sigma_{Ai} = (1 - \tau_{Ei})\sigma_{Bi}. \tag{7.11}$$

In addition to the calculation of postliquidation after-tax return, Exhibit 7.4 shows the postliquidation effective tax rate and standard deviation for each asset class.

Conclusion and Key Takeaways

An individual investor needs to decide not only on the overall asset allocation for the overall portfolio but also on how to allocate between taxable and tax-advantaged accounts (the asset location problem). It is common practice to address the asset allocation and location problems sequentially by performing two optimizations: one based on after-tax capital market assumptions (CMAs) leading to taxable asset allocation models, and one based on before-tax CMAs leading to tax-advantaged asset allocation models. In the next chapter, we present an approach to addressing the asset allocation and asset location problems simultaneously that uses both sets of CMAs, or equivalently, pretax CMAs and effective tax rates.

Appendix 7A. Indices Used to Represent Asset Classes

To represent the 10 asset classes, we used the indices listed in **Exhibit 7A.1**.

Covariance Matrix

We obtained monthly returns on the indices listed in Exhibit 7A.1 for the period February 2000 through December 2022. After combing the mid-cap and small-cap indices, we calculated the covariance matrix. We multiplied the matrix by 0.0012 to convert it to annual decimal units.

Tax Parameters

We derived the tax parameters for the equity asset classes from data provided by Morningstar Investment Management.

Exhibit 7A.1. Indices Used to Represent Asset Classes

Asset Class	Index
US Large-Cap Stocks	Morningstar US Large
US Mid/Small-Cap Stocks	0.76 × Morningstar US Mid + 0.24 × Morningstar US Small[66]
Global DM × US Stocks	Morningstar DM × US
Emerging Market Stocks	Morningstar EM
US Bonds	Morningstar US Core Bond
Inflation Linked Bonds	Morningstar US TIPS
Muni Bonds	Bloomberg Municipal
Global Bonds × US	Morningstar Global × US Core Bond
Cash	Morningstar USD One-Month

Appendix 7B. Modeling the Cost Basis

In this appendix, we drop the subscript "*i*" so that we use the following notation:

IR = the income return;

CG = the capital gains return;

τ_I = the blended income tax rate;

τ_{CG} = the blended capital gains rate; and

TO = the rate of turnover.

[66]The weights reflect the relative market capitalizations of the mid-cap and small-cap indices.

We model how the market value and cost basis evolve over time. Let:

MV_t = the market value at time t; and

CB_t = the cost basis at time t.

The market value of the portfolio evolves over time as follows:

$$MV_t = (1 + CG)MV_{t-1} + (1 - \tau_I)IR \cdot MV_{t-1} - \tau_{CG}TO((1 + CG)MV_{t-1} - CB_{t-1}). \quad (7B.1)$$

The first term on the right side of Equation 7B.1 is the growth in the market value of the existing assets. The second term is income net of taxes. The third term is capital gains taxes.

Cost basis evolves as follows:

$$CB_t = CB_{t-1} + (1 - \tau_I)IR \cdot MV_{t-1} - \tau_{CG}TO((1 + CG)MV_{t-1} - CB_{t-1}) + TO((1 + CG)MV_{t-1} - CB_{t-1}). \quad (7B.2)$$

The second, third, and fourth terms of Equation 7B.2 together give the change in the cost basis. The second term is income net of taxes. The third term is capital gains taxes. The fourth term is the change in the cost basis due to the replacement of old shares with new shares.

Both Equations 7B.1 and 7B.2 are linear on V_{t-1} and B_{t-1}. We can write them as a joint system of equations. Let:

$$A_{VV} = (1 - \tau_{CG}TO)(1 + CG) + (1 - \tau_I)IR, \quad (7B.3)$$

$$A_{VB} = \tau_{CG}TO, \quad (7B.4)$$

$$A_{BV} = (1 - \tau_I)IR + (1 - \tau_{CG})TO(1 + CG), \quad (7B.5)$$

$$A_{BB} = 1 - (1 - \tau_{CG})TO. \quad (7B.6)$$

We form the coefficients given in Equations 7B.3 through 7B.6 into a 2×2 matrix:

$$\mathbf{A} = \begin{bmatrix} A_{VV} & A_{VB} \\ A_{BV} & A_{BB} \end{bmatrix}. \quad (7B.7)$$

The matrix **A** gives the joint evolution of market and book value:

$$\begin{bmatrix} MV_t \\ CB_t \end{bmatrix} = \mathbf{A} \begin{bmatrix} MV_{t-1} \\ CB_{t-1} \end{bmatrix}. \quad (7B.8)$$

Given $\begin{bmatrix} MV_0 \\ CB_0 \end{bmatrix}$, we can repeatedly multiply by **A** to project the values of $\begin{bmatrix} MV_t \\ CB_t \end{bmatrix}$ for any t. Eventually, two values will converge: (1) the growth rates of MV_t and CB_t, and (2) the ratio CB_t/MV_t, which provides a scalar, λ, and a two-element vector, $\vec{x}$, such that:

$$\mathbf{A}\vec{x} = \lambda\vec{x}. \quad (7B.9)$$

Equation 7B.9 is called the *eigenvalue* problem in matrix algebra. The eigenvalue problem is to find a combination of λ and $\vec{x}$ that satisfy Equation 7B.9. λ is called an eigenvalue and $\vec{x}$ is called an *eigenvector*. To solve for the eigenvalue, we solve the following equation for λ:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0, \tag{7B.10}$$

where *det* signifies the determinant of a matrix.

From Equation 7B.10, we have the following:

$$(A_{VV} - \lambda)(A_{BB} - \lambda) - A_{VB}A_{BV} = 0. \tag{7B.11}$$

Let:

$$b = -(A_{VV} + A_{BB}), \tag{7B.12}$$

$$c = A_{VV}A_{BB} - A_{VB}A_{BV}. \tag{7B.13}$$

We can rewrite Equation 7B.11 as follows:

$$\lambda^2 + b\lambda + c = 0. \tag{7B.14}$$

We solve Equation 7B.14 using the quadratic formula:

$$\lambda = \frac{-b + \sqrt{b^2 - 4c}}{2}. \tag{7B.15}$$

λ is one plus the long-term growth rate of both market value and the cost basis. This is the long-term preliquidation after-tax expected return. which we denote μ_A. Hence:

$$\mu_A = \lambda - 1. \tag{7B.16}$$

Recall that *COST* denotes the long-term ratio of cost basis to beginning-of-period market value. To find *COST*, we set $\vec{x} = \begin{bmatrix} 1 \\ COST \end{bmatrix}$ in Equation 7B.9 and solve for *COST*. Thus, we have the following:

$$COST = \frac{\lambda - A_{VV}}{A_{VB}}. \tag{7B.17}$$

8. NET-WORTH OPTIMIZATION IN CONJUNCTION WITH LIFE-CYCLE MODELS[67]

Context

In this chapter, we review the basic MVO framework for asset allocation and how to frame it with expected utility theory to be consistent with the life-cycle models that we discussed in Chapter 6. We also show how to expand it using effective tax rates to simultaneously solve for optimal asset allocation and location. We then show how to further expand the model to consider an investor's balance sheet when deciding on the asset allocation and location of the financial assets. This expanded MVO model is the child model of the overall framework of this book in that it links multiperiod life-cycle models to tax-aware MVO.

Key Insights

- Based on the work of Levy and Markowitz (1979), we show how single-period expected utility can be well approximated using just expected return, variance, and the same risk tolerance parameter (θ or theta) that we introduced in chapter 3 and used in chapter 6. This allows us to use MVO to maximize expected utility and thus be consistent with life-cycle models.
- We extend MVO to solve not only for the optimal asset mix but also to *jointly* solve for the optimal location of each asset class between a taxable account and a tax-advantaged account. We refer to this as *joint MVO*. The end result is two separate policy portfolios optimized for asset location: a target for all taxable accounts and target for all tax-advantaged accounts.
- To solve the asset location problem, we need an effective tax rate for each asset class. We form these using the approach described in chapter 7.
- In the spirt of liability-relative optimization or surplus optimization, we further extend MVO to take into account human capital and liabilities from the investor's balance sheet, thus linking the child single-period asset allocation model to the parent life-cycle model. We refer to this extension of MVO as "net-worth optimization."[68]

In this chapter, we expand Harry Markowitz's MVO to jointly arrive at two separate target asset allocations: one for taxable accounts and one for tax-advantaged accounts. To our knowledge this is a new extension of MVO. The inputs for the single optimization include the effective tax rates developed in chapter 7. By applying the effective tax rates to the taxable account, but not to the tax-advantaged account, the two separate target asset allocations are simultaneously optimized for optimal asset location.[69]

We begin with a review of how expected utility can be approximated in a mean–variance model. We then expand MVO to jointly solve for asset location and asset allocation across two account types. Next, we connect this new joint asset location and allocation optimization (joint MVO) to the holistic individual balance

---

[67]This chapter is partially based on Kaplan (2020d).

[68]While liability-relative optimization has been in use since the late 1970s, Leibowitz (1987) is perhaps the earliest published account, although it is presented in a much more usable form in Sharpe (1990) and Sharpe and Tint (1990). More recent important pieces include Siegel and Waring (2004) and Waring (2004a, 2004b), all working within a tax-free institutional setting. Idzorek and Blanchett (2019) apply liability-relative optimization to the creation of asset allocation for individuals.

[69]An alternative and equivalent formulation is to apply the effective tax rates to the pretax expected returns and standard deviations to form a set of after-tax capital market assumptions, and apply the after-tax capital market assumptions to the assets in the taxable account and the pretax capital market assumptions to the assets in the tax-advantaged account.

sheet approach that permeates throughout our life-cycle models. We do this by combining joint MVO with net-worth optimization, which is an extension of asset-only MVO in which the optimizer is forced to a hold both a long position in an asset (or combination of assets) representing human capital, and a short position in an asset (or combination of assets representing) the value of the liabilities on the investor balance sheet.

Approximating Expected Utility in a Mean–Variance Model

In chapter 3, we explain the logic that Harry Markowitz uses to justify MVO. The expected utility of (one plus) a random return can be *approximated* by the expected return (μ) and the variance of return (σ^2) as follows:

$$U(\mu,\sigma^2) = u(1+\mu) + \frac{1}{2}u''(1+\mu)\sigma^2. \tag{8.1}$$

This means that total utility has two parts—an expected return (μ) part that increases utility and a variance of return (σ^2) part that decreases utility. We can say that utility is decreasing in variance because of diminishing marginal utility, which makes the term $u''(1+\mu)$ negative. Hence, variance of return is "bad." Because this approximation for expected utility was first introduced by Levy and Markowitz (1979), we refer to it as the Levy–Markowitz utility function. To link the Levy–Markowitz utility function to the risk tolerance parameter, in chapters 2, 3 and 6, we assume CRRA utility. Mathematically, the CRRA assumption means that:

$$u(1+\mu) = \begin{cases} \dfrac{\theta}{\theta-1}(1+\mu)^{\frac{\theta-1}{\theta}}, & \theta \neq 1 \\ ln(1+\mu), & \theta = 1 \end{cases} \tag{8.2}$$

where θ is the level of the investor's risk tolerance, theta, which is usually between 0 and 1 (i.e., between 0% and 100%). This means that marginal utility is declining at a rate that depends on the investor's risk tolerance.

For the mathematically inclined reader, in Equation 8.3, we state the second derivative of the CRRA utility function given in Equation 8.2 that multiplies variance in Equation 8.1. This is as follows:

$$u''(1+\mu) = \frac{-1}{\theta(1+\mu)^{\frac{1+\theta}{\theta}}}. \tag{8.3}$$

The results of MVO are almost always presented as an efficient frontier, which depicts the trade-off between standard deviation of return and expected return among efficient portfolios. The goal is to maximize expected return for a given level of risk. Markowitz sees the goal somewhat differently. Rather, he sees the goal of MVO as to maximize the Levy–Markowitz utility function. Markowitz said, "My basic assumption is that you act under uncertainty to maximize expected utility" (Markowitz, Savage, and Kaplan 2010). This does not change the composition of the efficient frontier, rather it determines which efficient mix is selected by the investor. To further the linkage to our parent life-cycle model, we take the utility maximization approach.

Mean–Variance Optimization without Taxes

To mathematically describe MVO without taxes, we use the following notation:

μ_i = the before-tax expected return on asset class i;

σ_i = the before-tax standard deviation on asset class i;

ρ_{ij} = the correlation of returns between asset classes i and j;

h_i = the allocation to asset class i; and

K = the number of asset classes.

In scalar notation and using the Levy–Markowitz utility function in which the investor's risk tolerance (θ or theta) is incorporated per Equation 8.2 to calculate the investor's utility per Equation 8.1, we can write the MVO problem as follows:

$$\max_{h_1,h_2,\cdots,h_K} \begin{array}{c} U\left(\sum_{i=1}^{K} h_i\mu_i, \sum_{i=1}^{K}\sum_{j=1}^{K} h_i h_j \sigma_i \sigma_j \rho_{ij}\right). \\ s.t. \sum_{i=1}^{K} h_i = 1, h_i \geq 0 \end{array} \tag{8.4}$$

This means the following:

1. The expected return of a portfolio is the weighted average of the expected returns of the asset classes, with the weights being the asset class portfolio weights.
2. The variance of the portfolio depends on the portfolio weights, the standard deviations of the returns on the asset classes, and the correlations of returns between the asset classes.
3. The utility of the portfolio increases as expected return increases and decreases as its variance increases.
4. The optimal portfolio is the one that maximizes utility subject to the constraints that all weights are nonnegative and sum to one.

Equation 8.4 can be written more succinctly using matrix notation. For readers familiar with matrix notation, we define the following:

$\boldsymbol{\mu}_{\mathbf{K}}$ = the vector of pretax expected returns;

$\mathbf{V}_{\mathbf{K}}$ = the covariance matrix of pretax asset class returns, the ij-element is $\sigma_i\sigma_j\rho_{ij}$;

$\mathbf{h}_{\mathbf{K}}$ = the vector of allocations to the asset classes; and

$\boldsymbol{\iota}_{\mathbf{K}}$ = a vector of K ones.

Using the Levy–Markowitz utility function in Equation 8.1 with these vectors and the covariance matrix, the MVO problem can now be written as follows:[70]

$$\max_{\mathbf{h}_{\mathbf{K}}} U(\mathbf{h}_{\mathbf{K}}^{\mathsf{T}}\boldsymbol{\mu}_{\mathbf{K}}, \mathbf{h}_{\mathbf{K}}^{\mathsf{T}}\mathbf{V}_{\mathbf{K}}\mathbf{h}_{\mathbf{K}})\, s.t.\, \mathbf{h}_{\mathbf{K}}^{\mathsf{T}}\boldsymbol{\iota}_{\mathbf{K}} = 1, \mathbf{h}_{\mathbf{K}} \geq 0. \tag{8.5}$$

Equation 8.5 is the same as Equation 8.4, just written more compactly.

In chapter 7, we derived pretax expected returns on a set of asset classes using the reverse optimization technique. In reverse optimization, we assume that a reference portfolio (which is often a market portfolio) is on the efficient frontier and therefore is optimal for investors who have an as yet unknown level of risk tolerance. Using a numerical analysis, we find what level of risk tolerance, θ, makes the reference

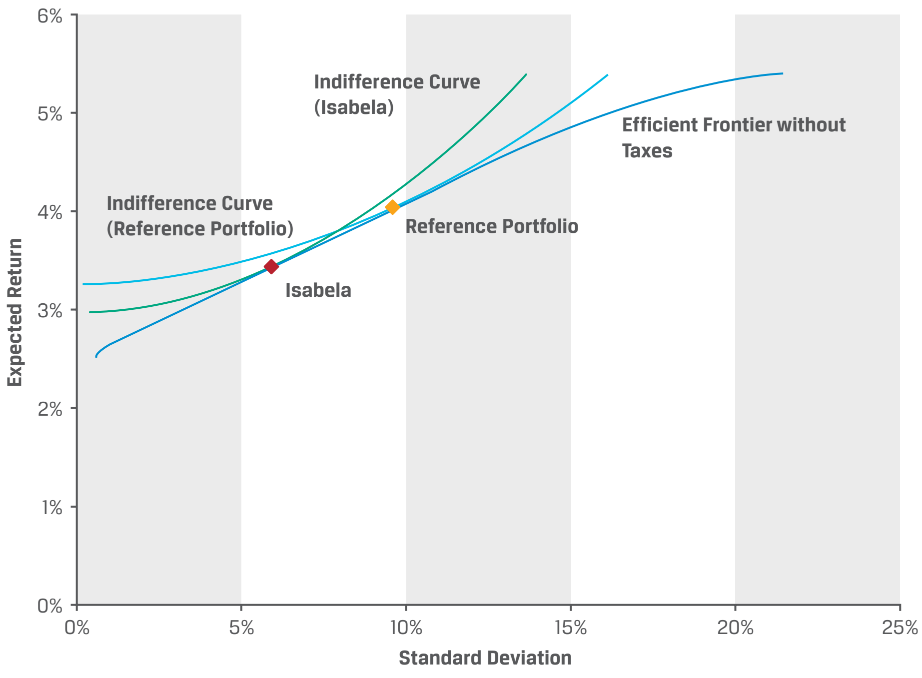
[70]As in other chapters, the superscript "T" is for transpose.

portfolio optimal. This turns out to be 56.6% (or equivalently, 0.566).[71] In **Exhibit 8.1**, we show the entire efficient frontier, the point on the frontier that represents the reference portfolio, and the indifference curve that shows that the reference portfolio maximizes the Levy–Markowitz utility function derived from the CRRA utility function with $\theta = 56.6\%$. We also show the point on the efficient frontier that maximizes the Levy–Markowitz utility function for Isabela, with $\theta = 35\%$.

To illustrate maximizing expected utility in the MVO framework, in Exhibit 8.1, we have included an indifference curve for the reference portfolio and for Isabela. Recall from chapter 3 that an indifference curve shows all combinations of expected return and standard deviation that result in the same level of utility. Indifference curves are upward sloping and increase at an increasing rate, showing that an investor is willing to take on additional risk to get additional expected return, but as we increase the level of risk, the increase in expected return needed to compensate for the risk increases.

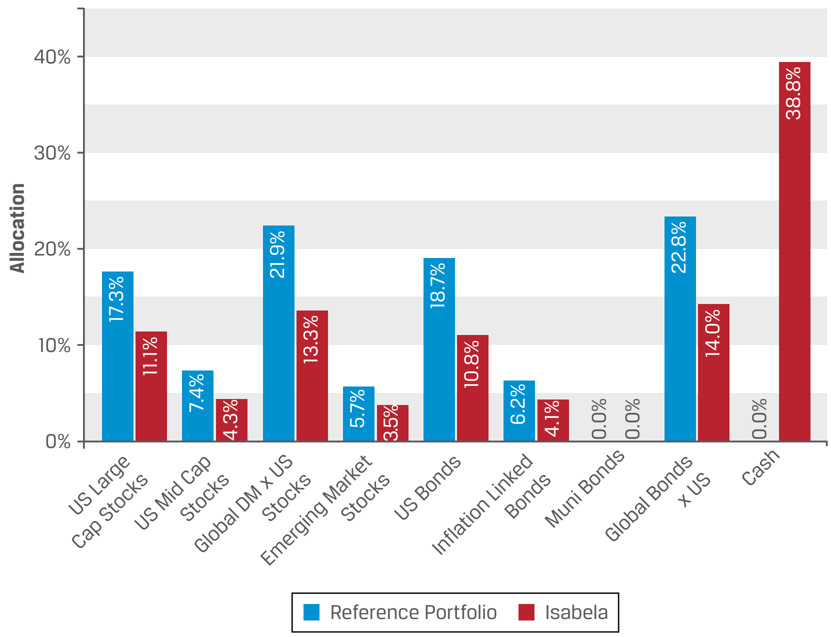
For any investor or level of risk tolerance, we could draw indifference curves that are above the efficient frontier as well as those that intersect it. For any given level of risk tolerance, however, a unique indifference curve is *tangent* to the efficient frontier, with the point of tangency being that of the optimal portfolio. This is where expected utility is the highest among efficient portfolios. Because we have assumed that the reference portfolio is an optimal portfolio, the point of tangency for the reference portfolio is the point of the reference portfolio. Similarly, because we have assumed that Isabela would invest in the portfolio that is optimal for her, the indifference curve for her is tangent to the efficient frontier. Notice that in Exhibit 8.1,

Exhibit 8.1. Risk and Expected Return without Taxes



[71]For the purpose of the example shown here, we started with a portfolio (the reference portfolio) that we assume is optimal, and then we find the corresponding value of the risk tolerance parameter. But when implementing the child model, the value of the risk tolerance parameter should be the same as is the parent model. The value should be based on investor preferences, discerned using the methods discussed in chapter 2.

Exhibit 8.2. The Reference Portfolio and Isabela's Optimal Portfolio when There Are No Taxes and Only Financial Wealth Is Taken into Account



even though each investor is indifferent to the other points on the indifference curve, those other points are all above the efficient frontier and thus are not feasible.

Exhibit 8.2 shows the asset class weights for the reference portfolio and for Isabela's portfolio. Note that the reference portfolio is 52.3% equity and Isabela's portfolio is 32.3% equity. These values are somewhat similar to the risk tolerance parameters that correspond to these portfolios, 56.6% and 35%, respectively. Next, we shall see that equity allocation can be quite different from risk tolerance when taxes and the investor's balance sheet are taken into account.

Extending Mean–Variance Optimization for the Joint Asset Allocation and Location Problem

We now present what we believe to be a new and novel extension to MVO that solves for account-specific multiple target asset allocations that simultaneously optimizes for asset allocation and tax-efficient asset location. The process is inherently personalized.

Suppose that the investor, in our example Isabela, has money spread across two account types, one taxable and one tax-advantaged. Let ϕ be the fraction of assets in the taxable accounts so that $1 - \phi$ is the fraction in the tax-advantaged accounts. We use the contemporaneous fraction based on the investor's current situation. In the case of Isabela, when she is 25 years old, as we will discuss, this is about 92.4%. The next year, when Isabela is 26 years old, the applicable fraction or split between taxable and tax-advantaged will likely be somewhat different.

Because of tax rules, money cannot be transferred between accounts. Money in the taxable account earns after-tax returns and money in the tax-advantaged account earns pretax returns. In chapter 7, we show how to derive an effective tax rate, τ_i, for asset class *i*. Effective tax rates allow us to convert pretax expected returns and standard deviations into their after-tax counterparts:

$$\mu_{Ai} = (1 - \tau_i)\,\mu_i, \tag{8.6}$$

$$\sigma_{Ai} = (1 - \tau_i)\sigma_i. \tag{8.7}$$

For example, if the expected return on an asset class were 6% and the effective take rate were 33.33%, one-third of returns goes to taxation leaving two-thirds for the investor. So, the after-tax expected return is 4%.

To incorporate effective tax rates into an expanded version of the MVO problem using matrix notation, we introduce a *K′K* matrix, $\mathbf{T_K}$, in which the *i*th diagonal is $1 - \tau_i$ and all off-diagonal elements are 0. This gives us a simple way of applying the effective tax rates in a compact notation. Let $\mathbf{h_K^A}$ and $\mathbf{h_K^B}$ be the asset allocation vectors of the taxable and tax-advantaged accounts, respectively. The expanded MVO problem is to jointly select $\mathbf{h_K^A}$ and $\mathbf{h_K^B}$ to maximize the expected utility of after-tax return over the portfolio as a whole. Hence, it not only solves for asset allocation but also for asset location. Thus, we refer to it as the *joint MVO* problem.[72]

We can write the joint MVO problem in matrix notation by extending the terms for expected return and variance in Equation 8.5 to include the impact of the effective tax rates on the returns of assets in the taxable account. We also need to have separate constraints for the taxable and tax-advantaged accounts. With these adjustments, the expanded MVO problem becomes:

$$\begin{gathered}\max_{\mathbf{h_K^A},\,\mathbf{h_K^B}} U((\mathbf{T_K h_K^A} + \mathbf{h_K^B})^{\mathbf{T}} \boldsymbol{\mu}_{\mathbf{K}}, (\mathbf{T_K h_K^A} + \mathbf{h_K^B})^{\mathbf{T}} \mathbf{V_K} (\mathbf{T_K h_K^A} + \mathbf{h_K^B})) \\ s.t.\, \mathbf{h_K^{AT}} \boldsymbol{\iota}_{\mathbf{K}} = \phi,\ \mathbf{h_K^{BT}} \boldsymbol{\iota}_{\mathbf{K}} = 1 - \phi,\ \mathbf{h_K^A} \geq 0,\ \mathbf{h_K^B} \geq 0.\end{gathered} \tag{8.8}$$

Equation 8.8 differs from Equation 8.5 in two ways:

1. The impact of allocations in the taxable account on returns is reduced because of taxes.
2. There are now two budget constraints, one for each account.

To illustrate how the joint MVO works, we estimated the effective tax rates for the asset classes in chapter 7 using the tax rates that we assumed for Isabela, namely 40% as the tax rate on ordinary income and 20% as the long-term capital gains rate. We assume that Isabela's assets are split as we have assumed in chapter 2, namely $250,000 in taxable assets and $20,500 in tax-advantaged assets. Hence, for Isabela, $\phi = (250{,}000/270{,}00) = 92.42\%$. In **Exhibit 8.3**, we trace out the efficient frontier in terms of after-tax expected return and standard deviation on the portfolio as a whole. We call the portfolio that maximizes the Levy–Markowitz utility function the *optimal joint portfolio*. To show how the optimal joint portfolio is determined for Isabela with her risk tolerance parameter of $\theta = 35\%$, we include her tangent indifference curve for the joint problem in Exhibit 8.3.[73]

Exhibit 8.4 shows the optimal joint portfolio with the allocations in the two accounts broken out. Note that the asset classes with the lowest effective tax rates, stocks and municipal bonds, are allocated to the taxable account. An asset class with the highest effective tax rate, global bonds, is allocated to the tax-advantaged account, albeit in a small amount because of the small size of the tax-advantaged account relative to the size of the taxable account.

---

[72]See Wilcox et al. (2006), Appendix A, for a similar approach to jointly solving the asset allocation and location problems.

[73]We define the point labeled "Optimal Joint Net-Worth Portfolio" later.

Exhibit 8.3. Risk and Expected Return with Taxable and Tax-Advantaged Accounts for Isabela

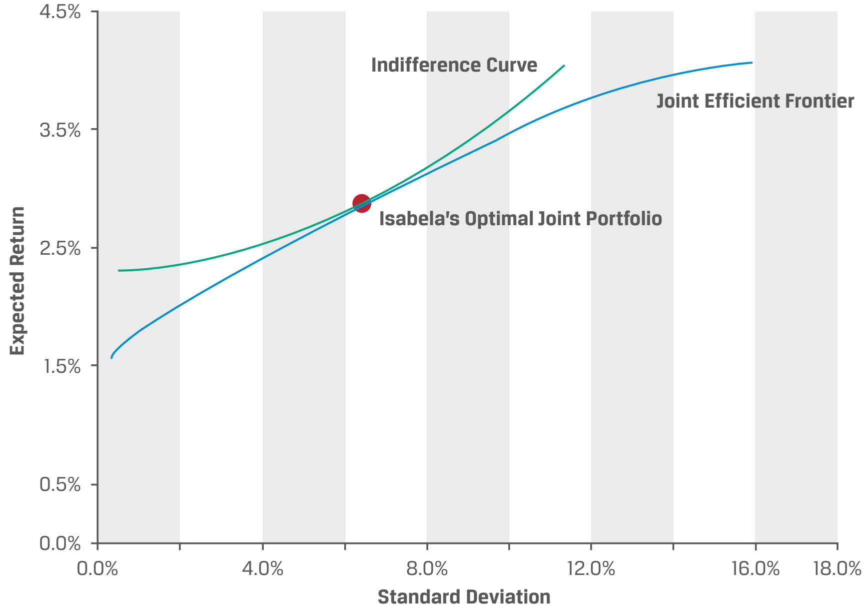


Exhibit 8.4. Optimal Asset Allocation and Location for Financial Wealth in Isolation for Isabela

Asset Class	Effective Tax Rate (%)	Allocation (%)
U.S. Large-Cap Stocks	25.7	24.1
U.S. Mid-Cap Stocks	26.6	
Global DM Ex-U.S. Stocks	26.8	20.9
Emerging Market Stocks	27.0	4.3
U.S. Bonds	40.0	
Inflation Linked Bonds	40.0	
Municipal Bonds	0.0	43.1
Global Bonds Ex-U.S.	40.0	7.6
Cash	40.0	

0% 10 20 30 40

Taxable Account   Tax-Advantaged Account

Also note that, when taking taxes into account, Isabela's optimal equity allocation is much higher than when taxes are not taken into account, namely about 49.3% for the former versus 32.3% for the latter. This is another consequence of the relative tax efficiency of stocks as compared to that of bonds.

The Child Model: The Joint Allocation Problem in a Net-Worth Framework

In this section, we build on the previous MVO extension and meld it with the investor's economic balance sheet from our life-cycle model in part I. More specifically, we assume that the investor's risk tolerance, θ or theta, applies to the investor's net worth, that is, total assets (financial wealth + human capital) minus liabilities. Our extension of MVO, *net-worth optimization,* explicitly includes *nontradeable human capital* and *nonchangeable liabilities* (both modeled as asset class exposures) in the joint optimization that solves for the multiple tax-efficient asset allocation targets for financial wealth.

We can now bring together net-worth optimization and joint MVO and connect them to our holistic life-cycle approach.

In chapters 4, 5, and 6, we discuss the concept of net worth (W), which is financial wealth (F) plus human capital (H) less the present value of future liabilities (L). Human capital is the present value of future income from all sources related to the investor's future earnings, including wages and social insurance. Liabilities consist of nondiscretionary spending, that is, spending on essentials, such as food, clothing, and shelter, as well as the present value of term life insurance premiums should the investor be planning on leaving a bequest. Net worth is the value of the investor's resources available for discretionary spending, such as the incremental cost of dining out and traveling for leisure.

As we discuss in chapter 4, we can model both H and L like assets. Their changes in value over time are largely due to their returns. Mathematically, we can define the return on net worth ($\tilde{R}_W$) as follows:

$$\tilde{R}_W = \frac{\hat{F}'}{\hat{W}'}\tilde{R}_F + \frac{\hat{H}'}{\hat{W}'}\tilde{R}_H - \frac{\hat{L}'}{\hat{W}'}\tilde{R}_L, \tag{8.9}$$

where $\tilde{R}_F$, $\tilde{R}_H$, and $\tilde{R}_L$ are the returns on financial wealth, human capital, and the value of liabilities, respectively. The primes (′) indicate that there are cash flow adjustments to capture the cash flows that occur just before returns are realized (see Appendix 8A).

Equation 8.9 links the output from a life-cycle model with a single-period MVO model. (This linkage is in addition to the linkage between the life-cycle models in chapter 6 and MVO through the risk tolerance parameter). The key to this linkage is that the investor makes allocation choices only for financial wealth and not for human capital or liabilities.[74] For clarity, we illustrate the connection between the life-cycle balance sheet and Equation 8.9 in **Exhibit 8.5**.

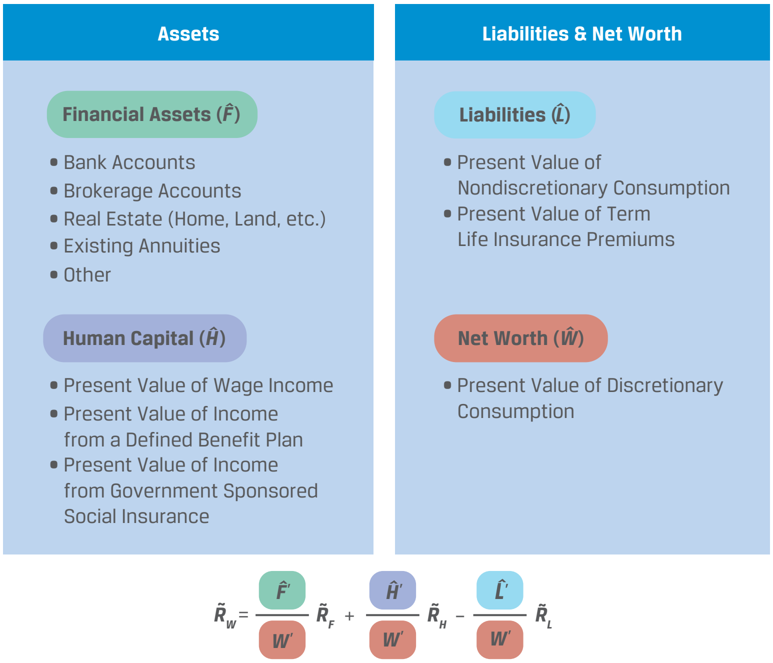
As we discuss in chapters 4 and 6, we generally model human capital and the value of liabilities as risky assets or portfolios of risky assets.[75] Because the cash flows of human capital and liabilities are largely

---

[74]People can always choose to change careers, but in the context of our model, we accept their current career and attempt to model the asset class-like characteristics of the associated income stream appropriately. Similarly, people can change what they consider their essential consumption, but we take it as given.

[75]As Exhibit 8.6 shows, we do include some cash in our asset allocation representations of human capital and liabilities. This is to model cash flows that we regard as certain. For human capital, we assume that Isabela receives the maximum match of $6,833 based on the rules of her 401(k) plan. For liabilities, we treat the term life premiums as certain.

Exhibit 8.5. Linking Life-Cycle Model to Expanded MVO Model through the Investor Balance Sheet



known in advance, human capital and the liabilities are both very bond-like. However, there could be some equity-like risks. For human capital, the investor's wages could be subject to the risk of economic downturns, which could be reflected in the stock market, leading to some correlation between the return on human capital ($\tilde{R}_H$) and the returns on equity asset classes. Similarly, because liabilities include housing costs that could be correlated with the stock market, there also could be correlation between $\tilde{R}_L$ and equity returns. **Exhibit 8.6** shows specifically how we model human capital and liabilities as asset class portfolios to reflect these ideas and how their allocations blend with the allocation of Isabela's financial wealth to determine the implied allocation of her net worth.

The expected return and standard deviation of return of net worth are related in part to the asset class weights of taxable and tax-advantaged financial assets. In Appendix 8B, we present these relationships mathematically. These relationships involve not only the assumptions about pretax asset class returns and effective taxes but also *all of the components of the investor's balance sheet*, and are thus the links between the life-cycle and single-period optimization models.

Mathematically, we denote these relationships as the functions $\mu_W(\cdot,\cdot)$ and $\sigma_W^2(\cdot,\cdot)$. This allows us to write the MVO problem for the net-worth joint asset allocation and location problem as follows:

$$\max_{\mathbf{h}_K^A,\mathbf{h}_K^B} U(\mu_W(\mathbf{h}_K^A,\mathbf{h}_K^B),\sigma_W^2(\mathbf{h}_K^A,\mathbf{h}_K^B))$$
$$s.t.\ \mathbf{h}_K^{AT}\iota_K=\phi,\ \mathbf{h}_K^{BT}\iota_K=1-\phi,\ \mathbf{h}_K^A\geq 0,\ \mathbf{h}_K^B\geq 0. \quad (8.10)$$

We call the portfolio that solves Equation 8.10 the *optimal joint net-worth portfolio*.

Exhibit 8.6. Asset Class Models of Human Capital and Liabilities for Isabela

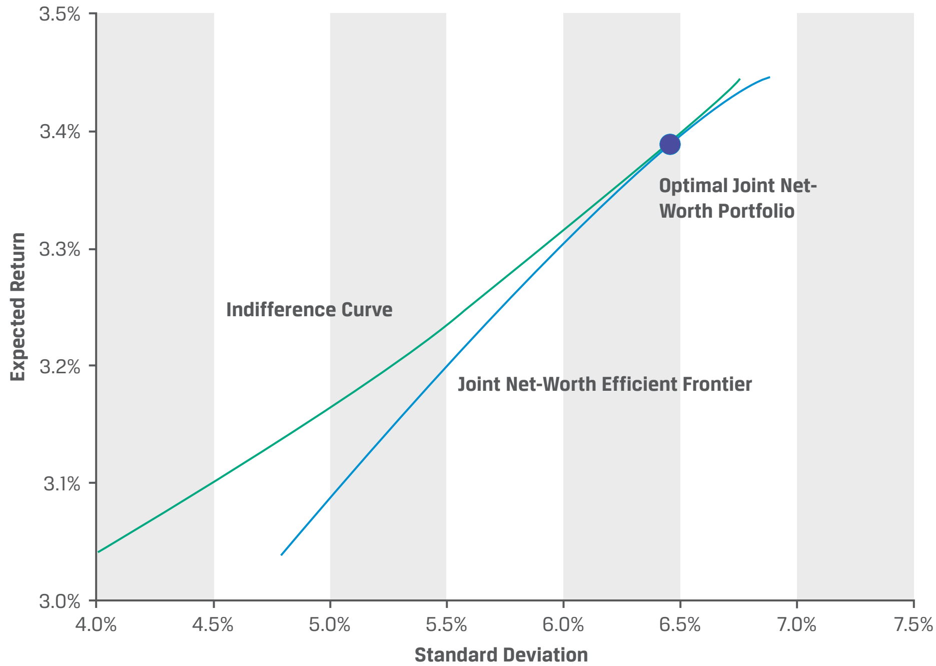
Asset Class	Financial Wealth			Human Capital	Liabilities	Net Worth
	Taxable	Tax Advantaged	Overall			
US Large-Cap Stocks	36.11%	0.00%	36.11%	9.36%	12.63%	11.01%
US Mid-/Small-Cap Stocks	0.00%	0.00%	0.00%	4.68%	0.00%	7.87%
Global DM Ex-US Stocks	45.12%	0.00%	45.12%	4.68%	0.00%	15.32%
Emerging-Markets Stocks	11.19%	4.13%	15.32%	0.00%	0.00%	2.53%
US Bonds	0.00%	0.00%	0.00%	37.46%	33.68%	34.48%
Inflation-Linked Bonds	0.00%	0.00%	0.00%	37.46%	37.89%	30.92%
Municipal Bonds	0.00%	0.00%	0.00%	0.0%	0.00%	0.00%
Global Bonds Ex-US	0.00%	0.00%	0.00%	0.0%	0.00%	0.00%
Cash	0.00%	3.45%	3.45%	6.35%	15.81%	−2.13%
Total Equity	92.92%	4.13%	96.55%	18.73%	12.63%	36.73%
Total	92.92%	7.58%	100.00%	100.00%	100.00%	100.00%
Fraction of Net Worth	16.49%			168.15%	84.65%	10.00%

To illustrate asset allocation and location in the net-worth framework, we consider the case of Isabela, who at age 25 has $250,000 in a taxable account and $20,500 in a tax-advantaged account. As we discussed in chapters 2 and 4, her human capital is $2,767,689 and the value of her liabilities is $1,392,064. We adjusted these figures by the cash flows discussed in Appendix 8A. From these values and cash flows, we calculated the fractions of net worth shown in the last row of Exhibit 8.6. We model the return on her human capital and the return on her human capital using the portfolio for these shown in Exhibit 8.6.

Exhibit 8.7 shows the net-worth optimization efficient frontier based on our assumption regarding Isabela. It also shows the optimal joint net-worth portfolio and the indifference curve that is tangent to the efficient frontier at the optimal joint net-worth portfolio.

In Exhibit 8.7, we see that the optimal portfolio for Isabela is fairly high on the efficient frontier. Going back to Exhibit 8.6, we can see why. Exhibit 8.6 shows the optimal joint net-worth portfolio with the allocations in the two accounts broken out, as well as the allocations for human capital, liabilities, and net worth. Although Isabela's risk tolerance parameter (θ) is relatively low (35%), the optimal allocation for her financial wealth is 96.6% equity, which seems quite high. As the last column of Exhibit 8.6 shows, however, her net worth is only 36.8% equity, which is more in line with her risk tolerance.

Exhibit 8.7. Risk and Expected Return in Net-Worth Optimization for Isabela



The example of Isabela shows why young investors should have a high equity allocation in their financial assets. As we discussed in chapter 4, taking a net-worth approach to asset allocation brings an investor's risk capacity front and center. Because their human capital is the dominant part of their net worth, young investors such as Isabela have a great deal of risk capacity and, therefore, should hold equity-centric portfolios even if their risk tolerance is low.

Conclusion and Key Takeaways

Constructing portfolios for individual investors is complicated by the fact that investors pay current taxes on returns in some accounts and can postpone or avoid taxes on returns in other accounts. This is further complicated by the fact that investors have income from sources outside of their portfolios, and by the fact that they have unavoidable expenses. In other words, investors have human capital and liabilities, which are elements of the investor balance sheet. In this chapter, we have presented some extensions to the standard mean–variance model to incorporate these complications, thus making it applicable to the problems faced by individual investors that include the existence of taxes and aspects of financial planning captured by life-cycle models. We also have illustrated how these extensions can have a large impact on asset allocation. In particular, we have shown why young investors with high-risk capacity should have equity-centric portfolios even if their risk tolerance is low.

In this chapter, we have assumed that the two account types are taxable and tax-advantaged accounts. In addition, there are at least two types of tax-advantaged accounts, including tax-deferred and tax-exempt accounts. Later, in chapter 11 in part III, we make the distinction between tax-deferred and tax-exempt accounts, bringing us to three types of accounts: taxable, tax-deferred, and tax-exempt accounts.

Appendix 8A. Cash Flow Adjustments to Values

The evolution over time of the values in the models we present includes cash flows. In this chapter, we indicate the cash flow adjustment with a prime (′). In this appendix, we spell out the cash flow adjustment for each value.

Financial Assets

The evolution of financial assets includes net savings which is exogenous income minus consumption and term life insurance premiums. So its cash flow–adjusted value is as follows:

$$\hat{F}_t' = \hat{F}_t + y_t - c_t - LI1_t B. \tag{8A.1}$$

This allows us to write the evolution of financial assets as follows:

$$\hat{F}_{t+1} = \frac{1+\tilde{R}_{Ft+1}}{\hat{q}_{t+1}^t}\hat{F}_t'. \tag{8A.2}$$

Human Capital

The evolution of human capital includes exogenous income. So, its cash flow–adjusted value is as follows:

$$\hat{H}_t' = \hat{H}_t - y_t. \tag{8A.3}$$

This allows us to write the evolution of financial assets as follows:

$$\hat{H}_{t+1} = \frac{1+\tilde{R}_{Ht+1}}{\hat{q}_{t+1}^t}\hat{H}_t'. \tag{8A.4}$$

Liabilities

The evolution of the value of liabilities includes nondiscretionary consumption and term life insurance premiums. So, its cash flow–adjusted value is as follows:

$$\hat{L}_t' = \hat{L}_t - \bar{c}_t - LI1_t B. \tag{8A.5}$$

This allows us to write the evolution of financial assets as follows:

$$\hat{L}_{t+1} = \frac{1+\tilde{R}_{Lt+1}}{\hat{q}_{t+1}^t}\hat{L}_t'. \tag{8A.6}$$

Net Worth

The evolution of net worth includes discretionary consumption. So, its cash flow–adjusted value is as follows:

$$\hat{W}_t' = \hat{W}_t - c_t + \bar{c}_t. \tag{8A.7}$$

This allows us to write the evolution of net worth as follows:

$$\hat{W}_{t+1} = \frac{1+\tilde{R}_{Wt+1}}{\hat{q}^t_{t+1}}\hat{W}'_t. \tag{8A.8}$$

Because $W = F + H - L$, we have the following:

$$\frac{1+\tilde{R}_{Wt+1}}{\hat{q}^t_{t+1}}\hat{W}'_t = \hat{F}_{t+1} = \frac{1+\tilde{R}_{Ft+1}}{\hat{q}^t_{t+1}}\hat{F}'_t + \frac{1+\tilde{R}_{Ht+1}}{\hat{q}^t_{t+1}}\hat{H}'_t - \frac{1+\tilde{R}_{Lt+1}}{\hat{q}^t_{t+1}}\hat{L}'_t. \tag{8A.9}$$

Therefore,

$$1+\tilde{R}_{Wt+1} = (1+\tilde{R}_{Ft+1})\frac{\hat{F}'_t}{\hat{W}'_t} + (1+\tilde{R}_{Ht+1})\frac{\hat{H}'_t}{\hat{W}'_t} - (1+\tilde{R}_{Lt+1})\frac{\hat{L}'_t}{\hat{W}'_t}. \tag{8A.10}$$

This simplifies to Equation 8.9.

Appendix 8B. Functions for the Expected Return and Standard Deviation of Return of Net Worth

Expected Return of Net Worth

From Equation 8.9, it follows that

$$E[\tilde{R}_W] = \frac{F}{W}E[\tilde{R}_F] + \frac{H}{W}E[\tilde{R}_H] - \frac{L}{W}E[\tilde{R}_L]. \tag{8B.1}$$

Let:

$\mathbf{h}^{\mathbf{H}}_{\mathbf{K}}$ = the asset allocation vector representing human capital; and

$\mathbf{h}^{\mathbf{L}}_{\mathbf{K}}$ = the asset allocation vector representing liabilities.

Then we have the following:

$$E[\tilde{R}_F] = (\mathbf{T}_{\mathbf{K}}\mathbf{h}^{\mathbf{A}}_{\mathbf{K}} + \mathbf{h}^{\mathbf{B}}_{\mathbf{K}})^{\mathbf{T}}\boldsymbol{\mu}_{\mathbf{K}}, \tag{8B.2}$$

$$E[\tilde{R}_H] = \mathbf{h}^{\mathbf{HT}}_{\mathbf{K}}\boldsymbol{\mu}_{\mathbf{K}}, \tag{8B.3}$$

$$E[\tilde{R}_L] = \mathbf{h}^{\mathbf{LT}}_{\mathbf{K}}\boldsymbol{\mu}_{\mathbf{K}}. \tag{8B.4}$$

Letting $\mu_H = E[\tilde{R}_H]$ and $\mu_L = E[\tilde{R}_L]$, from Equations 8B.1 and 8B.2, it follows that we can write $E[\tilde{R}_W]$ as a function of $\mathbf{h}^{\mathbf{A}}_{\mathbf{K}}$ and $\mathbf{h}^{\mathbf{B}}_{\mathbf{K}}$, $\mu_W(\cdot,\cdot)$, as follows:

$$E[\tilde{R}_W] = \mu_W(\mathbf{h}^{\mathbf{A}}_{\mathbf{K}}, \mathbf{h}^{\mathbf{B}}_{\mathbf{K}}) = \frac{F}{W}(\mathbf{T}_{\mathbf{K}}\mathbf{h}^{\mathbf{A}}_{\mathbf{K}} + \mathbf{h}^{\mathbf{B}}_{\mathbf{K}})^{\mathbf{T}}\boldsymbol{\mu}_{\mathbf{K}} + \frac{H}{W}\mu_H - \frac{L}{W}\mu_L. \tag{8B.5}$$

Variance of Net Worth

From Equation 8.9, it follows that

$$Var[\tilde{R}_W] = \left(\frac{F}{W}\right)^2 Var[\tilde{R}_F] + 2\frac{F}{W}\left(\frac{H}{W}Cov[\tilde{R}_H,\tilde{R}_F] - \frac{L}{W}Cov[\tilde{R}_L,\tilde{R}_F]\right) + \left(\frac{H}{W}\right)^2 Var[\tilde{R}_H] + \left(\frac{L}{W}\right)^2 Var[\tilde{R}_L] - 2\frac{H}{W}\frac{L}{W}Cov[\tilde{R}_H,\tilde{R}_L]. \tag{8B.6}$$

Then we have the following:

$$Var[\tilde{R}_F] = (\mathbf{T_K h_K^A} + \mathbf{h_K^B})^\mathbf{T}\, \mathbf{V_K}\, (\mathbf{T_K h_K^A} + \mathbf{h_K^B}), \tag{8B.7}$$

$$Var[\tilde{R}_H] = \mathbf{h_K^{HT} V_K h_K^H}, \tag{8B.8}$$

$$Var[\tilde{R}_L] = \mathbf{h_K^{LT} V_K h_K^L}, \tag{8B.9}$$

$$Cov[\tilde{R}_H,\tilde{R}_F] = \mathbf{h_K^{HT} V_K}\, (\mathbf{T_K h_K^A} + \mathbf{h_K^B}), \tag{8B.10}$$

$$Cov[\tilde{R}_L,\tilde{R}_F] = \mathbf{h_K^{LT} V_K}\, (\mathbf{T_K h_K^A} + \mathbf{h_K^B}), \tag{8B.11}$$

$$Cov[\tilde{R}_H,\tilde{R}_L] = \mathbf{h_K^{HT} V_K h_K^L}. \tag{8B.12}$$

Let:

$$\mathbf{c_K^H} = \mathbf{V_K h_K^H}, \tag{8B.13}$$

$$\mathbf{c_K^L} = \mathbf{V_K h_K^L}. \tag{8B.14}$$

Also let:

$$\sigma_H^2 = Var[\tilde{R}_H],$$

$$\sigma_L^2 = Var[\tilde{R}_L],$$

$$\sigma_{HL} = Cov[\tilde{R}_H,\tilde{R}_L].$$

From these definitions and from Equations 8B.6–8B.14, we can write $Var[\tilde{R}_W]$ a function of $\mathbf{h_K^A}$ and $\mathbf{h_K^B}, \sigma_W^2(\cdot,\cdot)$, as follows:

$$Var[\tilde{R}_W] = \sigma_W^2(\mathbf{h_K^A}, \mathbf{h_K^B}) = \left(\frac{F}{W}\right)^2 \left((\mathbf{T_K h_K^A} + \mathbf{h_K^B})^\mathbf{T}\, \mathbf{V_K}\, (\mathbf{T_K h_K^A} + \mathbf{h_K^B})\right) + 2\frac{F}{W}\left(\frac{H}{W}\mathbf{c_K^H} - \frac{L}{W}\mathbf{c_K^L}\right)^\mathbf{T} (\mathbf{T_K h_K^A} + \mathbf{h_K^B}) + \left(\frac{H}{W}\right)^2 \sigma_H^2 + \left(\frac{L}{W}\right)^2 \sigma_L^2 - 2\frac{H}{W}\frac{L}{W}\sigma_{HL}. \tag{8B.15}$$

PART III: GRANDCHILD MULTI-ACCOUNT PORTFOLIO CONSTRUCTION MODEL WITH NONFINANCIAL PREFERENCES

> This is a behavioral story, but it is not about irrational behavior.
> In economics, we take tastes as given, and make no judgments about them.
> —Eugene Fama (2014, p. 1482)

This quote is from Eugene Fama's Nobel Lecture (2014) and clearly demonstrates that the idea of "tastes" (nonpecuniary preferences) has become part of Fama's approach to understanding investor behavior and asset pricing, both of which have implications for personalized portfolio construction.

In part III, using outputs from the child asset location and allocation model developed in part II, we develop an extension to the alpha-tracking error optimization framework of Waring et al. (2000), which was developed to allocate capital to managers, that incorporates Fama's idea of tastes into portfolio construction. Our grandchild, personalized, multi-account, transaction cost–aware, tax-aware alpha-tracking error model implements the separate account-type-specific target asset allocations coming from the part II child model and takes the investor's nonpecuniary preferences (tastes) regarding securities into account. In chapter 2, we argued that these nonpecuniary preferences should be surveyed for in investor preference questionnaires.

Recall that in each period (such as a year), the parent life-cycle model from part I calculates values of the three distinct components of the investor's balance sheet (i.e., financial assets, human capital, and liabilities) and passes them on to the single-period child asset location and allocation model from part II. Then, the part II single-period child asset location and allocation model produces separate account-type-specific target asset allocations. Our part III grandchild model takes those account type–specific target asset allocations as well as the investor's nonpecuniary preferences, tax rates, and existing holdings and is designed to be run frequently as an ongoing tax-efficient, personalized investment management system.

Our grandchild model incorporates nonpecuniary preferences by including them in an objective function that is based on the investor objective function in the PAPM of Ibbotson et al. (2018) and Idzorek, Kaplan, and Ibbotson (2021, 2023). The PAPM makes the important assumption that each investor maximizes an objective function that accounts not only for their pecuniary views but also for their nonpecuniary preferences.

In chapter 9, we review the PAPM and its implications for asset prices and personalization. We specifically discuss the implications of both pecuniary and nonpecuniary aspects of ESG. We expand the mean–variance objective function to include an additional term that captures the benefit of having a portfolio that tilts toward characteristics that an investor likes and away from characteristics that an investor dislikes.

In chapter 10, based on the PAPM, we move from the expanded form of MVO with the new nonpecuniary preference term to an alpha-tracking error optimization model that also includes the new nonpecuniary preference term. Both of these expanded portfolio construction problems continue to capture all relevant pecuniary views (e.g., expected returns, standard deviations, correlations, after-fee forward-looking alphas), but they are expanded to incorporate nonpecuniary preferences.

In chapter 11, we present our complete grandchild model. We expand the alpha-tracking error optimization, with the new nonpecuniary preference term from chapter 10, to include multiple account types, multiple accounts, and taxes. It receives as inputs the separate asset location and asset allocation targets developed using the child model part II.

9. THE IMPACT OF NONPECUNIARY PREFERENCES, INCLUDING ENVIRONMENT, SOCIAL, AND GOVERNANCE, ON CAPITAL MARKETS

Context

In chapter 2, we developed a holistic investor profile that includes nonfinancial or nonpecuniary investor preferences and then touched on how one might measure those preferences. Since then, we have focused on creating models of rational investor behavior from a purely financial or pecuniary perspective. In this chapter, we introduce single-period optimization models with nonpecuniary preferences, especially preferences related to environment, social, and governance (ESG), into our framework. In this chapter, we focus on how nonpecuniary preferences influence how some investors form portfolios and thus affect market prices. We also consider how differing investor perceptions of the way ESG factors affect expected payouts on various assets cause different investors to view the same assets differently. This consideration leads to further portfolio personalization. In chapters 10 and 11, we show how to take nonpecuniary preferences into account when constructing a portfolio in practice. Importantly, the methodology for including nonpecuniary preferences in the personalized portfolio construction problem in chapters 10 and 11 emanates from the generalized equilibrium asset pricing model presented in this chapter.

Key Insights

- Standard asset pricing models, such as the CAPM, are missing two key ingredients that Fama and French (2007) call "disagreement" and "tastes." Disagreement refers to differences in investors' forecasts of the future payouts of assets, and tastes refer to nonpecuniary preferences in which investments are treated as goods with salient characteristics, other than expected return and risk, that influence purchasing decisions.
- ESG factors can have both pecuniary and nonpecuniary aspects. On one hand, different investors can differ in their assessments of the *pecuniary* impact of ESG factors on future asset payouts—for example, views on climate change, the impact of climate change, or how firms respond to it can help or hurt certain firms or industries. On the other hand, different investors can have different *nonpecuniary* preferences for the very same factors, perhaps preferring green energy companies regardless of their pecuniary view.
- The PAPM of Idzorek, Kaplan, and Ibbotson (2021, 2023) is a powerful generalization of the CAPM that takes disagreements and tastes into account. It can provide insight into the way disagreements and tastes affect asset prices and how investors form personalized portfolios based on these factors.
- The PAPM is well suited to model the impact of ESG factors on market prices and personalized investor portfolios because it takes into account both the pecuniary and nonpecuniary aspects of ESG.

Introduction

In this chapter, we begin by continuing to link our parent life-cycle model to single-period optimization models. We then introduce a new equilibrium asset pricing model, the PAPM. As we explain, the PAPM incorporates what Fama and French (2007) identified as the two missing ingredients from the CAPM: disagreement and tastes. We focus on the implications on asset prices of disagreement and tastes relative to asset prices from the CAPM, which excludes these two ingredients. This sets the stage for chapter 10, in which we move beyond the equilibrium implications on asset prices to focus on the implications for personalized portfolio construction.

From Life-Cycle Risk Tolerance to Single-Period Risk Aversion

In Chapter 8, we applied expected utility theory to portfolio selection by using MVO in an innovative and nonstandard manner. More specifically, we used the Levy–Markowitz approximation of expected utility, which we refer to as the Levy–Markowitz utility function. Because the Levy–Markowitz utility function is increasing in expected return, and decreasing in standard deviation, we can, in effect, maximize expected utility by finding the point on an MVO efficient frontier with the highest possible value for the Levy–Markowitz utility function, which is show as Equation 9.1.

We derive *our* version of the Levy–Markowitz utility function from the CRRA utility function. The CRRA utility function (and hence our Levy–Markowitz utility function) has a single parameter, θ (theta), which we call the risk tolerance parameter. Critically, this is the same theta used in our life-cycle models, and it is through the Levy–Markowitz utility function that we make this *new* bridge between life-cycle models and single-period optimization models. For any given value of θ, a point along the efficient frontier maximizes the Levy–Markowitz utility function. Conversely, for every point along the efficient frontier, there is a value of θ such that the portfolio represented by that point maximizes the Levy–Markowitz utility function.

We use the Levy–Markowitz utility function to link our child MVO model with parent life-cycle models, but it is rarely used in practice. Instead, the MVO problem is often formulated using risk-adjusted expected return (which we abbreviate RAER). Like the Levy–Markowitz utility function, RAER is increasing in expected return and decreasing in standard deviation, but it has a simpler form relative to Equations 8.1, 8.2, and 8.3 taken together. RAER can be written as a function in expected return and standard deviation, as follows:

$$RAER(\mu,\sigma) = \mu - \frac{\lambda}{2}\sigma^2, \tag{9.1}$$

where:

μ = expected return;

σ = standard deviation of return; and

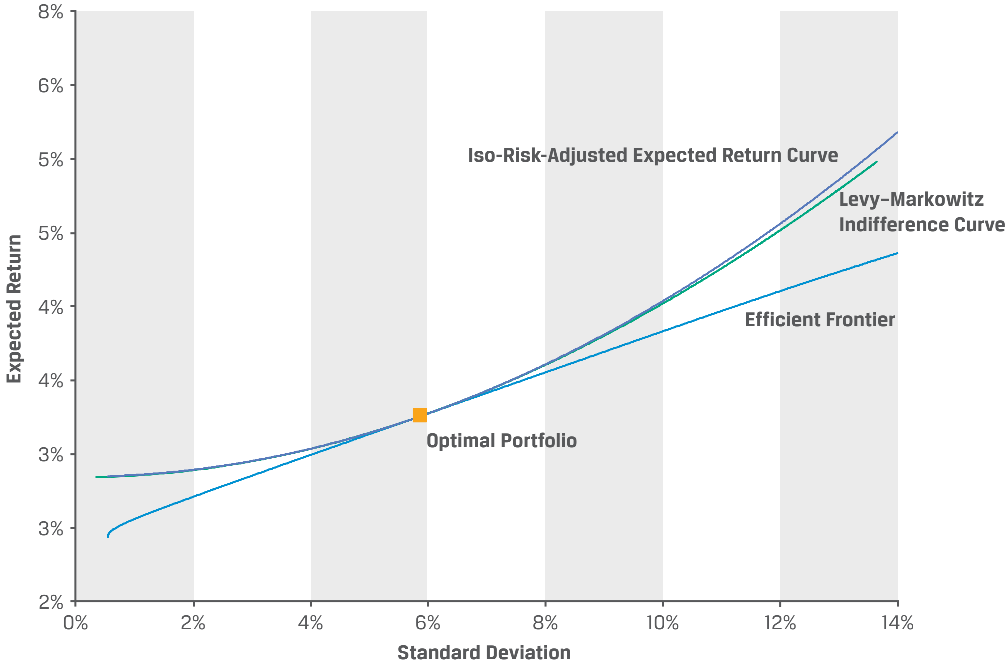
λ = what, in the context of RAER, we call the *risk aversion parameter*.

We formulate our grandchild MVO models in terms of RAER so we can take advantage of the computational machinery that has already been developed for MVO-type problems, known as quadratic programming problems. Also, as we shall see in this chapter, the RAER formulation allows us to develop the CAPM and PAPM in a straightforward manner.

Just as with θ, for any given value of λ (lambda), a point along the efficient frontier maximizes RAER. Conversely, for every point along the efficient frontier, there is a value of λ such that the portfolio represented by that point maximizes RAER.

Hence, every point along the efficient frontier has a corresponding value of θ and of λ. **Exhibit 9.1** illustrates this. The efficient frontier, optimal portfolio, and Levy–Markowitz indifference curve are from the previous

Exhibit 9.1. Finding the Optimal Portfolio with Levy–Markowitz Utility and Risk-Adjusted Expected Return Criteriaia



chapter (Exhibit 8.1) In Exhibit 9.1, we have added the "Iso-Risk-Adjusted Expected Return Curve."[76] This is the curve for which all combinations of expected return and standard deviation have the same RAER, with the value of λ being the value of the risk aversion parameter for which the optimal portfolio from Exhibit 8.1 is also the portfolio that maximizes RAER. Note that both the indifference curve and the iso-RAER curve are tangent to the efficient frontier at the optimal portfolio. Because the optimal portfolio is optimal under both criteria, it maximizes both the Levy–Markowitz utility function and RAER.

Exhibit 9.1 also shows how we can go from the risk tolerance parameter of our parent life-cycle models (θ) to the risk aversion parameter (λ). Given the value of θ, we maximize the Levy–Markowitz utility function. We then find the value of λ such that the optimal portfolio maximizes RAER. In Exhibit 9.1, $\theta = 35\%$ (Isabela's risk tolerance) and $\lambda = 2.72$. Because this mapping from θ to λ depends on the efficient frontier, it needs to be done with the frontier that is based on the same asset classes and capital market assets used in the child and grandchild models.

The Popularity Asset Pricing Model[77]

The CAPM remains the most influential model in finance, largely because of its elegant structure and powerful conclusions. The main conclusions of the CAPM are as follows: (1) all investors hold the market portfolio in combination with a risk-free asset (held long or short), making optimization unnecessary; and

[76]"Iso" means "the same," meaning that RAER is the same at all points on the curve.

[77]The remainder of this chapter is adapted from Idzorek, Kaplan, and Ibbotson (2021, 2023) and Kaplan (2021).

(2) the expected return in excess of the risk-free rate of each security is proportional to its systematic risk (beta), that is, the sensitivity of the return of the security to return on the market portfolio. These *conclusions*, however, depend heavily on the *assumptions* of the model. Changing the assumptions leads to very different conclusions.

Motivated in part by the shortcomings of the CAPM, in an academic article that is not well-known among practitioners, called "Disagreement, Tastes, and Asset Prices," Fama and French (2007) identify "disagreement" and "tastes" as two key ingredients missing from the CAPM that should affect asset prices. Disagreement refers to heterogeneous expectations. Tastes refer to investor preferences beyond desire for expected return and aversion to risk. Even though Fama and French (2007) identify two important ways to make the CAPM more realistic, they stopped short of developing an equilibrium asset pricing model that incorporates these improvements.

Disagreement and especially tastes are directly related to ESG investing. The topic of ESG has spawned a variety of papers putting forth special asset pricing models that incorporate ESG. These include Baker et al. (2020); Pástor, Stambaugh, and Taylor (2021); Pedersen, Fitzgibbon, and Pomorski (2021); and Zerbib (2019). Of course, investors care about a variety of characteristics beyond ESG, including liquidity, yield (the income part of return), taxes, faith-based values, and others. The PAPM is the generalized asset pricing model that encompasses the CAPM as well as these ESG-specific models, allowing for any number of asset characteristics and a wide range of investors with various expectations and tastes.

All else equal, investors seem to prefer a variety of nonpecuniary characteristics. For example, two companies are identical in every way except that one is more environmentally friendly; many investors would prefer the greener company. This increased demand for the greener firm, all else equal, raises the current price, decreasing expected returns.

The impact of these kinds of nonpecuniary preferences is easiest to see in the primary markets. If two companies with identical credit ratings are issuing identical bonds, the greener company will often have a slightly lower cost of capital in that it pays a reduced interest rate, because enough investors are willing to accept a slightly lower interest rate to invest in a greener firm. We think of this as a popularity premium or discount, depending on your perspective.

From a popularity perspective, characteristics that are nearly universally liked are in high demand (popular) and thus make the securities bearing these characteristics expensive, leading to lower expected returns. Conversely, characteristics that are nearly universally disliked are in low demand (unpopular) and thus make the securities bearing them inexpensive, leading to higher expected returns. We have found that this type of popularity-based explanation clarifies a wide variety of so-called premiums and anomalies. Along with Roger Ibbotson and James Xiong, we analyzed a wide variety of well-known premiums and anomalies in *Popularity: A Bridge between Classical and Behavioral Finance*, published by CFA Institute Research Foundation (Ibbotson et al. 2018). We found that many premiums and anomalies are consistent with a popularity-based explanation.

In addition to empirical evidence, Ibbotson et al. (2018) present the formal PAPM, in which investors can have preferences for nonpecuniary characteristics of securities. Later, Idzorek, Kaplan, and Ibbotson (2021, 2023) generalized the PAPM by including heterogenous expectations to create an adaptable CAPM that incorporates the two missing asset pricing ingredients identified by Fama and French (2007).

The PAPM and ESG

As noted earlier, the popularity of ESG investing has led to papers promoting specialty asset pricing models. The PAPM subsumes both the CAPM and a range of these newer ESG-specific models as special cases. In the PAPM, investors have divergent beliefs about expected returns and a variety of risk and nonrisk preferences, such as liquidity or ESG. Unfortunately, what it means to incorporate ESG into the investment process has caused considerable confusion. The two sides of ESG that must be kept distinct when building a portfolio are as follows:

- *Pecuniary ESG*. This is the impact that ESG factors have on the risk and expected return of securities issued by a company.[78]
- *Nonpecuniary ESG*. This is the extent to which investors find securities desirable for reasons other than risk and expected return. For example, investors may prefer stocks issued by "green" companies because of their personal values and concerns about the environment.[79]

In this chapter, we present a model for understanding how both pecuniary and nonpecuniary ESG can affect the way investors form portfolios in an equilibrium setting using an ESG-specific version of the PAPM. The model and example that we present are similar to those presented in Pedersen, Fitzgibbons, and Pomorski (2021, hereafter PFP).

Equilibrium with Pecuniary ESG Views and No Nonpecuniary Preferences

According the PAPM, investors can form portfolios based on pecuniary factors (risk and expected return) and possibly any number of nonpecuniary factors. Furthermore, investors can have different views regarding pecuniary factors. In the ESG version of the PAPM that we discuss here, we assume that investors have one of two pecuniary views: ESG-unaware or ESG-aware.[80] For now, we assume that no investors have nonpecuniary preferences, but we will introduce those who do into the model later in this chapter.

To model the impact of investors having different ESG views, we formed a simple model in which there are two stocks (ESG-positive and ESG-negative) and two investors (ESG-unaware and ESG-aware).

Both the CAPM and the PAPM are single-period models in which investors trade securities (stocks and cash) at the beginning of the period and receive the payouts of the stocks at the end of the period. In the ESG-unaware view, we assume that the two stocks have the same *expected* payout but differ in the standard deviations of their payouts as well as in their systematic risks (betas). We assume that the ESG-positive stock has both greater total and systematic risk, so that it is both riskier and has a greater expected return than the ESG-negative stock.

We assume that the ESG-positive stock is issued by a company with good ESG practices that contribute to its correct expected payout being greater than that in the ESG-unaware view. Similarly, we assume that the ESG-negative stock is issued by a company with poor ESG practices that contribute to its correct expected payout being less than that in the ESG-unaware view.[81] The ESG-aware view takes the ESG practices of both companies into account, while the ESG-unaware view ignores them leading to a less accurate estimate of expected payout. To keep the example symmetric, we subtract the same amount from the ESG-negative stock's correct expected payout as we add to the correct expected payout of the ESG-positive stock. Of course, the realized payouts of both stocks are the same for both investors.

---

[78]The Sustainalytics ESG Risk Rating, which is the basis for the Morningstar Sustainability Rating for Funds (the globe rating) is a pecuniary ESG rating.

[79]The Sustainalytics ESG Rating, which is distinct from the ESG Risk Rating, is a pecuniary ESG rating.

[80]We adopted this terminology from PFP.

[81]The assumptions regarding the relationship between ESG practices and expected payout are based on the assumptions of PFP. Alternatively, one could argue that good ESG practices are expensive and reduce payouts. The argument is that if a firm is internalizing environmental costs that other companies are externalizing, then it is at a disadvantage. For some investors it may be worth it.

In this example, we assume the investors have equal amounts of capital. We also assume that both investors have no nonpecuniary preferences and identical pecuniary preferences for risk and expected return (i.e., the same risk aversion parameter, λ in Equation 9.1). We start with a CAPM-like model, in which each investor *i* seeks to maximize risk-adjusted expected return but can have return expectations on each of the investments that differ from those of other investors:

$$Obj_i = \underbrace{\mathbf{h}_i^{\mathrm{T}}\boldsymbol{\mu}_i}_{\substack{\text{Portfolio}\\ \text{Expected}\\ \text{Excess Return}}} - \underbrace{\frac{\lambda_i}{2}\underbrace{\mathbf{h}_i^{\mathrm{T}}\mathbf{V}\mathbf{h}_i}_{\substack{\text{Portfolio}\\ \text{Variance}}}}_{\text{Penalty for Risk}}, \tag{9.2}$$

where:

Obj_i = investor *i*'s objective function (what investor *i* seeks to maximize);

$\mathbf{h}_i$ = investor *i*'s holdings (weights) on the investments ($M \times 1$ column vector);

$\boldsymbol{\mu}_i$ = investor *i*'s expected excess returns on the investments ($M \times 1$ column vector);

λ_i = investor *i*'s risk aversion coefficient;

$\mathbf{V}$ = variance–covariance matrix of the investment returns ($M \times M$ matrix); and

M = the number of investments.

In a CAPM world, solving a portfolio optimization problem is unnecessary. All investors share the same capital market assumptions (no disagreement) and would thus arrive at the same efficient frontier with the same capital market line identifying the agreed-upon Sharpe-maximizing portfolio or market portfolio. It is as if the market has done the optimization for you, and the market-cap weights are the optimal weights on risky assets for all investors.

The only degree of personalization in the process, then, relates to each investor's risk aversion coefficient, which dictates the degree to which the investor borrows or lends cash to arrive at a portfolio consistent with their risk appetite.

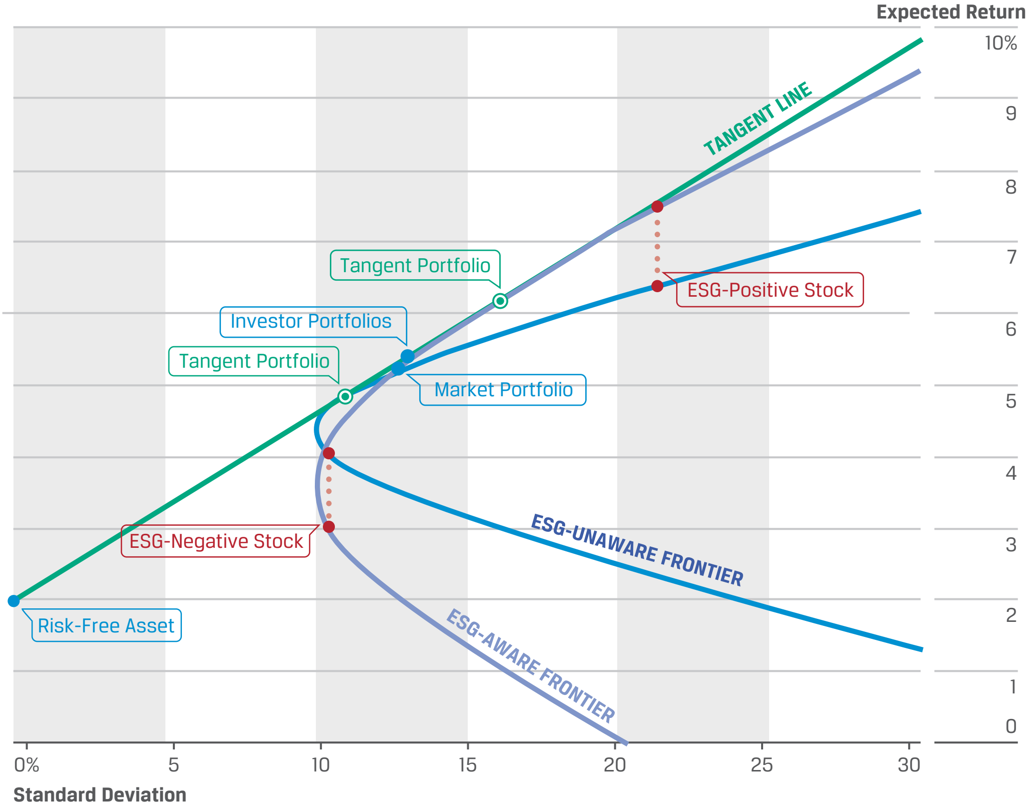
That is the standard or original CAPM. We now depart from it. If one allows for *disagreement*, as we have, it is necessary for each investor to maximize the risk-adjusted expected return in Equation 9.2 based on their view of the capital market assumptions (presumably arriving at an estimate of the efficient frontier and Sharpe-maximizing portfolio that differs from the estimates made by other investors). Each investor believes that their view is correct.

Exhibit 9.2 shows the expected returns and standard deviations of the stocks and investor portfolios in equilibrium under both the ESG-unaware and ESG-aware views. Under both views, the ESG-positive stock has the same standard deviation, but it has a higher expected return. Similarly, under both views, the ESG-negative stock has the same standard deviation but a lower expected return.

Taking all possible portfolio combinations of the two stocks (both long and short positions) under the ESG-unaware view, results in the ESG-unaware frontier (which is *not* the true frontier as it fails to incorporate pecuniary ESG factors). Similarly, taking all possible portfolio combinations of the two stocks under the ESG-aware view, results in the ESG-aware frontier (which is the true frontier as it correctly incorporates pecuniary ESG factors). Note how the ESG-aware frontier is higher and wider than the ESG-unaware frontier.

As in the standard CAPM, each investor holds a portfolio on a line that is tangent to their estimated frontier, emanating from the point that represents the risk-free asset (cash). Each of these portfolios is mean–variance efficient under the respective investor's view. As Exhibit 9.2 shows, in the model that we present, the frontiers under both views have the same tangent line. The portfolio represented by each point on the

Exhibit 9.2. Equilibrium with Different ESG Views and No Nonpecuniary Preferences



Source: Kaplan (2021).

line, however, depends on which view is in effect. In fact, the portfolios of both investors are on the same point, even though, as **Exhibit 9.3** shows, their compositions are different.

Interestingly, the market portfolio has the same expected return and standard deviation under both views. Because it is below the tangent line, it is not mean–variance efficient under either view.

Exhibit 9.3 provides the details on the portfolios shown in Exhibit 9.2. Under the ESG-unaware view, the tangent portfolio is about 66% in the ESG-negative stock. The investor with the ESG-unaware view holds a levered position of this tangent portfolio, going short about 19% in cash. In contrast, under the ESG-aware view, the tangent portfolio is about 72% in the ESG-positive stock. The investor with the ESG-aware view combines this tangent portfolio with about a 19% long position in cash, which offsets the short position in cash of the other investor.

In Exhibit 9.3, we have included the Sharpe ratios of each portfolio under the ESG-aware view.[82] The Sharpe ratio measures the mean–variance efficiency of a portfolio. Under the ESG-aware (correct) view, only the

[82]The Sharpe ratio of a portfolio is $(\mu - r_f)/\sigma$, where μ is the expected return of the portfolio, r_f is the risk-free rate, and σ is the standard deviation of return on the portfolio.

Exhibit 9.3. Details of Portfolios under Equilibrium with Different ESG Views and No Nonpecuniary Preferences

View	Portfolio	Portfolio Weights: ESG Positive Stock (%)	Portfolio Weights: ESG Negative Stock (%)	Cash (%)	Expected Return (%)	Standard Deviation (%)	Sharpe Ratio under ESG-Aware View
ESG-Unaware	Tangent	34.36	65.64	0.00	4.86	10.96	0.23
	Investor	40.88	78.09	−18.96	5.40	13.04	
ESG-Aware	Tangent	71.59	28.41	0.00	6.19	16.09	0.26
	Investor	58.01	23.03	18.96	5.40	13.04	
Both	Market	49.44	50.56	0.00	5.21	12.67	0.25

ESG-aware investor holds a mean–variance efficient portfolio. The ESG-unaware investor inadvertently holds an inefficient portfolio. Because the market portfolio is a blend of efficient and inefficient portfolios, it is inefficient.

Personalization Based on Different ESG Views and Nonpecuniary Preferences

To introduce nonpecuniary preferences into the model, we now assume that the following four investors all have the same level of capital:[83]

Investor 1. Holds the ESG-Unaware View and has no nonpecuniary ESG preference.

Investor 2. Holds the ESG-Unaware View and has a nonpecuniary ESG preference.

Investor 3. Holds the ESG-Aware View and has no nonpecuniary ESG preference.

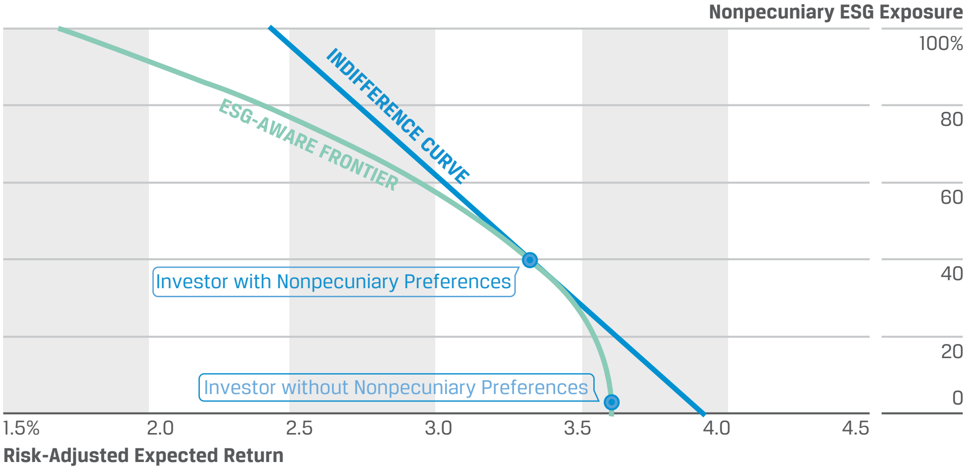
Investor 4. Holds the ESG-Aware View and has a nonpecuniary ESG preference.

There can be trade-offs between nonpecuniary ESG and pecuniary risk-adjusted expected return. Investors who have no nonpecuniary preferences seek to maximize risk-adjusted expected return, whereas those with pecuniary preferences seek to balance nonpecuniary ESG exposure and risk-adjusted expected return.

The inputs in Equation 9.2 (expected returns and the covariance matrix) are purely pecuniary in nature, because the preferences are purely pecuniary. They account for disagreement, but not for an investor's nonfinancial or nonpecuniary preferences, called "tastes" by Fama and French (2007). To account for an investor's nonpecuniary preferences, the PAPM expands to add a term to risk-adjusted expected return in the portfolio optimization problem to directly account for such preferences. Notice that, relative to Equation 9.2, Equation 9.3 includes this additional term:

[83]The PFP model has three types of investors who are like investors 1, 3, and 4. They do not include investor 2.

Exhibit 9.4. Pecuniary/Nonpecuniary ESG Frontier and Investor Decisions



Source: Kaplan (2021).

$$Obj_i = \underbrace{\mathbf{h}_i^T \boldsymbol{\mu}_i}_{\substack{\text{Portfolio}\\ \text{Expected}\\ \text{Excess Return}}} + \underbrace{\mathbf{h}_i^T \mathbf{C} \boldsymbol{\phi}_i}_{\substack{\text{Nonpecuniary}\\ \text{Benefit}}} - \underbrace{\frac{\lambda}{2} \underbrace{\mathbf{h}_i^T \mathbf{V} \mathbf{h}_i}_{\substack{\text{Portfolio}\\ \text{Variance}}}}_{\text{Penalty for Risk}}, \qquad (9.3)$$

where

$\mathbf{C}$ = exposure of investments to nonpecuniary characteristics ($M \times P$ matrix); and

$\boldsymbol{\phi}_i$ = investor *i*'s nonpecuniary preferences ($P \times 1$ column vector).

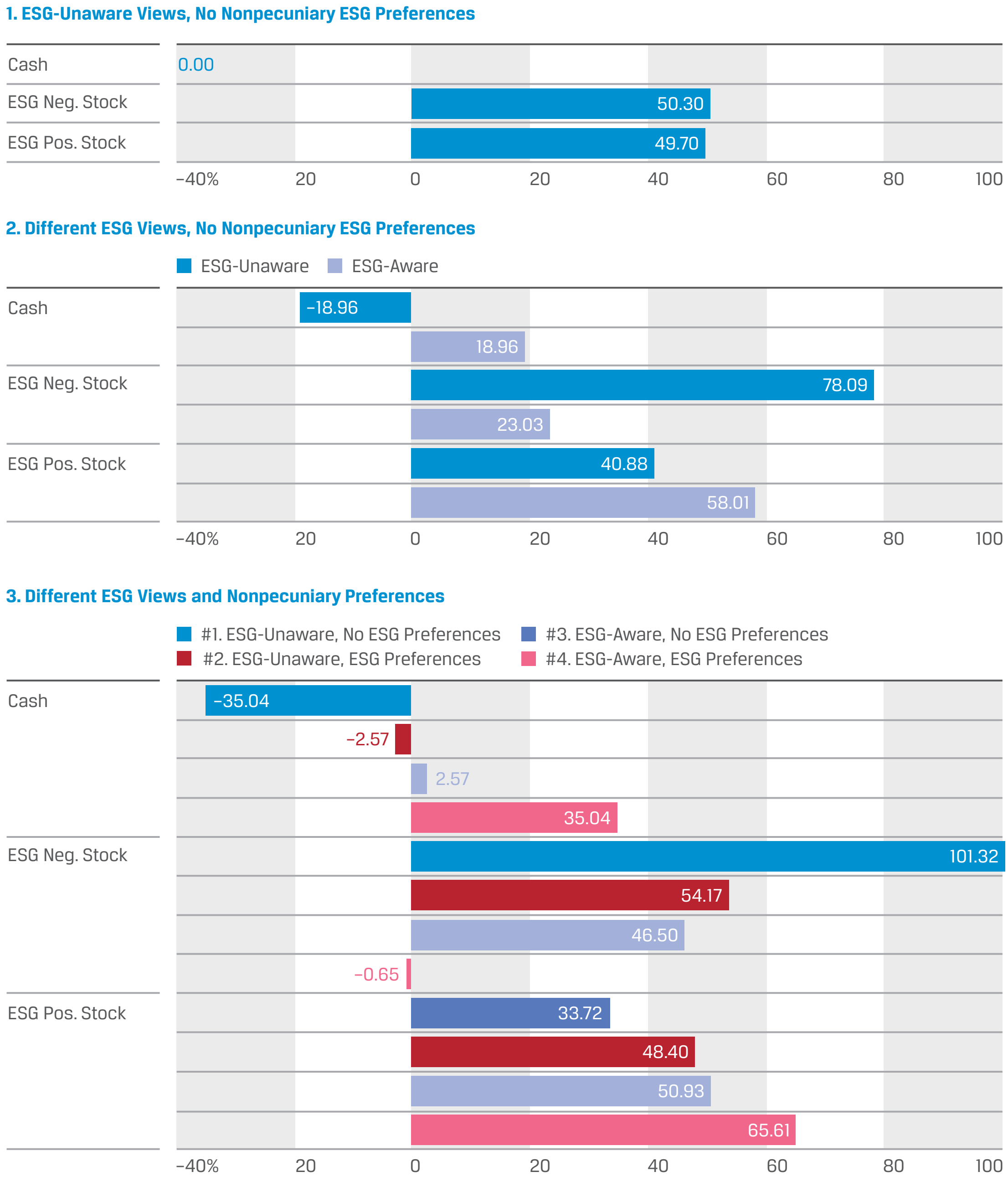
Exhibit 9.4 illustrates how this works. Based on the ESG-aware view, it shows the pecuniary/nonpecuniary frontier. Each point on this frontier gives the highest possible value of risk-adjusted expected return for a given level of nonpecuniary ESG exposure.[84] As this exhibit shows, investors who have no nonpecuniary preferences (investor 3) select whatever level of nonpecuniary ESG exposure goes with the portfolio with the highest level of risk-adjusted expected return. In contrast, investors with nonpecuniary preferences (investor 4), give up some risk-adjusted expected return to gain some nonpecuniary ESG exposure. To what extent they make this trade-off depends on (1) the curvature of the frontier and (2) the strength of their nonpecuniary preferences. In Exhibit 9.4, we have included an indifference curve for investor 4, which is a line that is tangent to the pecuniary/nonpecuniary frontier. The slope of this line shows how much risk-adjusted expected return investor 4 is willing to give up to gain nonpecuniary ESG exposure, based on their nonpecuniary preferences. The point of tangency with the frontier shows where they end up.

Exhibit 9.5 shows investor portfolios under three sets of assumptions:

1. Investors have ESG-unaware views and no nonpecuniary preferences.
2. Investors have different ESG views and no nonpecuniary preferences.
3. Investors have different ESG views and nonpecuniary preferences.

[84]PFP present a similar frontier but use the Sharpe ratio rather than risk-adjusted expected return as the pecuniary measure.

Exhibit 9.5. Investor Portfolios under Alternative ESG Assumptions



Source: Kaplan (2021).

The first set of assumptions gives us the CAPM with identical investors. The result is that all investors hold the market portfolio, which is about 50% in the ESG positive stock and 50% in the ESG negative stock. The second set of assumptions is given in the model that we presented in Exhibits 9.1 and 9.2. We have included it here to contrast it with the other models. The third set of assumptions is given in the model with four investors. Note how as we move from investor 1, to 2, to 3, to 4, the holdings on the ESG positive stock increase, and the holdings on the ESG negative stock decrease. This shows how both pecuniary ESG-awareness and nonpecuniary preferences affect the ESG exposure of a portfolio.

Conclusion and Key Takeaways

First and foremost, the PAPM is an equilibrium asset pricing model that generalizes the CAPM by incorporating what Fama and French (2007) identify as its two missing ingredients: disagreement and tastes. By incorporating pecuniary disagreement and nonpecuniary tastes, it moves us away from a world of purely pecuniary financial modeling into a world of personalization, albeit a world that continues to be based on the rich pedigree of the theory of rational behavior.

The PAPM is a powerful and flexible model that allows us to incorporate both differing economic pecuniary views and nonpecuniary preferences. It is therefore especially well-suited to address the impact on both pecuniary and nonpecuniary ESG factors in a single model to reveal how they affect asset prices and investor portfolios. The model that we present demonstrates the impact of *both* ESG views and preferences and, furthermore, the possible trade-offs between nonpecuniary ESG exposure and pecuniary risk-adjusted expected return that investors may need to make. Recognizing these distinct impacts of ESG views and preferences and the pecuniary/nonpecuniary trade-offs are the main lessons from the ESG version of the PAPM that we discussed in this chapter.

10. PERSONALIZED PORTFOLIO CONSTRUCTION:[85] SINGLE ACCOUNT

Context

In the previous chapter, we introduced the PAPM as a generalization of the CAPM that accounts for both (1) disagreement among investors about the financial or pecuniary prospects for securities (henceforth "heterogeneous expectations") and (2) nonfinancial or nonpecuniary investor preferences. Both (1) and (2) exist in the PAPM; however, in classical finance, they do not.

The PAPM ushers in a world of personalization in which investors build personalized portfolios that reflect their pecuniary views and nonpecuniary (nonfinancial) preferences. In this chapter, we recast and extend the portfolio maximization problem from one involving total return optimization to an alpha-tracking error optimization problem in which the goal is to implement a target asset allocation while reflecting the nonpecuniary preferences of the investor. In this chapter, we do not cover the added complexities of multiple account-*types*, multiple accounts, or taxes. We do that in the next chapter, in which we extend the single account example presented here to a multiple account setting in which there are accounts with different tax treatments.

Key Insights

- Alpha-tracking error optimization is a form of MVO across both active managers and index funds, where it is assumed that forward-looking alphas and tracking error (active risk) can be calculated for each manager. The target asset allocation is treated as an input and the key outputs are the weights or allocations to managers. Alpha-tracking error optimization thus finds the solution that maximizes forward-looking alpha for the entire portfolio for a given level of active risk.
- Consistent with the PAPM, as an extension of the Markowitz portfolio maximization problem, the alpha-tracking error optimization problem can be expanded to include a term that captures the investor's nonpecuniary preferences.
- We believe that including a nonpecuniary preference term directly in the objective function is better than imposing exclusionary constraints because doing so allows the optimizer to consider the various trade-offs involved in personalization, leading to personalized portfolios that tilt toward characteristics that the investor likes and away from characteristics they dislike.
- The absence of nonpecuniary preferences is simply a special case of the more general alpha-tracking error optimization problem that allows for nonpecuniary preferences.

In chapter 9, by incorporating pecuniary "disagreement" and nonpecuniary "tastes" into a generalized asset pricing model, the PAPM, we moved from a world of purely pecuniary financial modeling into a world of personalized portfolio construction. More specifically, we expanded the Markowitz MVO problem to include an additional term that reflects the investor's nonpecuniary preferences. This enables the objective function to tilt the solution toward characteristics that the investor likes and away from characteristics that the investor dislikes. Although one could always choose to include exclusionary constraints into the optimization, incorporating the nonpecuniary preference term directly in the objective function is consistent with both certain aspects of behavioral finance and the theory of rational behavior in which the optimizer makes decisions based on the investor's pecuniary views and nonpecuniary preferences.

---

[85]This chapter is adapted from Idzorek (2022) and Idzorek and Kaplan (2022).

Markowitz writes, "In our analyses the [portfolio weights] might represent individual securities or they might represent aggregates such as, say, bonds, stocks, and real estate" (1952, p. 91). Although the Markowitz framework can be applied to an opportunity set of individual *securities*, it is most commonly applied to an opportunity set of nonoverlapping *asset classes*. In fact, that is exactly what we do in chapter 8, but with extensions to account for taxes, asset location, human capital, and liabilities. MVO's standard form does not feature the concept of a benchmark. Rather, it is often used to *form* what one might call the strategic asset allocation or multi-asset-class *policy benchmark*.

In the presence of a benchmark, whether it is a single-asset class benchmark associated with a fund or a multi-asset-class policy benchmark created by a financial adviser, portfolio construction should be done following a benchmark-relative optimization approach. In such an approach, the benchmark is explicitly expressed as a list of constituents (individual securities, factor exposures, or asset class targets) along with a portfolio weight, or amount held, for each constituent. Regardless of where a benchmark comes from or how it is specified, the typical goal of passive management is to minimize tracking error relative to a benchmark. In contrast, the typical goal of active management is to outperform a benchmark, often subject to an active risk budget constraint. This leads to alpha-tracking error optimization in which, for a given level of forward-looking, after-fee *alpha*, the optimizer minimizes *tracking error* relative to the benchmark (subject to various constraints).[86] This type of fund-of-funds optimizer traces an alpha-tracking error efficient frontier from, at the left end, the expected after-fee alpha of the minimum tracking error portfolio to, at the right end, that of the maximum expected alpha portfolio.

Alpha-tracking error optimization of individual security holdings typically requires a multifactor model.[87] We refer to alpha-tracking error optimization relative to a policy benchmark in which the opportunity set of investments consists of funds or managers (e.g., mutual funds, ETFs, separate accounts) as fund-of-funds optimization. Waring et al. (2000) presents this type of optimization.[88] This framework is ideal for selecting funds to fulfill a diversified multi-asset-class asset allocation target. This chapter focuses on extending the fund-of-funds optimization framework to account for an investor's nonpecuniary preferences.

Alpha-Tracking Error Optimization

In the context of fund-of-funds optimization, Waring and Ramkumar (2008) present a method for estimating manager alphas as an extension of the *fundamental law of active management* and alpha forecasting framework of Grinold (1989, 1994). Based on assessments of the overall portfolio manager's skill in selecting funds, the skill of each fund manager, and the opportunity that each fund manager has for effectively applying skill, the Waring and Ramkumar model provides an estimate of the information ratio for each fund. For each fund, they multiply this estimated information ratio by an estimate of the fund-specific tracking error to arrive at an estimate of preexpense alpha. They then subtract each fund's expense ratio from its preexpense alpha to arrive at after-expense, tax-exempt estimated alpha.

In chapter 9, we introduced the PAPM and showed how the inclusion of heterogeneous expectations and nonpecuniary preferences leads to a direct extension and generalization of the mean–variance objective function. This generalization takes into account the idea that some investors derive benefits from holding a portfolio that tilts toward characteristics that they like and away from characteristics they dislike. Idzorek and Kaplan (2022) present fund-of-funds optimization in a *single* nontaxable account setting in which the

---

[86]Alpha-tracking error optimization stems from the separation of returns into systematic and idiosyncratic parts in the active expected return/active risk framework of Grinold and Kahn (2000).

[87]Examples of these types of optimizers would include those from firms like Barra, Northfield, Axioma, and Morningstar.

[88]Waring et al. (2000) uses the term "manager structure optimization" (MSO) to refer to alpha-tracking error optimization applied to an opportunity set consisting of funds or managers. What we call fund-of-funds optimization is the same thing. Other articles on fund-of-funds optimization include Baierl and Chen (2000), Stewart (2013), and Kaplan (2016, 2019).

objective function includes a nonpecuniary preference term based on Idzorek, Kaplan, and Ibbotson (2021, 2023) to create personalized fund-of-funds ESG portfolios for investors with different ESG preferences. We incorporate that same innovation here.

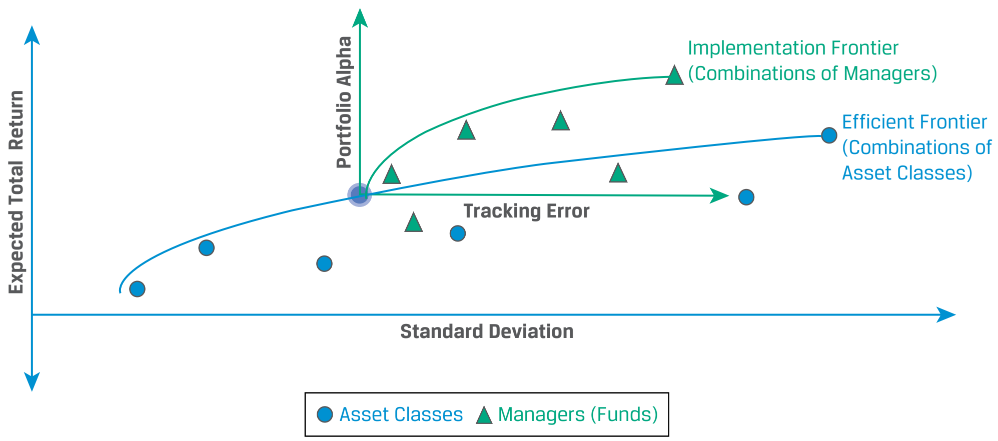
MVO, Fund-of-Funds Optimization, and Nonpecuniary Preferences

Exhibit 10.1, inspired by Waring et al. (2000), attempts to clarify the relationship between MVO and fund-of-funds optimization. In Exhibit 10.1, at the far left, the vertical axis represents the expected total return of asset allocations, and at the bottom, the horizontal axis represents the risk (standard deviation). This is mean–variance space, and everything corresponding to it is in blue. The blue circles represent the expected total return/risk points of various *asset classes*. The blue curve represents the mean–variance efficient frontier, in which each point on the efficient frontier is a combination of *asset classes* that maximizes expected *total* return for a given level of *total* risk.

We selected an arbitrary point along the efficient frontier to represent the desired "target strategic asset allocation policy benchmark." We could have chosen a mix that is off the frontier or thrown darts to arrive at the target. We recommend that separate account-type target asset allocations for taxable financial assets and for tax-advantaged (tax-deferred or tax-exempt) assets be set using the asset allocation and location model that we presented in chapter 8. Fund-of-funds optimization in this form accepts the benchmark as a given regardless of how it was created.

In Exhibit 10.1, originating from the policy benchmark is a secondary set of axes, in which the vertical axis corresponds to the expected excess return or alpha and the horizontal axis corresponds to the tracking error relative to the policy benchmark. This is alpha-tracking error space and everything corresponding to it is shown in green. Following Waring and Siegel (2003), we define the vertical axis as the expected portfolio *alpha* (the weighted sum of expected individual fund alphas) and the horizontal axis as the expected total *tracking error* relative to the policy portfolio (where tracking error comes from both asset allocation misfit and fund specific idiosyncratic risk). In the alpha-tracking error space in the graph, the green triangles represent the alphas of the *funds* and fund-specific tracking errors, while the green implementation

Exhibit 10.1. Mean–Variance Efficient Frontier versus Fund-of-Funds Optimization Implementation Frontier



efficient frontier corresponds to total tracking error relative to the policy portfolio versus the alpha of the portfolio.

Moving from this graphic illustration to formulas, in Equation 10.1 we present the Markowitz mean–variance objective function. Relative to Equation 9.2, in Equation 10.1 we drop the subscript "*i*" and use the subscript "*K*" to represent asset classes. The labels following Equation 10.1 explain the equation, which says that the objective is to find the holdings (allocations) to the different asset classes that maximize expected return minus a penalty for risk. Although Equation 10.1 could certainly be applied to individual securities, in this asset allocation context, we assume they are being applied to asset classes and we also typically assume no shorting and that the weights must sum to 1.[89]

$$Obj = \underbrace{\mathbf{h}_{\mathbf{K}}^{\mathbf{T}}\boldsymbol{\mu}_{\mathbf{K}}}_{\substack{\text{Asset}\\\text{Allocation}\\\text{Return}}} \underbrace{- \frac{\lambda_a}{2} \underbrace{\mathbf{h}_{\mathbf{K}}^{\mathbf{T}}\mathbf{V}_{\mathbf{K}}\mathbf{h}_{\mathbf{K}}}_{\substack{\text{Asset Allocation}\\\text{Variance}}}}_{\text{Penalty for Risk}}, \tag{10.1}$$

where:

Obj = investor's objective function (what the investor seeks to maximize);

$\mathbf{h}_{\mathbf{K}}$ = investor's holdings (weights) on the asset classes ($K \times 1$ column vector);

$\boldsymbol{\mu}_{\mathbf{K}}$ = investor's expected returns on the asset classes ($K \times 1$ column vector);

λ_a = risk aversion coefficient on asset allocation risk (1×1 scalar); and

$\mathbf{V}_{\mathbf{K}}$ = covariance matrix of the asset classes ($K \times K$ matrix).

We now move from asset classes to *investments* (e.g., funds/managers). Let:

$\mathbf{h}_{\mathbf{M}}$ = holdings (weights) to each fund/manager ($M \times 1$ column vector); and

$\mathbf{X}$ = exposure of investments to asset classes or factors ($M \times K$ matrix).

The asset class (or factor) exposure matrix ($\mathbf{X}$) identifies the exposure of each of the possible investments to the different asset class factors. This enables the calculation of the *effective asset allocation* of the portfolio. Using matrix math, the transpose of the exposure matrix of the investments to the (asset class factors) is multiplied by the list of holdings to each fund/manager, $\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}$.

Moving from Equation 10.1 to Equation 10.2, the decision variable changes. Rather than change the weights of different asset classes, the optimizer changes the weights of the different funds/managers. Thus, $\mathbf{h}_{\mathbf{K}}$ is replaced with $\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}$ (and thus, $\mathbf{h}_{\mathbf{K}}^{\mathbf{T}}$, with $\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{X}$), where we use the subscript "$_{\mathbf{M}}$" to distinguish a vector of holdings of managers ($\mathbf{h}_{\mathbf{M}}$) from a vector of asset class weights ($\mathbf{h}_{\mathbf{K}}$). To the degree that the factor model in question used to estimate $\mathbf{X}$ does not fully capture the total return of each fund and thus $\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}$ does not include the idiosyncratic return of each investment, we need two additional terms. First, we need a term to capture alpha or the expected excess return of the fund. Second, we need a term to capture idiosyncratic or residual risk associated with alpha. These terms are included at the far left and far right ends of the right-hand side of Equation 10.2, respectively:

$$Obj = \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\boldsymbol{\alpha}_{\mathbf{M}}}_{\text{Alpha}} + \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{X}\boldsymbol{\mu}_{\mathbf{K}}}_{\substack{\text{Asset}\\\text{Allocation}\\\text{Return}}} \underbrace{- \frac{1}{2}\lambda_a \underbrace{[\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}]^{\mathbf{T}}\mathbf{V}_{\mathbf{K}}[\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}]}_{\substack{\text{Asset Allocation}\\\text{Variance}}} - \frac{1}{2}\lambda_m \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{V}_{\alpha}\mathbf{h}_{\mathbf{M}}}_{\substack{\text{Residual}\\\text{Variance}}}}_{\text{Penalty for Risk}}, \tag{10.2}$$

[89]In chapter 8, we extended the asset allocation framework to account for human capital, liabilities, and taxes.

where:

$\boldsymbol{\alpha}_{\mathbf{M}}$ = alpha (expected excess return) of the fund/managers ($M \times 1$ column vector);

λ_m = risk aversion coefficient on fund-/manager-specific risk (1×1 scalar); and

$\mathbf{V}_{\alpha}$ = covariance matrix of fund-/manager-specific risk ($M \times M$ matrix).

In Equation 10.3, we extend Equation 10.2 to include nonpecuniary preferences by adding the nonpecuniary preference term in the objective function of the PAPM (see Equation 9.2):

$$Obj = \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\boldsymbol{\alpha}_{\mathbf{M}}}_{\text{Alpha}} + \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{X}\boldsymbol{\mu}_{\mathbf{K}}}_{\substack{\text{Asset}\\\text{Allocation}\\\text{Return}}} + \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{C}\boldsymbol{\phi}}_{\substack{\text{Nonpecuniary}\\\text{Benefit}}} \underbrace{- \frac{1}{2}\lambda_a \underbrace{[\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}]^{\mathbf{T}}\mathbf{V}_{\mathbf{K}}[\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}]}_{\substack{\text{Asset Allocation}\\\text{Variance}}} - \frac{1}{2}\lambda_m \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{V}_{\alpha}\mathbf{h}_{\mathbf{M}}}_{\substack{\text{Residual}\\\text{Variance}}}}_{\text{Penalty for Risk}}, \qquad (10.3)$$

where:

$\mathbf{C}$ = exposure of investments to nonpecuniary characteristics ($M \times P$ matrix); and

$\boldsymbol{\phi}$ = investor's nonpecuniary preferences ($P \times 1$ column vector).

Maximizing the objective function given in Equation 10.3 simultaneously solves for (1) the optimal weight assigned to each manager as well as (2) the optimal weight in each asset class. In practice, very few (if any) practitioners simultaneously solve for the optimal manager structure and optimal asset allocation. Rather, it is far more common to solve for the optimal asset allocation; formalize that as a target strategic asset allocation policy benchmark; and then, in a separate process, identify the optimal combination of funds/managers to implement the policy benchmark.

This moves us from a total return optimization framework to the alpha-tracking error optimization framework depicted by the second set of axes shown in Exhibit 10.1. These axes represent a benchmark-relative alpha-tracking error fund-of-funds optimization in which the target strategic asset allocation policy benchmark is an input rather than an output to the optimization. In the presence of a presumed efficient asset allocation target ($\mathbf{h}_{\mathbf{K}}$), as demonstrated in the appendix of Idzorek and Kaplan (2022) and in appendix 10A, Equation 10.3 reduces to Equation 10.4 in which the target asset allocation $\mathbf{h}_{\mathbf{K}}$ is an *input*. Here, the difference between the effective asset allocation ($\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}}$) and target asset allocation ($\mathbf{h}_{\mathbf{K}}^{\mathbf{T}}$) represents *active* asset class exposures and creates active asset class misfit risk.

$$Obj = \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\boldsymbol{\alpha}_{\mathbf{M}}}_{\text{Alpha}} + \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{C}\boldsymbol{\phi}}_{\substack{\text{Nonpecuniary}\\\text{Benefit}}} \underbrace{- \frac{1}{2}\lambda_a \underbrace{[\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}} - \mathbf{h}_{\mathbf{K}}]^{\mathbf{T}}\mathbf{V}_{\mathbf{K}}[\mathbf{X}^{\mathbf{T}}\mathbf{h}_{\mathbf{M}} - \mathbf{h}_{\mathbf{K}}]}_{\substack{\text{Asset Allocation}\\\text{Misfit Risk}}} - \frac{1}{2}\lambda_m \underbrace{\mathbf{h}_{\mathbf{M}}^{\mathbf{T}}\mathbf{V}_{\alpha}\mathbf{h}_{\mathbf{M}}}_{\substack{\text{Residual}\\\text{Variance}}}}_{\text{Penalty for Active Risk}}. \qquad (10.4)$$

Equation 10.4 is equivalent to the alpha-tracking error objective function put forth in Waring et al. (2000), with one important addition, the inclusion of the nonpecuniary benefit term, similar to the one in Cooper et al. (2016). This term allows the alpha-tracking error framework to simultaneously evaluate the pecuniary alpha-tracking error trade-off and the nonpecuniary benefit of using investments that tilt toward liked characteristics and away from disliked characteristics.

Equation 10.4 shows the two different contributors to active risk or total tracking error: (1) asset class misfit risk caused by an investment-specific portfolio with an "effective asset allocation" that does not match the target and (2) investment-specific residual risk. Equation 10.4 allows for two different types of risk aversion to the two different sources of tracking error. In this context, a particular investor may or may not care where the tracking error comes from; if they do not care, these would have the same value.

To create the complete implementation frontier depicted in Exhibit 10.1, we minimize tracking error, holding the target level of alpha fixed at various values, ranging from the alpha of the minimum tracking error portfolio to the highest feasible level of alpha, subject to the constraints on the problem. To solve the problem for a specific investor, we use the investor's risk aversion coefficients and maximize the objective function given in Equation 10.4, subject to the constraints on the problem. Fund-of-funds optimization is usually carried out with no shorting and a budget constraint in which the sum of the *manager* holdings/weights must sum to 100%.

In the presence of the budget and non-negativity constraints, in Equation 10.4, $\mathbf{h}_{\mathbf{M}}$ is simply a list of allocations to all available managers/funds, in which the weight or amount allocated to each manager is a percentage between 0% and 100% and the sum of all allocations equals 100%. In practice, when one solves the optimization problem, 0% is often allocated to a number of available managers in favor of large allocations to the best and most suitable funds. As one moves along the implementation frontier, the allocations change and eventually, at the far right of the frontier, in the absence of other limiting constraints, 100% is allocated to the single fund with the highest expected alpha.

Inherent in this setup is the assumption that the optimization is taking place within a single account (with no trading costs) or, if it is occurring across multiple accounts, that all money is completely fungible with no taxes or trading costs—thus the distinctions between accounts can be ignored. In chapter 11, we will see that, in the world of individual investors, money is not completely fungible across accounts or account types because of various tax rules and trading costs. These complexities motivate us to build a new model that explicitly considers these complicating and economically meaningful trade-offs.

Personalized Preference-Based Portfolios

Following Idzorek and Kaplan (2022), we use Equation 10.4 to create personalized portfolios for six hypothetical investors, each of whom potentially has different preferences for six nonpecuniary characteristics. (One could include any number of nonpecuniary preferences.)

The six nonpecuniary characteristics assumed in this example are as follows: gender equality, green energy, board diversity, alcohol, tobacco, and guns. In general, we would say that many investors like gender equality, green energy, and board diversity and that many investors dislike alcohol, tobacco, and guns. Of course, some investors have different preferences, such as *liking* alcohol, tobacco, or guns. Although we have focused on ESG-oriented characteristics, this generalized framework works for any type and any number of nonpecuniary characteristics (e.g., liquidity, dividends, home country) and investor preferences for them.

We recognize that different investors frequently have different pecuniary views on how different characteristics, including pecuniary ESG characteristics, may influence risk and expected return, resulting in different portfolios; in this example, our focus is to control for pecuniary views and highlight the differences driven by nonpecuniary preferences. Moreover, to isolate and to focus on the impact of differing nonpecuniary preferences, we assume that each of the six investors has the same risk tolerance, the same target strategic asset allocation or policy benchmark, the same aversion to tracking error relative to the policy benchmark, the same opportunity set of nine available funds (including a money market fund), and the same expectations for the nine different funds (including alpha, residual risk, and asset class exposures). We have standardized these parameters in the spirit of a controlled experiment, but in the real world, we envision investment management professionals having optimization tools that enable them to build personalized portfolios for their clients regardless of risk tolerance, the target policy portfolio, and the opportunity set of investments. To that end, the spreadsheet used to create this example is available as part of the supplementary materials to this book.

Starting from the *shared-pecuniary-perspective*, in **Exhibit 10.2**, we identify the pecuniary views that the six investors all share related to the funds' after-fee expected alphas, residual risks, and respective exposures to the asset classes used to define the policy benchmark (the target asset allocation). All six investors have the same target policy benchmark: 35% US Equity, 20% Developed Markets ex-US Equity, 10% Emerging Market Equity, 20% US Bonds, 10% Non-US Bonds, and 5% Cash. The target policy benchmark could be the result of solving either Equations 10.1 or the asset location and allocation problem described in chapter 8.

To complete our view the different attributes of the investment options, in **Exhibit 10.3** we identify the non-pecuniary characteristics embedded in each of the funds.

Looking across both Exhibit 10.2 and Exhibit 10.3, notice that the investor can use four equity funds, four bond funds, and one money market fund to implement the target asset allocation. Fund A and Fund E are global index ETFs, with a slightly negative alpha (due to expenses) and low residual risk. To draw out this intuition, we assume that the two global passive funds provide a 20% exposure to each of the six nonpecuniary characteristics. The rest of the funds have 300 basis points of residual risk. Fund B and Fund F are *impact* funds offering higher exposures to three impact themes, both of which have a moderately positive alphas of 20 basis points. Fund C and Fund G are anti-vice funds offering lower exposures to alcohol, tobacco, and guns as well as moderately positive alphas of 20 basis points. Fund D and Fund H seek out high exposures to alcohol, tobacco, and guns, both of which have more positive alphas of 40 basis points.

We now turn to the nonpecuniary preference of the six investors for the six different nonpecuniary characteristics.

To make informed preference-based trade-offs, one needs to quantify what it means to like or dislike, or to love or hate, a characteristic. For a given investor, we need to estimate how much expected return they

Exhibit 10.2. Shared Pecuniary View on Funds

			Asset Class Exposures					
Fund	Alpha	Residual Risk	US Equity	DM ex-US Equity	EM Equity	US Bonds	Non-US Bonds	Cash
Fund A – Global Equity ETF	–0.05%	0.10%	55.0%	30.0%	15.0%	0.0%	0.0%	0.0%
Fund B – Impact Equity	0.20%	3.00%	55.0%	30.0%	15.0%	0.0%	0.0%	0.0%
Fund C – Anti-Vice Equity	0.20%	3.00%	55.0%	30.0%	15.0%	0.0%	0.0%	0.0%
Fund D – Vice Equity	0.40%	3.00%	55.0%	30.0%	15.0%	0.0%	0.0%	0.0%
Fund E – Global Bond ETF	–0.05%	0.10%	0.0%	0.0%	0.0%	52.5%	35.0%	12.5%
Fund F – Impact Bond	0.20%	3.00%	0.0%	0.0%	0.0%	52.5%	35.0%	12.5%
Fund G – Anti-Vice Bond	0.20%	3.00%	0.0%	0.0%	0.0%	52.5%	35.0%	12.5%
Fund H – Vice Bond	0.40%	3.00%	0.0%	0.0%	0.0%	52.5%	35.0%	12.5%
Fund I – Money Market	0.00%	0.00%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
Target Asset Allocation			35.0%	20.0%	10.0%	20.0%	10.0%	5.0%

Exhibit 10.3. Nonpecuniary Fund Characteristics and Exposures

	Gender Equality	Green Energy	Board Diversity	Alcohol	Tobacco	Guns
Fund A – Global Equity ETF	20%	20%	20%	20%	20%	20%
Fund B – Impact Equity	35%	55%	40%	25%	35%	15%
Fund C – Anti-Vice Equity	15%	15%	15%	0%	0%	5%
Fund D – Vice Equity	15%	10%	5%	35%	55%	40%
Fund E – Global Bond ETF	20%	20%	20%	20%	20%	20%
Fund F – Impact Bond	35%	55%	40%	25%	35%	15%
Fund G – Anti-Vice Bond	15%	15%	15%	0%	0%	5%
Fund H – Vice Bond	15%	10%	5%	35%	55%	40%
Fund I – Money Market	0%	0%	0%	0%	0%	0%

would be willing give up to either (1) increase their exposure to a characteristic they like or (2) decrease their exposure to a characteristic they dislike. The exercise is similar to and every bit as challenging as estimating an investor's risk tolerance. In the spirit of revealed risk preferences (Samuelson 1938, 1948), in practice and as somewhat demonstrated in chapter 2, we would attempt to estimate the trade-offs a given investor would make through a series of iterative, interactive trade-off questions and use them to come up with the investor's nonpecuniary preference parameters, ϕ.

For our example, we assume six investors with a variety of nonpecuniary preferences. Investor 1 has no nonpecuniary preferences. Investors 2 and 3 both prefer impact investing, although they differ in the degree of preference—that is, investor 2 "likes" impact investing and investor 3 "loves" impact investing. Investor 4 is a faith-based investor and has a relatively strong "dislike" of alcohol and a moderate "dislike" of tobacco. Investor 5 disdains guns. Investor 6 has multiple preferences, liking the first three characteristics and disliking the last three characteristics.

Exhibit 10.4 identifies the six investors and the assumed quantification of their nonpecuniary preferences. More specifically, each column identifies ϕ in Equations 10.4 for a given investor. Note that, if an investor does not have a nonpecuniary preference (for which they are willing to sacrifice some level of expected return), the corresponding element of the vector is zero. As constructed in this example, if an investor likes a characteristic, the value is positive, and if they dislike a characteristic, the value is negative.

To develop insights around the pecuniary and nonpecuniary trade-offs inherent in the objective function, we have purposely made the pecuniary assumption that the alphas of the funds with desirable characteristics are relatively less attractive than the funds with less sought-after characteristics. This is just an example and may or may not correspond with one's intuition. We want to emphasize that actual investors are likely to have different pecuniary views—in other words, they "disagree"—and we would encourage investors to apply the framework based on their own pecuniary views.

As a side note on the measurement of characteristics in practice, we believe it is best to start by measuring the characteristics at the individual security level. Then, fund-level characteristics can be calculated by

Exhibit 10.4. Nonpecuniary Preferences of Investors

	Investor 1: No Preferences	Investor 2: Likes Impact	Investor 3: Loves Impact	Investor 4: Faith Driven	Investor 5: Anti-Guns	Investor 6: Mixed Preferences
Characteristics	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
Gender Equality	0.00%	0.75%	1.50%	0.00%	0.00%	0.25%
Green Energy	0.00%	0.75%	1.50%	0.00%	0.00%	0.25%
Board Diversity	0.00%	0.75%	1.50%	0.00%	0.00%	0.25%
Alcohol	0.00%	0.00%	0.00%	–1.50%	0.00%	–0.25%
Tobacco	0.00%	0.00%	0.00%	–0.75%	0.00%	–0.25%
Guns	0.00%	0.00%	0.00%	0.00%	–2.00%	–0.25%

taking a weighted average in which the weights are based on the fund's current individual security holdings. Although this approach can introduce other complexities, measuring characteristics in this manner enables one to potentially contemplate an opportunity set that includes both individual securities and pooled investment vehicles.[90] By their nature, the characteristics of individual securities are typically more extreme than those of diversified funds.

Exhibit 10.5 contains each of the six investors' optimal personalized portfolio as well as a well of additional information. Panels A, B, C, and D contain the allocations to the different funds, various portfolio statistics, the effective asset allocation of each portfolio, and each portfolio's exposure to the nonpecuniary characteristics.

Starting with Panel A, despite having the same pecuniary inputs, the six different investors arrive at different optimal portfolios because of their differing nonpecuniary preferences. Investor 1, with no nonpecuniary preferences, has a large allocation to Funds D and H given their high alphas. Investor 2, who "likes" impact, mostly buys the impact funds (B and F) with a moderate allocation to Fund D with its superior alpha. Investor 3, who "loves" impact, forgoes the superior alpha of Fund D, investing entirely in the two impact funds. The portfolios of investor 4 (who is faith driven) and investor 5 (who is anti-gun) are quite similar. This is due to the limited investment options in this example. They both end up investing heavily in the two anti-vice funds (Fund C and Fund G), although for different reasons. Investor 4 is seeking to avoid alcohol and tobacco, whereas investor 5 is seeking to avoid guns. Given the limited opportunity set, Fund C and Fund G are the only funds that enable one to avoid these exposures.

Moving to the portfolio summary statistic in Panel B, investor 1 with no nonpecuniary preferences has the highest alpha portfolio and lowest tracking error. Investor 3 who "loves" impact investing has the most extreme portfolio, allocating 100% to the two impact funds (Funds B and F), and ends up with the highest amount of tracking error yet the highest value of the objective function. Panel D displays each portfolio's exposure to the six nonpecuniary characteristics. To make comparisons a bit easier, **Exhibit 10.6** displays

[90]Idzorek (2022) considers the complexities of alpha-tracking error optimization when either or both the opportunity set of available investments, or the specification of the policy benchmark includes both individual securities and pooled investment vehicles.

Exhibit 10.5. Optimal Personalized Portfolios

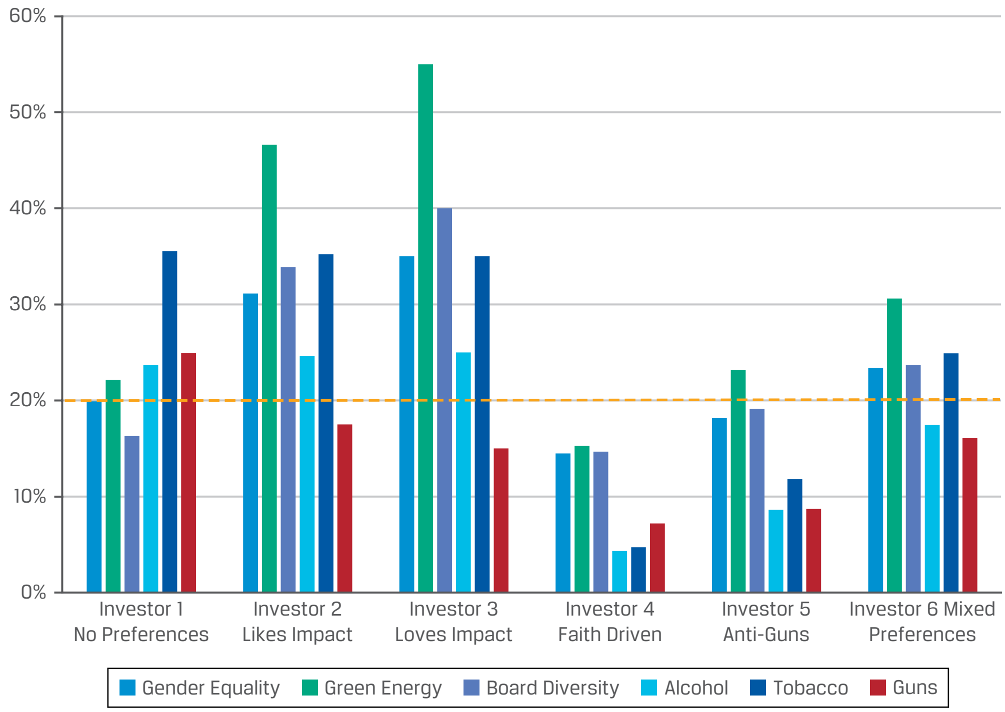
	Investor 1: No Preferences	Investor 2: Likes Impact	Investor 3: Loves Impact	Investor 4: Faith Driven	Investor 5: Anti-Guns	Investor 6: Mixed Preferences
Panel A: Allocation to Funds						
Fund A. Global Equity ETF	6.2%	0.0%	0.0%	16.8%	5.4%	7.0%
Fund B. Impact Equity	15.4%	46.0%	64.1%	3.9%	21.6%	23.9%
Fund C. Anti-Vice Equity	15.4%	6.7%	0.0%	43.2%	34.0%	21.6%
Fund D. Vice Equity	27.8%	12.1%	0.0%	0.0%	3.1%	12.4%
Fund E. Global Bond ETF	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Fund F. Impact Bond	7.6%	34.6%	35.9%	0.0%	4.2%	16.3%
Fund G. Anti-Vice Bond	7.6%	0.0%	0.0%	21.9%	16.6%	14.0%
Fund H. Vice Bond	20.0%	0.7%	0.0%	0.0%	0.0%	4.8%
Fund I. Money Market	0.0%	0.0%	0.0%	14.2%	15.2%	0.0%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Panel B: Portfolio Statistics						
Alpha	0.28%	0.23%	0.20%	0.13%	0.16%	0.22%
Nonpecuniary Term	0.00%	0.84%	1.95%	–0.10%	–0.17%	0.05%
Tracking Error	1.26%	1.78%	2.21%	1.51%	1.38%	1.23%
Total Objective Function	0.14%	0.78%	1.71%	–0.18%	–0.18%	0.13%
Panel C: Effective Asset Allocation						
US Equity	35.7%	35.6%	35.3%	35.2%	35.2%	35.7%
DM ex-US Equity	19.5%	19.4%	19.2%	19.2%	19.2%	19.5%
EM Equity	9.7%	9.7%	9.6%	9.6%	9.6%	9.7%
US Bonds	18.5%	18.5%	18.8%	11.5%	10.9%	18.5%
Non-US Bonds	12.3%	12.3%	12.6%	7.7%	7.3%	12.3%
Cash	4.4%	4.4%	4.5%	16.9%	17.8%	4.4%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

(*continued*)

Exhibit 10.5. Optimal Personalized Portfolios (*continued*)

	Investor 1: No Preferences	Investor 2: Likes Impact	Investor 3: Loves Impact	Investor 4: Faith Driven	Investor 5: Anti-Guns	Investor 6: Mixed Preferences
Panel D: Nonpecuniary Exposures						
Gender Equality	19.9%	31.1%	35.0%	14.5%	18.2%	23.4%
Green Energy	22.1%	46.6%	55.0%	15.3%	23.2%	30.6%
Board Diversity	16.3%	33.9%	40.0%	14.7%	19.1%	23.7%
Alcohol	23.7%	24.6%	25.0%	4.3%	8.6%	17.5%
Tobacco	35.6%	35.2%	35.0%	4.7%	11.8%	24.9%
Guns	24.9%	17.5%	15.0%	7.2%	8.7%	16.1%

Exhibit 10.6. Portfolio Exposures to the Six Nonpecuniary Characteristics

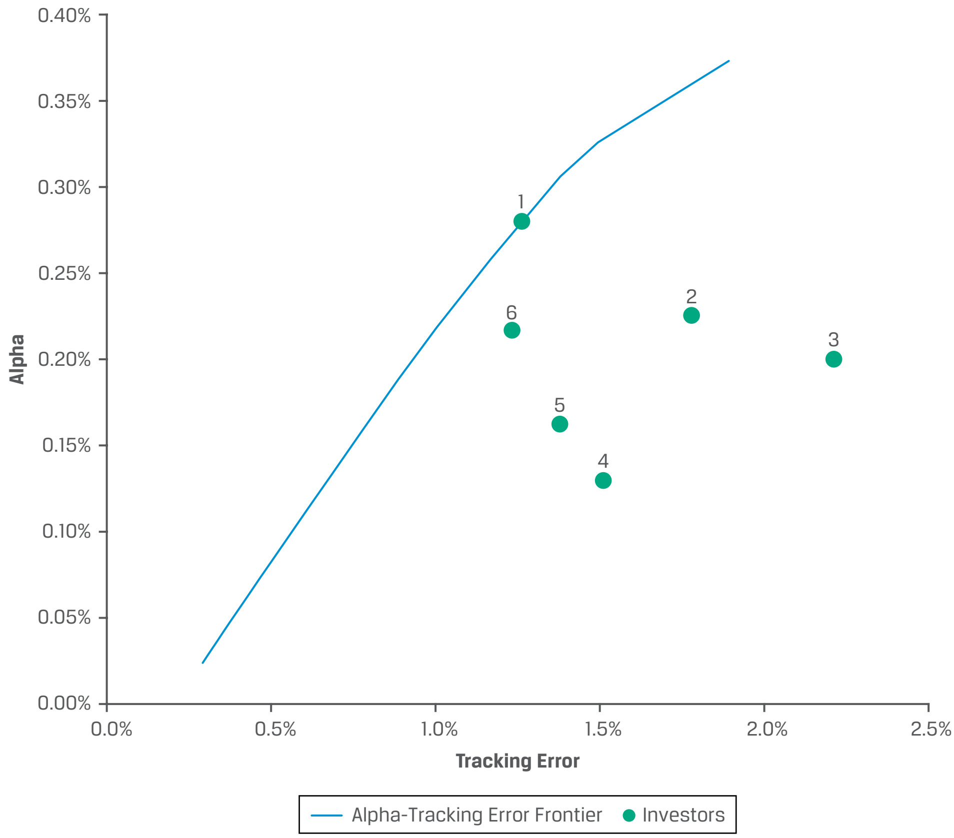


the exposures graphically. Recall that, in this quasi-controlled experiment, a market-based allocation leads to a 20% exposure to each of the nonpecuniary characteristics. This is shown as a dashed yellow line.

Investor 1, with no nonpecuniary preferences, ends up with a portfolio tilted toward the three characteristics that are generally considered to be less desirable from a nonpecuniary perspective. Relative to the market exposure of 20%, the five investors with various nonpecuniary preferences all end up with portfolios that tend to tilt toward the characteristics that they like and away from the characteristics that they dislike (to the degree that they have such preferences). In Exhibit 10.6, we see that investor 2 and investor 3 end up with similar exposures, except that investor 3's tilt toward impact themes is larger. In Panel B of Exhibit 10.5, relative to investor 2, we see that investor 3 is willing to accept a lower alpha and take on more tracking error in pursuit of a personalized portfolio that aligns with their passionate values.

To bring this somewhat full circle and return to the bigger picture of the simultaneous desire for positive alpha and low tracking error coupled with a desire for a personalized portfolio that reflects one's nonpecuniary preferences, **Exhibit 10.7** presents the actual alpha-tracking error frontier (rather than the stylized alpha-tracking frontier presented in Exhibit 10.1). Notice that only investor 1, who does not have any nonpecuniary preferences, ends up with a portfolio on the frontier. The rest of the investors are willing to give up some level of alpha to have a personalized portfolio that reflects their nonpecuniary preferences.

Exhibit 10.7. Alpha-Tracking Error Implementation Frontier



Conclusion and Key Takeaways

In this chapter we moved beyond MVO by expanding the alpha-tracking error optimization to also include a nonpecuniary preference term in the investor's maximization problem. Such a process enables advisers to construct personalized investment-specific portfolios that include all of the pecuniary inputs inherent in the alpha-tracking error optimization while simultaneously considering the investor's nonpecuniary preferences.

This type of "tilting" approach, in which the investor's maximization problem simultaneously considers the benefit of higher expected returns versus tracking error along with the benefit of nonpecuniary investment characteristics on equal terms, allows the optimizer to find the optimal personalized solution for a given investor. This process not only is more elegant than exclusionary-based approaches but also leads to solutions that dominate exclusionary-based approaches.

In the spirt of a controlled experiment, we demonstrated how investors who agree on all aspects of the problem from a pecuniary perspective may, and typically do, arrive at different portfolios as a result of their nonpecuniary preferences.

Appendix 10A. Derivation of Alpha-Tracking Error Objective Function

Let us restate Equation 10.3 using a single risk aversion coefficient, λ:

$$Obj = \mathbf{h}_M^T \alpha + \mathbf{h}_M^T \mathbf{X} \mu_K + \mathbf{h}_M^T \mathbf{C} \phi - \frac{\lambda}{2} [\mathbf{h}_M^T \mathbf{X} \mathbf{V}_K \mathbf{X}^T \mathbf{h}_M + \mathbf{h}_M^T \mathbf{V}_\alpha \mathbf{h}_M]. \quad \text{(A10.1)}$$

The first variance term can be decomposed into three terms:

- misfit risk, which is the variance of the difference between the effective asset mix and the target asset allocation;
- twice the covariance of the difference between the effective asset mix and the target asset allocation and the target asset allocation; and
- the variance of the target asset allocation.

Hence:

$$Obj = \mathbf{h}_M^T \alpha + \mathbf{h}_M^T \mathbf{X} \mu_K + \mathbf{h}_M^T \mathbf{C} \phi - \frac{\lambda}{2} \begin{bmatrix} (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K)^T \mathbf{V}_K (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K) \\ + 2(\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K)^T \mathbf{V}_K \mathbf{h}_K + \mathbf{h}_K^T \mathbf{V}_K \mathbf{h}_K + \mathbf{h}_M^T \mathbf{V}_\alpha \mathbf{h}_M \end{bmatrix}. \quad \text{(A10.2)}$$

Suppose that the target asset allocation is the solution to the following:[91]

$$\max_{\mathbf{h}_K} U_A = \mathbf{h}_K^T \mu_K - \frac{\lambda}{2} \mathbf{h}_K^T \mathbf{V}_K \mathbf{h}_K. \quad \text{(A10.3)}$$

[91]This problem does not have any constraints because it assumes that a risk-free asset can be held long or short. The allocation to the risk-free asset is $1 - \sum_{i=1}^{K} h_{Ki}$.

The first-order condition for this problem is as follows:

$$\boldsymbol{\mu}_K = \lambda \mathbf{V}_K \mathbf{h}_K. \tag{A10.4}$$

Substituting $\boldsymbol{\mu}_K$ in Equation A10.2 for the right-hand side of Equation A10.4, we have the following:

$$Obj = \mathbf{h}_M^T \boldsymbol{\alpha} + \lambda \mathbf{h}_M^T \mathbf{X} \mathbf{V}_K \mathbf{h}_K + \mathbf{h}_M^T \mathbf{C} \boldsymbol{\phi} - \frac{\lambda}{2} \begin{bmatrix} (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K)^T \mathbf{V}_K (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K) \\ + 2(\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K)^T \mathbf{V}_K \mathbf{h}_K + \mathbf{h}_K^T \mathbf{V}_K \mathbf{h}_K + \mathbf{h}_M^T \mathbf{V}_\alpha \mathbf{h}_M \end{bmatrix}. \tag{A10.5}$$

Looking at the second term and the covariance term, we see that $\lambda \mathbf{h}_M^T \mathbf{X} \mathbf{V}_K \mathbf{h}_K$ cancels out, leaving the following:

$$Obj = \mathbf{h}_M^T \boldsymbol{\alpha} + \mathbf{h}_M^T \mathbf{C} \boldsymbol{\phi} - \frac{\lambda}{2} [(\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K)^T \mathbf{V}_K (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K) - \mathbf{h}_K^T \mathbf{V}_K \mathbf{h}_K + \mathbf{h}_M^T \mathbf{V}_\alpha \mathbf{h}_M]. \tag{A10.6}$$

We can treat the variance of the target asset mix as given, so we can drop it, leaving the utility function in alpha, nonpecuniary preferences, and tracking error:

$$Obj = \mathbf{h}_M^T \boldsymbol{\alpha} + \mathbf{h}_M^T \mathbf{C} \boldsymbol{\phi} - \frac{\lambda}{2} \left[(\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K)^T \mathbf{V}_K (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K) + \mathbf{h}_M^T \mathbf{V}_\alpha \mathbf{h}_M \right]. \tag{A10.7}$$

We can assign different levels of risk aversion to asset allocation risk and to residual risk:

$$Obj = \mathbf{h}_M^T \boldsymbol{\alpha} + \mathbf{h}_M^T \mathbf{C} \boldsymbol{\phi} - \frac{\lambda_a}{2} (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K)^T \mathbf{V}_K (\mathbf{X}^T \mathbf{h}_M - \mathbf{h}_K) - \frac{\lambda_m}{2} \mathbf{h}_M^T \mathbf{V}_\alpha \mathbf{h}_M. \tag{A10.8}$$

This is Equation 10.4.

11. PERSONALIZED MULTI-ACCOUNT OPTIMIZATION[92]

Context

The formation and linking of our parent, child, and grandchild models is nearly complete. Chapters 5 and 6 presented our parent life-cycle models. In chapter 7, we demonstrated how to estimate effective tax rates for asset classes to account for taxes. In chapter 8, we jointly solved a MVO asset allocation and location problem, creating account-type asset allocation targets. Based on a holistic individual balance sheet, we then extended the MVO framework to take human capital and liabilities into account when solving the joint asset location and allocation problem as a child model under a parent life-cycle model. In chapters 9 and 10, respectively, we extended MVO and alpha-tracking error optimization (grandchild model) to include a term for the investor's nonpecuniary preferences, albeit in a single account setting without taxes. In this chapter, we unite all of that work into our grandchild model: a single comprehensive optimization framework that implements personalization.

Key Insights

1. We put forth a first-of-its-kind optimization procedure that simultaneously optimizes across different accounts and account types with different tax treatments.
2. In addition to the benefits of standard alpha-tracking error optimization, solving this expanded maximization problem has unexpected benefits, allowing it to serve as a(n):
    - nonpecuniary preference optimizer
    - asset *location* optimizer
    - rollover/reverse rollover optimizer
    - smart transition management optimizer
    - smart rebalancing optimizer
    - tax loss harvesting optimizer
    - new money deployment optimizer
    - withdrawal optimizer

For wealth advisers and planners, an important decision is how much money to invest in each of the available investments across multiple accounts. Key challenges include when to replace an existing fund, when to move money from one fund to another, when to use passive versus active funds to maximize after-fee expected alpha, when to roll over money from one account to another (e.g., 401(k) plan to IRA), and when to rebalance. A taxable account has additional challenges, such as where to locate different funds in different types of accounts (tax-exempt, tax-deferred, or taxable); how to minimize taxes; how to minimize trading costs; how and when to transition a new client away from current investments to more appropriate ones; and how to personalize a portfolio based on an investor's nonpecuniary preferences, such as a portfolio that tilts toward ESG-friendly firms. Techniques for answering these questions in a cohesive manner that emanate from a theoretically sound starting point are, until now, lacking. In this chapter, we build on elements from chapters 7, 8, 9, and 10 to propose a personalized multi-account, tax-efficient alpha-tracking

[92]The chapter is based on Idzorek (2022).

error optimization framework that simultaneously answers these questions. More specifically, we expand the single-account alpha-tracking error objective function from chapter 10 to a multi-account setting in which tax-aware, account type–specific inputs from chapter 7 are used and tax-aware account-type-specific asset allocation targets from chapter 8 are assumed. We also introduce two new terms into the objective function to minimize trading costs and to harvest tax losses.

We begin by restating Equation 10.4 as Equation 11.1:

$$Obj = \underbrace{\mathbf{h}_\mathbf{M}^\mathbf{T}\boldsymbol{\alpha}_\mathbf{M}}_{\text{Alpha}} + \underbrace{\mathbf{h}_\mathbf{M}^\mathbf{T}\mathbf{C}\boldsymbol{\phi}}_{\substack{\text{Nonpecuniary}\\ \text{Benefit}}} \underbrace{-\frac{1}{2}\lambda_a \underbrace{[\mathbf{X}^\mathbf{T}\mathbf{h}_\mathbf{M} - \mathbf{h}_\mathbf{K}]^T \mathbf{V}_\mathbf{K} [\mathbf{X}^\mathbf{T}\mathbf{h}_\mathbf{M} - \mathbf{h}_\mathbf{K}]}_{\substack{\text{Asset Allocation}\\ \text{Misfit Risk}}} - \frac{1}{2}\lambda_m \underbrace{\mathbf{h}_\mathbf{M}^\mathbf{T}\mathbf{V}_\alpha\mathbf{h}_\mathbf{M}}_{\substack{\text{Residual}\\ \text{Variance}}}}_{\text{Penalty for Active Risk}}, \tag{11.1}$$

where:

Obj = the objective function (what the investor seeks to maximize);

$\mathbf{h}_\mathbf{M}$ = investor's holdings (weights) to each fund/manager ($M \times 1$ column vector);

$\boldsymbol{\alpha}_\mathbf{M}$ = alpha (expected excess return) of the fund/managers ($M \times 1$ column vector);

$\mathbf{C}$ = exposure of investments to nonpecuniary characteristics ($M \times P$ matrix);

$\boldsymbol{\phi}$ = investor's nonpecuniary preferences ($P \times 1$ column vector);

λ_a = risk aversion parameter on asset allocation (misfit) risk (1×1 scalar);

$\mathbf{X}$ = exposure of investments to asset classes or factors ($M \times K$ matrix);

$\mathbf{h}_\mathbf{K}$ = target asset allocation ($K \times 1$ column vector);

$\mathbf{V}_\mathbf{K}$ = variance-covariance matrix of the asset classes ($K \times K$ matrix);

λ_m = risk aversion coefficient on fund-/manager-specific risk (1×1 scalar);

$\mathbf{V}_\alpha$ = covariance matrix of fund-/manager-specific risk ($M \times M$ matrix); and

M = the number of investments.

Inherent in this setup is the assumption that the optimization is taking place within a single account (with no trading costs) or, if it is occurring across multiple accounts, that all money is completely fungible with no taxes or trading costs and thus the distinction between accounts can be ignored. In the world of individual investors, money is not completely fungible across accounts or account types given various tax rules and penalties, and there are trading costs. These considerations motivate us to build a new model to explicitly consider these complicating and economically meaningful trade-offs.

Multiple Accounts and Account Types[93]

Multiple accounts with different tax treatments and available investments result in a complicated web of choices with real-world tax implications. Successfully navigating this web leads to lower taxes. The relative quality and costs of available investments may vary across different accounts. Moreover, the ability to look across accounts while considering tax efficiency allows one to select the best funds. Similarly, for investors who want additional personalization based on their nonpecuniary preferences, the ability to contemplate a wider range of available investments from across their different accounts may allow for greater personalization while concurrently considering investment quality, costs, and tax-efficiency.

[93]Adapted from Idzorek (2023).

In the United States, the following opportunities for tax-aware, multi-account portfolio management can add significantly better after-tax performance:

- various evolving tax rules related to different account types (tax-exempt, tax-deferred, and taxable),
- differences in the taxation of short-term versus long-term capital gains,
- different tax rates for qualified versus nonqualified dividends,
- different tax rates for income versus capital appreciation, and
- the ability to offset taxable gains with taxable losses.

Additionally, for many investors, the prospect of a lower federal income tax rate in retirement creates the opportunity to add value through tax-aware investment decisions.

The multi-account optimization framework that we propose creates better after-tax performance in four distinct manners:

- through separate asset allocation policy portfolios that are specific to the account-types;
- through separate fund-specific input estimates that account for fund-specific tax efficiency depending on the account-type (tax-exempt, tax-deferred, or taxable) in which it is held,
- through tax loss harvesting, and
- by only selling any of the investor's current holdings if it is in the economic interest of the investor; thus, the optimizer serves as a new type of tax-efficient transition management optimizer.

As we describe in this chapter, fund-of-funds optimization has a single set of change variables—the weights to the possible managers ($\mathbf{h}_M$). The reality for many investors is that their wealth is spread across multiple accounts, often with different available investment options in different accounts, different tax treatments, and different trading costs. In many cases, the money in the different accounts cannot easily be moved between accounts or account types. To make the problem more tractable, we make several assumptions:[94]

- All investment options are available in fractional shares.
- Money is transferrable only between equivalent account types (taxable to taxable, tax-deferred to tax-deferred, and tax-exempt to tax-exempt).
- For each account, a separate process filters the available investments to a manageable number of best options and makes only that subset available for allocation.

Target Asset Allocation Policy Benchmark(s)

To create separate target asset location and allocation asset class targets, we leverage the model we present in chapters 7 and 8. Just as it is with fund-of-funds optimization *without* taxes, with multiple accounts *with* taxes, there is a target asset allocation benchmark; however, when we have multiple accounts, we can specify separate account-*type*-specific target asset allocation policy benchmarks. One can choose to use a single target asset allocation for all account types, in which tax location is driven primarily through the trade-offs represented by different account type inputs. *Or*, preferably, one can use different account-type-specific target policy asset allocations, in which tax location is driven by *both* account-type-specific inputs and account-type-specific asset allocation targets. The same fund should have two different after-fee sets of inputs: (1) a set for when the fund is held in a tax-advantaged account (tax-exempt or tax-deferred account) and (2) a set for when the fund is held in a taxable account.

---

[94]We can relax most of these assumptions or impose stricter constraints, depending on the realities of a given jurisdiction or situation.

In chapter 8, we assumed two account types: taxable and tax-advantaged. Now we assume the investor's wealth is spread across multiple account types and specify three different vectors of asset allocation targets: one for tax-exempt accounts $\mathbf{h}_{K,E}$, one for tax-deferred accounts $\mathbf{h}_{K,D}$, and one for the taxable accounts $\mathbf{h}_{K,T}$, all of which are scaled by total *tax-adjusted* wealth. That is, the elements of $\mathbf{h}_{K,E}$ must sum to the fraction of the overall portfolio (on an after-tax basis) in tax-exempt assets; the elements of $\mathbf{h}_{K,D}$ must sum to the fraction of the overall portfolio (on an after-tax basis) in tax-deferred assets; and the elements of $\mathbf{h}_{K,T}$ must sum to the fraction of assets (on an after-tax basis) in taxable assets. To state all holdings as a fraction of the overall portfolio on an after-tax basis, the pretax amount in the *tax-deferred* account must be multiplied by one minus the income tax rate projected to be in effect at the time of withdrawal. This step converts the pretax amounts to after-tax dollars and thus makes them comparable to the dollars in the tax-exempt and taxable accounts.

In the absence of taxes, the strategic asset allocation should be the asset allocation that maximizes the objective function presented in Equation 10.1 or that emerges from the asset-only version of the model presented in chapter 8 without a taxable account. Similarly, when there are tax-exempt, tax-deferred, and taxable accounts, the strategic asset allocation and location should be the set of allocation vectors ($\mathbf{h}_{K,E}$, $\mathbf{h}_{K,D}$, and $\mathbf{h}_{K,T}$) that maximizes after-tax risk-adjusted expected return in a multi-account setting, as follows:

$$Obj_{SAA} = \underbrace{(\mathbf{h}_{K,E} + \mathbf{h}_{K,D} + \mathbf{T}_K \mathbf{h}_{K,T})^T \boldsymbol{\mu}_K}_{\text{Asset Allocation Expected Return}} - \underbrace{\frac{\lambda}{2} \underbrace{(\mathbf{h}_{K,E} + \mathbf{h}_{K,D} + \mathbf{T}_K \mathbf{h}_{K,T})^T \mathbf{V}_K (\mathbf{h}_{K,E} + \mathbf{h}_{K,D} + \mathbf{T}_K \mathbf{h}_{K,T})}_{\text{Asset Class Variance}}}_{\text{Penalty for Risk}}. \quad (11.2)$$

As in chapter 8, $\mathbf{T}_K$ is a diagonal matrix with one minus the *effective tax rate* of *each asset class* along the diagonal (we will further discuss effective tax rates). For those who are interested in simultaneously solving for *asset allocation* and *manager structure*, in appendix 11A, we develop a generalized, fully functional, multi-account, tax-aware objective function for simultaneously solving for both. In appendix 11B, we extend that to include a liability.

Taxes and Return Generation Process

In chapter 7, we discussed how to summarize the impact of taxes on *asset class* returns with effective tax rates. We extend that discussion to include the effective tax rates for *specific investments* (funds or managers). As before, prior to accounting for the impact of taxes on investment returns, we first need to consider the underlying tax-free return generation process for a fund. In Equation 11.1, the base case alpha-tracking error objective function is based on the following return generation process:

$$\tilde{R}_j = \alpha_j + \mathbf{x}_j^T \tilde{\mathbf{R}}_k + \tilde{u}_j, \quad (11.3)$$

where:

$\tilde{R}_j$ = the pretax *realized* total return on fund j;

α_j = the alpha of fund j (jth element of $\boldsymbol{\alpha}_M$);

$\mathbf{X}_j$ = K-element vector of the asset class exposures of fund j;

$\tilde{\mathbf{R}}_k$ = the vector of *realized* asset class returns; and

$\tilde{u}_j$ = the realized residual with standard deviation ω_j.

This tax-free return generation process is the basis for our starting base case *tax-exempt* parameters. Using the additional subscript "$_E$" for exempt, we have the following:

$\boldsymbol{\alpha}_{\mathbf{M,E}}$ = alpha (expected excess return) of the funds/managers when held in a *tax-exempt* account ($M_E \times 1$ column vector);

$\mathbf{X_E}$ = exposure matrix identifying the asset class exposures of each fund/manager when held in a *tax-exempt* account ($M_E \times K$ matrix); and

$\mathbf{V}_{\boldsymbol{\alpha},\mathbf{E}}$ = covariance matrix of fund-/manager-specific risk when held in a *tax-exempt* account (excess returns) ($M_E \times M_E$ matrix).

Moving to a tax-deferred account setting, the return generation process remains the same. When the money is withdrawn from a tax-deferred account, however, it will be taxed based on the individual's marginal income tax rate at the time of the withdrawal. The parameters for the tax-deferred account-type are the same as those for the tax-exempt account-type, although we track them separately should changes in tax rates lead to a difference in the future. Using the additional subscript "$_\mathbf{D}$" for deferred, we have the following:

$\boldsymbol{\alpha}_{\mathbf{M,D}}$ = alpha (expected excess return) of the funds/managers when held in a *tax-deferred* account ($M_D \times 1$ column vector);

$\mathbf{X_D}$ = exposure matrix identifying the asset class exposures of each fund/manager when held in a *tax-deferred* account ($M_D \times K$ matrix); and

$\mathbf{V}_{\boldsymbol{\alpha},\mathbf{D}}$ = covariance matrix of fund-/manager-specific risk when held in a *tax-deferred* account (excess returns) ($M_D \times M_D$ matrix).

To move from a *tax-free* return generation process to an *after-tax* return generation process associated with investments held in a *taxable* account, one needs to know (or estimate) the effective tax rate of each *investment* in question. Thus, let τ_j^E be the effective tax rate of fund j so that the portion of the return that is left after taxes is $1-\tau_j^E$. For after-tax returns, we have the following:

$$(1-\tau_j^E)\tilde{R}_j = (1-\tau_j^E)\alpha_j + (1-\tau_j^E)\mathbf{x}_\mathbf{j}^\mathbf{T}\tilde{\mathbf{R}}_\mathbf{k} + (1-\tau_j^E)\tilde{u}_j. \qquad (11.4)$$

Based on Equations 11.3 and 11.4, we can start with tax-free fund parameters and calculate investor-specific and fund-specific after-tax parameters for the funds when they are available in a taxable account.

Previously, we defined $\mathbf{T_K}$ as the diagonal matrix with one minus the *effective tax rate* of *each asset class* along the diagonal. We now define $\mathbf{T_M}$ as the diagonal matrix with one minus the *effective tax rate* of *each manager (or fund)* along the diagonal. Multiplying each of the three fund-specific parameters by $\mathbf{T_M}$ in Equations 11.5, 11.6, and 11.7 moves us from pretax parameters to after-tax or taxable account parameters:

$$\boldsymbol{\alpha}_{\mathbf{M,T}} = \mathbf{T_M}\boldsymbol{\alpha}_{\mathbf{M,E}}, \qquad (11.5)$$

$$\mathbf{X_T} = \mathbf{T_M X_E}, \qquad (11.6)$$

$$\mathbf{V}_{\boldsymbol{\alpha},\mathbf{T}} = \mathbf{T_M V}_{\boldsymbol{\alpha},\mathbf{E}}\mathbf{T_M}, \qquad (11.7)$$

where:

$\boldsymbol{\alpha}_{\mathbf{M,T}}$ = expected alpha (excess return) of the funds/managers when held in a *taxable* account ($M_T \times 1$ column vector);

$\mathbf{X_T}$ = exposure matrix identifying the asset class exposures of each fund/manager when held in a *taxable* account ($M_T \times K$ matrix); and

$\mathbf{V}_{\alpha,T}$ = covariance matrix of fund-/manager-specific risk when held in a *taxable* account (excess returns) ($M_T \times M_T$ matrix).

The number of available investment options in each account and account type will likely differ as will the degree of overlap, with some investment options likely appearing in multiple accounts.

We can stack the tax-exempt, tax-deferred, and taxable parameters:

$$\hat{\boldsymbol{\alpha}}_{\mathbf{M}} = \begin{bmatrix} \boldsymbol{\alpha}_{\mathbf{M,E}} \\ \boldsymbol{\alpha}_{\mathbf{M,D}} \\ \boldsymbol{\alpha}_{\mathbf{M,T}} \end{bmatrix}, \tag{11.8}$$

$$\hat{\mathbf{X}} = \begin{bmatrix} \mathbf{X_E} & 0 & 0 \\ 0 & \mathbf{X_D} & 0 \\ 0 & 0 & \mathbf{X_T} \end{bmatrix}, \tag{11.9}$$

$$\hat{\mathbf{V}}_{\alpha} = \begin{bmatrix} \mathbf{V}_{\alpha,\mathbf{E}} & \mathbf{V}_{\alpha,\mathbf{E,D}}^{\mathbf{T}} & \mathbf{V}_{\alpha,\mathbf{E,T}}^{\mathbf{T}} \\ \mathbf{V}_{\alpha,\mathbf{E,D}} & \mathbf{V}_{\alpha,\mathbf{D}} & \mathbf{V}_{\alpha,\mathbf{D,T}}^{\mathbf{T}} \\ \mathbf{V}_{\alpha,\mathbf{E,T}} & \mathbf{V}_{\alpha,\mathbf{D,T}} & \mathbf{V}_{\alpha,\mathbf{T}} \end{bmatrix}, \tag{11.10}$$

where within $\hat{\mathbf{V}}_{\alpha}$ the off-diagonal submatrices ($\mathbf{V}_{\alpha,\mathbf{E,D}}$, $\mathbf{V}_{\alpha,\mathbf{E,T}}$, and $\mathbf{V}_{\alpha,\mathbf{D,T}}$) consist of all zeros unless the same fund appears in different accounts. If the same fund is in the tax-exempt and tax-deferred account the element is ω_j^2. If the same fund is in either the tax-exempt account or tax-deferred account as well as the taxable account, the element is $(1-\tau_j^E)\omega_j^2$, where τ_j^E is the effective tax-rate of the *fund* in question.

We stack the three account-type-specific asset allocation targets in the same manner:

$$\hat{\mathbf{h}}_{\mathbf{K}} = \begin{bmatrix} \mathbf{h}_{\mathbf{K,E}} \\ \mathbf{h}_{\mathbf{K,D}} \\ \mathbf{T_K h_{K,T}} \end{bmatrix}. \tag{11.11}$$

We also stack the corresponding holdings vectors:

$$\hat{\mathbf{h}}_{\mathbf{M}} = \begin{bmatrix} \mathbf{h}_{\mathbf{M,E}} \\ \mathbf{h}_{\mathbf{M,D}} \\ \mathbf{h}_{\mathbf{M,T}} \end{bmatrix}. \tag{11.12}$$

A Multi-Account Objective Function

Equation 11.13 is an expanded version of the single account personalized fund-of-funds optimization objective function (Equation 11.1), in which we have substituted our "stacked" parameters that allow for multiple account-type-specific target asset allocations that incorporate asset location:

$$Obj = \hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\boldsymbol{\alpha}}_{\mathbf{M}} + \hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\mathbf{C}\boldsymbol{\phi} - \frac{1}{2}\lambda_a\,[\hat{\mathbf{X}}^{\mathbf{T}}\hat{\mathbf{h}}_{\mathbf{M}} - \hat{\mathbf{h}}_{\mathbf{K}}]^{\mathbf{T}}\hat{\mathbf{V}}_{\mathbf{K}}[\hat{\mathbf{X}}^{\mathbf{T}}\hat{\mathbf{h}}_{\mathbf{M}} - \hat{\mathbf{h}}_{\mathbf{K}}] - \frac{1}{2}\lambda_m\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{V}}_{\alpha}\hat{\mathbf{h}}_{\mathbf{M}}. \tag{11.13}$$

The only undefined parameter at this point is $\hat{\mathbf{V}}_K$. In this setup, recalling that $\hat{\mathbf{X}}$ has already incorporated the effective tax rate of the different managers held in taxable accounts when arriving at the effective asset allocation of the taxable accounts, the covariance matrix of asset class returns $\mathbf{V}_K$ does not need further adjustment. However, given all of the expanded or stacked vectors, we need to expand the dimensions of the covariance matrix of asset class returns to correspond to the dimension of $\hat{\mathbf{h}}_K$. We experimented with and used two potential specifications:

$$\hat{\mathbf{V}}_K = \begin{bmatrix} \mathbf{V}_K & \mathbf{V}_K & \mathbf{V}_K \\ \mathbf{V}_K & \mathbf{V}_K & \mathbf{V}_K \\ \mathbf{V}_K & \mathbf{V}_K & \mathbf{V}_K \end{bmatrix}, \tag{11.14a}$$

$$\hat{\mathbf{V}}_K = \begin{bmatrix} \mathbf{V}_K & 0 & 0 \\ 0 & \mathbf{V}_K & 0 \\ 0 & 0 & \mathbf{V}_K \end{bmatrix}. \tag{11.14b}$$

In Equation 11.14b, the 0s represent submatrices of the stacked-offset-covariance-matrix $\hat{\mathbf{V}}_K$ in which all elements are 0.

Equation 11.14a leads to the *correct* estimate of asset allocation misfit risk. All else equal, when used in Equation 11.13, it leads to a manager structure with the lowest overall asset allocation misfit risk; however, it also can lead to significant asset allocation misfit risk within each account type, which is not optimal when $\mathbf{h}_{K,E}$, $\mathbf{h}_{K,D}$, and $\mathbf{h}_{K,T}$ reflect tax-efficient asset location targets.

Equation 11.14b leads to an *incorrect* estimate of asset allocation misfit risk (understates total asset allocation misfit risk). All else equal, when used in Equation 11.13, Equation 11.14b leads to a manager structure with a somewhat higher overall asset allocation misfit risk; however, it typically leads to low amounts of asset allocation misfit risk within each account type (which is beneficial from a tax-efficiency and asset location perspective).

Because minimizing account-type-specific asset allocation misfit risk results in greater tax efficiency, in practice, we frequently use Equation 11.14b when solving for the optimal manager structure and use Equation 11.14a to provide the correct estimate of overall asset allocation misfit risk.

In the following sections, we build out the inputs and gradually add to the objective function of Equation 11.13, thus moving toward a more comprehensive and powerful objective function for personalized portfolio construction.

Account-Type and Investor-Specific After-Tax Parameters

As Equation 11.4 shows, a fund's effective tax rate translates a tax-free expected total return into an after-tax expected total return. Kaplan (2020c) and chapter 7 demonstrate how an effective tax rate is calculated for an *asset class* from the following data:

- the investor's marginal tax rate on ordinary income;
- the investor's marginal tax rate on long-term capital gains;
- the division of total return between income and capital gains;

- the division of income between the part taxed at the ordinary income rate and the part taxed at the long-term capital gains rate (qualified dividends in the United States);
- the division of capital gains between the part that is taxed at the ordinary income rate (short-term in the United States) and the part that is taxed at the long-term capital gains rate; and
- the rate of turnover of the investment.

We use the same approach to develop effective tax rates for individual *funds*, with some modifications to account for the unique fund-specific features.

Marginal tax rates should be investor specific. Hence, the entire exercise of asset location at the asset class level and after-tax fund-of-funds optimization should be *personalized*. That is, a prebuilt set of taxable asset allocation models or a predetermined list of funds using generic or average tax rates are likely to be suboptimal relative to an optimized solution using an individual's tax rates.

Following Kaplan (2020c) and chapter 7, we estimate expected total returns and expected income returns for a set of asset class indices. We use these values to calculate an expected total return and expected income return of the *asset class* indices. Because a fund can have exposures to multiple asset classes, we use the asset-class-weighted averages of expected total return and expected income return to come up with the division of total return between income return and capital gains for a given fund.

Using data from Morningstar Direct, we also calculate the division between *qualified* and *nonqualified* income and the division between *long-term* and *short-term* capital gains for each Morningstar fund category. We do this for all active funds and passive funds. Because each asset class index can be linked to a Morningstar Category and each fund is assigned a Morningstar Category, we use these data to calculate both asset class and fund parameters. In the same manner, we calculate rates of *turnover* for each Morningstar Category from the turnover rates of the funds. We use these turnover rates to calculate turnover for each asset class. For each fund, we use the *fund-specific turnover*.

In the stacked vector of weights of the current and possible managers $\hat{\mathbf{h}}_{\mathbf{M}}$, the same manager may appear multiple times. In the corresponding stacked vector of expected alphas $\hat{\boldsymbol{\alpha}}_{\mathbf{M}}$, and the other fund inputs, we must simply use the appropriate set of inputs for a given account type.[95] Each set of parameters captures our assessment of pretax excess return generation. If the fund is held in a taxable account, the parameters also simultaneously capture the degree to which the fund is or is not tax-efficient as well as the inherent tax-efficiency of the asset classes in which the fund invests. Account-type inputs are one of the ways that we can simultaneously account for manager skill and tax-efficiency and tax location within a single, personalized optimization.

Account Level and Account-Type Budget Constraints

Account level and account-type constraints are a critical part of the process. Thus far, we have focused on the stacked inputs based on account type. These can and should be further broken down for each account. In fact, we expand each account-specific vector to include separate entries for different tax lots—for example, the "$_{L1}$" and "$_{L2}$" subscript modifiers in the final column of Equation 11.15:[96]

[95]A variety of potential methods can be used for coming up with starting after-fee, pretax expected alphas. Without going into the details, our approach follows Waring and Ramkumar (2008) and is parameterized based on information from Morningstar's analyst ratings. This approach requires an estimate of skill for each manager as well as other variables.

[96]A tax lot is a set of shares (of a company or fund) all of which have the same cost basis and purchase date.

$$\hat{\mathbf{h}}_{\mathbf{M}} = \begin{bmatrix} \mathbf{h}_{\mathbf{M,E}} \\ \mathbf{h}_{\mathbf{M,D}} \\ \mathbf{h}_{\mathbf{M,T}} \end{bmatrix} = \begin{bmatrix} h_{m,E,1} \\ \vdots \\ h_{m,E,?} \\ h_{m,D,1} \\ \vdots \\ h_{m,D,?} \\ h_{m,T,1,L1} \\ h_{m,T,1,L2} \\ \vdots \\ h_{m,T,?,L1} \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} \text{Individual Fund Allocations in Tax-Exempt} \\ \left.\vphantom{\begin{matrix} a \\ b \\ c \\ d \end{matrix}}\right\} \text{Individual Fund Allocations in Tax- Deferred} \\ \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} \text{Individual Fund Allocations in Taxable with } \textit{Tax Lots}. \end{matrix} \quad (11.15)$$

Accounting for different individual accounts enables us to apply account-level budget constraints, account-type budget constraints, or combinations of the two, all based on the degree to which money is or is not fungible. Accounting for different tax lots enables optimal tax loss harvesting and allows tax lots with the smallest gains to be sold first.

To help make the application of account-level and account-type constraints more concrete, as we build up this framework, we simultaneously build out an example to illustrate various points. Because of space considerations, we do not incorporate nonpecuniary preferences into the example and refer readers to Idzorek and Kaplan (2022) and chapter 10.

In **Exhibit 11.1**, we present separate asset allocation targets for the three account types: $\mathbf{h}_{\mathbf{K,E}}$, $\mathbf{h}_{\mathbf{K,D}}$, and $\mathbf{h}_{\mathbf{K,T}}$ along with the estimated effective tax rate of each asset class (it is one minus these rates that form the diagonal of $\mathbf{T}_{\mathbf{K}}$). These separate asset allocation targets are inputs into our overall process and ideally would be the output of maximizing Equation 11.2.

In this example, the investor in question has a tax-deferred account balance of \$287,500 across two accounts, which has been multiplied by 1–20% = 0.8, where 20% is the assumed individual tax rate at the time of the withdrawals. After this restatement, the investor currently has a total of \$500,000 (\$70,000 in a tax-exempt account, \$230,000 across two tax-deferred accounts after the restatement, and \$200,000 in a taxable account). The total restated value of \$500,000 (not \$557,500) serves as the denominator for converting dollars into percentages. If one were maximizing the objective function in Equation 11.2, these balances would form the basis of account-type constraints.

In this example, we have chosen to work with a granular asset allocation specification to help highlight the personalized effective tax rate of different asset classes and show how that relates to asset location. In the effective tax rate column of Exhibit 11.1, the highest tax rates are colored red and the lowest tax rates are colored green. In the next two columns showing expected returns, the color scheme is the opposite—the highest expected returns are in green and lowest expected returns are in red. Notice that the most tax-efficient asset classes have significant weights in the taxable target, whereas the least tax-efficient asset classes have significant weights in the two qualified accounts. The asset classes with the highest expected returns are located in the tax-exempt target. The final column displays the overall or summed asset allocation, which happens to correspond to a prototypical 60% equity and 40% fixed-income asset allocation.

Exhibit 11.2, which we will refer to often, contains a great deal of important information. Notice that Account 1 is a qualified tax-exempt Roth IRA, Account 2 is a qualified (tax-deferred) 401(k), Account 3 is a qualified (tax-deferred) IRA, and Account 4 is a taxable account. Given that Accounts 2 and 3 are really the same account type—that is, a tax-deferred account—money in these two accounts could be fungible. We return to this shortly.

Exhibit 11.1. Account-Type Asset Allocation Targets

Asset Class	Personalized Effective Tax Rate	Unconditional Expected Pretax Total Return	Unconditional Expected Post-Tax Total Return	Asset Allocation Targets			Total Target
				Tax-Exempt $h_{K,E}$	TaxDeferred $h_{K,D}$	Taxable $h_{K,T}$	
US Large Growth	13.7%	9.5%	8.2%	0.0%	0.0%	9.8%	9.8%
US Large Value	13.9%	9.5%	8.2%	0.0%	0.0%	10.0%	10.0%
US Mid Growth	14.2%	10.4%	9.0%	0.0%	0.0%	4.2%	4.2%
US Mid Value	13.9%	10.4%	8.9%	1.0%	0.0%	3.3%	4.3%
US Small Growth	14.4%	11.8%	10.1%	2.9%	0.0%	0.0%	2.9%
US Small Value	14.1%	11.8%	10.2%	3.0%	0.0%	0.0%	3.0%
REITs	25.3%	10.2%	7.6%	2.6%	0.0%	0.0%	2.6%
Non-US Developed Equity	14.2%	9.6%	8.2%	0.0%	6.3%	8.5%	14.8%
Emerging Market Equity	15.1%	10.2%	8.7%	2.2%	3.5%	0.0%	5.7%
Commodities	24.5%	4.8%	3.6%	0.0%	2.6%	0.0%	2.6%
US Short-Term Bonds	32.2%	3.7%	2.5%	0.0%	1.6%	0.0%	1.6%
US Intermediate-Term Bonds	32.4%	3.9%	2.6%	0.0%	7.7%	0.0%	7.7%
US Long-Term Bonds	34.8%	5.1%	3.3%	0.0%	17.4%	0.0%	17.4%
Short-Term Inflation Linked Bonds	19.3%	3.7%	3.0%	0.0%	2.4%	0.0%	2.4%
Long-Term Inflation Linked Bonds	22.5%	4.8%	3.7%	0.0%	2.0%	0.0%	2.0%
High Yield	40.0%	6.9%	4.2%	0.8%	2.3%	0.0%	3.1%
Non-US Bonds	28.9%	4.5%	3.2%	0.0%	5.0%	0.0%	5.0%
Emerging Market Bonds	35.5%	7.2%	4.7%	0.0%	0.0%	0.0%	0.0%
Cash	35.0%	2.3%	1.5%	0.0%	0.9%	0.0%	0.9%
			Sum	12.6%	51.6%	35.9%	
			Weighted Post-Tax Return	10.8%	5.7%	8.4%	

Exhibit 11.2. Sample Case: Pre-Optimization

(1) Accounts/Managers	(2) Alpha (Qualified) α_M	(3) Alpha (Taxable) α_M	(4) Specific Tracking Error	(5) Current	(6) Cost Basis	(7) Days Held	(8) Cost per Trade	(9) Manager Holdings $h_{M,t=-1}$	(10) Comments on Investment Options
Account 1. Tax-Exempt									
Employer Stock (Large Cap)	0.00%	N/A	19.6%	$0	N/A	N/A	$20	0.0%	High specific risk.
Passive US Equity ETF	–0.03%	N/A	1.1%	$20,000	N/A	N/A	$20	4.0%	Not well aligned with tax-exempt target.
Active Global Equity	0.20%	N/A	1.3%	$50,000	N/A	N/A	$20	10.0%	Good alpha; aligned with target.
Emerging Market Equity	–0.48%	N/A	2.7%	$0	N/A	N/A	$20	0.0%	Poor alpha; aligned with target.
High Alpha Active Bond Fund	0.33%	N/A	0.7%	$0	N/A	N/A	$20	0.0%	Excellent alpha; not aligned with target.
Passive Bond ETF	–0.04%	N/A	0.4%	$0	N/A	N/A	$20	0.0%	Not well aligned with tax-exempt target.
Weighted Subtotal:	0.13%		**Subtotal:**	$70,000					
Account 2. Tax-Deferred 401(K)									
Employer Stock (Large Cap)	0.00%	N/A	19.6%	$75,000	N/A	N/A	$0	15.0%	High specific risk; no tax consequence to sell.
High Cost US Equity Fund	–1.05%	N/A	2.4%	$25,000	N/A	N/A	$0	5.0%	Poor alpha; not aligned with target.
High Cost Non-US Equity Fund	–1.28%	N/A	3.0%	$0	N/A	N/A	$0	0.0%	Poor alpha; not aligned with target.
High Cost Bond Fund	–1.63%	N/A	0.7%	$0	N/A	N/A	$0	0.0%	Poor alpha; aligned with target.
Active Bond Fund	–0.01%	N/A	0.5%	$10,000	N/A	N/A	$0	2.0%	Reasonable alpha; aligned with target.
Weighted Subtotal:	–0.24%		**Subtotal:**	$110,000					

(continued)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Accounts/Managers	Alpha (Qualified) α_M	Alpha (Taxable) α_M	Specific Tracking Error	Current	Cost Basis	Days Held	Cost per Trade	Manager Holdings $h_{M,t=-1}$	Comments on Investment Options
Account 3. Tax-Deferred IRA									
Employer Stock (Large Cap)	0.00%	N/A	19.56%	$0	N/A	N/A	$10	0.0%	High specific risk; no tax consequence to sell.
Passive US Equity ETF	–0.03%	N/A	1.07%	$0	N/A	N/A	$10	0.0%	Reasonable alpha; not aligned with target.
Passive Global Equity ETF	–0.08%	N/A	0.85%	$40,000	N/A	N/A	$10	8.0%	Reasonable alpha; not aligned with target.
High Alpha Bond Fund	0.33%	N/A	0.72%	$80,000	N/A	N/A	$10	16.0%	Excellent alpha; aligned with target.
Passive Bond ETF	–0.04%	N/A	0.39%	$0	N/A	N/A	$10	0.0%	Reasonable alpha; aligned with target.
Weighted Subtotal:	0.19%		**Subtotal:**	$120,000					
Account 4. Taxable									
Employer Stock (Large Cap)—Lot 1	N/A	0.00%	17.86%	$30,000	$25,000	480	$25	6.0%	High specific risk; significant tax consequence to sell.
Employer Stock (Large Cap)—Lot 2	N/A	0.00%	17.86%	$20,000	$2,500	2571	$25	4.0%	High specific risk; moderate tax consequence to sell.
Passive US Equity ETF	N/A	–0.03%	0.95%	$75,000	$70,000	1113	$25	15.0%	Reasonable alpha; moderate tax consequence to sell.
Passive Global Equity ETF	N/A	–0.07%	0.75%	$0	N/A	N/A	$25	0.0%	Reasonable alpha; aligned with target.
High Cost Bond Fund	N/A	–1.10%	0.49%	$75,000	$76,000	464	$25	15.0%	Poor alpha; poor alignment; harvestable tax loss.
Passive Bond ETF	N/A	–0.03%	0.26%	$0	N/A	N/A	$25	0.0%	Reasonable alpha; poor alignment.
Weighted Subtotal:		–0.42%	**Subtotal:**	$200,000					
			Total	$500,000					

For now, let us focus on Exhibit 11.2, columns 2 and 3, which contain the expected alphas. Notice that each fund has two possible alpha estimates: one for when it is held in a qualified account and one for when it is held in a taxable account. In this example, the tax-exempt Roth IRA, the tax-deferred IRA, and the taxable account have similar investment options, whereas the 401(k) investment options are somewhat different.

To keep the example in Exhibit 11.2 manageable, we have assumed a relatively small number of additional funds; in practice, however, the number of available funds will usually be much greater. The additional managers with 0% weights in the various accounts are available to be allocated to within a given account but currently are not receiving an allocation. We have assumed considerable overlap in the available funds across accounts, but that need not be the case.

Adjacent to Exhibit 11.2, column 9, we have placed blue brackets to represent the various account-level and account-type constraints. In this case, the investment-specific allocations in Account 1 must sum to 14% or $70,000. Accounts 2 and 3 are both tax-deferred accounts, and if money is fungible between the two accounts, we want to constrain their combined allocations to sum to 46% or $230,000. Corresponding to the adjustment made at the asset class level, the $230,000 in investment balances has already been reduced from a combined tax-deferred balance of $287,500 by making the 20% adjustment for taxes described earlier. This puts the investment balances in Exhibit 11.2 on an equivalent after-tax footing. If the money is not fungible or if it is fungible in only one direction, one would use constraints to reflect these conditions. The allocations in Account 4 must sum to 40% or $200,000. This example would result in three linear constraints:

- The sum of weights to managers in Account 1 must equal 14%.
- The sum of weights to managers in Accounts 2 and 3 must equal 46%.
- The sum of weights to managers in Account 4 must equal 40%.

If specified correctly, these constraints sum to 100%, thus making the overall budget constraint of 100% redundant. For clarity, relative to single-account alpha-tracking error optimization, the single budget constraint of 100% is replaced by a combination of account-level and account-type budget constraints.

As illustrated in Exhibit 11.2, in taxable accounts with different tax lots, we find it critical to include a separate entry for each tax lot, effectively treating the allocation to each lot as an individual decision variable.[97] Notice that in Account 4, the taxable account, we assume that the investor has two different tax lots of employer stock. The expected alphas and the investment-specific tracking errors are the same for both lots, but the current value of Lot 2 is significantly above its cost basis.

We could have numerous tax lots for multiple securities across multiple accounts. We return to taxes and tax-loss harvesting shortly. For now, in addition to illustrating how to incorporate different tax lots into the decision-making process, we want to illustrate that the fund-of-funds optimization framework can work with individual securities.[98] To the degree that there may be additional account types, perhaps with different tax lots and different types of investments, our hope is that this provides enough of a blueprint that the creators of systems for practitioners could expand the framework accordingly.[99]

---

[97]For the purpose of calculating fixed trading costs, we need to keep track of which holdings have been parsed into separate entries within the list of current/possible managers ($\mathbf{h}_M$). With fixed trading costs, there is only one cost regardless of the number of lots.

[98]Idzorek (2022) investigates the challenges of optimizing an opportunity set consisting of individual securities and funds.

[99]Another area for future research is to expand the framework put forth here to incorporate goal-specific asset allocation targets and corresponding goal-specific constraints.

Finally, in Exhibit 11.2, column 9, we include the additional subscript $_{t=-1}$ to denote that this is the current, *pre-optimization* weights or holdings ($\hat{\mathbf{h}}_{\mathbf{M},\mathbf{t}=-1}$). Later, when necessary, we use $_{t=0}$ to denote current *postoptimization* weights or holdings ($\hat{\mathbf{h}}_{\mathbf{M},\mathbf{t}=0}$).

A Rollover Optimization System

A somewhat unexpected benefit of this multi-account framework is its ability to serve as an ongoing rollover optimization system. With some effort, one can transfer money between equivalent accounts. Large retirement plans often have attractive funds with lower expense ratios than typical investment retirement accounts, while conversely, the investment options in typical IRAs can be less expensive than the investment options offered by smaller employer-sponsored retirement plans. If the employer-sponsored defined contribution plans allows it, one can roll money into the plan from an IRA, something referred to as a "reverse rollover."

In the example, at this point it is unclear if the investor should roll money from the 401(k) into the tax-deferred IRA, or vice versa. The optimal answer may even be a *partial* rollover.

Moving to nonpecuniary preferences, most 401(k) line-ups are relatively limited and often lack specialty funds (e.g., sector or ESG-oriented funds); thus, for investors with strong nonpecuniary preferences, a typical IRA rollover account platform may allow for a greater level of personalization. Simply focusing on the relative attractiveness, from a pecuniary or nonpecuniary perspective, of available funds in these two accounts in isolation is an oversimplification of the problem. To make an optimal decision, one must consider the complete picture, simultaneously contemplating things like trading costs, tax-efficient location, current tax consequences, and the characteristics of the available investments across all accounts. It is in this way, and through appropriate account constraints, that the multi-account fund-of-funds optimizer becomes an objective, *holistic* rollover optimization system.

Transaction Costs

All other things equal, transaction costs are "bad" and should be included in the alpha-tracking error optimization problem as a penalty. The original fund-of-funds optimizer from Waring et al. (2000) does not model transaction costs. Across different accounts and for different types of funds within a single account (mutual fund versus ETF versus separate account), transaction costs can vary significantly. Trading fees can be a fixed dollar amount (e.g., $25 per trade), variable (e.g., one-fifth of a basis point per amount traded), or a mixture of the two. If the monetary value of the accounts is small relative to the transaction costs, transaction costs can have a significant impact on the optimal investment allocations. As the monetary value of the accounts grows large relative to the transaction costs, transaction cost will have very little impact on the optimal investment allocations.[100]

Ideally, for all current and possible investments, the applicable fee structure will be known, and dynamically changing the allocation to a given manager in $\hat{\mathbf{h}}_{\mathbf{M}}$ will cause the associated total transaction cost to be dynamically calculated. Thus, as the optimizer adjusts the weights to the current and possible managers $\hat{\mathbf{h}}_{\mathbf{M}}$, the total transaction costs are simultaneously calculated and included in an expanded objective function as a penalty. To do this, one must know the current holdings, not only as weights but also as monetary values (e.g., pounds, euros, yuan).[101]

---

[100]Incorporating transaction costs has the advantage of limiting (or eliminating) trades that have little expected benefit and thus helps to avoid unnecessary turnover.

[101]We leave the complexity of multiple accounts across multiple countries and currencies to future research, although it would seem to be a direct extension of the methods developed here.

In this chapter, we choose to operate in "percent return" space rather than "monetary value" space, and so we translate monetary values/costs into percentages. This is done by dividing the monetary cost by the *total* value of all accounts, with an important caveat. Depending on one's use case, some practitioners may convert everything into monetary values.

Exhibit 11.3 builds on Exhibit 11.2, introducing the notion of a "new" or *postoptimization* set of holdings that can be compared to the "starting" holdings, enabling us to calculate the difference. When additional clarity is needed, we continue to add the subscripts "$_{t=-1}$" and "$_{t=0}$" to represent pre-optimization versus postoptimization holdings. For the continuing example, we assume fixed per-trade costs of $20, $0, $10, and $25 in Accounts 1, 2, 3, and 4, respectively. Dividing the per-trade costs by $500,000 translates them into percentage costs of 0.004%, 0%, 0.002%, and 0.005%. These results are given in Exhibit 11.3, columns 6 and 7. Note that trading multiple tax lots of the same investment within a single account results in only one trade cost.

There are various trade cost types. For a *fixed* dollar trade cost, we define the transaction cost for the *m*th manager and the *a*th account in percentages as follows:

$$TC_{m,a,\%} = \text{Cost in Dollars/Value of All Accounts in Dollars.} \tag{11.16a}$$

We define an investment with a *variable* trading cost as follows:

$$TC_{m,a,\%} = \text{Variable Fee} \times \text{Transacted Dollars/Value of All Accounts in Dollars.} \tag{11.16b}$$

We define an investment with a fixed and variable component as follows:

$$TC_{m,a,\%} = (\text{Fixed Fee} + (\text{Variable Fee} \times \text{Transacted Dollars}))/\text{Value of All Accounts in Dollars.} \tag{11.16c}$$

Finally, for each possible entry within $\hat{\mathbf{h}}_{\mathbf{M}}$, we assume that there are two corresponding vectors: a vector of changes relative to the starting holdings and a vector of dynamically updated transaction costs with the same dimensions ($\mathbf{TC}_{\mathbf{M,A,\%}}$). Recall that the same investment may be owned in multiple accounts with multiple tax lots or may be available for purchase in multiple accounts; therefore, it may appear multiple times in $\hat{\mathbf{h}}_{\mathbf{M}}$. Because both buys and sells result in costs, we focus on the absolute value of the change to make sure all costs are accounted for. The vector of absolute changes in the allocation to the available managers is as follows:

$$\left|\hat{\mathbf{h}}_{\Delta\mathbf{M}}\right| = \left|\hat{\mathbf{h}}_{\mathbf{M,t}=0} - \hat{\mathbf{h}}_{\mathbf{M,t}=-1}\right|, \tag{11.17}$$

where:

$\hat{\mathbf{h}}_{\mathbf{M,t}=0}$ = postoptimization holdings (weights) to each fund/manager ($M \times 1$ column vector); and

$\hat{\mathbf{h}}_{\mathbf{M,t}=-1}$ = pre-optimization holdings (weights) to each fund/manager ($M \times 1$ column vector).

And from this, we have the following:

$\mathbf{TC}_{\mathbf{M,A,\%}}$ = dynamical updated transaction cost in percent for each fund/manager ($M \times 1$ column vector).

Looking ahead to the continuing example in Exhibit 11.3, column 2 is $\hat{\mathbf{h}}_{\mathbf{M,t}=0}$, column 3 is $\hat{\mathbf{h}}_{\mathbf{M,t}=-1}$, column 4 is $\left|\hat{\mathbf{h}}_{\Delta\mathbf{M}}\right|$, and column 5 is $\mathbf{TC}_{\mathbf{M,A,\%}}$.

We incorporate transactions costs in the objective function with an additional penalty term in Equation 11.18. We allow for an aversion coefficient to transaction costs (λ_{TC}) which, depending on other choices one makes, could also be interpreted as a scaling mechanism. Because these transaction costs directly subtract from alpha, a reasonable option is to set λ_{TC} to 1. We elaborate on the interpretation and role of the different lambdas shortly.

$$Obj = \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}} \hat{\boldsymbol{\alpha}}_{\mathbf{M}}}_{\text{Alpha}} + \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}} \mathbf{C} \boldsymbol{\phi}}_{\text{Nonpecuniary Benefit}} \underbrace{- \frac{1}{2} \lambda_a \underbrace{[\hat{\mathbf{X}}^{\mathbf{T}} \hat{\mathbf{h}}_{\mathbf{M}} - \hat{\mathbf{h}}_{\mathbf{K}}]^T \hat{\mathbf{V}}_{\mathbf{K}} [\hat{\mathbf{X}}^{\mathbf{T}} \hat{\mathbf{h}}_{\mathbf{M}} - \hat{\mathbf{h}}_{\mathbf{K}}]}_{\text{Asset Allocation Misfit Risk}} - \frac{1}{2} \lambda_m \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}} \hat{\mathbf{V}}_{\alpha} \hat{\mathbf{h}}_{\mathbf{M}}}_{\text{Residual Variance}}}_{\text{Total Active Risk Penalty}} - \lambda_{TC} \underbrace{\sum_{m=1}^{M} \mathbf{TC}_{\mathbf{M,A\%}}}_{\text{Transaction Costs}}, \quad (11.18)$$

where λ_{TC} is the aversion coefficient to transaction costs.

Taxes Due to Transactions

In most cases, investors with taxable accounts will have already allocated money to various investments. In the example in Exhibit 11.2, column 5 shows the current value of existing holdings, and for the holdings in the taxable account, column 6 shows the corresponding cost basis for each holding/tax lot. The current value of these existing investments may be higher or lower than their respective cost basis for tax purposes, and thus we want to introduce a term to capture the potential taxable gains and taxable losses. As presented here, the term is not about future tax-efficiency; rather, it is about harvesting immediate tax losses and avoiding current taxes.

Ignoring any broader factors influencing the timing of taxable events, if (1) a holding's current value is below the cost basis; (2) the benefit of realizing a loss exceeds the transaction costs; and (3) there are reasonable substitute funds (e.g., expected alpha, style weights, nonpecuniary exposures), it is likely in the investor's economic interest to harvest the loss and redeploy the money with one of the available substitute funds. One could seek to use (1) a nearly equivalent fund (e.g., a small-cap value fund for another small-cap value fund); (2) a less exact substitute fund (e.g., global equity fund for a US large-cap-centric fund); or (3) simply remain in a cash equivalent investment for the required number of days to avoid violating a wash rule.[102] With the proper setup, the optimizer will determine the appropriate fund-specific investments ($\hat{\mathbf{h}}_{\mathbf{M}}$). To the degree that there are rules prohibiting one from repurchasing the same or nearly the same security within a given time period, one must create logic/methods to avoid such violations (e.g., a wash sale).

Conversely and hopefully more typically, a holding's current value will exceed the cost basis and, when sold, will result in a realized gain. In different countries and through time, tax rules and tax brackets change, so a tax-aware optimizer needs to incorporate these items and be updated appropriately.

In this context, an investor would want to realize a taxable gain for the following primary reasons: the fund in question is meaningfully contributing to total tracking error (via misfit or fund-specific risk); the expected (after-fee, after-tax) alpha of the fund is significantly worse than other available funds; or changes in the investor's nonpecuniary preferences would cause them to derive greater benefit from a fund with different nonpecuniary characteristics. We begin to encounter some of the limitations of a single-period framework in a multiperiod world. The benefits of moving to a different fund that either significantly decreases tracking

[102]A wash sale is one followed by a purchase of an identical or effectively identical security with a given time period (e.g., 30 days), and any resulting loss disallowed for tax purposes.

error, increases expected alpha, or is better aligned with the investor's nonpecuniary preferences are likely to extend beyond a single period. We return to multiperiod effects shortly. The tax cost incurred for a transaction is as follows:[103]

$$\text{Dollar Impact} = \text{Gain (or Loss)} \times \text{Tax Rate}. \tag{11.19}$$

The applicable tax rate should correspond to the specific gain (or loss); in that way, short-term gains are more likely to be avoided.[104] As before, we can then move from dollar/monetary space to percent/return space by dividing the dollar impact by the value of all accounts, recalling that tax-deferred balances have been multiplied by one minus the expected future income tax rate at the time of the withdrawals. For a given manager holding or position:

$$\mathbf{TAX}_{\mathbf{M,A,\%}} = \frac{\mathbf{RG}_{\Delta\mathbf{M,\$}} \times \text{Tax Rate}}{\textit{Value of all accounts}}, \tag{11.20}$$

where:

$\mathbf{TAX}_{\mathbf{M,A,\%}}$ = dynamically updated tax impact in percent due to changes in existing fund/managers ($M_T \times 1$ column vector); and

$\mathbf{RG}_{\Delta\mathbf{M,\$}}$ = dynamically updated realized gains (or losses) in dollars due to changes in $\hat{\mathbf{h}}_{\mathbf{M}}$.

To make this more concrete, let us look at two possible trades for funds held for more than one year, by an investor with an applicable tax rate of 30%, and a current account value of $500,000. Let us assume the optimizer completely sells off Investment A with a cost basis of $25,000 for $30,000. This results in a $5,000 profit before taxes, additional taxes for the investor of $1,500, and an entry in $\mathbf{TAX}_{\mathbf{M,A,\%}}$ of 30 basis points or 0.30% ($1,500/$500,000). The optimizer also completely sells off Investment B with a cost basis of $152,000 for $150,000. This results in a harvested loss of $2,000, which will decrease the amount of taxes the investor must pay in the subsequent tax year by $600, and becomes an entry in $\mathbf{TAX}_{\mathbf{M,A,\%}}$ of −12 basis points or −0.12% ($600/$500,000). Notice that it is entirely possible that tax losses exceed taxable gains as the optimizer changes $\hat{\mathbf{h}}_{\mathbf{M}}$.

To arrive at the total tax implication in percent/return space, we simply sum the elements of $\mathbf{TAX}_{\mathbf{M,A,\%}}$. Equation 11.21 expands the objective function to include the impact of realized taxable gains and losses. Once again, we include a coefficient that represents an aversion to paying taxes that could be used to reflect preferences or to serve as a scaling factor. As with transaction costs, because the realized tax impact directly subtracts from alpha, a logical choice is to set λ_{TAX} to 1. Note that it is possible that, if there are significant harvested losses, the sum of the values of $\mathbf{TAX}_{\mathbf{M,A,\%}}$ could be negative and, thus, rather than serving as a penalty for taxes to be paid, it could serve as a benefit from harvesting losses. In this way, regularly rerunning this type of optimization serves as a system to harvest tax losses.

[103]We place no limits on how negative the dollar amount of taxes can be, even though, in reality, there are limits on how much loss can be applied in one year. We assume the part that cannot be applied is carried over to future years. We are assuming that all losses are netted against income from other sources. Because of the very small amount of losses that can be netted against labor income (currently $3,000 in any one year in the United States), the income from other sources is assumed to come from other investments.

[104]Additionally, it is helpful to be aware of the number of days before a short-term gain would become a long-term gain. In many cases, it is advantageous to explicitly constrain the realization of short-term gains.

$$Obj = \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\boldsymbol{\alpha}}_{\mathbf{M}}}_{\text{Alpha}} + \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\mathbf{C}\boldsymbol{\phi}}_{\substack{\text{Nonpecuniary}\\ \text{Benefit}}} \underbrace{-\frac{1}{2}\lambda_a \underbrace{[\hat{\mathbf{X}}^{\mathbf{T}}\hat{\mathbf{h}}_{\mathbf{M}} - \hat{\mathbf{h}}_{\mathbf{K}}]^T \hat{\mathbf{V}}_{\mathbf{K}} [\hat{\mathbf{X}}^{\mathbf{T}}\hat{\mathbf{h}}_{\mathbf{M}} - \hat{\mathbf{h}}_{\mathbf{K}}]}_{\text{Asset Allocation Misfit Risk}} - \frac{1}{2}\lambda_m \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{V}}_{\alpha}\hat{\mathbf{h}}_{\mathbf{M}}}_{\substack{\text{Residual}\\ \text{Variance}}}}_{\text{Total Active Risk Penalty}}$$
$$- \lambda_{TC} \underbrace{\sum_{m=1}^{M} \mathbf{TC}_{\mathbf{M,A\%}}}_{\substack{\text{Transaction}\\ \text{Costs}}} - \lambda_{TAX} \underbrace{\sum_{m=1}^{M} \mathbf{TAX}_{\mathbf{M,A\%}}}_{\substack{\text{Tax}\\ \text{Costs}}}, \tag{11.21}$$

where λ_{TAX} is the aversion coefficient to realized taxes.

Equation 11.21 presents a fully functional, multi-account, tax-aware, transaction cost–aware fund-of-funds optimization objective function. Regularly rerunning this type of multi-account fund-of-funds optimization serves as an integrated tax loss harvesting tool and rollover optimization system. A closely related advantage is that this periodic optimization also serves as an optimal rebalancing system. It is superior to rebalancing systems that simply return the portfolio to a set of prespecified investment weights, such as those in a model portfolio. With each reoptimization, all of the various trade-offs inherent in Equation 11.21 are reconsidered. In the case of rebalancing, as the current funds produce returns, the amount of asset class misfit risk, fund-specific tracking error, and tax consequences evolve. Furthermore, rebalancing (reflecting decisions about how and where to rebalance) may incur transaction costs. Reoptimizing finds the optimal solution, simultaneously deciding how to allocate to all of the available managers in a tax-efficient manner, whether to harvest tax losses, when to roll/move money between the same account types, and how best to reflect the investor's nonpecuniary preferences. If and when new funds become available or existing funds are no longer available (e.g., a fund is removed from a 401(k) plan line-up), the multi-account fund-of-funds optimization provides the optimal solution.

Additional Considerations and Finishing the Example

With the personalized optimization objective function complete, in this section, we examine several additional benefits, use cases, and model extensions. With some of these additional benefits in mind, we complete the working example.

A Better Form of Transition Management

A common challenge faced by many wealth advisers onboarding new clients with existing investments is knowing how to transition the legacy portfolio to reflect the new adviser's recommendation. A standard approach, in the spirit of dollar cost averaging, is to create a time-based schedule that sells off the legacy portfolio holdings to purchase the recommended portfolio while spreading out realized taxable gains. Relative to this naïve approach, one of the key advantages of the multi-account fund-of-funds optimization is that it considers the merits of the legacy holdings and will sell them only if it improves the solution. From this perspective, the framework serves as a new type of transition management optimizer. As we shall see in our example, the optimal solution typically retains some of the investments in the taxable account.

Optimally Deploying New Money and Withdrawing Money

Investors periodically add money to accounts, others withdraw money, and some do both. Equation 11.21 and the framework presented here provides an optimal solution.

For asset accumulators, most workers make regular contributions to their defined-contribution retirement plan/account established by their employer. These contributions do not typically show up as cash; rather, in most cases, the money is immediately deployed or invested based on the investor's fund elections, or if they made no elections, the default. In either case—cash contributions or predetermined elections—rerunning the optimizer with updated constraints would produce the new optimal allocations and corresponding trade instructions.

For asset decumulators, withdrawals can be made from predetermined accounts or account types. If the withdrawal is from a predetermined account, the optimizer produces the new optimal allocations given the remaining account balance after the withdrawal. This is done by setting a constraint such that the total allocations in the account are equal to the account balance after the withdrawal. If the withdrawal is from a specific account type, the optimizer would decide from which account(s) of that account type to take the withdrawal, based on the investment options, trading costs, and tax costs (if the withdrawal is from taxable accounts) associated with each account. This is done by setting constraints such that the total allocations in all accounts of the type are equal to the total of those account balances after the withdrawal. Additionally, the constraints are such that the total allocations in each of those accounts are less than or equal to the account balance before the withdrawal.

Withdrawals can be made from predetermined accounts or account types. Based on any specified accounts or account types, the optimizer determines the optimal source of the withdrawal.

Contributions to some accounts and withdrawals from other accounts can also occur in the same optimization. This includes the special and relatively frequent case in which one is contributing to an employer-sponsored defined contribution plan (presumably to take advantage of a match) and simultaneously needing to withdraw funds from a specified account. In these cases, the optimal manager structure is determined based on all the trade-offs in the objective function.

Accounting for Multiperiod Implications within a Single-Period Framework

Although this is a single-period framework, to the degree that changing the fund allocations leads to better long-term alphas, better long-term tax location, lower long-term tracking error, or more desirable nonpecuniary characteristics, we would expect these benefits to persist for multiple years.[105] Conversely, the realized transaction costs and realized tax obligations do not persist for multiple years. As such, it may be desirable to reduce or discount the impact of these two terms. Alternatively, rather than reducing the impact of these two terms, one may want to compound the benefits of a better alpha or more efficient tax location over multiple periods.

To clarify this, let us assume that a fund has two share classes: share class A has an annual expense ratio of 30 basis points and share class B has an annual expense ratio of 20 basis points. Furthermore, the amounts invested are $10,000 in share class A and $0 in share class B, and the cost to trade is $10 per trade. Moving from share class A to share class B requires two trades—a sell and a buy—for a total trade cost of $20. Given the current amount of $10,000, moving from share class A to share class B will reduce the annual amount paid for fund expenses by $10. So, if the investment horizon is one year, as it is in the single-period annual framework, it does not make sense to incur $20 of trading costs to save $10 in annual fund expenses. If the time horizon is longer than two years, it would make sense to move to the less expensive share class. This same logic applies to realizing a sizable tax cost and either (1) moving to a more tax-efficient fund held within a taxable account or (2) relocating tax-inefficient asset classes/investments into tax-preferred accounts. When viewed through a single-period lens, the change may not make sense, but it makes sense when the benefits compound beyond a single period.

[105]Although we were unaware of it during the writing of earlier drafts of this paper, my desire to adapt a single-period framework to account for multi-period implications is similar in spirit to the approach described in DiBartolomeo (2012).

We explore two potential approaches for accounting for these multiperiod implications.

Lambda Adjustment Approach

As mentioned earlier, in the Markowitz framework, an investor's risk aversion coefficient, or lambda (λ), identifies the rate at which the investor is willing to make the trade-off between risk and expected return. It thus identifies the appropriate point on the efficient frontier for that investor. We have a more generalized use for the term *lambda*. In our interpretation, a lambda might collectively represent preference/aversion, scaling, and/or a weighting for a given penalty. Notice that in Equation 11.21, we included lambdas for each of the elements that detract from the objective function: an asset-allocation misfit lambda (λ_a), a fund-specific tracking error lambda (λ_m), a transaction-cost lambda (λ_{TC}), and a tax lambda (λ_{TAX}). Like the controls on an audio equalizer, an investor can dial in, or adjust, these based on their unique preferences.

Therefore, to the degree that one believes that the benefits of better after-fee alphas, more efficient tax location, lower tracking error, or more desirable current nonpecuniary characteristics will persist into the future, one may want to discount or decrease the impact of transaction costs and taxes when calculating the objective function (Obj_P). Unfortunately, this involves a bit more art than science. One way to think about this is in terms of the approximate number of years over which one believes the benefits of changes to the manager structure will persist and then to divide or amortize the immediate impact of transaction costs and tax implications over that number of years.

Although transaction costs are likely to be small, we can modify Equations 11.16a, 11.16b, and 11.16c in the following manner to decrease their influence on the objective function (Obj_P):

$$TC_{m,a,\%} = \frac{\text{Cost in Dollars}}{\text{Value of All Accounts in Dollars}} / \text{Years of Impact}, \tag{11.22a}$$

$$TC_{m,a,\%} = \frac{\text{Variable Fee} \times \text{Transacted Dollars}}{\text{Value of All Accounts in Dollars}} / \text{Years of Impact}, \tag{11.22b}$$

$$TC_{m,a,\%} = \frac{\text{Variable Fee} + \text{Variable Fee} \times \text{Transacted Dollars}}{\text{Value of All Accounts in Dollars}} / \text{Years of Impact}. \tag{11.22c}$$

If one year is the period over which the impact is estimated to occur, there is no difference between the corresponding formulas, Equations 11.16a and 11.22a, Equations11.16b and 11.22b, and Equations 11.16c and 11.22c, respectively. Earlier, we said that λ_{TC} and λ_{TAX} should be the same and that it would be logical to set them equal to 1. Equations 11.16a, 11.16b, and 11.16c relative to Equations 11.22a, 11.22b, and 11.22c are equivalent to dividing by the estimated years of impact. This suggests the following setup:

$$\text{Dollar Impact} = \text{Gain (or Loss)} \times \text{Tax Rate/Years of Impact}. \tag{11.23}$$

As before, we can then move from the dollar/monetary space to the percent/return space by dividing the dollar impact by the value of all accounts and then further dividing by the years of expected impact. For a given fund/manager holding or position, we have the following:

$$TAX_{m,a,\%} = \frac{\left(\dfrac{RG_{\Delta m,\$} \times \text{Tax Rate}}{\textit{Value of All Accounts}} \right)}{\text{Years of Impact}}. \tag{11.24}$$

And across all holdings, we have the following:

$$\mathbf{TAX}_{\mathbf{M,A,\%}} = \frac{\left(\dfrac{\mathbf{RG}_{\Delta\mathbf{M,\$}} \times \text{Tax Rate}}{\textit{Value of All Accounts}}\right)}{\text{Years of Impact}}. \quad (11.25)$$

Discounted Cash Flow Approach

If one has more precise estimates of the future, such as when funds are likely to be sold, changes in the tax rates/tax bracket, the value of an investment relative to the cost basis, and an appropriate discount rate for the time value of money, one can be more precise. More specifically, the impact of alphas can be estimated, future transaction costs can be expressed as net present values, and future realized capital gains can be expressed as net present values. Even if one is estimating with a margin of error, it may be a useful exercise to carry out such estimates.

An investor may own a fund in their taxable account with a current value greater than the basis. Under one scenario, the investor will never need to realize the taxable gain and, under US law, will be able to pass the fund on to their heirs, in which case the tax on the gain is forgiven. A more common occurrence for many investors, and in line with the heuristic that retirees should draw down taxable accounts before drawing down qualified assets, is that the investor will need to sell taxable investments, thus incurring taxable gains. Equation 11.21 ignores the future tax consequence of selling an investment in the future. To the degree that one can forecast the time of that sale, the gain, future tax rate/bracket, and an appropriate time value of money discount rate, one can be more precise.

Additional Potential Constraints

In addition to the constraints already discussed, practitioners may want to impose a variety of additional constraints on the optimization.

- Cap on Realized Capital Gains: At times, it may be useful to constrain the amount of realized capital gains (or taxes to be paid).
- Timing of Short-Term Gains: Typically, short-term gains are taxed at a higher rate than long-term gains. To the degree that a short-term gain is approaching the time at which it will be considered a long-term gain, it may make sense to delay or constrain such transactions.
- Asset Class Group Constraints: This relates more to what we think of as aesthetics. For instance, if the policy portfolio is a 50/50 mix of equities and fixed-income asset classes, and if rolling up the effective asset classes leads to some other values, such as 53/47, this may bother some investors. It can be controlled through an asset class group constraint.
- Fund Minimums and Maximums: Beyond the 0% to 100% range, one may choose a specific minimum and maximum for one or more of the current/available investments. The motivation could be to satisfy a minimum allowable investment threshold, to avoid overallocating to any single investment, or to avoid trading with a given manager.

Implementation Challenges

The creation of the multi-account optimization system presented in this chapter is complicated and novel. Most advisers and wealth managers will need to use commercial software systems that overcome the vast number of implementation challenges. Some of those challenges include having inputs for the universe of available funds, including explicit alpha forecasts for funds held in qualified accounts and taxable accounts, style exposure estimates, and fund-specific tracking error estimates. The gathering of investor-specific information, such as information on each tax lot (cost basis, date of purchase) requires information to be manually inputted or gathered via an electronic connection. The system requires updated price data for all of the investments held in taxable accounts. What types of connections will the system have with different accounts—will it simply push trade instructions, or will it have an interactive connection? How and when will trading occur? ETFs trade intraday, yet mutual funds trade at the end of the day—thus, the optimal answer may change during the day.

Finishing the Example

Returning to the relatively simple case presented earlier, the working example has largely been covered. We now elaborate on the optimal investment allocations derived by maximizing the objective function in Equation 11.21 using the lambda-based amortization approach to account for costs that are incurred to produce benefits that are expected to persist for multiple years.

In this example, we calculate the vector of transaction costs $\mathbf{TC}_{\mathbf{M,A,\%}}$ and the vector of tax consequences $\mathbf{TAX}_{\mathbf{M,A,\%}}$ under the assumption that the benefits from changes would last 25 years.

Starting with Account 1, the tax-exempt account, the optimizer sold off the Passive US Equity ETF to purchase more of the Active Global Equity option, likely because of its superior alpha and to reduce tracking error relative to the tax-exempt targets ($\mathbf{h}_{\mathbf{K,E}}$).

Moving to Accounts 2 and 3, the tax-deferred 401(k) and the tax-deferred IRA, recall that the money in these two accounts is being treated as fungible. The Employer Stock was completely sold, and the majority of the equity funds were sold with the proceeds being allocated to the High Alpha Active Bond fund to ensure that they would be more closely aligned with the tax-deferred asset allocation targets ($\mathbf{h}_{\mathbf{K,D}}$). In this case, all of the money was rolled out of the 401(k) account and into the tax-deferred IRA.

In Account 4, the taxable account, Lot 1 of the Employer Stock was completely sold off, creating a realized gain. Lot 2 of the Employer Stock was mostly sold. Holding large amounts of Employer Stock was undesirable because, in this example, it contributes significantly to investment-specific tracking error. The optimizer completely sold off the Taxable-Active Bond Fund for two reasons: (1) it was harvesting the tax loss and (2) the alpha of the fund when held in a taxable account was unattractive, thus holistically the fund-of-funds optimizer decided to hold taxable bonds in a qualified account. The money in the taxable account was primarily allocated to the Passive Global Equity ETF with its relatively attractive alpha and alignment to the taxable target. Perhaps the biggest surprise is the allocation to the Passive Bond Fund. Presumably, its low investment-specific tracking error, relative to other funds, contributed to its small but positive allocation.

Exhibit 11.4 compares the starting or pre-optimization effective asset allocation and the ending or post-optimization effective asset allocation as well as the asset allocation misfits relative to the account-type target policy asset allocations. In all cases, the amount of misfit risk was reduced.

If we were to rerun the optimizer, with no changes in the inputs, there would be no changes in the outputs. On a daily basis, the most likely input to change is the values of the holdings. A down market might present the opportunity to harvest multiple losses. Changing values of the holdings will also change the amount of total tracking error, including the part caused by asset allocation misfits and the part caused by fund-specific residual risk. To the degree that the optimizer can use the tax-deferred account with no trading costs or consequences to increase the value of the objective function, it will likely do so. If at some point the benefit to trading overcomes the trading costs and/or any potential taxes, the optimizer will determine when and how it makes sense to do so. Likewise, if new or better managers become available, a manager becomes unavailable, investment fees change, alpha estimates change, tax rates change, the investor's nonpecuniary preferences change, or any other inputs change, the optimizer will recommend the new, optimal allocations based on all of the updated inputs.

Exhibit 11.3. Sample Case: Postoptimization

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Accounts/Managers	Manager Holdings $h_{M,t=-1}$	Manager Holdings $h_{M,t=0}$	Change $h_{\Delta M}$	Absolute Change $\lvert h_{\Delta M}\rvert$	Transaction Cost $TC_{M,\$}$	Transaction Cost $TC_{M,\%}$	Realized Gain/Loss $RG_{M,\$}$	Realized Gain/Loss $RG_{M,\%}$	Taxes $TAX_{M,\%}$
Account 1. Tax-Exempt									
Employer Stock (Large Cap)	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
Passive US Equity ETF	4.0%	0.0%	−4.0%	4.0%	$20	0.004%	N/A	N/A	N/A
Active Global Equity	10.0%	14.0%	4.0%	4.0%	$20	0.004%	N/A	N/A	N/A
Emerging Market Equity	0.0%	0.0%	0.0%	0.0%	$20	0.004%	N/A	N/A	N/A
High Alpha Active Bond Fund	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
Passive Bond ETF	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
Subtotal:	14.0%	14.0%	0.0%	8.0%					
Account 2. Tax-Deferred 401(k)									
Employer Stock (Large Cap)	15.0%	0.0%	−15.0%	15.0%	$0	0.000%	N/A	N/A	N/A
High Cost US Equity Fund	5.0%	0.0%	−5.0%	5.0%	$0	0.000%	N/A	N/A	N/A
High Cost Non-US Equity Fund	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
High Cost Bond Fund	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
Active Bond Fund	2.0%	0.0%	−2.0%	2.0%	$0	0.000%	N/A	N/A	N/A
Subtotal:	22.0%	0.0%	−22.0%	22.0%					

(*continued*)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Accounts/Managers	**Manager Holdings $h_{M,t=-1}$**	**Manager Holdings $h_{M,t=0}$**	**Change $h_{\Delta M}$**	**Absolute Change $\|h_{\Delta M}\|$**	**Transaction Cost $TC_{M,\$}$**	**Transaction Cost $TC_{M,\%}$**	**Realized Gain/Loss $RG_{M,\$}$**	**Realized Gain/Loss $RG_{M,\%}$**	**Taxes $TAX_{M,\%}$**
Account 3. Tax-Deferred IRA									
Employer Stock (Large Cap)	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
Passive US Equity ETF	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
Passive Global Equity ETF	8.0%	8.0%	0.0%	0.0%	$10	0.002%	N/A	N/A	N/A
High Alpha Bond Fund	16.0%	38.0%	22.0%	22.0%	$10	0.002%	N/A	N/A	N/A
Passive Bond ETF	0.0%	0.0%	0.0%	0.0%	$0	0.000%	N/A	N/A	N/A
Subtotal:	24.0%	46.0%	22.0%	22.0%					
Account 4. Taxable									
Employer Stock (Large Cap), Lot 1	6.0%	0.0%	–6.0%	6.0%	$12.5	0.0025%	$5,000	1.00%	0.15%
Employer Stock (Large Cap), Lot 2	4.0%	2.8%	–1.2%	1.2%	$12.5	0.0025%	$8,696	1.74%	0.26%
Passive US Equity ETF	15.0%	10.1%	–4.9%	4.9%	$25	0.005%	$1,437	0.29%	0.04%
Passive Global Equity ETF	0.0%	26.1%	26.1%	26.1%	$25	0.005%	$0	0.00%	0.00%
High Cost Bond Fund	15.0%	0.0%	–15.0%	15.0%	$25	0.005%	($1,000)	–0.20%	–0.03%
Passive Bond ETF	0.0%	1.05%	1.1%	1.1%	$25	0.005%	$0	0.00%	0.00%
Subtotal:	40.0%	40.0%	0.0%	54.3%					
Expected Alpha	–0.16%	0.13%							
Total Tracking Error	5.97%	1.63%							
Asset Allocation Misfit Risk	3.64%	1.49%							
Investment-Specific Risk	4.73%	0.65%							
Portfolio Utility	–2.60%	–0.33%							

Exhibit 11.4. Sample Case: Effective Asset Allocations

Asset Class	Effective Asset Allocations								
	Tax-Exempt $h_{K,E}$			Tax-Deferred $h_{K,D}$			Taxable $h_{K,T}$		
	Target	Pre	Post	Target	Pre	Post	Target	Pre	Post
US Large Growth	0.0%	2.4%	1.6%	0.0%	20.6%	1.5%	9.8%	13.4%	9.0%
US Large Value	0.0%	2.8%	2.6%	0.0%	2.2%	1.3%	10.0%	3.6%	5.9%
US Mid Growth	0.0%	1.3%	1.1%	0.0%	3.1%	0.9%	4.2%	2.0%	3.9%
US Mid Value	1.0%	2.6%	2.6%	0.0%	1.2%	1.2%	3.3%	2.6%	4.9%
US Small Growth	2.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
US Small Value	3.0%	1.2%	1.6%	0.0%	0.0%	0.0%	0.0%	0.4%	0.4%
REITs	2.6%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%	0.0%
Non-US Developed Equity	0.0%	1.6%	2.1%	6.3%	3.0%	3.0%	8.5%	0.3%	8.1%
Emerging Market Equity	2.2%	0.5%	0.7%	3.5%	0.9%	0.9%	0.0%	0.0%	2.5%
Commodities	0.0%	0.2%	0.2%	2.6%	0.1%	0.1%	0.0%	0.1%	0.3%
US Short-Term Bonds	0.0%	0.0%	0.0%	1.6%	12.2%	26.2%	0.0%	7.4%	0.5%
US Intermediate-Term Bonds	0.0%	0.0%	0.0%	7.7%	0.3%	0.1%	0.0%	0.4%	0.2%
US Long-Term Bonds	0.0%	0.0%	0.0%	17.4%	4.7%	10.1%	0.0%	2.6%	0.2%
Short-Term Inflation-Linked Bonds	0.0%	0.0%	0.0%	2.4%	1.7%	3.3%	0.0%	0.7%	0.0%
Long-Term Inflation-Linked Bonds	0.0%	0.0%	0.0%	2.0%	0.0%	0.0%	0.0%	0.2%	0.0%
High Yield	0.8%	0.0%	0.0%	2.3%	0.0%	0.0%	0.0%	0.5%	0.0%
Non-US Bonds	0.0%	0.0%	0.0%	5.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Emerging Market Bonds	0.0%	0.0%	0.0%	0.0%	1.3%	2.8%	0.0%	1.1%	0.0%
Cash	0.0%	0.0%	0.0%	0.9%	0.2%	0.0%	0.0%	0.3%	0.1%
Sum	12.6%	12.6%	12.6%	51.6%	51.6%	51.6%	35.9%	35.9%	35.9%
Misfit Risk versus Account-Type Target		0.88%	0.86%		4.08%	1.31%		2.19%	0.38%

Conclusion and Key Takeaways

Allocating money to actual investments is a critical, and often regulated, fiduciary endeavor that is unfortunately dominated by ad hoc practices rather than repeatable, theoretically sound methods. Alpha-tracking error optimization, serving as our grandchild model, provides such a method, complementing MVO-based child models for asset location and allocation and thus enabling investors to implement a target asset allocation with a set of available investments.

In this chapter, we present a new alpha-tracking error optimization framework for a multiple account setting in which the multiple accounts can include accounts with different tax treatments, existing holdings with different cost bases, different opportunity sets of investments, different investment-specific trading costs, and additional preferences and constraints. In doing so, we simultaneously solve for asset location by placing tax-inefficient investments in tax-exempt and tax-deferred accounts, and placing tax-efficient investments in taxable accounts. We develop a fully functional, multi-account, tax-aware, transaction cost-aware alpha-tracking error objective function that can be used to solve for the optimal allocations in a single optimization. The optimizer simultaneously considers account-type-specific alphas for different managers, the implications of owning a manager in a particular account type, transaction costs, and tax consequences. Additionally, by including a nonpecuniary preference term in the objective function, the optimizer can create a new level of personalization in which the optimal portfolio reflects an investor's unique nonpecuniary preferences, such as those related to ESG.

Frequently rerunning this type of optimization serves as an integrated tax loss harvesting tool, an ongoing transition management tool, a smart portfolio rebalancing tool, and rollover tool. In contrast to most other approaches to tax-loss harvesting, transition management, rebalancing, and rollovers, this multi-account optimizer considers all of the relative trade-offs, including transactions costs, the merits of legacy holdings, tax consequences, the amount of tracking error, asset location, opportunities to allocate to better funds, and the opportunity to create a fundamentally more personalized portfolio based on the investor's nonpecuniary preferences. Such a framework fills a clear gap in the tool kit of most wealth managers who are assisting investors with multiple accounts. The framework can also help power holistic, multi-account digital advice solutions. Finally, as a multi-account optimizer in which the accounts could be housed at any number of firms, this type of optimizer enables a form of *distributed* or *decentralized* financial advice.

Appendix 11A. Total Return Fund-of-Funds Optimization and Joint Asset Allocation Optimization with Multiple Accounts

Depending on the context, it may or may not be desirable to work in a total return space. It is most likely necessary if one would like to simultaneously set asset allocation policy and the manager structure as described in the appendix of Waring et al. (2000). A notable exception is found if one accepts the target asset allocation as optimal in Equation 11A.6, which simplifies to Equation 11.21.

Focusing on the expected return side of the objective function in Equation 11A.1, a new and critical term is the vector of asset class expected returns, which we denote $\boldsymbol{\mu}_K$. We also include alpha ($\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\boldsymbol{\alpha}}_{\mathbf{M}}$) and the total expected return from the asset class exposure ($\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}}\boldsymbol{\mu}_K$). We continue to include a term for nonpecuniary investor preferences ($\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\mathbf{C}\phi$).

$$\underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{K}}}_{\substack{\text{Asset allocation}\\ \text{expected return}}} + \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\boldsymbol{\alpha}}_{\mathbf{M}}}_{\substack{\text{Alpha}\\ \text{expected}\\ \text{return}}} + \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\mathbf{C}\boldsymbol{\phi}}_{\substack{\text{Nonpecuniary}\\ \text{preferences}}}. \tag{11A.1}$$

Moving from the good to the bad in a total return setting, items that decrease the objective function are (1) the risk of the policy portfolio, (2) asset class misfit risk, (3) investment-specific risk, (4) trading costs, and (5) taxes. The only new item in this list is the first one—that is, the risk of the asset allocation of the policy portfolio. We now explore the roles of these terms, how they relate to each other, and how to weight them appropriately.

The Risk from Asset Class Exposures

In total return space, asset allocation risk is as follows:

$$\frac{1}{2}\lambda_b[\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}}]\hat{\mathbf{V}}_{\mathbf{K}}[\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}}]^{\mathbf{T}}. \tag{11A.2}$$

If one is simultaneously attempting to set asset allocation policy and allocate investments, λ_b and λ_m would likely have similar values.

Investment-Specific Tracking Error

Specific tracking error from investments remains the same:

$$\frac{1}{2}\lambda_m\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{V}}_{\alpha}\hat{\mathbf{h}}_{\mathbf{M}}. \tag{11A.3}$$

As the optimizer programmatically adjusts the allocation to the available investments ($\mathbf{h}_{\mathbf{M}}$), the amount of specific tracking error evolves.

Trading Costs

The calculation of trading costs remains the same:

$$\lambda_{TC}\sum_{m=1}^{M}\mathbf{TC}_{\mathbf{M,A,\%}}. \tag{11A.4}$$

$\mathbf{TC}_{\mathbf{M,A,\%}}$ is a vector or list of transaction costs resulting from changes to the available managers ($\mathbf{h}_{\mathbf{M}}$). As the optimizer programmatically adjusts the allocation to the available managers ($\mathbf{h}_{\mathbf{M}}$), the amount of trading costs evolves. We continue to use the "%" subscript to indicate that monetary amounts are being expressed as a percentage of the entire portfolio.

Taxes

The calculation of taxes remains the same:

$$\lambda_{TAX}\sum_{m=1}^{M}\mathbf{TAX}_{\mathbf{M,A,\%}}. \tag{11A.5}$$

Recall that $\mathbf{TAX}_{\mathbf{M,A,\%}}$ is a vector or list of tax costs or savings resulting from changes to the available managers ($\mathbf{h}_\mathbf{M}$).

In total return space, the objective function is as follows:

$$Obj = \hat{\mathbf{h}}_\mathbf{M}^\mathbf{T}\hat{\mathbf{X}}\hat{\boldsymbol{\mu}}_\mathbf{K} + \hat{\mathbf{h}}_\mathbf{M}^\mathbf{T}\hat{\boldsymbol{\alpha}}_\mathbf{M} + \hat{\mathbf{h}}_\mathbf{M}^\mathbf{T}\mathbf{C}\boldsymbol{\phi} - \frac{1}{2}\lambda_b[\hat{\mathbf{h}}_\mathbf{M}^\mathbf{T}\hat{\mathbf{X}}]\hat{\mathbf{V}}_\mathbf{K}[\hat{\mathbf{h}}_\mathbf{M}^\mathbf{T}\hat{\mathbf{X}}]^\mathbf{T} - \frac{1}{2}\lambda_m\hat{\mathbf{h}}_\mathbf{M}^\mathbf{T}\hat{\mathbf{V}}_\alpha\hat{\mathbf{h}}_\mathbf{M} - \lambda_{TC}\sum_{m=1}^{M}\mathbf{TC}_{\mathbf{M,A,\%}} - \lambda_{TAX}\sum_{m=1}^{M}\mathbf{TAX}_{\mathbf{M,A,\%}}. \quad (11A.6)$$

To the degree that transaction costs are becoming *de minimis*, arguably the hardest item to scale appropriately remains the realized tax consequences, which could limit current manager structure changes that would ultimately be beneficial in the long run. Equation 11A.6 is a generalized, fully functional, multi-account, tax-aware objective function capable of simultaneously solving for asset allocation and manager structure.

Appendix 11B. Liability-Relative, Total Return, Fund-of-Funds Optimization and Joint Asset Allocation Optimization with Multiple Accounts

An important extension of what we might call "asset-only" MVO is to make the approach more holistic or complete by recognizing that the investor's total portfolio consists of both assets and liabilities and to include both in a single optimization. Leibowitz (1987) is perhaps the earliest published account, although it is presented in a much more usable form in Sharpe (1990) and Sharpe and Tint (1990). More recently, important pieces include Siegel and Waring (2004) and Waring (2004a, 2004b), all working within a tax-free institutional setting. Idzorek and Blanchett (2019) apply liability-relative optimization to the creation of asset allocations for individuals, although it does not contemplate different tax treatments. Chapter 8 extends liability-relative optimization in an asset allocation context to jointly consider taxable and qualified accounts. We incorporate the joint asset allocation and location liability-relative optimization techniques into the single, multipurpose optimization framework.

In liability-relative optimization, or surplus optimization as it is often called, the optimizer is constrained to hold an asset, or combination of assets representing the systematic characteristics of the liability, short, and then in the presence of that liability, finds the optimal combinations of assets. Almost all portfolios exist to pay for something that can be thought of as a liability, and as such, the liability is the inescapable real-life benchmark.

In Equation 11B.1, we expand Equation 11A.6 to include the liability $\hat{\mathbf{h}}_\mathbf{L}$, restating or rescaling the portfolio assets based on the liability. We continue to use the "stacked" multiple-account-type variables and notation, and in the spirit of multiple goals or account-type-specific liabilities, we also included a stacked liability vector $\hat{\mathbf{h}}_\mathbf{L}$. For simplicity, we use a single overall asset-to-liability funding ratio, but we could have expanded this to include account-type-specific funding ratios.

$$
\begin{aligned}
Obj = {} & \underbrace{\frac{A_0}{L_0}\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\boldsymbol{\alpha}}_{\mathbf{M}}}_{\text{Alpha}} + \underbrace{\frac{A_0}{L_0}\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}}\hat{\boldsymbol{\mu}}_{\mathbf{k}} - \hat{\mathbf{h}}_{\mathbf{L}}^{\mathbf{T}}\hat{\boldsymbol{\mu}}_{\mathbf{k}}}_{\substack{\text{Expected Return of}\\ \text{Effective Asset Allocation}\\ \text{Relative to Liability}}} + \underbrace{\frac{A_0}{L_0}\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}}\mathbf{C}\vec{\boldsymbol{\phi}}_{\mathbf{i}} - [\hat{\mathbf{X}}\hat{\mathbf{h}}_{\mathbf{L}}]^{T}\,\mathbf{C}\boldsymbol{\phi}}_{\substack{\text{Nonpecuniary Preferences}\\ \text{minus Nonpecuniary}\\ \text{Characteristics of Liability}}} - \frac{1}{2}\lambda_a \underbrace{\left[\frac{A_0}{L_0}\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}} - \hat{\mathbf{h}}_{\mathbf{L}}^{\mathbf{T}}\right]\hat{\mathbf{V}}_{\mathbf{K}}\left[\frac{A_0}{L_0}\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{X}} - \hat{\mathbf{h}}_{\mathbf{L}}^{\mathbf{T}}\right]^{T}}_{\substack{\text{Risk of Effective Asset Allocation}\\ \text{Relative to Liability}}} \\
& - \frac{1}{2}\lambda_m\left(\frac{A_0}{L_0}\right)^2 \underbrace{\hat{\mathbf{h}}_{\mathbf{M}}^{\mathbf{T}}\hat{\mathbf{V}}_{\boldsymbol{\alpha}}\hat{\mathbf{h}}_{\mathbf{M}}}_{\text{Residual Risk}} - \lambda_{TC}\left(\frac{A_0}{L_0}\right)\underbrace{\sum_{m=1}^{M}\mathbf{TC}_{\mathbf{M,A\%}}}_{\text{Transaction Costs}} - \lambda_{TAX}\left(\frac{A_0}{L_0}\right)\underbrace{\sum_{m=1}^{M}\mathbf{TAX}_{\mathbf{M,A\%}}}_{\text{Tax Impact}},
\end{aligned}
\tag{11B.1}
$$

where:

$\hat{\mathbf{h}}_{\mathbf{L}}$ = holdings (weights) of the liability ($K \times 1$ column vector);

A_0 = value of assets at time 0; and

L_0 = value of liabilities at time 0.

12. AN END-TO-END EXAMPLE AND CALLS FOR ACTION

Context

With our three-stage model complete, we now conclude with the example that we have presented throughout this book. We also make several calls for action.

We have covered an enormous amount of material, ultimately putting forth a system of coherent and powerful tools for providing optimal lifetime advice. Embracing and leveraging this system requires work and change from researchers, practitioners, software creators, and regulators. In the main part of this final chapter, we focus on our hypothetical investor, Isabela, demonstrating how to use the three-stage multilevel model with an emphasis on the parent life-cycle model and the asset allocation and asset location *net-worth optimization* of the child model. We conclude the chapter and the book with eight key points and corresponding calls for action.

End-to-End Example

At the end of chapter 1, we introduced a hypothetical investor, Isabela, who is working with a financial planner, Paula. Throughout the book, we have periodically checked in on Isabela, incorporating her specific situation into many of the examples. We have also assumed that her financial planner, Paula, is using a state-of-the-art financial planning and investment management system based on the three-stage model and concepts presented in this book. In this chapter, we consolidate those various check-ins on Isabela to demonstrate the three-stage model.

In practice, the models would be run at least annually, but we jump through time rerunning the model when Isabela is 25, 45, and again when she is 65. We focus on the parent and child models.

Isabela, Age 25

Parent Model

When we first checked in on Isabela, she was 25 years old. Recall that a year earlier Isabela received her master's degree in marine biology and started working at a large scientific-research-oriented aquarium, earning $75,000 per year (after taxes). Each year, she should contribute $13,667 to her employer's retirement plan to receive the maximum employer match of $6,833, decreasing her taxable income and locating money in a tax-deferred account. To the degree that she is saving too much, initially this is offset by withdrawing from her taxable account to smooth her consumption. After one year of working and contributing to her employer-sponsored defined contribution (DC), tax-deferred retirement plan (coupled with matching employer funds), she had a tax-deferred account balance of $20,500. Isabela also has $250,000 in a taxable brokerage account resulting from the sale of her grandmother's home that she (along with her two siblings) had inherited a year earlier when her 97-year-old grandmother passed away. Inheriting the $250,000 and facing uncertainty around how to invest across her accounts was part of what motivated Isabela to seek out her financial planner, Paula.

Paula captures Isabela's financial information and preferences, all of which seamlessly feed into Paula's financial planning and ongoing investment management software system.

Exhibit 12.1. Financial Preferences for the Life-Cycle Model, Isabela, Age 25

Financial Preferences	Qualitative Assessment	Numeric Input
Impatience for Consumption: Subjective Discount Rate (ρ or rho)	**Patient:** Isabela is patient, wanting to live somewhat frugally now in hopes of a higher living standard later in life.	2%
Preference for Smooth Consumption: EOIS (η or eta)	**Moderate:** Because she plans ahead, Isabela is willing to have moderate interruptions to her consumption.	50%
Risk Tolerance (θ or theta)	**Low:** Isabela has a somewhat low tolerance for risk.	35%
Flexibility of Consumption versus Bequest: Intergenerational Elasticity (γ or gamma)	**Low:** Isabela has low flexibility when it comes to her desire to have both a moderate standard of living and her ability to leave a bequest.	25%
Importance of Consumption versus Bequest: Strength of Bequest Motive (ϕ or phi)	**Moderate:** Isabela prefers a moderate standard of living and would like to plan to leave a moderate bequest.	1.5%

From chapter 2, Paula's state-of-the-art financial planning system provides a system for evaluating Isabela's financial preferences, including Isabela's risk tolerance. Fortunately, the interactive system that Paula used for evaluating financial preferences was much more elegant than the sample questions provided in chapter 2. We summarize Isabela's key financial preferences in **Exhibit 12.1**. These are the key financial preferences that drive the parent life-cycle model.

Moving from financial preferences to a holistic view of Isabela's financial circumstances, **Exhibit 12.2** presents Isabela's balance sheet. Paula focuses her planning practice on the individual balance sheet and her client's overall financial health, as measured by Isabela's net worth.

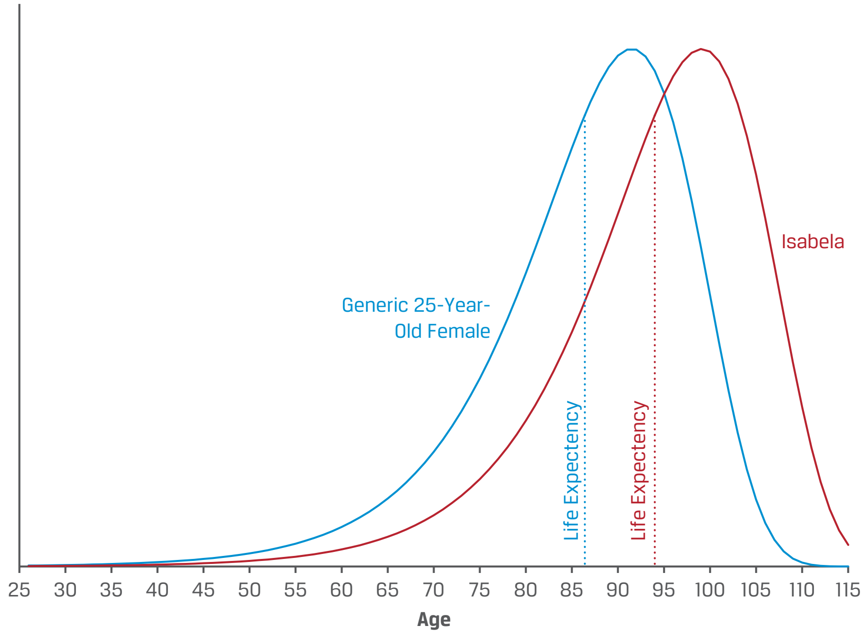
Switching from assets to liabilities, based on information from Isabela, Paula estimates Isabela's nondiscretionary spending (e.g., rent, food) at $40,000 per year. Using the methods presented in chapter 4, Paula's financial planning software modeled Isabela's human capital as 20% equities and 80% fixed income and her nondiscretionary consumption liability as 15% equities and 85% fixed income.

The various net present values displayed in the balance sheet incorporate Paula's belief that Isabela is likely to live longer than the typical 25-year-old woman. As we mentioned earlier, in talking with Isabela, Paula discovered that Isabela's paternal grandfather had died two years earlier at 99, both of her maternal grandparents were still alive at ages 92 and 94, respectively, and two of her paternal great-aunts are still alive at ages 100 and 102. Based on this information, within the financial planning software, Paula indicated that Isabela was likely to live longer than the default life expectancy for a 25-year-old woman. From the Gompertz model that we describe in chapter 3, with our standard set of parameters, the default is age 86.4. Paula overrides the default life expectancy with a personalized estimate of age 94 to reflect the high longevity in Isabela's family. **Exhibit 12.3** shows the impact of raising the life expectancy on the probability distribution of age of death. Importantly, all of Paula's calculations that involve the probability of being alive are personalized for Isabela.

Exhibit 12.2. Isabela's Individual Balance, Age 25, Actual

Assets		Liabilities and Net Worth	
Financial Wealth	$270,500	Liabilities	$1,392,064
Taxable	$250,000	Due to Nondiscretionary Consumption	$1,171,977
Tax-Advantaged	$20,500	Due to Life Insurance	$220,087
Human Capital	$2,767,689	Net Worth	$1,646,126

Exhibit 12.3. Probability Distribution of Age of Death: Generic versus Personalized (Isabela)



Based on Isabela's preferences, current account balances, and salary information, Paula's financial planning software uses the life-cycle model from chapter 6 as the basis for the holistic advice process.

The parent life-cycle model provides immediately implementable advice as well as a big picture view of the lifetime plan and a look at what the future might hold, including the following:

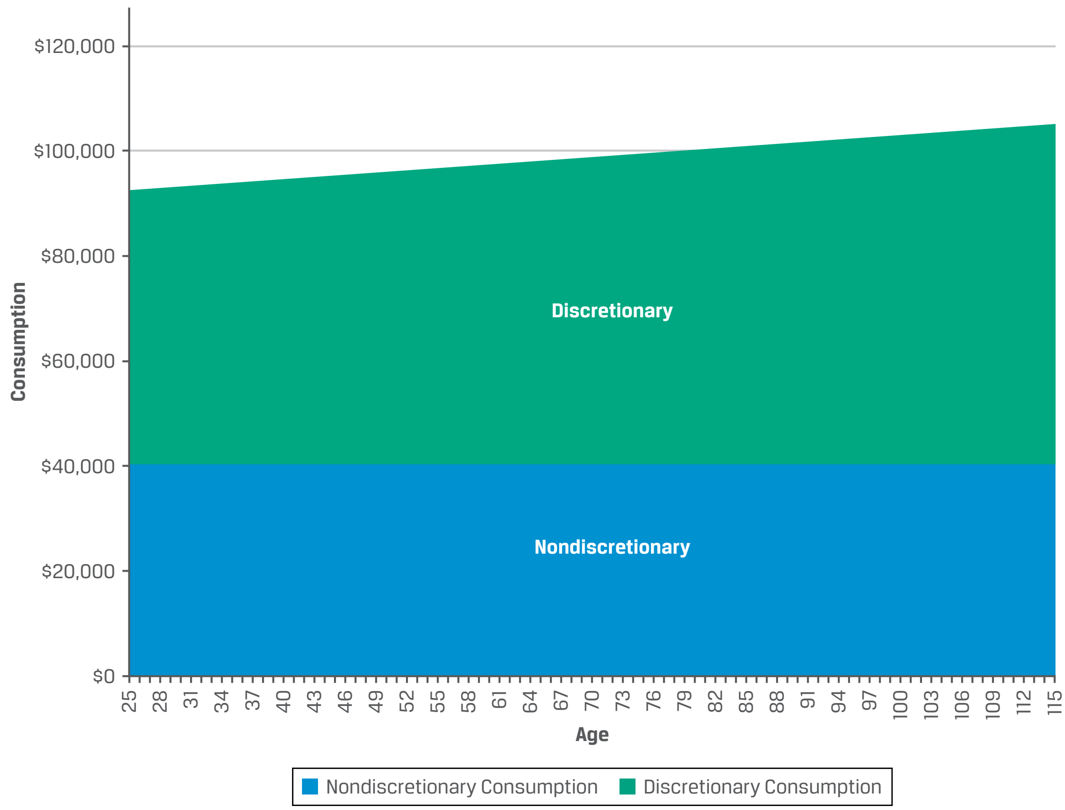
- total consumption schedule
- savings rate schedule
- net-worth estimate
- human capital and financial capital estimates

Paula begins by showing Isabela what her possible lifetime consumption schedule might look like. Paula notes that, with the plan she developed, Isabela will be able to more than pay for her nondiscretionary spending. Importantly, because Isabela will eventually annuitize a portion of her wealth, she will never run out of money, should enjoy an improving standard of living throughout her lifetime (higher amounts of discretionary consumption), and will be able to leave an approximate real bequest of $1,000,000. In other words, in the big picture, Isabela has an actionable plan that meets her needs and wants.

To help Isabela understand the possible evolution of consumption over her lifetime, Paula shows Isabela **Exhibit 12.4**, which projects her nondiscretionary consumption and discretionary consumption over Isabela's lifetime in the absence of risk.

Isabela's desire to leave a bequest is baked into the plan. Based on the values of the parameters related to the desire to leave a bequest that Paula estimates for Isabela, the software calculates a target bequest of just over $1,000,000.[106] For now, we assume that the target bequest is $1,000,000. At age 25, Isabela does not have enough financial assets to leave a $1,000,000 bequest. Looking back at Exhibit 12.2, her current financial assets (age 25) are worth $270,500, resulting in a $729,500 bequest gap.

Exhibit 12.4. Projected Nondiscretionary and Discretionary Consumption



[106]As we mention in chapter 6, Isabela's optimal bequest is $1,157,671. The balance sheets throughout this chapter reflect this as the bequest target.

The balance sheet entry "Due to Life Insurance" under "Liabilities" in Exhibit 12.2 measures something a bit different. We have chosen to show the present value of term life premiums (calculated using a real discount rate of 2.5%) on the full bequest (over $1,000,000) that Isabela would have to pay if her financial assets were not available to fulfill her bequest. (This would be the case if she annuitizes all of her financial assets.) The purpose of this calculation is to track the economic impact of guaranteeing her desired bequest on net worth and, therefore, on discretionary consumption. As we will see in balance sheets for ages 45, 65, and 85, it grows larger over time, getting closer and closer to the size of the bequest.

Returning to the actionable model presented in chapter 5, each year until about age 51, Isabela will purchase term life insurance to cover the shrinking gap between her target bequest and her unannualized financial assets. (This shrinking amount is not shown on the balance sheet where we have chosen to display the full value of guaranteeing the bequest).

The evolution of Isabela's regular (nonannuity) financial wealth in a world without market risk is shown in blue in **Exhibit 12.5**. The solid green line shows the $1,000,000 real bequest target. With Isabela's current financial wealth at $275,500, she will purchase term life insurance until approximately age 51 when her nonannuity financial wealth reaches the $1,000,000 real bequest target. Then, at the expected retirement age of 65 and based on our assumption that immediate fixed-payout annuities become available for purchase, Isabela then annuitizes the part of her financial wealth that is in excess of her $1,000,000 real bequest target.

Exhibit 12.5. Evolution of Isabela's Regular Financial Wealth, Annuity Wealth, and Term Life Insurance

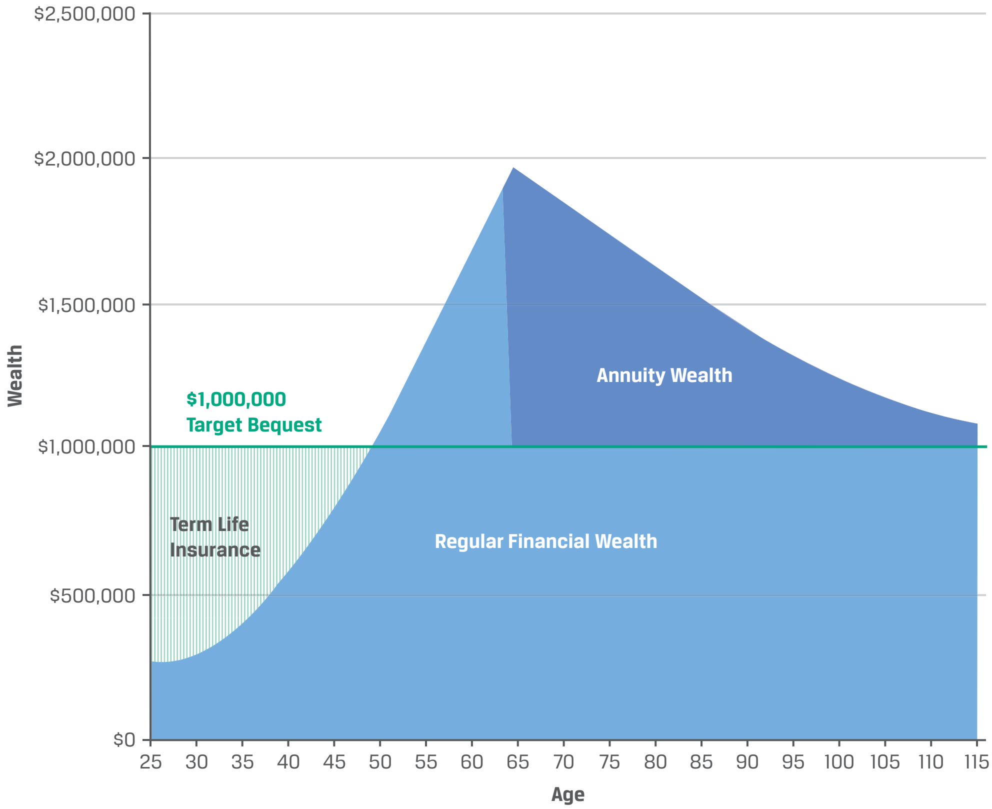


Exhibit 12.5 shows how things would evolve in a riskless world for Isabela. Things such as market returns, Isabela's earnings, unexpected expenditures, and her ability to save all contribute to uncertainty. The complete life-cycle model accounts for that uncertainty.

For engaged clients such as Isabela, Paula likes to present her clients with a variety of additional charts that facilitate a discussion around uncertainty and how the plan might evolve moving forward. Given Paula's focus on the balance sheet, she likes to emphasize the different components (financial assets, human capital, and liabilities) and show how they form an interconnected system that leads to the client's net worth. Based on economic theory, Paula focuses on net worth as the primary indicator of overall financial well-being. Hence, she likes to present her clients with a picture of how net worth is expected to evolve as well as an optimistic scenario and pessimistic scenario, given the plan (see **Exhibit 12.6**). The optimistic scenario shows net worth at the best 75th percentiles of forecasted outcomes, and the pessimistic scenario shows net worth at the worst 25th percentiles of forecasted outcomes.

Although many of Paula's clients find the expected decrease in net worth intuitive, some do not. As such, Paula frequently likes to show an additional chart, such as **Exhibit 12.7**, that illustrates how the client's human capital (in green) and financial capital (in blue) are expected to evolve over time. Even though Exhibit 12.7 focuses on human capital and financial assets, the "Optimistic Scenario" and the "Pessimistic Scenario" are keyed off of the best 75th percentile of net worth and the worst 25th percentile of net worth.

A key aspect of Paula's holistic financial planning process centers on presenting and educating Isabela about her current financial health as represented by her holistic individual balance sheet depicted in Exhibit 12.2.

Given the stochastic nature of the life-cycle model, the software uses Monte Carlo simulation to create output related to the future distribution of wealth, income, and spending. The life-cycle model enables one

Exhibit 12.6. Evolving Distribution of Net Worth, Isabela, Age 25

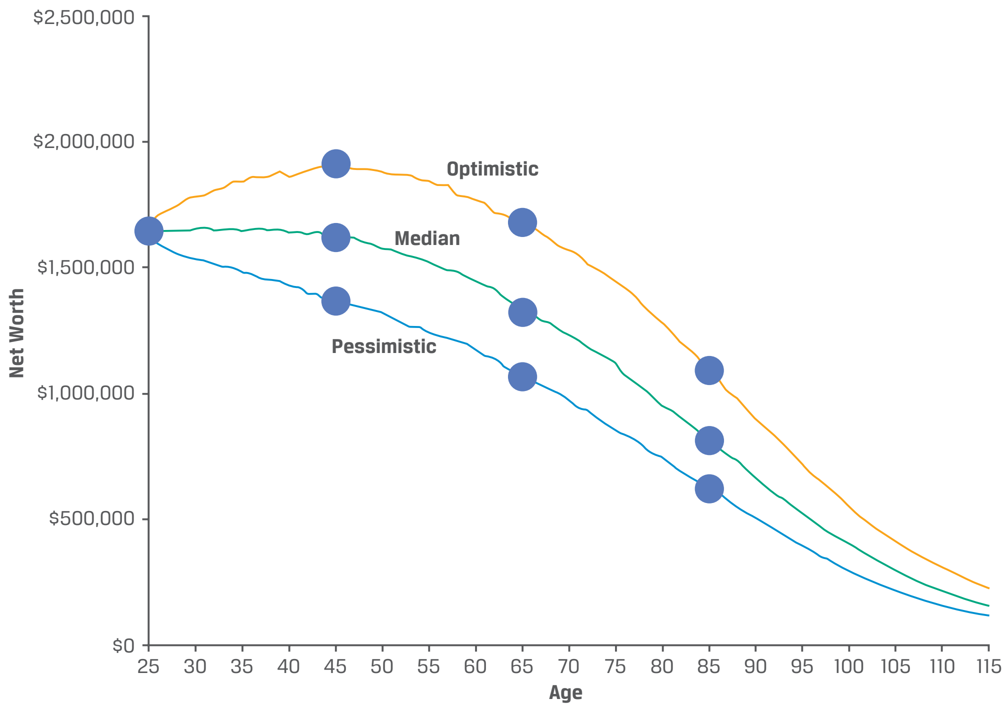
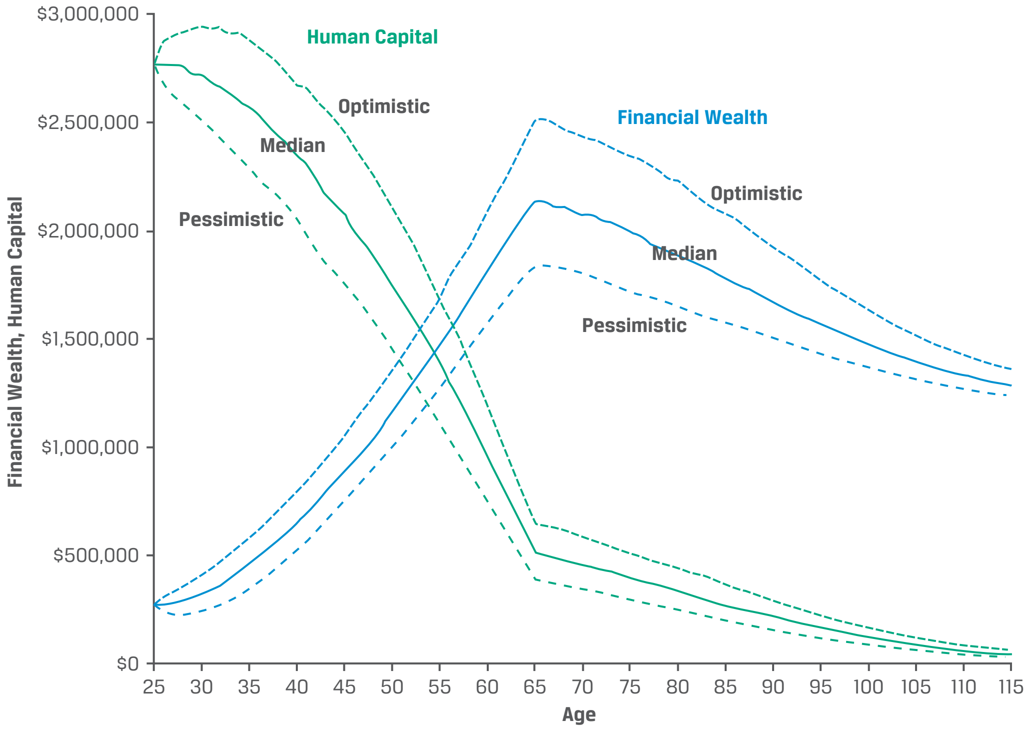


Exhibit 12.7. Evolving Distribution of Human Capital (with Mortality Weighting) and Financial Capital, Isabela, Age 25



to estimate what Isabela's individual balance sheet could look like under various conditions, at *any* point in the future, for *any* combination of the realized values of the various stochastic variables (e.g., market returns, salary changes).

After explaining the current individual balance sheet, Paula presents Isabela with a lifetime plan. An important element of that plan is a forecast of Isabela's individual balance sheet at age 65. **Exhibit 12.8** contains Isabela's *forecasted* individual balance sheet at age 65 based on median outcomes for her net worth.

Exhibit 12.8. Isabela's Individual Balance: Forecast for Age 65 at Age 25, *Median* Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$2,132,176	Liabilities	$1,372,629
Taxable	$687,092	Due to Nondiscretionary Consumption	$800,985
Tax-Advantaged	$1,445,083	Due to Life Insurance	$571,644
Human Capital	$562,683	Net Worth	$1,322,229

Exhibit 12.9. Isabela's Individual Balance: Forecast for Age 65 at Age 25, *Optimistic* Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$2,458,101	Liabilities	$1,318,033
Taxable	$553,172	Due to Nondiscretionary Consumption	$746,389
Tax-Advantaged	$1,904,929	Due to Life Insurance	$571,644
Human Capital	$537,849	Net Worth	$1,677,917

Exhibit 12.10. Isabela's Individual Balance: Forecast for Age 65 at Age 25, *Pessimistic* Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$1,773,071	Liabilities	$994,707
Taxable	$0	Due to Nondiscretionary Consumption	$423,063
Tax-Advantaged	$1,773,071	Due to Life Insurance	$571,644
Human Capital	$288,403	Net Worth	$1,066,768

Throughout Paula's discussions with Isabela, Paula emphasizes that the future is uncertain and that things could go better than expected, but they also could be worse. As such, Paula sometimes likes to show forecasts of the age 65 balance sheet that correspond to the optimistic and pessimistic scenarios presented as **Exhibit 12.9** and **Exhibit 12.10**, respectively.

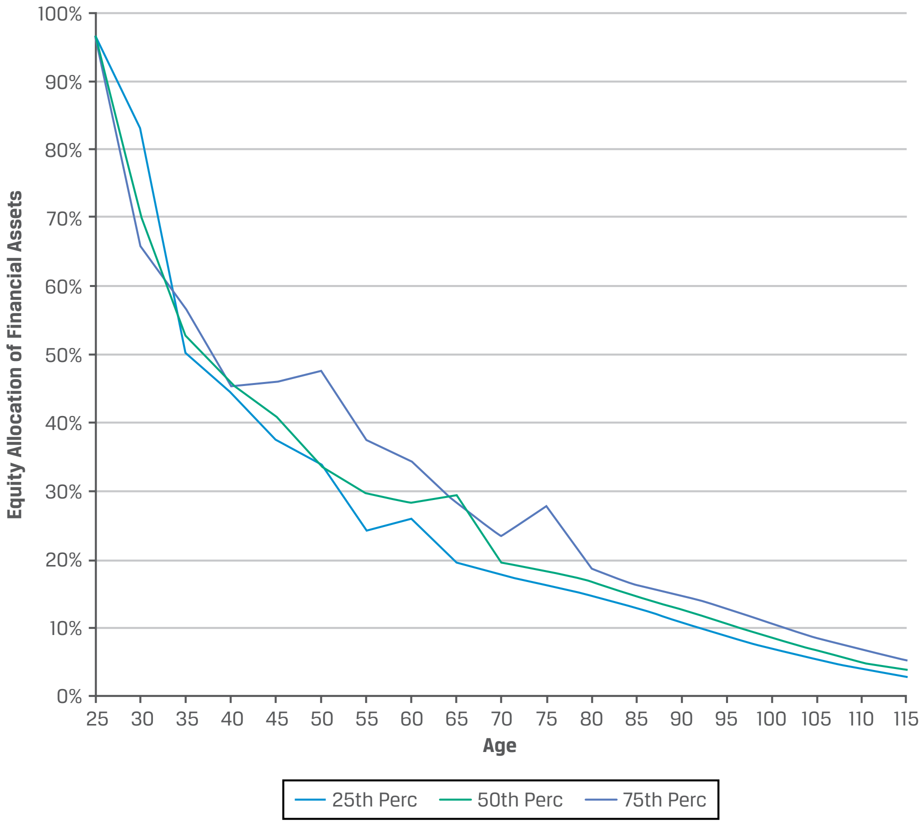
Paula also likes to present her clients with a picture of how the asset allocation for their financial assets may evolve across their lifetimes. Isabela's lifetime equity allocations for financial assets are presented in **Exhibit 12.11**. Again, in addition to the median, it shows the allocations associated with the optimistic and pessimistic scenarios keyed off of net worth.

Because it is advantageous to take advantage of the company match and to locate assets in the tax-advantage account, it is in Isabela's interest to purposely withdraw from her taxable account allowing her tax-advantaged assets to grow tax-deferred. Isabela is not a typical 25-year-old in that having a $250,000 taxable balance is relatively uncommon.

Before moving from the parent life-cycle model to the child net-worth optimization model, we highlight some of the headline recommendations for Isabela for the current year:

- Save $13,667 in her DC plan so as to receive the maximum matching employer contribution of $6,833, which would bring the total contribution to $20,500. However, because Isabela seeks to smooth her consumption between the current year and future years, this is too much savings. So, to smooth consumption, she withdraws from her taxable account.

Exhibit 12.11. Evolving Lifetime Equity Allocations for Financial Assets Projected at Age 25, Different Scenarios



- Based on Isabela's circumstances and preferences, she should try to leave a $1,000,000 bequest. Thus, she should purchase approximately $730,000 of term life insurance to cover her bequest gap, but perhaps slightly more to account for the riskiness of her financial assets.
- Invest her financial assets in a mix consisting of approximately 96% equity/4% fixed-income asset classes.

Isabela's risk tolerance of 35%, the estimated human capital of $2,767,689, and the estimated total liability of $1,392,064 all feed into the child, asset allocation and asset location, and net-worth optimization.

Child Model

Based on the output from the parent model, Paula, the planner, now uses the part of the financial planning software based on chapters 7 and 8 to determine the recommend detailed asset allocation targets for Isabela's taxable account of $250,000 and DC account of $20,500, which collectively represent her target asset allocation for her financial assets. Based on the net-worth optimization, Paula's recommended asset allocation targets are displayed in **Exhibit 12.12**.

Exhibit 12.12. Isabela's Target Asset Allocations and Net-Worth Asset Allocation, Age 25

Asset Class	Financial Wealth			Human Capital	Liabilities	Net Worth
	Taxable	Tax-Advantaged	Overall			
US Large-Cap Stocks	36.1%	0.0%	36.1%	9.4%	12.6%	11.0%
US Mid-/Small-Cap Stocks	0.0%	0.0%	0.0%	4.7%	0.0%	7.9%
Global DM ex-US Stocks	45.1%	0.0%	45.1%	4.7%	0.0%	15.3%
Emerging-Markets Stocks	11.2%	4.1%	15.3%	0.0%	0.0%	2.5%
US Bonds	0.0%	0.0%	0.0%	37.5%	33.7%	34.5%
Inflation-Linked Bonds	0.0%	0.0%	0.0%	37.5%	37. 9%	30.9%
Municipal Bonds	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Global Bonds ex-US	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Cash	0.0%	3.5%	3.5%	6.4%	15.8%	–2.1%
Total Equity	92.9%	4.1%	96.6%	18.7%	12.6%	36.7%
Total	92.9%	7.6%	100.0%	100.0%	100.0%	100.0%
Fraction of Net Worth	16.5%			168.2%	84.7%	100.0%

Note: DM = developed markets.

Notice that overall, the asset allocation of Paula's net worth is 36.7% equities and 63.3% fixed income, which roughly corresponds with her risk tolerance of 35%. Focusing just on her *financial assets*, she is allocated almost entirely to equities (96.6%).

The detailed asset allocations for taxable assets and tax-deferred assets then feed into the grandchild, multi-account, alpha-tracking error optimization model.

Grandchild Model

To save space, we are going to skip over the details of the grandchild model. Ideally, Paula's financial planning and investment management software system would automatically receive the detailed asset allocation targets as inputs into the grandchild model. Then, it would automatically manage the portfolios based on settings that Paula established on behalf of Isabela. Additionally, any of Isabela's nonfinancial or nonpecuniary preferences, determined during the holistic investor profiling exercise, would feed into the grandchild model as additional inputs. For example, Isabela is very concerned about global warming and the health of the ocean, and she would like these values reflected in her portfolio, and in keeping with these preferences, she is willing to receive a slightly lower return to hold a portfolio that reflects her values.

The grandchild multi-account tracking error optimization is designed to be run automatically and frequently (e.g., daily, weekly, monthly). Each time it is run, it considers all of the various trade-offs, seeks the best possible funds for implementing the target asset allocations, harvests tax losses, manages tracking error, and incorporates any trading costs.

Paula's financial planning system is integrated with an individual retirement account platform that offers a wide variety of low-cost, high-quality funds. Based on the relative quality of the investments options offered through Isabela's DC plan, the grandchild model determines if it is in Isabela's best interest to roll all or part of her current DC account balance to the IRA platform. If Isabela had money in an IRA, the grandchild model would also determine whether a reverse rollover from the IRA into the corporate sponsored DC plan was optimal.

Isabela, Age 45

Parent Model

We now fast forward 20 years into the future and Isabela is now age 45. We assume that she has continued to work with Paula the planner to update her plan each year. For simplicity and space considerations, we assume that Isabela's financial preferences depicted in Exhibit 12.1 remain unchanged. As before, Paula anchors the planning process around Isabela's individual balance sheet. **Exhibit 12.13** shows the balance sheet that we project for Isabela at age 45 in the median scenario.[107]

Based on the running of the parent life-cycle model at age 45, the advice at age 45 in the median scenario is as follows:

- Save a real $13,667 that year in her DC plan to receive the maximum employer matching contribution of $6,833, which would bring the total contribution to $20,500. To smooth her consumption between the current year and future years, she should save more by adding to her investments in her taxable account.
- Based on Isabela's approximately $1,000,000 real bequest target, she should purchase approximately $29,000 of term life insurance to cover her remaining bequest gap, but perhaps slightly more to account for the riskiness of her financial assets.
- She should invest her financial assets in a mix consisting of approximately 41% equity/59% fixed-income asset classes (see **Exhibit 12.14**).

Exhibit 12.13. Isabela's Individual Balance, Age 45: Median Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$970,927	Liabilities	$1,291,889
Taxable	$362,226	Due to Nondiscretionary Consumption	$933,385
Tax-Advantaged	$608,701	Due to Life Insurance	$358,504
Human Capital	$1,954,537	Net Worth	$1,633,574

[107]For the purposes of this illustration, we use the median of simulated balance sheets for year 20, where in year 0, the investor is 25 years old.

Exhibit 12.14. Isabela's Target Asset Allocations and Net-Worth Asset Allocation, Age 45: Median Scenario

Asset Class	Financial Wealth			Human Capital	Liabilities	Net Worth
	Taxable	Tax-Advantaged	Overall			
US Large-Cap Stocks	19.1%	0.0%	19.1%	9.4%	10.8%	14.4%
US Mid-/Small-Cap Stocks	0.00%	0.0%	0.0%	4.7%	0.0%	5.5%
Global DM ex-US Stocks	17.5%	0.0%	17.5%	4.7%	0.0%	16.5%
Emerging-Markets Stocks	0.7%	3.8%	4.5%	0.0%	0.0%	2.8%
US Bonds	0.0%	0.0%	0.0%	37.8%	28.9%	21.1%
Inflation-Linked Bonds	0.0%	0.0%	0.0%	37.8%	37.9%	18.3%
Municipal Bonds	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Global Bonds ex-US	0.0%	7.9%	7.9%	0.0%	0.0%	5.0%
Cash	0.0%	51.0%	51.0%	5.6%	27.8%	16.4%
Total Equity	37.3%	3.8%	41.1%	18.9%	10.8%	39.2%
Total	37.3%	62.7%	100.0%	100.0%	100.0%	100.0%
Fraction of Net Worth	62.7%			116.9%	79.6%	100.0%

Looking forward to age 65, we have new optimistic, median, and pessimistic projections of Isabela's individual balance sheet. These are presented as **Exhibits 12.15, 12.16, and 12.17**, respectively. The liability entries may feel a bit unintuitive, but recall that the optimistic, median, and pessimistic scenarios are keyed off of net worth. Furthermore, because human capital and liabilities are both bond-like they are highly correlated with one another.

Exhibit 12.15. Isabela's Individual Balance: Forecast for Age 65 at Age 45, *Median* Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$2,056,632	Liabilities	$1,067,589
Taxable	$254,772	Due to Nondiscretionary Consumption	$495,945
Tax-Advantaged	$1,801,860	Due to Life Insurance	$571,644
Human Capital	$367,613	Net Worth	$1,356,657

Exhibit 12.16. Isabela's Individual Balance: Forecast for Age 65 at Age 45, *Optimistic* Scenario

Assets		**Liabilities and Net Worth**	
Financial Wealth	$2,378,052	Liabilities	$1,283,493
Taxable	$422,627	Due to Nondiscretionary Consumption	$711,849
Tax-Advantaged	$1,955,425	Due to Life Insurance	$571,644
Human Capital	$530,947	Net Worth	$1,625,506

Exhibit 12.17. Isabela's Individual Balance: Forecast for Age 65 at Age 45, *Pessimistic* Scenario

Assets		**Liabilities and Net Worth**	
Financial Wealth	$1,918,500	Liabilities	$1,263,626
Taxable	$407,516	Due to Nondiscretionary Consumption	$691,982
Tax-Advantaged	$1,510,984	Due to Life Insurance	$571,644
Human Capital	$489,683	Net Worth	$1,144,557

Child Model

As before, the output from the life-cycle model as well as the holistic balance sheet (Isabela's current estimated human capital and liability) feed into the child, asset allocation and asset location, and net-worth optimization. The detailed asset allocations for Isabela's taxable and tax-deferred accounts as well as her net-worth allocation are presented in Exhibit 12.14.

The cash allocations are notable. The 27.8% cash allocation of liabilities reflects "Liabilities Due to Life Insurance" in the balance sheet shown in Exhibit 12.13 because we treat term life insurance premiums as being certain. The 51% in cash under Financial Wealth in part defeases the life insurance liability. It can be interpreted as part of a nonannuitized low risk investment earmarked for the bequest. We have chosen not to impose any constraints on the optimization beyond the corresponding budget constraints of the two accounts. In practice, we could see advisers and planners choosing to impose additional constraints, such as limiting cash to 5%.

Grandchild Model

The updated taxable and tax-advantaged asset allocation targets from Exhibit 12.17 continue to feed into the grandchild, multi-account alpha-tracking error optimization.

Isabela, Age 65

Parent Model

We now fast forward another 20 years into the future and Isabela is now age 65, and as planned, has just retired. Additionally, Paula the planner also retired some years earlier and Isabela is now working with Peter the planner using the same financial planning system that Paula had used. Isabela has continued to work with Paula, and now Peter, updating her plan each year. As before, for simplicity and space considerations, we assume that Isabela's financial preferences depicted in Exhibit 12.1 remain unchanged. We also assume that her balance sheet at age 65 is the balance sheet shown in Exhibit 12.15 (the median scenario forecast of age 65 back when Isabela was 45).

The advice based on the life-cycle model has changed somewhat.

- Isabela is no longer working and saving, and her income for consumption will come from Social Security and annuitized assets.
- She has about $2,000,000 in financial assets, which is almost twice her bequest. So, she no longer needs actual life insurance. (We continue to show the "Due to Life Insurance" construct on the balance sheet to track the economic impact of a guaranteed bequest.)
- She will receive Social Security of $19,325 and will receive additional income from variable payout annuities.[108]
- Isabela's annuitized assets will be invested in a mix consisting of approximately 24.3% equity/75.7% fixed-income asset classes.

In a manner similar to what Paula used to do, Peter likes to provide his clients with a median, optimistic, and pessimistic view of the future. **Exhibits 12.18, 12.19, and 12.20** present forecasted age 85 balance sheets for the median, optimistic, and pessimistic scenarios, respectively.

Exhibit 12.18. Isabela's Individual Balance: Forecast for Age 85 at Age 65, *Median* Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$1,759,243	Liabilities	$1,158,559
Taxable	$0	Due to Nondiscretionary Consumption	$313,142
Tax-Advantaged	$1,759,243	Due to Life Insurance	$845,417
Human Capital	$229,988	Net Worth	$830,672

[108]For simplicity, we have ignored that it might be beneficial to delay Social Security.

Exhibit 12.19. Isabela's Individual Balance: Forecast for Age 85 at Age 65, *Optimistic* Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$1,887,942	Liabilities	$1,102,455
Taxable	$0	Due to Nondiscretionary Consumption	$257,038
Tax-Advantaged	$1,887,942	Due to Life Insurance	$845,417
Human Capital	$196,364	Net Worth	$981,851

Exhibit 12.20. Isabela's Individual Balance: Forecast for Age 85 at Age 65, *Pessimistic* Scenario

Assets		Liabilities and Net Worth	
Financial Wealth	$1,615,674	Liabilities	$1,067,387
Taxable	$0	Due to Nondiscretionary Consumption	$221,970
Tax-Advantaged	$1,615,674	Due to Life Insurance	$845,417
Human Capital	$161,712	Net Worth	$709,999

Child Model

As before, the output from the life-cycle model as well as the holistic balance sheet (31.4% equity/68.6% fixed-income asset allocation for financial assets, Isabela's estimated human capital, and Isabela's nondiscretionary liability) feed into the child, asset allocation and asset location, and net-worth optimization. The detailed asset allocations for Isabela's taxable and tax-deferred accounts as well as her net-worth allocation at age 65 in the median scenario are presented in **Exhibit 12.21**.

Exhibit 12.21. Isabela's Target Asset Allocations and Net-Worth Asset Allocation, Age 65: Median Scenario

Asset Class	Financial Wealth			Human Capital	Liabilities	Net Worth
	Taxable	Tax-Advantaged	Overall			
US Large-Cap Stocks	12.4%	0.0%	12.4%	10.0%	7.00%	16.2%
US Mid-/Small-Cap Stocks	0.0%	1.9%	1.9%	5.0%	0.0%	4.2%
Global DM ex-US Stocks	0.0%	7.7%	7.7%	5.0%	0.0%	13.2%
Emerging-Markets Stocks	0.0%	2.3%	2.3%	0.0%	0.0%	3.5%
US Bonds	0.0%	9.8%	9.8%	40.0%	18.6%	10.9%
Inflation-Linked Bonds	0.0%	6.6%	6.6%	40.0%	20.9%	4.1%
Municipal Bonds	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Global Bonds ex-US	0.0%	9.1%	9.1%	0.0%	0.0%	14.0%
Cash	0.0%	50.2%	50.2%	0.0%	53.5%	33.9%
Total Equity	12.4%	11.9%	24.3%	20.0%	7.00%	37.1%
Total	12.4%	87.6%	100.0%	100.0%	100.0%	100.0%
Fraction of Net Worth	153.4%			27.00%	80.4%	100.0%

Grandchild Model

The updated taxable and tax-deferred asset allocation targets from Exhibit 12.21 continue to feed into the grandchild, multi-account alpha-tracking error optimization.

End-to-End Wrap

Although elements of Paula's (and later Peter's) hypothetical financial planning and investment management system already exist, we are unaware of such a system. We believe the methods presented in this book provide a blueprint for a system capable of revolutionizing financial planning as we know it. With PhD programs, such as those of Kansas State University, Ohio State University, Texas Tech University, and University of Georgia, producing skilled people with doctorates in financial planning, the number of people capable of building, improving, and using such system is reaching a critical mass. With this in mind, we conclude with our key points and calls for action.

Eight Key Points and Corresponding Calls for Action

We conclude with eight key points and corresponding calls for action.

1. A Comprehensive Normative Theoretical Framework for Lifetime Advice

All too often financial planners are forced to rely on *ad hoc* frameworks and heuristics rather than the powerful theories of economics, finance, and insurance. Steeped in rigorous theory, we created a comprehensive and actionable framework for providing optimal financial advice.

- Some practitioners interpret behavioral finance and its numerous examples of irrational investor behavior and decision making as a reason to dismiss the value of, and lessons from, models that provide the optimal solution. Behavioral finance does not diminish the value of a comprehensive normative framework for providing optimal financial advice.
- Based on an understanding of investor behavior (irrational and otherwise), practitioners should find ways to coach and nudge investors toward optimal, holistic advice.

2. The Importance of Life-Cycle Finance

Life-cycle finance solves a first-order problem, while investment-only advice solves a second-order problem.

- Practitioners need to embrace and elevate life-cycle finance as the guiding framework for financial advice.
- Curriculum creators need to change curriculums not only to include life-cycle finance but also to frame it appropriately as the most important element of personal finance.

3. Moving beyond Risk Tolerance and a Risk Profile to an Investor Profile

The industry is largely focused on risk tolerance and the investor's risk profile. These are subsets of a more holistic investor profile, which includes a variety of additional financial preferences and nonfinancial preferences.

- Practitioners should develop a complete investor profile that accounts for additional financial preferences and nonfinancial preferences.
- Researchers need to develop methods for estimating the full range of investor preferences and the creators of software tools for advisers and wealth managers need to incorporate such methods into their software.

4. A Holistic Individual Economic Balance Sheet Approach

A holistic approach based the investor's economic balance sheet is superior to myopic account-specific and investment-centric approaches. The only way to have a complete picture of an investor's financial situation is through a holistic individual balance sheet.

- Practitioners need tools for producing economic balance sheets for their clients.
- The creators of software tools for advisers need to add functionality, making it easy to estimate an investor's economic balance sheet.
- Investors should expect to receive periodic individual balance sheets on an annual basis.

5. Coherent and Consistent Financial Planning Systems

Regardless of the scope of application, the various elements of a financial planning system need to work together in a coherent and consistent fashion.

- Practitioners need to review their financial planning systems, especially if they are coming from multiple vendors, to ensure consistency with one another.
- Software creators need to make sure their systems are coherent and when they are not, need to make that clear.

6. Making Asset Allocation a Dynamic Process That Responds to the Changing Circumstances of the Investor

The combined effect of our parent and child models is that asset allocation is a personalized dynamic process. Target-date funds are a poor substitute for this process.

7. Inspired Regulations

Well-meaning regulation and policies often have the unintended consequences of leading advisers and wealth managers to act narrowly when developing and recommending a portfolio.

- Regulations, such as those of the Canadian Securities Administrators and European Securities and Markets Authority, that inadvertently contribute to matching specific investments to investors without a holistic perspective, should be revisited.
- Similarly, the home offices of networks of advisers, if they mandate processes for their advisers that inadvertently contribute to matching specific investments to investors without a holistic perspective, should be revisited.

8. Joining Life-Cycle and Single-Period Models

To our knowledge, this book presents the most complete effort, so far, to connect models of life-cycle finance with single-period optimization models. While we believe what we have put forth is relatively complete and powerful, we hope we are opening the door for a new path of research.

- Our aspiration is that researchers embrace and improve this path!

REFERENCES

Arrow, K. J. 1965. "The Theory of Risk Aversion." In *Aspects of the Theory of Risk Bearing*, ed. Yrjo Jahnssonin Saatio, 90–109. Chicago: Markham Publ. Co.

Baker, Malcolm, Daniel Bergstresser, George Serafeim, and Jeffrey Wurgler. 2020. "Financing the Response to Climate Change: The Pricing and Ownership of U.S. Green Bonds." 1 August. Older version of paper available at http://papers.ssrn.com.

Becker, Gary S. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 3rd ed. Chicago, IL: University of Chicago Press. doi:10.7208/chicago/9780226041223.001.0001

Baierl, Gary, and Peng Chen. 2000. "Choosing Managers and Funds: Maximizing Your Alpha without Sacrificing Your Target." *Journal of Portfolio Management* 26 (2): 47–53. doi:10.3905/jpm.2000.319745

Bellman, Richard. 1957. *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Bengen, William P. 1994. "Determining Withdrawal Rates Using Historical Data." *Journal of Financial Planning* 7 (4): 171–80.

Blay, Kenneth A., and Harry M. Markowitz. 2016. "Tax-Cognizant Portfolio Analysis: A Methodology for Maximizing After-Tax Wealth." *Journal of Investment Management* 14 (1): 26–64.

Blanchett, David M., and Paul D. Kaplan. 2013. "Alpha, Beta, and Now… Gamma." *Journal of Retirement* 1 (2): 29–45. doi:10.3905/jor.2013.1.2.029

Blanchett, David M., and Philip U. Straehl. 2015. "No Portfolio Is an Island." *Financial Analysts Journal* 71 (3): 15–33.

Bodie, Zvi, Dennis McLeavey, and Laurence B. Siegel, eds. 2008. *The Future of Life-Cycle Savings and Investing*. 2nd ed. Charlottesville, VA: Research Foundation of CFA Institute. https://www.cfainstitute.org/-/media/documents/book/rf-publication/2008/rfv2008n1.ashx.

Bodie, Zvi, Robert C. Merton, and William F. Samuelson. 1992. "Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model." *Journal of Economic Dynamics and Control* 16: 427–49. doi:10.1016/0165-1889(92)90044-F

Bodie, Zvi, Jonathan Treussard, and Paul Willen. 2008. "The Theory of Optimal Life-Cycle Savings and Investing." In *The Future of Life-Cycle Saving and Investing*. 2nd ed., ed. Zvi Bodie, Dennis McLeavey, and Laurence Siegel, 19–37. Charlottesville, VA: Research Foundation of CFA Institute. Available at www.cfapubs.org/toc/ rf/2007/2007/3.

Brayman, Shawn, Michael Finke, Ellen Bessner, John Grable, Paul Griffin, and Rebecca Clement. 2015. "Current Practices for Risk Profiling in Canada and Review of Global Best Practices." Prepared for the Investor Advisory Panel of the Ontario Securities Commission. PlanPlus, Inc. Acquired by Morningstar.

Charupat, Narat, Huaxiong Huang, and Moshe A. Milevsky. 2012. *Strategic Financial Planning over the Lifecycle*. New York: Cambridge University Press. doi:10.1017/CBO9780511807336

Cooper, Lisette, Jeremy Evnine, Jeff Finkelman, Kate Huntington, and David Lynch. 2016. "Social Finance and the Postmodern Portfolio: Theory and Practice." *Journal of Wealth Management* 18 (4): 9–21. doi:10.3905/jwm.2016.18.4.009

Daga, Ankul, Timothy Smart, and David C. Pakula. 2023 (forthcoming). "The Multi-Goal Framework: Why Practitioners Have Not Adopted the Lifecycle Model-Yet." *Journal of Retirement* 10 (3). doi:10.3905/jor.2023.10.3.071

Dammon, Robert M., Chester S. Spatt, and Harold H. Zhang. 2004. "Optimal Asset Location and Allocation with Taxable and Tax-Deferred Investing." *Journal of Finance* 59 (3): 999–1037. doi:10.1111/j.1540-6261.2004.00655.x

Deaton, Angus S. 2005. "Franco Modigliani and the Life Cycle Theory of Consumption." Princeton University, March. doi:10.2139/ssrn.686475

De Finetti, Bruno. 2006. "'The Problem of Full-Risk Insurances,' Translated by Luca Barone." *Journal of Investment Management* 4: 19–43.

Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (4): 937–69. doi:10.2307/1913778

Fama, Eugene F. 1970. "Multiperiod Consumption-Investment Decisions." *American Economic Review* 60 (1): 163–74.

Fama, Eugene F. 2014. "Two Pillars of Asset Pricing." *American Economic Review* 104 (6): 1467–85. doi:10.1257/aer.104.6.1467

Fama, Eugene F., and Kenneth R. French. 2007. "Disagreement, Tastes, and Asset Pricing." *Journal of Financial Economics* 83 (3): 667–89. doi:10.1016/j.jfineco.2006.01.003

Fidelity Magellan Fund Prospectus. 2019. 30 May. https://institutional.fidelity.com/app/proxy/content?literatureURL=/B-CMAG.PDF

Fisher, Irving. 1930. *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*. New York: Macmillan.

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press. doi:10.1515/9780691188485

Friedman, Milton. 1970. "A Friedman Doctrine: The Social Responsibility of Business Is to Increase Its Profits." *New York Times Magazine*, September 13.

Grinold, Richard C. 1989. "The Fundamental Law of Active Management." *Journal of Portfolio Management* 15 (3): 30–37. doi:10.3905/jpm.1989.409211

Grinold, Richard C. 1994. "Alpha Is Volatility Times IC Times Score." *Journal of Portfolio Management* 20 (4): 9–16. doi:10.3905/jpm.1994.409482

Grinold, Richard, and Ronald Kahn. 2000. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. New York: McGraw-Hill.

Habib, F., H. Huang, and M. A. Milevsky. 2017. "Approximate Solutions to Retirement Spending Problems and the Optimality of Ruin." Working paper, 31 March. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2944125. doi:10.2139/ssrn.2944125

Huang, Jennifer. 2008. "Taxable and Tax-Deferred Investing: A Tax-Arbitrage Approach." University of Texas at Austin, McCombs Research Paper Series No. FIN-04-07. doi:10.1093/rfs/hhn064

Ibbotson, Roger G., Thomas M. Idzorek, Paul D. Kaplan, and James X. Xiong. 2018. *Popularity: A Bridge between Classical and Behavioral Research*. Charlottesville, VA: CFA Institute Research Foundation.

Ibbotson, Roger G., Moshe A. Milevsky, Peng Chen, and Kevin X. Zhu. 2007. *Lifetime Financial Advice: Human Capital, Asset Allocation, and Insurance*. Charlottesville, VA: Research Foundation of CFA Institute.

Idzorek, Thomas M. 2007. "A Step-by-Step Guide to the Black–Litterman Model." In *Forecasting Expected Returns in the Financial Markets*, ed. Stephen Satchell, 17–38. London, UK: Academic Press. doi:10.1016/B978-075068321-0.50003-0

Idzorek, Thomas M. 2022. "Personalized Portfolio Optimization with Multiple Tax-Exempt, Tax-Deferred, and Taxable Accounts." Working Paper.

Idzorek, Thomas M. 2023. "Personalized Multiple Account Portfolio Optimization." *Financial Analysts Journal* 79 (3): 155–70. doi:10.1080/001518X.2023.2212581

Idzorek, Thomas M., and David M. Blanchett. 2019. "LDI for Individual Portfolios." *Journal of Investing* 28 (1): 31–54. doi:10.3905/joi.2019.28.1.031

Idzorek, Thomas M., and Paul D. Kaplan. 2022. "Forming ESG-Oriented Portfolios: A Popularity Approach." *Journal of Investing*, 30th Anniversary Special Issue 31 (4): 63–75. doi:10.3905/joi.2022.31.4.063

Idzorek, Thomas M., and Paul D. Kaplan. Kaplan, and Roger G. Ibbotson. 2021. "The Popularity Asset Pricing Model." Morningstar working paper, October 25. SSRN: https://ssrn.com/abstract=3451554 or doi:10.2139/ssrn.3451554

Idzorek, Thomas M., Paul D. Kaplan, and Roger G. Ibbotson. 2023. "The CAPM, APT, and PAPM." Working paper, September 8. SSRN: https://ssrn.com/abstract=4566414 or doi:10.2139/ssrn.4566414

iShares Core U.S. Aggregate Bond ETF Summary Prospectus. 2019. 28 June. https://www.ishares.com/us/literature/summary-prospectus/sp-ishares-core-us-aggregate-bond-etf-2-28.pdf

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–92. doi:10.2307/1914185

Kahneman, Daniel, and Mark W. Riepe. 1998. "Aspects of Investor Psychology." *Journal of Portfolio Management* 24 (4): 52–65. doi:10.3905/jpm.1998.409643

Kaplan, Paul D. 1986. "Asset Prices, Finite Horizons, and Proxies for the Intertemporal Marginal Rate of Substitution." PhD diss., Department of Economics, Northwestern University.

Kaplan, Paul D. 2006. "Asset Allocation with Annuities for Retirement Income Management." *Journal of Wealth Management* 8: 27–40.

Kaplan, Paul D. 2014. "Clearing Up the Great Confusion." *Morningstar*, October/November.

Kaplan, Paul D. 2015a. "When Will I Go?" *Morningstar*. February/March.

Kaplan, Paul D. 2015b. "In Search of the Rational Investor." *Morningstar*, April/May.

Kaplan, Paul D. 2020a. "A Kernel of Truth." *Morningstar*, Spring.

Kaplan, Paul D. 2020b. "A Better Retirement Spending Rule for Everyone." *Morningstar*, Q2.

Kaplan, Paul D. 2020c. "Solving the Asset Location Problem, Part I." *Morningstar*, Q3.

Kaplan, Paul D. 2020d. "Solving the Asset Location Problem, Part II." *Morningstar*, Q4.

Kaplan, Paul D. 2021. "The Two Sides of ESG Investing." *Morningstar*, Q2.

Kaplan, Paul D. 2022a. "Me Versus the Kids." *Morningstar*, Q2.

Kaplan, Paul D. 2022b. "Immediate Variable Annuities: Real and Imagined." *Morningstar*, Q3.

Kaplan, Paul D, and David M. Blanchett. 2020. "A Dynamic Spending Rule for All Retirees." Morningstar working paper, May.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money*. London, UK: Palgrave Macmillan.

Klement, Joachim, ed. 2018. *Risk Profiling and Tolerance: Insights for the Private Wealth Manager*. Charlottesville, VA: CFA Institute Research Foundation.

Leibowitz, Martin L. 1987. "Pension Asset Allocation through Surplus Management." *Financial Analysts Journal* 43 (2): 29–40. doi:10.2469/faj.v43.n2.29

Leibowitz, Martin L., and Roy D. Henriksson. 1988. "Portfolio Optimization within a Surplus Framework." *Financial Analysts Journal* 44 (2): 43–51. doi:10.2469/faj.v44.n2.43

Levy, Haim, and Harry M. Markowitz. 1979. "Approximating Expected Utility by a Function of Mean and Variance." *American Economic Review* 69 (3): 308–17.

Lewis, Michael. 2004. *Moneyball*. New York: W. W. Norton.

Lucas, Robert E., Jr. 1978. "Asset Prices in an Exchange Economy." *Econometrica* 46 (6): 1429–45. doi:10.2307/1913837

Markowitz, Harry M. 1952. "Portfolio Selection." *Journal of Finance* 7: 77–91.

Markowitz, Harry. 1959. *Portfolio Selection: Efficient Diversification of Investments*. Hoboken, NJ: John Wiley.

Markowitz, Harry M. 1999. "The Early History of Portfolio Theory: 1600–1960." *Financial Analysts Journal* 55 (4): 5–16. doi:10.2469/faj.v55.n4.2281

Markowitz, Harry M., and Kenneth A. Blay. 2014. *Risk-Return Analysis, Volume 1*. New York: McGraw-Hill.

Markowitz, Harry M., Sam L. Savage, and Paul D. Kaplan. 2010. "What Does Harry Markowitz Think?" *Morningstar Advisor*, April/May.

Merton, Robert C. 1969. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case." *Review of Economics and Statistics* 51 (3): 247–57. doi:10.2307/1926560

Merton, Robert C. 1971. "Optimum Consumption and Portfolio Rules in a Continuous-Time Model." *Journal of Economic Theory* 3 (4): 373–413. doi:10.1016/0022-0531(71)90038-X

Merton, Robert C. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41 (5): 867–87. doi:10.2307/1913811

Milevsky, Moshe A. 2006. *The Calculus of Retirement Income*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511753855

Milevsky, Moshe A. 2012a. *The 7 Most Important Equations for Your Retirement*. Toronto: Wiley.

Milevsky, Moshe A. 2012b. *Are You a Stock or Bond? Identify Your Own Human Capital for a Secure Financial Future, Updated and Revised*. Upper Saddle River, NJ: FT Press.

Milevsky, Moshe A., and Huaxiong Huang. 2011. "Spending Retirement on Planet Vulcan: The Impact of Longevity Risk Aversion on Optimal Withdrawal Rates." *Financial Analysts Journal* 67 (2): 45–58. doi:10.2469/faj.v67.n2.2

Milevsky, Moshe A., and Chris Robinson. 2005. "A Sustainable Spending Rate without Simulation." *Financial Analysts Journal* 61 (6): 89–100. doi:10.2469/faj.v61.n6.2776

Modigliani, Franco. 1966. "The Life Cycle Hypothesis of Saving, the Demand for Wealth and the Supply of Capital." *Social Research* 33 (2): 160–217.

Nevins, Daniel. 2004. "Goals-Based Investing: Integrating Traditional and Behavioral Finance." *Journal of Wealth Management* 6 (4): 8–23. doi:10.3905/jwm.2004.391053

Parker, Franklin J. 2021. "A Goals-Based Theory of Utility." *Journal of Behavioral Finance* 22 (1): 10–25. doi:10.1080/15427560.2020.1716359

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor. 2021. "Sustainable Investing in Equilibrium." *Journal of Financial Economics* 142 (2): 550–71. doi:10.1016/j.jfineco.2020.12.011

Pedersen, Lasse Heje, Shaun Fitzgibbons, and Lukasz Pomorski. 2021. "Responsible Investing: The ESG-Efficient Frontier." *Journal of Financial Economics* 142 (2): 572–97. doi:10.1016/j.jfineco.2020.11.001

Pimco Total Return Fund Summary Prospectus. 2019. 31 July (supplemented 11 December).

Pratt, J. W. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* 32 (1/2): 122–136. doi:10.2307/1913738

Roy, A. D. 1952. "Safety First and the Holding of Assets." *Econometrica* 20 (3): 431–49. doi:10.2307/1907413

Samuelson, Paul A. 1938. "A Note on the Pure Theory of Consumers' Behaviour." *Econometrica* 5 (17): 61–71. doi:10.2307/2548836

Samuelson, Paul A. 1948. "Consumption Theory in Terms of Revealed Preference." *Econometrica* 15 (60): 243–53. doi:10.2307/2549561

Samuelson, Paul A. 1969. "Lifetime Portfolio Selection by Dynamic Stochastic Programming." *Review of Economics and Statistics* 51 (3): 239–46. doi:10.2307/1926559

Sharpe, William F. 1974. "Imputing Expected Security Returns from Portfolio Composition." *Journal of Financial and Quantitative Analysis* 9 (3): 463–72. doi:10.2307/2329873

Sharpe, William F. 1990. "Asset Allocation." Chap. 7 in *Managing Investment Portfolios: A Dynamic Process*, ed. John L. Maginn and Donald L. Tuttle. New York: Warren, Gorham & Lamont.

Sharpe, William F., and Lawrence G. Tint. 1990. "Liabilities—A New Approach." *Journal of Portfolio Management* 16 (2): 5–10. doi:10.3905/jpm.1990.409248

Shefrin, Hersh, and Meir Statman. 2000. "Behavioral Portfolio Theory." *Journal of Financial and Quantitative Analysis* 35 (2): 127–51. doi:10.2307/2676187

Siegel, Laurence B. 2008. Foreword to *The Future of Life-Cycle Savings and Investing*, 2nd edition, ed. Zvi Bodie, Dennis McLeavey, and Laurance B. Siegel Charlottesville VA: CFA Institute Research Foundation.

Siegel, Laurence B., and M. Barton Waring. 2004. "TIPS, the Dual Duration, and the Pension Plan." *Financial Analysts Journal* 60 (5): 52–64. doi:10.2469/faj.v60.n5.2656

Stewart, Scott D. 2013. *Manager Selection*. Charlottesville, VA: CFA Institute Research Foundation.

Straehl, Philip U., Robert ten Brincke, and Carlos Gutierrez Mangas. 2023. "No DI Portfolio Is an Island." Working paper, December 8.

Thaler, Richard. 1985. "Mental Accounting and Consumer Choice." *Marketing Science* 4 (3): 199–214. doi:10.1287/mksc.4.3.199

Tobin, James. 1958. "Liquidity Preferences as Behavior towards Risk." *Review of Economic Studies* 25 (2): 65–86. doi:10.2307/2296205

Vanguard 500 Index Fund Prospectus. 2019. 26 April. https://www.vanguard.com/pub/Pdf/p040.pdf

Von Neumann, John, and Oskar Morgenstern. (1944) 1967. *Theory of Games and Economic Behavior*. 3rd ed. Reprint, Princeton, NJ: Princeton University Press.

Waring, M. Barton. 2004a. "Liability-Relative Investing." *Journal of Portfolio Management* 30 (4): 8–20. doi:10.3905/jpm.2004.8

Waring, M. Barton. 2004b. "Liability-Relative Investing II: Surplus Optimization with Beta, Alpha, and an Economic View of the Liability." *Journal of Portfolio Management* 31 (1): 40–53. doi:10.3905/jpm.2004.443318

Waring, M. Barton, and Laurence B. Siegel. 2003. "The Dimensions of Active Management." *Journal of Portfolio Management* 29 (3): 35–51. doi:10.3905/jpm.2003.319882

Waring, M. Barton, and Laurence B. Siegel. 2015. "The Only Spending Rule Article You Will Ever Need." *Financial Analysts Journal* 71 (1): 91–107. doi:10.2469/faj.v71.n1.2

Waring, M. Barton, and Laurence B. Siegel. 2018. "What Investment Risk Means to You, Illustrated: Strategic Asset Allocation, the Budget Constraint, and the Volatility of Spending during Retirement." *Journal of Retirement* 6 (2): 7–26. doi:10.3905/jor.2018.1.041

Waring, M. Barton, and Sunder R. Ramkumar. 2008. "Forecasting Fund Manager Alphas: The Impossible Just Takes Longer." *Financial Analysts Journal* 64 (2): 65–80. doi:10.2469/faj.v64.n2.12

Waring, M. Barton, and Duane Whitney. 2009. "An Asset-Liability Version of the Capital Asset Pricing Model with a Multi-Period Two-Fund Theorem." *Journal of Portfolio Management* 35 (4): 111–30. doi:10.3905/JPM.2009.35.4.111

Waring, M. Barton, Duane Whitney, John Pirone, and Charles Castille. 2000. "Optimizing Manager Structure and Budgeting Manager Risk." *Journal of Portfolio Management* 26 (3): 90–104. doi:10.3905/jpm.2000.319719

Yaari, Menahem E. 1965. "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer." *Review of Economic Studies* 32 (2): 137–50. doi:10.2307/2296058

Wilcox, Jarrod, Jeffrey E. Horvitz, and Dan DiBartolomeo. 2006. *Investment Management for Taxable Private Investors*. Charlottesville, VA: CFA Institute Research Foundation.

Zerbib, Olivier David. 2019. "A Sustainable Capital Asset Pricing Model (S-CAPM): Evidence from Green Investing and Sin Stock Exclusion." 13 October. Available at SSRN. doi:10.2139/ssrn.3455090.

Named Endowments

CFA Institute Research Foundation acknowledges with sincere gratitude the generous contributions of the Named Endowment participants listed below.

Gifts of at least US$100,000 qualify donors for membership in the Named Endowment category, which recognizes in perpetuity the commitment toward unbiased, practitioner-oriented, relevant research that these firms and individuals have expressed through their generous support of CFA Institute Research Foundation.

Ameritech
Anonymous
Robert D. Arnott
Theodore R. Aronson, CFA
Asahi Mutual Life Insurance Company
Batterymarch Financial Management
Boston Company
Boston Partners Asset Management, L.P.
Gary P. Brinson, CFA
Brinson Partners, Inc.
Capital Group International, Inc.
Concord Capital Management
Dai-Ichi Life Insurance Company
Daiwa Securities
Mr. and Mrs. Jeffrey Diermeier
Gifford Fong Associates
John A. Gunn, CFA
Investment Counsel Association of America, Inc.
Jacobs Levy Equity Management
Jon L. Hagler Foundation
Long-Term Credit Bank of Japan, Ltd.
Lynch, Jones & Ryan, LLC
Meiji Mutual Life Insurance Company
Miller Anderson & Sherrerd, LLP
John B. Neff, CFA
Nikko Securities Co., Ltd.
Nippon Life Insurance Company of Japan
Nomura Securities Co., Ltd.
Payden & Rygel
Provident National Bank
Frank K. Reilly, CFA
Salomon Brothers
Sassoon Holdings Pte. Ltd.
Scudder Stevens & Clark
Security Analysts Association of Japan
Shaw Data Securities, Inc.
Sit Investment Associates, Inc.
Standish, Ayer & Wood, Inc.
State Farm Insurance Company
Sumitomo Life America, Inc.
T. Rowe Price Associates, Inc.
Templeton Investment Counsel Inc.
Frank Trainer, CFA
Travelers Insurance Co.
USF&G Companies
Yamaichi Securities Co., Ltd.

Senior Research Fellows

Financial Services Analyst Association

For more on upcoming CFA Institute Research Foundation publications and webcasts, please visit www.cfainstitute.org/research/foundation.