Investing for the Long Run

Andrew Ang

DOI:10.1093/acprof:oso/9780199959327.003.0004

Abstract and Keywords
The foundation of long-term investing is to rebalance to fixed asset positions, which are determined in a one-period portfolio choice problem where the asset weights reflect the investor’s attitude toward risk. Rebalancing is a counter-cyclical strategy that buys low and sells high. It worked well even during the Great Depression of the 1930s and the Lost Decade of the 2000s. Rebalancing goes against investors’ behavioral instincts and is also a short volatility strategy.

Keywords: portfolio choice, rebalancing, Merton, dynamic programming, buy-and-hold, hedging demand, counter-cyclical investing, rebalancing bands, short volatility, liability hedging, diversification return, variance drain, Kelly (1956) rule

Chapter Summary
The foundation of long-term investing is to rebalance to fixed asset positions, which are determined in a one-period portfolio choice problem where the asset weights reflect the investor’s attitude toward risk. Rebalancing is a counter-cyclical strategy that buys low and sells high. It worked well even during the Great Depression of the 1930s and the Lost Decade of the
2000s. Rebalancing goes against investors’ behavioral instincts and is also a short volatility strategy.

1. Stay the Course?
In April 2009, just after the worst of the financial crisis, Amy Harrison, an independent investment advisor, prepared to meet with her client, Amelia Daniel.¹ Harrison had been introduced to Daniel three years earlier. At that time, Daniel had just sold her medical information company, Daniel Health Systems, for $10 million cash. Daniel had also recently divorced. She felt that both the liquidity event and the end of a chapter in her personal life would allow her to start afresh on new, smaller ventures.

After their first meeting, Harrison drafted an investment policy statement (IPS) for Daniel that:

1. Described Daniel’s understanding of risk and set her risk tolerance,
2. Identified Daniel’s intermediate and long-term goals and her preferences and constraints,
3. Crafted a long-term investment plan,
4. Served as a reminder of guidelines to be used in achieving her goals, and
5. Defined the investment and monitoring process.

(p.114) Harrison’s first year of working with her new client had gone smoothly. It took some convincing, however, for Daniel to follow Harrison’s advice. Like many entrepreneurs, Daniel had built up her wealth by holding a concentrated portfolio, essentially all in her own company. But Harrison’s advice was rooted in diversification and optimal asset allocation based on reducing risk and maximizing return. Daniel had essentially no liabilities; her parents were well off, and she had no children and no plans for having any. She lived modestly, and her expenses were covered by the salary the acquiring company was paying her to stay on as a consultant. Given her entrepreneurial background, Daniel was comfortable taking risk and had a long-term investment focus. Thus, Harrison recommended that most of Daniel’s portfolio be evenly split between a myopic (growth or market-oriented) investment portfolio and a long-term hedging demand (opportunistic) portfolio. The myopic portfolio consisted of liquid U.S. and international equities and high-yield bonds. The opportunistic portfolio consisted of some direct private equity
investment in a friend’s company (representing 10% of
Daniel’s total wealth), and investment vehicles (private equity
funds and hedge funds) that allowed fund managers to time
the market and take on factor risks unavailable in traditional
index funds (see chapter 14).

Daniel’s portfolio suffered terribly in 2008. Financial markets
around the globe plummeted, and, like many investors, Daniel
watched her portfolio take a beating. Equity returns were
down 30% to 50% around the world in 2008. Daniel’s portfolio
lost 30%. Her direct private equity investment was wiped out.
By April 2009, while the economy was in recession, there was
a sense that the markets were no longer in free fall. Daniel
was still very concerned about the state of her portfolio.
Fortunately, she didn’t need the wealth to support her
standard of living. Nor did Daniel have immediate liquidity
needs that required drawing down the capital in her
investment portfolio. Daniel was still single but was now in a
relationship. She felt there was some way to go before she
would consider getting married. Although there were no plans
in the immediate future to have children, she was worried that
her greatly reduced portfolio would diminish the legacy that
she could leave them if she had any. Daniel thought that her
IPS and her asset allocation needed a “total overhaul.”

Harrison knew this was going to be a difficult meeting. On the
one hand, perhaps some of Daniel’s attitude was an irrational
overreaction to market conditions. On the other hand, perhaps
Daniel had genuinely become more risk averse, and the advice
that Harrison gave in 2007 was no longer valid. “People
always think they have more risk tolerance when things are
going well,” as Harrison said. Should Daniel stay the course or
revise her IPS and transition to a less risky portfolio?

In this chapter we discuss portfolio choice over long horizons
and how an investor can dynamically change her portfolio in
response to changing returns and investment opportunities.
The theory behind dynamic portfolio choice was (p.115)
formulated initially by Paul Samuelson (1969), who won the
won the Nobel Prize in 1997 with Myron Scholes, one of the
creators of the Black–Scholes (1973) option pricing model, for
the valuation of derivatives. As we’ll see shortly, the solution
to the dynamic portfolio choice problem is intimately related to
derivative valuation, and both employ the same economic concepts and solution techniques.

2. The Dynamic Portfolio Choice Problem
An investor facing a dynamic portfolio choice problem has a long horizon, say ten years, and can change her portfolio weights every period. A period could be one year, which is common for retail investors meeting with their financial planners for an annual tune up, or one quarter, which is common for many institutional investors. For high-frequency traders, a period could even be even shorter than one minute. The portfolio weights can change each period in response to time-varying investment opportunities as the investor passes through economic recessions or expansions, in response to the horizon approaching (as she approaches retirement, say), and potentially in response to how her liabilities, income, and risk aversion change over time. In this section, we abstract from the last of these considerations and assume that she has no liabilities and no income and is (fortuitously) given a pile of money to invest. (We introduce liabilities in section 3 and consider income in chapter 5.) We also assume her risk aversion and utility function remain constant.
2.1. Dynamic Trading Strategies

At the beginning of each period $t$, the investor chooses a set of portfolio weights, $x_t$. Asset returns are realized at the end of the period $t+1$, and the portfolio weights chosen at time $t$, $x_t$, with the realized asset returns lead to the investor’s wealth at the end of the period, $W_{t+1}$. The wealth dynamics follow

\[(4.1)\]

$$W_{t+1} = W_t(1 + r_{p,t}(x_t)),$$

where wealth at the beginning of the period, $W_t$, is increased or decreased by the portfolio return from $t$ to $t+1$, $r_{p,t}(x_t)$ and this is a function of the asset weights chosen at the beginning of the period, $x_t$.

I illustrate this in Figure 4.1 for a dynamic horizon problem over $T=5$ periods. At the beginning of each period the investor chooses portfolio weights, $x_t$. These weights, together with realized asset returns, produce her end of period wealth, $W_{t+1}$, following equation (4.1). The procedure is repeated every period. The sequence of weights over time, $\{x_t\}$, is called a dynamic trading strategy. It can (p.116) potentially change due to pre-determined variables, like investor constraints or liabilities changing, or due to time-varying investment returns, like booms versus busts.

The investor wishes to maximize the expected utility of end of period wealth at time $T$ by choosing a dynamic series of portfolio weights:

\[(4.2)\]

$$\max_{\{x_t\}} \mathbb{E}[U(W_T)]$$

subject to constraints. Some examples of constraints are that an investor may not be able to short (this is a positivity constraint so $x_t \geq 0$), may not be able to lever (so the portfolio weight is bounded, $0 \leq x_t \leq 1$), or can only sell a certain portion of her portfolio each period (this is a turnover constraint).
Although the portfolio weights $x_{t\tau}$ are, of course, only implemented at time $t + \tau$, the complete set of weights $\{x_t\}$ from $t$ to $T-1$ is chosen at time $t$, the beginning of the problem. The set of optimal weights can be quite complicated: they may not only vary through time as the horizon approaches, but they may vary by state. For example, $x_t$ could take on two values at time $t$: hold 50% in equities if we are in a recession and 70% if we are in a bull market. The complete menu of portfolio strategies across time and states is determined at the start.

Thus, the optimal dynamic trading strategy is completely known from the beginning, even though it changes through time: as asset returns change, the strategy optimally responds, and as utility and liabilities change, the strategy optimally responds.

For the remainder of this chapter, we work with constant relative risk aversion (CRRA) utility (see chapter 2):

$$U(W) = E\left[\frac{W^{1-\gamma}}{1-\gamma}\right],$$

where $W$ is the investor’s wealth at the end of the period and $\gamma$ is her risk aversion coefficient. In what follows, I will omit the $1-\gamma$ term in the denominator of equation (4.3) because maximizing $E[W^{1-\gamma}(1-\gamma)]$ is exactly the same as maximizing $E[W^{1-\gamma}]$. CRRA is locally mean-variance so the risk aversion $\gamma$ has the same meaning in mean-variance utility, $U^{MV}$ (see chapter 3):

$$U^{MV} = E(r_p) - \frac{\gamma}{2} \text{var}(r_p),$$

where $r_p$ is the portfolio return. The unconstrained solution to both the CRRA utility and mean-variance utility problem with one risky asset and one risk-free asset paying $r_f$ is:

$$x^* = \frac{\mu - r_f}{\sigma^2},$$

where the asset has expected return $\mu$ and volatility $\sigma$. The investor holds $x^*$ in the risky asset and $(1-x^*)$ in the risk-free asset. We developed this solution in detail in chapters 2 and 3. In this sense, CRRA and mean-variance utility are equivalent.
Since the optimal one-period weight is a mean-variance solution in equation (4.5), it nests as special cases many popular choices like equal weights, risk parity, market weights, and constant risk exposures considered in chapter 3. Thus, equation (4.5) does not necessarily mean a full-blown mean-variance solution—which we know does poorly. In the rest of the chapter, you can interpret the optimal one-period weight as any mean-variance optimal portfolio chosen by the investor in a single-period problem.

2.2. Dynamic Programming

The dynamic portfolio choice problem is an optimal control problem. It is solved by dynamic programming—the same technique used to control nuclear power plants, send rockets to the moon, and value complicated derivative securities. (I admit, the last of these examples certainly feels much less impressive than the first two.) Portfolio choice turns out to be rocket science—literally. The reader not interested in the math behind representing long-run investing as a series of short-run investment problems can skip directly to section 2.3.

Long-horizon wealth is a product of one-period wealth:

\[(4.6)\]
\[W_{t+1} = W_t(1 + r_{p,t})(1 + r_{m,t}) \]

from equation (4.1), and we can apply CRRA utility (equation (4.3)) to each one-period wealth term. Apply CRRA expected utility to long-horizon wealth, we have a series of one-period CRRA utility problems:

\[(4.7)\]
\[E[U(W_{t+1})] = E[U(W_t)E[U(1 + r_{p,t+1})U(1 + r_{m,t+1})]]\]

Since \(U(W)\) appears outside the expectation in equation (4.7), it does not matter whether we start with $1 or with $1 million—the portfolio weights do not depend on the size of the initial wealth, which is the wealth homogeneity property (p.118) (see chapter 2). Let’s normalize \(W_t\) to be one so we don’t have to worry about it. The portfolio returns, \(r_{p,t+1}\) in equations (4.6) and (4.7) depend on the portfolio weights chosen at the beginning of the period, \(x_t\), as equation (4.1) emphasizes. Thus, we can write equation (4.7) as

\[(4.8)\]
\[E[U(W_{t+1})] = E[U(W_t)E[U(1 + r_{p,t+1}(x_t))U(1 + r_{m,t+1}(x_{t+1}))...U(1 + r_{p,m}(x_{m}))]]\]
\[= E[U(1 + r_{p,t+1}(x_t))U(1 + r_{m,t+1}(x_{t+1}))U(1 + r_{p,m}(x_{m}))] \]
Figure 4.2 sketches an outline of the dynamic programming solution. Let’s start at the end, at \( t+4 \) to \( t+5 \), where the investor chooses portfolio weights to maximize expected utility at the terminal horizon \( T = t + 5 \). This is Panel A of Figure 4.2. This is a static one-period problem, and, for CRRA utility without (p.119) constraints, this is identical to the one-period mean-variance problem that we covered in chapter 3. The solution for a single risky asset with expected return \( \mu \) and volatility \( \sigma \), with a risk-free rate of \( r_f \) is given in equation (4.5), and we denote it by \( x^*_t \), where the asterisk means that the portfolio weight is optimal. The investor holds \( x^*_t \) in equities and \( (1 - x^*_t) \) in risk-free bonds. In principle, this portfolio weight can depend on what expected return and volatility are prevailing at \( t+4 \) (say, the economy is booming or in bust).
The maximum utility obtained at \( t+4 \) is:

\[(4.9)\]

\[V_{t+4} = E[U(1 + r_{t+5})]\]

where the portfolio return from \( t+4 \) to \( t+5 \), \( r^*_{t+5} \), is a function of the optimal portfolio weight chosen at \( t+4 \), \( x^*_{t+4} \), so \( r^*_{t+5} = r^*_{t+5}(x^*_{t+4}) \). The maximum utility \( V_{t+4} \) in equation \((4.9)\) is called the **indirect utility**, and it potentially differs across economic conditions prevailing at time \( t+4 \).

Having solved the last period’s problem, let us turn to the problem two periods before the end. At \( t+3 \), we need to solve both the portfolio weights at \( t+3 \) and \( t+4 \), which are \( x_{t+3} \) and \( x_{t+4} \), respectively:
But, we already solved the last period problem and found the optimal portfolio weight at $t+4$, $x_{t+4}^*$ for any outcome at $t+3$. This allows us to write the problem two periods before the end as the problem from $t+3$ to $t+4$, plus the problem with the known solution that we solved from $t+4$ to $t+5$:

\begin{equation}
\max_{\{x_{t+3}, r_{t+3}\}} \mathbb{E}[U(1+r_{t+4}d(x_{t+3}))U(1+r_{t+4}d(x_{t+3}))]
= \max_{x_{t+3}} \mathbb{E}[U(1+r_{t+4}d(x_{t+3}))U(1+r_{t+4}d(x_{t+3}))]
\end{equation}

The first equality in equation (4.11) substitutes the last period’s solution into the time $t+3$ problem. This now leaves just one portfolio weight at $t+3$, $x_{t+3}$, to solve. The second equality says that this problem is a standard single-period problem, except that it involves the indirect utility $V_{t+4}$ but we know everything about the indirect utility and the optimal strategies at $t+4$ from solving the last period’s problem (equation (4.9)). We can solve the problem in equation (4.11) as a one-period problem, and we denote the optimal weight at $t+3$ as $x_{t+3}^*$. It has the same as the one-period solution in equation (4.5), except that we adjust equation (4.5) for the optimized strategies adopted at $t+4$ that are captured by the indirect utility, $V_{t+4}$. Panel B of Figure 4.2 shows this pictorially.

Given the known solution at $t+4$, we use the optimal portfolio weight at $t+4$ to solve the portfolio weight at $t+2$. The recursion is applied once more to the $t+2$ problem having solved the $t+3$ and $t+4$ problems. Equation (4.11) also shows the origin of the name “indirect utility” because the indirect utility from the previous problem, at $t+4$, enters the direct utility from the current problem, at $t+3$.

Solving the problem two periods before the end gives us the optimal $t+3$ portfolio weight, $x_{t+3}^*$. We compute the maximum utility at $t+3$, which is the indirect utility at $t+3$:

\begin{equation}
V_{t+3} = \mathbb{E}[U(1+r_{t+4}d(x_{t+3}^*))V_{t+4}]
\end{equation}

Panel C of Figure 4.2 shows the recursion applied once more to the $t+2$ problem having solved the $t+3$ and $t+4$ problems. Again, the $t+2$ optimization is a one-period problem. After solving the $t+2$ problem, we continue backward to $t+1$ and then finally to the beginning of the problem, time $t$. Dynamic
programming turns the long-horizon problem into a series of one-period problems (following equation (4.11)). Dynamic programming is an extremely powerful technique, and Samuelson won the Nobel Prize in 1970 for introducing it into many areas of economics. Monetary policy (see chapter 9), capital investment by firms, taxation and fiscal policy, and option valuation are all examples of optimal control problems in economics that can be solved by dynamic programming. In continuous time, the value function is given by a solution to a partial differential equation called the Hamilton–Jacobi–Bellman equation. A more general form is called the Feynman–Kac theorem, widely used in thermodynamics. These are the same heavy-duty physics and mathematics concepts used in controlling airplanes and ballistic missiles. Portfolio choice is rocket science.

2.3. Long-Horizon Investing Fallacies

The important lesson from the previous section on dynamic programming is not that you should hire a rocket scientist to do portfolio choice (although there are plenty of ex-rocket scientists working in this area), but that dynamic portfolio choice over long horizons is first and foremost about solving one-period portfolio choice problems. Viewing the dynamic programming solution of long-horizon portfolio choice this way demolishes two widely held misconceptions about long-horizon investing.

Buy and Hold Is Not Optimal

A long-horizon investor never buys and holds. The buy-and-hold problem is illustrated in Figure 4.3: the investor chooses portfolio weights at the beginning of the period and holds the assets without rebalancing over the entire long-horizon problem. The buy-and-hold problem treats the long-horizon problem as a single, static problem. Buy-and-hold problems are nested by the dynamic portfolios considered in the previous section; they are a special case where the investor’s optimal choice is to do nothing. Buy and hold is dominated by the optimal dynamic (p.121) strategy that trades every period. Long-horizon investing is not to buy and hold; long-horizon investing is a continual process of buying and selling.
There is much confusion in practice about this issue. The World Economic Forum, for example, defined long-term investing as “investing with the expectation of holding an asset for an indefinite period of time by an investor with the capability to do so.” Long-horizon investors could, but in almost no circumstances will, buy and hold an asset forever. They dynamically buy and sell those assets over time.

The buy-and-hold confusion is also partly due to the popular sentiment generated by Jeremy Siegel’s famous book, *Stocks for the Long Run*, first published in 1994. This book is often described as the “buy-and-hold bible.” Siegel makes a case for sticking to a long-run allocation to equities. If this allocation is constant, then it is maintained by a constant rebalancing rule. Investors increase their share of equities after equities have done poorly to maintain this long-run, constant share. Long-run investors never buy and hold; they constantly trade.
Long-Term Investing Is Short-Term Investing

Another popular misconception about long-term investing is that by having a long-term investment horizon, long-run investors are fundamentally different from myopic, short-term investors. Some, like Alfred Rappaport (2011), suggest that long-run investors should act totally differently from short-term investors. The dynamic programming solution shows this to be blatantly false. Dynamic programming solves the long-horizon portfolio choice problem as a series of short-term investment problems. That is, *long-run investors are first and foremost short-run investors*. They do everything that short-run investors do, and they can do more because they have the advantage of a long horizon. The effect of the long horizon enters through the indirect utility in each one-period optimization problem (see equation (4.11)). I am not suggesting that long-run investors should engage in “short termism,” the myopic behavior that often befalls short-term corporate managers and short-term investors.⁵ The dynamic programming solution (p.122) suggests that, to be a successful long-run investor, you should start off being a successful short-run investor. After doing this, take on all the advantages that the long horizon gives you.

I now discuss one important case where there is no difference between long-run investors and short-run investors. This case happens to be the most empirically relevant and is the foundation of any long-term investment strategy.

2.4. Rebalancing

Suppose that returns are not predictable, or the investment opportunity set is independent and identically distributed (i.i.d.). The i.i.d. assumption is very realistic. Asset returns are hard to predict, as chapter 8 will show. A good way to think about the i.i.d. assumption is that asset returns are like a series of coin flips, except coins can only land heads or tails, and returns can take on many different values. The coin flip is i.i.d. because the current probability of a head or tail does not depend on the series of head or tails realized in the past. The same is true for asset returns: when asset returns are i.i.d., in every period returns are drawn from the same distribution that is independent of returns drawn in previous periods. Under i.i.d. returns, assets are glorified coin flips.
With i.i.d. returns and a fixed risk-free rate, the dynamic portfolio problem becomes a series of identical one-period problems, as shown in Figure 4.4. If returns are not predictable, then the long-horizon portfolio weight is identical to the myopic portfolio weight. Put another way, with i.i.d. returns, there is no difference between long-horizon investing and short-horizon investing: all investors are short term, and it does not matter what the horizon is. We can write this as:

\[(4.13)\]

The short-run weight is the myopic portfolio weight in equation (4.5). It is stated in terms of CRRA utility, but more generally it is the portfolio weight of (p. 123) a one-period utility problem using most of the utility functions that we covered in chapter 2. All investors—whether they have short or long horizons—act like short-term investors in the i.i.d. world because returns are not predictable; the long-run investor faces a series of independent coin flips each period. The optimal strategy is to manage portfolio risk and return in each period, treating each period’s asset allocation as a myopic investment problem. The optimal holding then is the myopic, short-run weight. The investor needs to rebalance back to this weight to avoid any one asset dominating in her portfolio for her given level of risk aversion.

If the optimal dynamic strategy is actually a myopic strategy, is the rebalancing strategy in Figure 4.4 different from a buy-and-hold strategy as shown in Figure 4.3? Absolutely. The dynamic problem is a series of one-period problems, and it involves rebalancing back to the same portfolio weight. The buy-and-hold problem involves doing nothing once the investor
has bought at the beginning of the period. To rebalance back to the same weight, the investor has to trade each period.

Consider the simplest case of stocks and risk-free bonds. To maintain a fixed portfolio weight in stocks, an investor must invest counter-cyclically. If equity has done extremely well over the last period, equities now are above target and it is optimal to sell equity. Thus, the investor sells stocks when stocks have done well. Conversely, if equity loses money over the last period relative to other assets, equities have shrunk as a proportion of the total portfolio. The equity proportion is too low relative to optimal, and the investor buys equity. Thus, rebalancing buys assets that have gone down and sells assets that have gone up. This rebalancing is irrelevant to a myopic investor, because the myopic investor is not investing anymore after a single period. Rebalancing is the most basic and fundamental long-run investment strategy, and it is naturally counter-cyclical. An important consequence of optimal rebalancing is that long-run investors should actively divest from asset classes or even stocks that have done well and they should increase weights in asset classes or stocks that have low prices. Thus, rebalancing is a type of value investing strategy (see chapter 7): long-run investors are at heart value investors.

Rebalancing is optimal under i.i.d. returns, but it turns out to be advantageous when returns exhibit mean reversion or are predictable. If expected returns vary over time, prices are low because future expected returns are high—as our investor Daniel experienced during the 2008 financial crisis; prices of many risky assets plummeted, but their future expected returns from 2008 onwards were high. Rebalancing buys assets that have declined in price, which have high future expected returns. Conversely, rebalancing sells assets that have risen in price, which have low future expected returns.6

2.5. Rebalancing in Practice

Rebalancing over 1926-1940

Figure 4.5 illustrates rebalancing from 1926 to 1940, which includes the Great Depression. The companion Figure 4.6 shows rebalancing over 1990 to 2011, which includes the financial crisis and the Great Recession. In each case I rebalance to a position of 60% U.S. equities and 40% U.S.
Treasury bonds and use data from Ibbotson Associates. Rebalancing occurs at the end of every quarter.

(p.125) Figure 4.5 starts off with $1 at the beginning of January 1926. The dashed line represents a 100% bond position, which rises steadily. A 100% stock position is shown in the dotted line, and the stock wealth is relatively volatile. Stocks rise though the 1920s and reach a peak of $2.93 at the end of August 1929. Then the Great Depression hits with a vengeance. Stocks markets crash and remain depressed into the early 1930s. Stocks hit a minimum of $0.49 in May 1932. Stocks begin a slow climb upward from this point and end in December 1940 at $1.81, which is below the cumulated bond position of $2.08 at that time. The solid line in Figure 4.5, Panel A, shows the rebalanced 60%/40% position. It is much less volatile than the 100% stock position so, while it does not rise as much until 1929, it also does not lose as much during the early 1930s. The 60%/40% strategy ends at December 1940 at $2.46.

Rebalancing is beneficial during the early twentieth century because it counter-cyclically cuts back on equities as they were peaking in 1929 and adds equities when they were at their lowest point in the early 1930s. Panel B of Figure 4.5 shows the rebalanced strategy, which goes back to 60%/40%
at the end of each quarter, versus a buy-and-hold strategy, which starts off at 60%/40% at the beginning of the sample and then fluctuates only according to how bond and stock returns vary. The rebalanced strategy, by design, hovers around the 60% equity proportion. There are some deviations because the strategy is not continuously rebalanced, but overall the rebalanced strategy is less risky because it does not allow equities to rise or fall to dangerously high or low levels. In terms of utility, the rebalanced strategy attains the optimal balance of stocks and bonds for the investor’s risk aversion. But, as an added benefit, rebalancing is counter-cyclical. In contrast, the equity holding in the buy-and-hold strategy was very high in early 1929 (when stock prices are high and expected returns low), right before stocks crash in October 1929. The buy-and-hold equity weight was very low in 1932, right before stock prices pick up (stock prices are low, and expected returns are high).

**Rebalancing over 1990–2011**

Figure 4.6 does a similar exercise for the 1990–2011 period. In Panel A, we start with $1 invested at the beginning of January 1990. The bond position is shown in the dashed line. During 2008, bond prices suddenly spiked as there was a flight to quality when Lehman Brothers failed, but overall the series is relatively stable. The ending bond position at December 2011 is $7.12. The equity position in the dotted line shows two large peaks and declines: the bull market of the late 1990s followed by the bursting of the Internet bubble in the early 2000s and the rise in equity prices during the early to mid-2000s followed by the financial crisis in 2007 and 2008. The ending equity position at December 2011 is $6.10. Like Figure 4.5, the solid line shows returns of a rebalanced 60%/40% strategy where the rebalancing occurs at the end of every quarter. This dynamic strategy is less volatile, by holding fewer equities, than the 100% equity position, and ends up doing better at December 2011, at $7.41, than either than full stock or bond strategy.
Panel B of Figure 4.6 shows the proportion invested in equities. The 60%/40% rebalanced strategy is optimal for the investor as it rebalances the equity position so that the risk of a single asset does not dominate. It also takes advantage of counter-cyclical investing. The buy-and-hold strategy shown in the dashed line loads up on equities, peaking at 2000, just as equities hit the post-bubble period. The equity proportion is also high right before the 2008 financial crisis. In contrast, the rebalanced strategy actively buys low-priced equities in late 2008 benefiting from the upward movement in prices (low prices, high expected returns) in 2009.

The standard 60%/40% strategy outperforms a 100% bond or 100% stock strategy over the 1926–1940 period (Figure 4.5) and over the 1990–2011 period (Figure 4.6). You should not take away that rebalancing will always outperform 100% asset positions—it won’t. In small samples, anything can happen. But I show below that, under certain conditions, rebalancing will always outperform a buy-and-hold portfolio given sufficient time, resulting in a rebalancing premium. The main takeaway from the figures is to understand why rebalancing works for the investor: it cuts back on assets that do well so
that they do not dominate in the portfolio. The investor rebalances so that the asset mix is optimal for her risk aversion every period. The 1926–1940 and 1990–2011 samples highlight an additional benefit of rebalancing: it is counter-cyclical, buying when prices are low and selling when prices are high.

**Investment Policy Statement**

Figures 4.5 and 4.6 may look impressive—but, in practice, rebalancing is hard. It involves buying assets that have lost value and selling those that have risen in price. This goes against human nature. Investors tend to be very reluctant to invest in assets that have experienced large losses. Investors are just as reluctant to relinquish positions that have done extremely well. How many investors can buy an asset because it has lost money? How many institutions can take capital away from traders because they have been successful and give it to their colleagues who have underperformed? The natural tendency of investors is to be *pro-cyclical*, whereas rebalancing is *counter-cyclical*.

Good financial advisors like Harrison, who is helping Daniel, play an important role in counteracting the pro-cyclical tendencies of individual investors. Maymin and Fisher (2011) argue that this is one of the areas where a financial advisor can add most value for a client. The IPS plays an important role in this process. Harrison as a financial advisor was right to insist on the IPS as the foundation of her advisor–client relationship with Daniel.

The IPS helps the investor to be *time consistent*: the investor has made decisions in written form, in consultation with the investment advisor, and in doing so lays out a game plan. The IPS allows investors to stick to that game plan. Charles Ellis, an indefatigable advocate for the investor, says in *Investment Policy*, a book originally published in 1987 with its original concepts reiterated in many of his other books:

**(p.128)** The principal reason for articulating long-term investment policy explicitly and in writing is to enable the client and the portfolio manager to protect the portfolio from ad hoc revisions of sound long-term policy, and to help them hold to long-term policy when short-
term exigencies are most distressing and the policy is most in doubt.

Medical directives, especially for the mentally ill, often take the form of Ulysses contracts, named for the wily Greek who, en route home from the Trojan War, commanded his crew to bind him to the mast of their ship so he could resist the Sirens’ song. The IPS is the Ulysses contract of an individual investor and helps him not to overreact. Shlomo Benartzi, a behavioral finance expert, is a keen advocate of using an IPS as a Ulysses strategy. He says, “Pre-commitment to a rational investment plan is important, because the intuitive impulse to act otherwise is strong.”

Ann Kaplan, a partner at Circle Wealth Management (and the donor of my chaired professorship), specializes in private wealth management. While the IPS is the capstone of her relationship with clients, she cautions that it takes effort to write, maintain, and reevaluate. “The challenge,” she says, is to “translate the individual/family’s multiple goals, changing cash flow needs and time horizons with the constraints of their current level of assets, tax status, investment biases, and psycho/social dynamics into a realistic and actionable investment policy.”

Take the challenge: having and hewing to an IPS is worth it.

**Do Investors Rebalance?**
Yes, but incompletely. Calvet, Campbell, and Sodini (2009) examine Swedish households. Data on Swedish asset holdings are very nearly complete because Swedes pay taxes on both income and wealth. Swedish households have a “surprisingly large propensity to rebalance,” in the words of the authors. Wealthy, educated investors tend to hold more diversified portfolios and also tend to rebalance more actively. While there is active rebalancing, there is some inertia so that investors do not completely reverse the passive, buy-and-hold changes in their portfolios. In contrast, Brunnermeier and Nagel (2008) show that, for U.S. households, inertia is the dominant factor determining asset allocation rather than rebalancing. (Maybe Swedish households are smarter, on average, than American ones.) Households start with a fixed allocation and then the asset weights evolve as a function of
realized gains and losses on the portfolio. Rebalancing does occur but sluggishly.

**p.129** Institutional investors often fail to rebalance. While many pension funds and foundations resorted to panic selling and abandoned rebalancing during 2008 and 2009, CalPERS stands out in its failure. CalPERS’s equity portfolio shrank from over $100 billion in 2007 to $38 billion in 2009. CalPERS did the opposite of counter-cyclical rebalancing: it invested pro-cyclically and sold equities right at their lowest point—precisely when expected returns were highest. While part of CalPERS’s problems in failing to rebalance stemmed from inadequate risk management, particularly of liquidity risk, CalPERS also failed to buy stocks when they were cheap because of structural misalignments between board members and the delegated fund manager. These are *agency problems*, and I discuss them in chapter 15. CalPERS did have a *statement of investment policy*, the institutional version of an individual investor’s IPS, but this did not help CalPERS to rebalance during the financial crisis. CalPERS either didn’t believe or optimally use its IPS.

In contrast to CalPERS, the Norwegian sovereign wealth fund rebalanced during 2008 and 2009. It bought equities at low prices from those investors like CalPERS who sold at the wrong time. Norway had its own version of Ulysses bound to the mast: the Ministry of Finance and parliament decided on a rebalancing rule, rather than having committees make rebalancing decisions. Adopting a rule, which was automatically implemented by the fund manager (Norges Bank Investment Management), ensured that the rebalancing decisions were not left to a committee whose members could be swayed by panic, fear, or hubris.

**Rebalancing Bands**

There are some technical considerations in implementing a rebalancing strategy. The theory presented has rebalancing occurring at regular time intervals: in Figure 4.5 and 4.6, rebalancing is done quarterly. But, in practice, if the equity portfolio weight is 61% at the end of a quarter, should the investor rebalance that small 1% back to a 60% target given transaction costs?
State-of-the-art rebalancing practices involve *contingent* rebalancing, rather than *calendar* rebalancing. Optimal rebalancing strategies trade off the utility losses of moving away from optimal weights versus the transaction costs from rebalancing. If the benefits of rebalancing outweigh the cost of doing so, then it is an optimal time to rebalance, and rebalancing becomes a contingent event.

Rebalancing bands are often used, set around optimal targets. The optimal rebalancing target may be 60% equities, for example, with bands set at 55% and 65%. A move outside the band triggers rebalancing. The bands are a function of transaction costs, liquidity, asset volatility, and minimum transaction sizes. When transaction costs are large or asset volatility is high, the bands are (p.130) wider. The first paper to derive optimal rebalancing bands was Constantinides (1979), and since then many variations have been developed.

The basic rebalancing model is shown in Panel A of Figure 4.7, where the horizontal axis indicates the evolution of an asset class weight. There is a single band around a target weight. If the asset weight lies within the band, the investor does not trade. As soon as the asset weight goes outside the bands, the investor rebalances to target. Other authors suggest rebalancing to the edge of the band. Whether you rebalance to the target or to the edge depends on whether the transaction costs are fixed exchange fees (rebalance to target) or proportional like brokerage fees and taxes (rebalance to the edge).

Panel B of Figure 4.7 presents a more sophisticated rebalancing strategy with two bands surrounding the target weight. There is no trade if the portfolio lies within the outer band. But, if the portfolio breeches the outer band, then the investor rebalances back to the inner band. Institutional investors often use derivatives to synthetically rebalance, which in many cases have lower transaction costs.
In my opinion all of these technical considerations in rebalancing are precisely that—technical. The most important thing is to rebalance.

2.6. Opportunistic Strategies

We’ve shown that rebalancing is the foundation of any long-term strategy and applies under i.i.d. returns. In addition, if returns are predictable, then there are further benefits from a long-term horizon. I call these opportunistic strategies.

When expected returns and volatilities change over time, the optimal short-run weight changes. In equation (4.5), we can put subscript \( t \) on the means and standard deviations, \( \mu_r \) and \( \sigma_r \), respectively, of an asset indicating that these are conditional estimates at time \( t \) of expected returns and volatilities over \( t \) to \( t+1 \). The risk-free rate is likely to vary as well (note that the risk-free rate is known at the beginning of the period, so the risk-free rate from \( t \) to \( t+1 \) is denoted as \( r_f(t) \)).
The time-varying short-run weight in equation (4.5) now becomes

\[
\text{Short-RunWeight}(t) = \frac{1}{V} \frac{\mu_r - r_f(t)}{\sigma_r^2}.
\]

Under time-varying, predictable returns, the optimal long-run strategy comprises the time-varying short-run strategy plus an opportunistic portfolio:

\[
\text{Long-RunWeight}(t) = \text{Short-RunWeight}(t) + \text{OpportunisticWeight}(t).
\]
The time-varying short-run weight is given in equation (4.14) and is called the myopic portfolio. The opportunistic weight is called the hedging demand by Merton, who chose the name because the hedging demand portfolio hedges against changes in the investment opportunity set. I prefer to think of it as how the long-run investor can opportunistically take advantage of time-varying returns.

Tactical and Strategic Asset Allocation

Campbell and Viceira (2002) interpret the Merton–Samuelson portfolios in equation (4.15) as:

\[
\text{Long-RunWeight}(t) = \text{Long-RunMyopicWeight} + \text{OpportunisticWeight}(t).
\]

16

17
where we split the short-run weight in equation (4.15) into two parts: the average, long-run myopic weight and a deviation from the constant rebalancing weight. We can interpret the first two terms in equation (4.16) as *long-run fixed weights* and *tactical asset allocation*, respectively. The overall long-run weight in equation (4.16) is sometimes called *strategic allocation*.

(p.132) The first term in equation (4.16) is the average value of the short-run weight in equation (4.14):

\[(4.17) \quad \text{Long - Run Myopic Weight} = \frac{1}{\sqrt{\bar{X}^2}} \frac{\bar{X} - \bar{r}}{\bar{\sigma}^2},\]

where the mean and volatility of the asset are at steady state levels denoted by bars above each variable. This can be interpreted as the equivalent of the constant rebalancing weight in the i.i.d. case.

The short-run weight represents tactical asset allocation and is how a short-run investor responds to changing means and volatilities. Tactical asset allocation then comprises the constant rebalancing weight plus a temporary deviation from the rebalancing rule (the first two terms in equation (4.16)).

Strategic asset allocation is the long-run weight and is the sum of the long-run fixed weight, tactical asset allocation, and the opportunistic weight. It is the optimal strategy for a long-term investor. As expected from the dynamic programming solution to long-term portfolio choice, long-run investors do everything that short-run investors do (tactical asset allocation), plus they can act opportunistically in a manner that their short-run cousins cannot. Thus, strategic asset allocation is the sum of all three terms in equation (4.16).

**Characterizing Long-Run Opportunistic Portfolios**

Computing the precise form of the long-run opportunistic portfolio can be difficult. But insight can be obtained on opportunistic weights without wading through rocket science. There are two determinants of the opportunistic weight. The first is investor specific. Just like the myopic portfolio weight depends on the risk tolerance of an investor so does the opportunistic portfolio. But now the investor’s horizon plays a role. Second, the opportunistic weights depend on asset-specific properties of how returns vary through time. The interaction between the investor’s horizon and the time-
varying asset return properties is crucial. This makes sense: an asset that has a low return today but will mean-revert gradually back over many years to a high level is unattractive to someone with a short horizon. Only a long-horizon investor can afford the luxury of waiting. Similarly, some assets or strategies can be very noisy in the short run, but over the long run volatility mean-reverts, and the risk premiums of these assets (p.133) manifest reliably only over long periods. Such strategies are also unattractive for short-run investors, but investors with long horizons can afford to invest in them.

Viewed broadly, the opportunistic portfolio for long-run investing also represents the ability of long-run investors to profit from periods of elevated risk aversion or short-term mispricing. In rational asset pricing models, prices are low because the average investor’s risk aversion is high and investors bid down prices to receive high expected returns. If a long-horizon investor’s risk aversion remains constant, then he can take advantage of periods with low prices. In behavioral frameworks, prices can be low because of temporary periods of mispricing. These can also be exploited by a long-term investor who knows that prices will return to fair values over the long run. While the simple rebalancing strategy is counter-cyclical and has a value tilt, some of the best opportunistic strategies are even more counter-cyclical and strongly value oriented. Crises and crashes are periods of opportunity for truly long-run investors. As Howard Marks (2011), a well-known value investor, explains with great clarity: “The key during a crisis is to be (a) insulated from the forces that require selling and (b) positioned to be a buyer instead.” That’s what rebalancing forces the investor to do.

There have been debates in the academic literature on how large these hedging demand, long-run opportunistic effects really are. In a major paper, Campbell and Viceira (1999) estimate hedging demands to be very large that are easily double the average total demand for stocks by short-run investors. In Campbell and Viceira’s model, the portfolio weight in equities for a long-term investor would have varied from -50% to close to 400% from 1940 to the mid-1990s. However, Brandt (1999) and Ang and Bekaert (2002), which appeared around the same time as Campbell and Viceira’s paper, estimate small hedging demands. The long-run opportunistic demands depend crucially on how predictable returns are and the model used to capture that predictability.
In chapter 8, I show that the evidence for predictability is weak, so I recommend that both the tactical and opportunistic portfolio weights be small in practice. Opportunistic hedging demands become much smaller once investors have to learn about return predictability or when they take into account estimation error.\(^{22}\)

A system of predictable equity returns that has been widely studied in the portfolio choice literature is the Stambaugh (1999) system, where stock returns are driven by a valuation ratio like the dividend or earnings yield. The valuation ratio is a convenient instrument to capture time-varying expected returns. As dividends yields drop (or equity prices rise), future expected returns increase. The (\textbf{p.134}) dividend yield itself also persistently varies over time.\(^{23}\) Under the Stambaugh system, the long-term opportunistic portfolios are positive and increase with horizon.\(^{24}\) This is shown in Figure 4.8 where the short-run, myopic weights and the total long-run weights increase as expected returns increase and investment opportunities become more attractive. Long-run investors actually are leveraged versions of short-run investors: if short-run investors want to buy when expected returns are high, long-run investors will buy more. Opportunistic investing allows long-term investors to exploit predictability even more than short-run investors do.\(^{25}\)

The Norwegian sovereign wealth fund, following a constant rebalancing rule during 2008 and 2009, bought equities when prices were low and expected returns were high. Had the fund taken advantage of long-run, opportunistic strategies, it would have bought even more equities. Thus, the rebalancing rule functions as a conservative lower bound for “buying low and selling high.” I advise you to concentrate on rebalancing first before focusing on opportunistic strategies. Simple rebalancing is itself
counter-cyclical; long-run opportunistic investing in the Stambaugh model is much more aggressively counter-cyclical. If you cannot rebalance, which already involves buying assets that are falling in price (relative to other assets), then there is no way that you can implement opportunistic long-run investing. When returns follow the Stambaugh model, opportunistic long-run investing involves buying even more of the assets that have fallen in price than what simple rebalancing suggests. I also recommend that opportunistic portfolios should be modest: taking into account estimation error, combined (p.135) with the overall very weak predictability in data (see chapter 8), any realistic application of Figure 4.8 considerably flattens both the time-varying short-run and opportunistic weights as a function of expected returns.

3. Rebalancing is Short Volatility
Rebalancing is an option strategy and, in particular, a short volatility strategy. (Option traders would call rebalancing a negative gamma strategy.) This is not well known, although at some level should not be surprising for the reader steeped in financial theory because the same method used in section 2 to solve long-horizon portfolio choice problems (dynamic programming) is used to value options (where it is called backward induction).26 Showing how rebalancing is mechanically a short volatility strategy gives us deeper insights into what long-run investors are gaining and losing from rebalancing. Nothing is free after all, at least not in economic theory.

3.1. Example
This example is highly stylized and simple but conveys enough to see rebalancing as a collection of options.

Suppose that a stock follows the binomial tree given in Figure 4.9, Panel A. Each period the stock can double, with probability 0.5, or halve starting from an initial value of $S = 1$. There are two periods, so there are three nodes in the tree. At maturity, there are three potential payoffs of the stock: $S_{uu} = 4$, $S_{ud} = S_{du} = 1$, and $S_{dd} = 0.25$, which have probabilities of 0.25, 0.5, and 0.25, respectively. In addition, the investor can hold a risk-free bond that pays 10% each period.
Let us first consider a buy-and-hold strategy that starts out with 60% equities and 40% in the risk-free asset. (We know buy and hold is not optimal for the long-run investor from section 2.) At the end of the first period, the wealth of this investor can increase or decrease to (4.18)

\[
W_u = 0.6 \times 2.0 + 0.4 \times 1.1 = 1.6400
\]
\[
W_d = 0.6 \times 0.5 + 0.4 \times 1.1 = 0.7400
\]

which is shown by branching of the binomial tree in Figure 4.9, Panel B. In equation (4.18), the return on the stock is either \(2 - 1 = 100\%\) if we go into the upper branch or \(0.5 - 1 = -50\%\)
if we go into the lower branch. In the upper node at time 1, the proportion of the buy-and-hold portfolio held in equities is \(0.6 \times 2.0 / 1.64 = 73.17\%\) and the proportion of the portfolio in equities in the (p.136) (p.137) lower node is \(0.6 \times 0.5 / 0.74 = 40.54\%\). For the last two nodes at time 2, the final wealth for the buy-and-hold strategy is

\[
W_{uu} = 1.6400 \times (0.7317 \times 2.0 + 0.2683 \times 1.1) = 2.8840, \\
orW_{ud} = 1.6400 \times (0.7317 \times 0.5 + 0.2683 \times 1.1) \\
= 1.0840 = W_{du}, \\
which is the same as W_{du} = 0.7400 \times (0.4054 \times 2.0 + 0.5946 \times 1.1) \\
= 1.0840 = W_{ud} \\
orW_{dd} = 0.7400 \times (0.4054 \times 0.5 + 0.5946 \times 1.1) = 0.6340. 
\]

This is shown in Figure 4.9, Panel B at the end of the binomial tree.

Now consider the optimal rebalanced strategy, which rebalances at time 1 back to 60% equities and 40% bonds. The end of period wealth at time 1 is exactly the same as equation (4.18). The final wealth at time 2 for the rebalanced strategy is

\[
W_{uu} = 1.6400 \times (0.6 \times 2.0 + 0.4 \times 1.1) = 2.6896, * 3pt \\
orW_{ud} = 1.6400 \times (0.6 \times 0.5 + 0.4 \times 1.1) = 1.2136 = W_{du} * 3pt \\
whichis the same as W_{du} = 0.7400 \times (0.6 \times 2.0 + 0.4 \times 1.1) = 1.2136 = W_{ud} * 3pt \\
orW_{dd} = 0.7400 \times (0.6 \times 0.5 + 0.4 \times 1.1) = 0.5476. 
\]

This set of payoffs is shown in Panel C of Figure 4.9.

The last panel D of Figure 4.9 plots the payoffs of the buy-and-hold strategy (equation (4.19)) and the rebalanced strategy (equation (4.20)) as a function of the stock value at maturity time 2. The buy-and-hold, un-rebalanced strategy is shown in the dashed straight line. The gains and losses on the buy-and-hold position are linear, by construction, in the stock price. The payoffs of the rebalanced strategy, in contrast, are concave over the stock price. Rebalancing adds more wealth to the investor if the stock price returns to 1.0 at maturity (1.2136 for the rebalanced vs. 1.0840 for the buy and hold for \(S_{uu} - S_{dd} = 1\)). This is offset by the rebalancing strategy underperforming the buy-and-hold strategy when the ending stock values are low (\(S_{dd} = 0.25\)) or high (\(S_{uu} = 4\)).

This nonlinear pattern of the rebalancing strategy can be equivalently generated by short option positions. The strategy
sells out-of-the-money call and put options and hence is short volatility.

Suppose there is a European call option with strike $3.6760 maturing at time 2. This call option has the following payoffs at time 2:

\[
\begin{align*}
C_{out} &= \max(4.0000 - 3.6760, 0) = 0.3240, \times 3\text{pt} \\
\text{or } C_{adj} &= \max(1.0000 - 3.6760, 0) = 0 - C_{out} = 0 \times 3\text{pt} \\
\text{or } C_{sl} &= \max(0.2500 - 3.6760, 0) = 0
\end{align*}
\]

(p.138) The value of this call option at time 0 is $0.0428.\text{27}

There is also a European put option with strike $0.4660 maturing at time 2. This put option is worth $0.0643 at time 0 and has the following payoffs at time 2:

\[
\begin{align*}
P_{out} &= \max(0.4660 - 4.0000, 0) = 0, \times 3\text{pt} \\
\text{or } P_{adj} &= \max(0.4660 - 4.0000, 0) = 0 = P_{sl} \\
\text{or } P_{sl} &= \max(0.4660 - 0.2500, 0) = 0.2160
\end{align*}
\]

Now compare the following strategies:
<table>
<thead>
<tr>
<th>Strategy</th>
<th>time 0</th>
<th>$S_{ud} = 4$</th>
<th>$S_{ud} = S_{du} = 1$</th>
<th>$S_{dd} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sell Put</td>
<td>+0.0643</td>
<td>0</td>
<td>0</td>
<td>−0.2160</td>
</tr>
<tr>
<td>Sell Call</td>
<td>+0.0428</td>
<td>−0.3240</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Buy Bonds</td>
<td>−0.1071</td>
<td>0.1296</td>
<td>0.1296</td>
<td>0.1296</td>
</tr>
<tr>
<td>Buy-and-Hold Strategy</td>
<td>1.0000</td>
<td>2.8840</td>
<td>1.0840</td>
<td>0.6340</td>
</tr>
<tr>
<td>Short Volatility + Bonds + Buy and Hold</td>
<td>1.0000</td>
<td>2.6896</td>
<td>1.2136</td>
<td>0.5476</td>
</tr>
<tr>
<td>Rebalanced Strategy</td>
<td>1.0000</td>
<td>2.6896</td>
<td>1.2136</td>
<td>0.5476</td>
</tr>
</tbody>
</table>
The table lists the values today in the column labeled “time 0” and the payouts of the various strategies at “time 2.” The time 2 payouts are contingent on the stock values at time 2; hence, there are three columns representing the stock values $S_{uu} = 4$, $S_{ud} = S_{du} = 1$, and $S_{dd} = 0.5$ at time 2.

Consider the first set of strategies. Selling a put today means money comes in (+ sign) with a put premium of $0.0643$. If the stock price is low ($S_{dd} = 0.5$) at time 2, then the investor must pay out (– sign) an amount of $0.2160$. Likewise, selling a call today means money comes in with a call premium of $0.0428$. The investor must make a payout to the person buying the option of $0.3240$ if the stock increases at time 2 ($S_{uu} = 4$). The investor also purchases $0.1071$ of bonds at time 0. The purchase means a cash outflow, so there is a negative sign. At time 2, these bonds are worth $0.1071 \times (1.1)^2 = 0.1296$ at time 2.

Finally, we have the payoffs of the buy-and-hold strategy starting with $1$ invested at time 0. The payoffs of the strategy in each state of the world for the buy-and-hold strategy are listed in equation (4.19).

If we add the short call, the short put, the long bond position, and the buy-and-hold strategy, we get a value of $1$ today at time 0, with identical payoffs to (p.139) the rebalancing strategy at time 2 (which are listed in equation (4.20)). That is, a short volatility position that is financed by bonds together with the buy-and-hold strategy is identical to the rebalanced strategy. Hence, rebalancing is a short volatility strategy.

In Figure 4.9, Panel D, the buy-and-hold strategy is the completely passive straight line. The rebalancing strategy is an active strategy that transfers payoffs from the extreme low and high stock realizations ($S_{uu}$ and $S_{dd}$) to the middle stock realization ($S_{ad} = S_{da}$). Rebalancing does this by selling when stock prices are high and buying when stock prices are low. Short volatility positions do exactly the same. A call option can be dynamically replicated by a long stock position and a short bond position. This buys equity when stock prices rise and sells equity when stock prices falls. A short call option does the opposite: a short call position is the same as selling when equity prices rise and buying when they fall. Likewise, a short put is also dynamically replicated by selling equity when
prices rise and buying when prices fall. These are exactly the same actions as rebalancing.

3.2. Interpretation

What is the market value of rebalancing? In this two-period binomial example, the action of rebalancing relative to the buy-and-hold strategy can be replicated by selling a call, selling a put, and investing in bonds. This has value:

\[
\text{ShortCall} + \text{ShortPut} + \text{LongBonds} = 0.0643 + 0.0428 - 0.1071 = 0.
\]

That is, the action of rebalancing is assigned a zero market value. The market does not value rebalancing.

The optimal rebalancing strategy is a partial equilibrium strategy. Not everyone can rebalance. For every institution like Norway buying equities during the darkest periods of the financial crisis, there were institutions like CalPERS who couldn’t wait to shed their risky equity allocations. CalPERS’s losses in failing to rebalance represent Norway’s gains from successful rebalancing. Put simply, for every buyer, there must be a seller. In equilibrium, it is impossible for everyone to simultaneously sell or buy.\(^{28}\) Rebalancing is not valued by the market. In fact, consistent with this fact, the average investor who holds the market portfolio does not rebalance: the market itself is buy and hold!\(^ {29}\)

\(^{(p.140)}\) The benefit to rebalancing is investor specific. Moving the payoffs from the extreme stock positions back to the center (as in Figure 4.9, Panel D) is optimal for the investor because it cuts back on risk. In our example, the 60% equity/40% bond portfolio turns out to be optimal for an investor with a \(\gamma = 0.51\) degree of risk aversion. A certainty equivalent calculation (see chapter 2) reveals that he needs to be compensated 0.29 cents for each dollar of initial wealth for being forced to do the buy-and-hold strategy instead of optimally rebalancing.\(^ {30}\) The long-term investor values rebalancing because it reduces her risk and increases her utility. The market does not because there must be other investors who are not rebalancing to take the other side.

Because rebalancing is short volatility, it automatically earns the volatility risk premium. In our example, volatility is constant (the stock volatility is equal to 0.75), but in reality volatility varies over time. Volatility is a risk factor and earns a
negative risk premium. An investor collects the volatility risk premium by selling options or by being short volatility. I discuss this further in chapter 7.

Viewing rebalancing as a short volatility strategy in moving the payoffs to the center, increasing the losses during extreme low markets, and underperforming the buy-and-hold strategy during extreme high markets makes clear that rebalancing profits from reversals. This is one reason that rebalancing performed well during 1926–1940 and 1990–2011 in Figure 4.5 and 4.6, respectively. When there are strong reversals from the steepest crashes, rebalancing works wonders. We experienced strong reversals after the Great Depression and more recently after the Great Recession and financial crisis in 2008–2009. Conversely, if reversals do not occur, such as in permanent bull or permanent bear markets, then rebalancing will underperform the buy-and-hold strategy.

The fact that rebalancing is equivalent to a short volatility strategy is equivalent to the statement, in the words of Antti Ilmanen, my fellow advisor to the Norwegian sovereign wealth fund, “rebalancing is short regime changes.” Take the extreme case where a regime change occurs and permanently kills equity markets: rebalancing performs poorly because it adds equities as prices decline, and then equity prices are permanently lower. The opposite extreme case is a regime change (p.141) so that stocks permanently go into a bull market. Rebalancing also underperforms a buy-and-hold strategy because rebalancing would have sold stocks that only keep going up.

Permanent regime changes sometimes occur, but they are rare. I fully agree with Reinhart and Rogoff (2011) that people too often think, “This time is different.” Two examples of true regime changes where “these times really were different” are the changing shape of the yield curve pre-and post-1933 and the pricing of out-of-the-money put options pre- and post-1987. Pre-1933 the yield curve was downward sloping, compared to its now (post-1933) normal upward-sloping shape. Academics think going off the gold standard plays a role in explaining the change (see Wood (1983)). Implied option volatilities were symmetric across strikes in the pre-1987 sample (called the implied volatility smile). After the 1987 crash, there has been significant negative skewness in implied volatilities (see Rubinstein (1994)), with option
volatilities for low strikes much larger than option volatilities for high strikes. The implied volatility smile turned into an implied volatility smirk. The financial crisis in 2007 and 2008 was not a regime change. The European sovereign debt crisis since 2009 is also not a regime change. True permanent regime changes are rare.

The fact is that rebalancing is short volatility, and thus short regime changes, means that you must practice rebalancing on broad asset classes (or across factors, see chapter 14) that are extremely unlikely to undergo permanent regime change. Russian equities in 1900 were a large market and then totally disappeared, with investors receiving zero, less than two decades later. But global equities are still around more than a century later and are likely to be here for a long time, perhaps until capitalism disappears. Russian bonds also disappeared during the Russian Revolution, but global bonds did not. Global equities and global bonds have been and will continue to be with us for a long time. Rebalance with the tried and true.\(^{32}\)

4. Liability Hedging

4.1. Liability Hedging Portfolio

Few investors are without liabilities. Even investors lacking explicit liabilities (like Norway), at least over the short term, often have implicit liabilities through stewardship expectations of stakeholders. Liabilities can be fixed, like loan payments; variable but steady, like pension costs; or contingent one-off payments, like a person’s death benefit.

When liabilities are introduced, the optimal portfolio strategy has three components:

\[
\text{Long - Run Weight}(t) = \frac{\text{Liability Hedge}(t)}{\text{Investment Portfolio}(t)} + \text{Short - Run Weight}(t) + \text{Opportunistic Weight}(t).
\]

The investment portfolio is exactly the same as the nonliability case that we examined in sections 2 and 3: the optimal policy is to rebalance under i.i.d. returns and when returns are predictable, the optimal short-run portfolio changes over time, and the long-run investor has additional opportunistic strategies. The liability hedging portfolio is the portfolio that best ensures the investor can meet those liabilities. We solve
for it by holding asset positions that produce the highest correlation with the liabilities.\textsuperscript{33}

There are several special cases of optimal liability hedging portfolios:

1. Cash flow matching or immunization.
   This involves constructing a perfect match of liability outflows each period. You immunize each liability cash flow by holding bonds of appropriate maturities.
2. Duration matching.
   If liabilities can be summarized by a single interest rate factor, which is common for pension liabilities, then the liabilities can be offset by an asset portfolio with the same duration.\textsuperscript{34}
3. Liability-driven investing.
   This aims to construct a portfolio of risky assets that best meets the liability obligations. It is also common in pension fund management and was introduced by Sharpe (1992). It is related to and often used synonymously with . . .
   This is a more general case than duration matching. In asset-liability matching, dimensions other than just duration are used to match liability characteristics with assets, including liquidity, sensitivity to factors besides only interest rates, and horizon.

The Merton–Samuelson advice of long-horizon asset allocation extended to liabilities is, first, to meet the liabilities and then to invest the excess wealth over the present value of liabilities in the same style as sections 2 and 3, using the myopic market portfolio and the opportunistic long-horizon portfolio.

\textbf{(p.143)} For a long-horizon investor, U.S. Treasuries are usually neither a risk-free asset nor the optimal liability-hedging asset. If the investment horizon exceeds the longest available maturity of the risk-free bond, which is the case for some sovereign wealth funds and family offices, then investors do not have access to a risk-free asset. Furthermore, many investors have liabilities denominated in \textit{real}, not \textit{nominal}, terms. But even long-horizon real bonds are not the optimal liability-hedging asset if there are other factors. For pension plans, these include longevity risk, economic growth, and credit risk. Individual investors may face inflation risks, like
for medical care and college tuition, that are not adequately reflected in general CPI inflation. The liability hedging portfolio emphasizes what types of assets (or more broadly what kinds of factors, see chapter 14) pay off to meet the worst times of the investor, in terms of when and how the liabilities come due. If liabilities increase when credit spreads narrow, for example, as they do for pension funds, then the liability-hedging portfolio must hold large quantities of assets that are sensitive to credit risk.

What if you can’t meet the liabilities in the first place? Sadly, this condition applies to many investors today, especially public pension funds. CalPERS, for example, only had a funding ratio (the ratio of assets to actuarial liabilities) of 65% at June 30, 2010. Strictly speaking, the Merton–Samuelson asset allocation advice outlined in sections 2 and 3 applies only after the liabilities can be met, both in terms of the present value of the liabilities and after the liability-hedging portfolio has been constructed. If assets are not sufficient to meet current liabilities, then the investor must face the fact that default will happen in some states of the world. Portfolios can be constructed to minimize this probability, but avoiding insolvency requires a different optimization than the maximization of utility examined in equation (4.2). In certain cases, it may be optimal for the investor to engage in risk-seeking behavior if the assets are far enough below the value of the liabilities. It is the Hail Mary pass; you have nothing to lose, and you are likely to go bankrupt anyway.35

4.2. Popular Investment Advice

The three types of portfolios for long-term investors

1. Liability-hedging portfolio;
2. Short-run, or myopic, market portfolio; and
3. Long-run opportunistic, or long-term hedging demand, portfolio

(p.144) that are derived in the Merton-Samuelson dynamic trading context accord well with the advice given by some financial advisors. A practitioner framework developed by Ashvin Chhabra (2005), who for a time managed the endowment of the Institute for Advanced Study in Princeton, suggests that an individual investor create three buckets:
1. Protective portfolio, which covers “personal” risk. The portfolio is designed to minimize downside risk and is a form of safety first (see chapter 2). The maxim is: “Do not jeopardize the standard of living.”

2. Market portfolio, which is a balance of “risk and return to attain market-level performance from a broadly diversified portfolio” and is exposed to market risk.

3. Aspirational portfolio, which is designed to “take measured risk to achieve significant return enhancement.” Aspirational risk is a property of an investor’s utility function and is a desire to grow wealth opportunistically to reach the next desired wealth target.

This looks very much like the Merton–Samuelson advice. Chhabra’s buckets correspond to the three Merton–Samuelson portfolios:

1. Protective portfolio = Liability-hedging portfolio;
2. Market portfolio = Short-run portfolio; and
3. Aspirational portfolio = Long-run opportunistic portfolio.

There are some small differences between Merton–Samuelson and Chhabra. Chhabra advocates mostly safe, fixed-income assets for the protective portfolio, while the concept of the Merton liability-hedging portfolio recognizes that U.S. Treasuries may not be safe and sometimes are extremely risky in terms of meeting liability commitments. But the overall concepts of Chhabra are similar to Merton and Samuelson’s theory.

Thus, some financial planners have been advocating Merton–Samuelson dynamic portfolio choice theory even though they have not been exposed to the original Nobel Prize-winning papers written in the 1960s and 1970s. The difference is that the full (rocket science) glory of formal portfolio choice leads to quantitative solutions (equations can be numerically solved by rocket scientists to give portfolio weights when analytical solutions are not available), economic rigor, and some deep insights linking dynamic portfolio choice with option strategies to understand when and how long-run advice will do well or poorly.
5. Rebalancing Premium
Long-horizon investing is not complete without a final discussion of the rebalancing premium. This goes under a variety of names including the diversification (p.145) return, variance drain, growth-optimal investing, volatility pumping, and the Kelly criterion or Kelly rule, named after John Kelly (1956), an engineer who worked at Bell Labs. The term “diversification return” was introduced by Booth and Fama (1992) and is probably the best-known term in finance, whereas the Kelly rule and volatility pumping are better known in mathematics. I prefer not to use Booth and Fama’s terminology because there is a difference between diversification in a single period and rebalancing, which earns a premium over time. Diversification gets you a benefit in one period, but this diversification benefit dies out if you do not rebalance. The rebalancing premium only exists for a long-horizon investor, and he can collect it by rebalancing to constant weights every period. I use the term “rebalancing premium” to emphasize that the premium comes from rebalancing, not from diversification.

5.1. Rebalancing Beats Buy-and-Hold over the Long Run
Suppose that the price of each underlying asset is stationary; that is, the price of each asset tends to hover around a fixed range and never goes off to infinity. Holding 100% positions in each asset never gives you increasing wealth. But, a rebalanced portfolio does give you wealth that increases over the long run to infinity (wealth increases exponentially fast). Furthermore, by rebalancing to a fixed constant weight each period, an investor can generate wealth that increases over time, and any such rebalancing strategy will eventually beat the best buy-and-hold portfolio. This seems like magic: Erb and Harvey (2006) call it “turning water into wine” and Evstigneev and Schenk-Hoppé (2002) call it going from “rags to riches.”

Mathematically, this is not quite as impressive as Jesus’ first miracle at the wedding at Cana. It arises as a consequence of compounding. We can see this in equation (4.7), where long-term wealth is a product of arithmetic returns, \((1+r_1)(1+r_{t-1})(1+r_{t-2})\ldots\), rather than a sum of arithmetic returns, \((1+r_1)(1+r_{t-1})(1+r_{t-2})\ldots\neq 1+r_1+r_{t-1}+r_{t-2}+\ldots\). The compounding of products gives rise to many nonlinearities over time, which are called Jensen’s terms, and the effect of the nonlinear terms...
increases over time. The entire rebalancing premium is due to Jensen’s terms, and in fact the whole diversification return and Kelly rule literature can be viewed as a paean to Jensen’s inequality.

(p.146) Jensen’s terms arise because of the difference between geometric returns, which take into account the compounding over the long run, and arithmetic returns, which do not compound. In a one-period setting, geometric and arithmetic returns are economically identical; they are simply different ways of reporting increases or decreases in wealth. Thus, there is no rebalancing premium for a short-run investor. Over multiple periods, the difference between geometric and arithmetic returns is a function of asset volatility, specifically approximately \( \frac{1}{2} \sigma^2 \), where \( \sigma \) is the volatility of arithmetic returns. The greater the volatility, the greater the rebalancing premium. As this manifests over time, only long-term investors can collect a rebalancing premium.

For U.S. stocks, the rebalancing premium a long-run investor can earn is approximately \( \frac{1}{2} (0.2)^2 \approx 1\% \). Erb and Harvey (2006) estimate a rebalancing premium of around 3.5% in commodities. These are significant premiums for simple, automatic rebalancing. In his 2009 book, David Swensen, the superstar manager of Yale University’s endowment, emphasizes that rebalancing plays an important role in his practice of investment management, especially in the daily rebalancing of Yale’s liquid portfolio. He refers to a “rebalancing bonus” arising from maintaining a constant risk profile.

So, is rebalancing optimal not only because it reduces risk, but also because it provides a “free lunch” in the form of a rebalancing premium? Not so fast. In section 3, I showed that rebalancing has no value in the market by interpreting rebalancing as an option strategy. The rebalancing premium seems too good to be true—and in fact it is. Rebalancing is a short volatility strategy that does badly compared to buy and hold when asset prices permanently continue exploding to stratospheric levels or permanently implode to zero and disappear. Rebalancing is short a regime change. The crucial assumption behind the rebalancing premium is that the assets over which you rebalance continue to exist. If there are assets that experience total irreversible capital destruction, then rebalancing leads to buying more assets that eventually
disappear—this is wealth destruction, not wealth creation. The rebalancing premium can only be collected for assets that will be around in the long run, so rebalance over very broad asset classes or strategies: global equities, global sovereign bonds, global corporate bonds, real estate, commodities, and so on, rather than individual stocks or even individual countries.

5.2. The Very Long Run

In the very long run, the portfolio that maximizes wealth is a rebalanced portfolio that holds constant asset weights that maximize the rebalancing premium. This strategy maximizes long-run growth and is called the Kelly rule. It is obtained by finding the portfolio that maximizes one-period log returns. Since this portfolio maximizes (very) long-run wealth, it is also called the *optimal growth portfolio*.

The Kelly optimal growth portfolio dominates all other portfolios with a sufficiently long time span. So for the very long-run investor, should we hold the optimal growth portfolio if it maximizes very long-run wealth? No. This was settled by Samuelson in the 1970s, but the question is raised periodically by the unconvinced. Samuelson wrote a cute paper in 1979, entirely written in words of one syllable, entitled “Why We Should Not Make Mean Log of Wealth Big though Years to Act Are Long,” to answer this question. (Not surprisingly, limiting yourself to words of one syllable makes a paper quite hard to read.)

In a one-period model, you can maximize the portfolio growth rate by holding the portfolio that maximizes expected log returns. But do you have a log utility function? Probably not. You trade off risk and return differently than a log investor and are better off holding a portfolio optimized for your own risk aversion and utility function. Over the long run, you will outperform by following the Kelly rule. But, there is risk in doing so, and you might not be able to tolerate this risk. Furthermore, the long run in the Kelly rule could be very, very long. And as Keynes famously observed, in the long run, we are all dead.

In summary, follow the Merton–Samuelson advice and not the Kelly rule. Find your optimal one-period portfolio holdings over broad asset classes or strategies. Rebalance back to fixed weights. This is optimal with i.i.d. returns, and it will earn you
a rebalancing premium. If you can forecast returns well, you also have a long-run opportunistic portfolio available.

6. Stay the Course? Redux
Daniel has suffered large losses and is feeling skittish about sticking to her long-term plan. But she is fortunate that she has no immediate liabilities and her income, which covers her expenses, is relatively safe. Yet the losses seem to have changed her tolerance for risk. She has told her financial planner, Harrison, that her IPS needed a “total overhaul” and she could not afford such big losses going forward.

According to the long-run investment advice from Merton and Samuelson, Harrison should advise Daniel to rebalance. Rebalancing to fixed weights (or exposures) is optimal when returns are not predictable. Even though returns are predictable in reality, the amount of predictability is very small. This makes rebalancing the foundation of the long-run strategy. The small extent to which returns are predictable can be exploited by a long-run investor through an opportunistic portfolio. Daniel should stay the course and rebalance.

Rebalancing, however, goes against human nature because it is counter-cyclical. It is difficult for individuals to buy assets that have crashed and to sell assets that have soared. Part of Harrison’s job as an investment advisor is to counteract these behavioral tendencies. The IPS can help by functioning as a precommitment device—a Ulysses contract—in preventing Daniel from overreacting and abandoning a good long-term plan.

But perhaps Daniel’s risk aversion has truly changed. The classical assumptions, which we used in sections 2 to 4, are that risk preferences are stable and unaffected by economic experiences. This is not true in reality. Malmendier and Nagel (2011) show that investors who experienced the searing losses of the Great Depression became permanently more risk averse and were far less willing to invest in stocks than younger investors who did not experience such large losses and economic hardships. They show further that after the recessions of the late 1970s and early 1980s, young investors who only experienced the market’s low returns during these periods were more risk averse and held fewer equities and more bonds than older investors who had experienced the high returns of the stock market during the 1950s and 1960s. Thus,
life experience does influence the extent to which investors are willing to take financial risks. But Malmendier and Nagel show that what changes after large losses is not so much investors’ risk preferences but investors’ expectations. People tend to lower their expectations about future returns as a response to searing losses rather than changing their utility functions.

If Daniel has truly become more risk averse, then rebalancing back to the old portfolio pre-2007 is no longer valid, and Daniel has to work with her financial advisor to come up with a new IPS. Changes should be made deliberately and carefully. The worst portfolio decisions often result from pure panic; many ad-hoc changes lead to pro-cyclical behavior. Otherwise, the dynamic portfolio choice advice is to stay the course. Rebalance.

Notes:
(1) This is based on the case “Stay the Course? Portfolio Advice in the Face of Large Losses,” Columbia CaseWorks, ID #110309.

(2) Strictly speaking, the CRRA weight applies in continuous time or when the interval is very small. See Merton (1971).


(5) Jeremy Stein (1988) and other authors show that short-termism can arise as a rational response to incentives and leads to underinvestment and misvalued firms.

(6) I return to predictability in asset returns in chapter 8 and counter-cyclical factor investing in chapter 14.

(7) Kydland and Prescott wrote a famous paper in 1977 showing how to implement time-consistent monetary policy. It won them a Nobel Prize in 2004.

(9) Ulysses contracts have been used in medical treatments for mental disorders since the 1970s. Early references are Culver and Gert (1981) and Winston et al. (1982).


(11) Taken from Kaplan’s Private Wealth Management presentation to my Asset Management class on Sept. 19, 2012.


(13) The original Merton (1969, 1971) theory is presented in continuous time so rebalancing happens at every instant.

(14) The bands expand at an approximate rate of the cube root of the transaction costs. For example, for a 5% transaction cost, the bands are approximately \((0.05/0.001)^{1/3} \approx 3.7\) times larger than a 1% transaction cost. This is shown theoretically by Goodman and Ostrov (2010). Solutions for correlated multiple assets generally need to be obtained numerically; see Cai, Judd, and Xu (2013).


(17) Equation (4.15) was originally formulated by Samuelson (1969) and Merton (1969, 1971). The opportunistic weight, or hedging demand, for an investor with log utility (CRRA utility with \(\gamma = 1\)) is zero. Intuitively, a log investor maximizes log returns, and long-horizon log returns are simple sums of one-period returns. Since the portfolio weight is freely chosen each period, the sum is maximized by maximizing each individual term in the sum. That is, a log investor with a long horizon is always a short-run investor.

(18) The term “strategic asset allocation” is much abused in the industry and is often used as an excuse not to rebalance to long-run asset weights. The term itself was introduced by Brennan, Schwartz, and Lagnado (1997).
You might need to hire a rocket scientist after all to compute long-term portfolio weights. See Campbell and Viceira (2002), Brandt (2009), Avramov and Zhou (2010), and Wachter (2010) for literature summaries. The highly technical reader is encouraged to look at Duffie (2001). Cochrane (2013a) shows that you can avoid computing intertemporal hedging demands if you stick to long-term payoffs, but this does not necessarily help if the long-term payoffs are not directly traded.

See Ang and Kjær (2011).

See chapter 7 for a discussion of rational and behavioral determinants of risk premiums.

See Brandt et al. (2005) and Pástor and Stambaugh (2012a).

Chapter 8 shows this is a good system to capture predictability. Although predictability is generally weak, the best predictor variables tend to be valuation ratios.

This system is used by Campbell and Viceira (1999, 2002), for example, and is generalized by Pástor and Stambaugh (2009, 2012a).

Opportunistic demands are not always positive, as Liu (2007) shows for different models of predictability.


This can be valued using risk-neutral pricing. The call value at time 0 is \( \frac{q^{0.124}}{(1.0)^2} = 0.0428 \), where \( q \) is the risk-neutral probability given by \( \frac{1+0.5}{2} = 0.4 \). For an introduction to risk-neutral option pricing, see Bodie, Kane, and Marcus (2011). The put value in equation (4.22) is worth \( \frac{(1-q)^{0.216}}{(1.0)^2} = 0.0643 \) at time 0.

Because of this, Sharpe (2010) advocates long-horizon investors to use “adaptive” asset allocation strategies that just drift up and down with the market instead of actively rebalancing. These are dominated, strictly, by rebalancing with i.i.d. returns as shown in section 2. See also Perold and Sharpe (1988).
(29) In the CAPM and in multifactor models, which we cover in chapter 7, the average investor holds the market portfolio. The average investor does not rebalance. Individual investors can rebalance only if some investors do not. Investors following portfolio insurance strategies, created by Hayne Leland and Mark Rubinstein in 1976, which sell stocks as prices decrease—the opposite of rebalancing (see Rubinstein and Leland (1981)). Kimball et al. (2011) develop a model of equilibrium rebalancing. Chien, Cole, and Lustig’s (2012) equilibrium model also has some investors who rebalance and others who do not.

(30) The optimal utility for rebalancing is
\[ \frac{1}{1+\gamma} \left[ 0.5^2 \times (2.6896)^{1+\gamma} + 2 \times (0.5^2 \times 1.2136)^{1+\gamma} + (0.5^2 \times 0.5476)^{1+\gamma} \right] = 2.3303 \]
and the optimal utility for buy and hold is
\[ \frac{1}{1+\gamma} \left[ 0.5^2 \times (2.8840)^{1+\gamma} + 2 \times (0.5^2 \times 1.0840)^{1+\gamma} + (0.5^2 \times 0.6340)^{1+\gamma} \right] = 2.3270. \]
The certainty equivalent compensation required by the investor to do buy-and-hold investing instead of optimal rebalancing is
\[ \left( \frac{2.3303}{2.3270} \right)^{1+\gamma} - 1 = 0.29 \text{ cents per dollar of initial wealth.} \]
Notice that these are the only calculations where we actually use the real-world 0.5 probability of an upward move in the tree. All the option valuations are done using risk-neutral probabilities.

(31) Recurring regime changes, in contrast, are common. Examples include recessions versus expansions, bear markets versus bull markets, high volatility versus low volatility periods, and, more generally, bad times versus good times. Recurring regime changes are well described by regime-switching processes introduced into economics and finance by Hamilton (1989). See Ang and Timmermann (2012) for a summary.

(32) If asset returns follow Markov processes, then you want to rebalance over assets or strategies that are recurrent.

(33) It was also introduced by the grandfather of portfolio choice theory, Merton, in 1993.

(34) Duration is exposure to the interest rate level factor, which is the most important factor in fixed income investments. See chapter 9.

(35) Ang, Chen, and Sundaresan (2013) demonstrate that this behavior is optimal in a liability-driven investment context with downside risk. Andonov, Bauer, and Cremers (2012) show
that U.S. public pension funds, which have discount rates based on the earning rates of their assets, have incentives to take more risk. They show that funds that are more underfunded have invested proportionately more in risky assets.

(36) For growth-optimal investing, see Latané (1959) and Messmore (1995) for the variance drain terminology. Luenberger (1997) introduced the term “volatility pumping.” A nice collection of papers in the literature is MacLean, Thorp, and Ziemba (2011).

(37) See also Willenbrock (2011) who differentiates between diversification as being necessary to give you different weights over one period but not sufficient to earn the rebalancing premium over multiple periods.

(38) This is true also for modest transaction costs, as shown by Dempster, Evstigneev, and Schenk-Hoppé (2009).

(39) Named after the Danish mathematician Johan Jensen.

(40) The arithmetic return $r$ represents $(1+r)$ at the end of the period. The same amount can be expressed as a geometric return, $g$, where $(1+r)=\exp(g)$. The means of the arithmetic return and the geometric return are related by $E(r) = E(g) - \frac{1}{2}\sigma^2$, where $\sigma$ is the volatility of $r$. This relation holds exactly for log-normal distributions. For more details, see the appendix.

(41) The formal mathematical statement is that there exists a number $M(W)$ that depends on current wealth $W$ such that $\Pr(W_T < W)$ using the Kelly rule is strictly less than $\Pr(W_T < W)$ using any other portfolio for all $T > M(W)$. Thus as $T \to \infty$, the Kelly rule dominates any other rule.


(43) See Hertwig et al. (2004).