Tactical Asset Allocation, Parameter Uncertainty, and Intertemporal Hedging: An Exploratory Study for the Swiss Stock Market

David Rey

First version: March 2004
Current version: October 2004

This research has been carried out within the NCCR FINRISK projects on “Asset Pricing and Portfolio Management”.
Tactical Asset Allocation, Parameter Uncertainty, and Intertemporal Hedging: An Exploratory Study for the Swiss Stock Market

David Rey*

University of Basel (WWZ)

October 7, 2004

Abstract

This paper examines how the evidence of stock market predictability affects optimal portfolio choice for buy-and-hold and dynamic investors with different planning horizons. As in Barberis (2000), particular attention is paid to estimation risk, i.e., uncertainty about the true values of the predictive regression parameters. The empirical analysis for the Swiss stock market shows that there is enough predictability in returns to make short-term buy-and-hold investors time the market. However, optimal weights for long-term buy-and-hold investors are generally not very sensitive to the initial value of the predictive variable, particularly when parameter uncertainty is taken into account. In general, there is no horizon effect and the intertemporal hedging demand is empirically negligible, too. In general, dynamic investors who follow a dynamic rebalancing strategy do not hold more stocks and do not time the market more aggressively than myopic buy-and-hold investors. The Swiss stock market does not provide an intertemporal hedge to changes in investment opportunities. Consequently, common definitions for tactical asset allocation strategies as myopic short-term strategies seem to be empirically justified.

*Department of Finance, WWZ University of Basel, Holbeinstrasse 12, 4051 Basel, Switzerland; Phone: +41-61-267 33 07, Mail: david.rey@unibas.ch. Second preliminary version. The valuable comments of Manuel Amman, Stefan Beiner, and Heinz Zimmermann are gratefully acknowledged. Financial support from the National Center of Competence in Research “Financial Valuation and Risk Management” (NCCR FINRISK) is gratefully acknowledged. The NCCR FINRISK is a research program supported by the Swiss National Science Foundation.
This paper examines how the evidence of stock market predictability affects optimal portfolio choice for buy-and-hold and dynamic investors with different planning horizons. As in Barberis (2000), particular attention is paid to estimation risk, i.e., uncertainty about the true values of the predictive regression parameters. The empirical analysis for the Swiss stock market includes a number of predictive variables found important in previous studies of return predictability. Irrespective of whether we account for parameter uncertainty or not, we find that there is enough predictability in returns to make short-term buy-and-hold investors time the market. However, optimal weights for long-term buy-and-hold investors are generally not very sensitive to the initial value of the predictive variable. There is no horizon effect. The only exception is when we consider the dividend-price ratio as predictive variable, ignore parameter uncertainty, and restrict the analysis to the first subsample from 1975 to 1988. Over this time period, buy-and-hold investors allocate substantially more to stocks, the longer their planning horizon. After incorporating estimation risk, however, the horizon effect diminishes even in this case. Long-horizon investors who ignore it may overallocate to stocks by a sizeable amount. Thus, the weak statistical significance of return predictability makes it important to take parameter uncertainty into account. Indeed, recently, from 1989 to 2002, any evidence of stock market predictability by the dividend-price ratio seems to have completely disappeared. Finally, we show that when a tactical asset allocation strategy is motivated by return predictability, the intertemporal hedging demand is empirically negligible. In general, dynamic investors who follow a dynamic rebalancing strategy do not hold more stocks and do not time the market more aggressively than myopic buy-and-hold investors. The exception is again the case of the dividend-price ratio (and, to a much lesser degree, the earnings-price ratio) as predictive variable and if the analysis is restricted to the early subsample (irrespective of whether parameter uncertainty is taken into account or not). But for the remaining variables parameterizing expected excess returns, the stock market does not provide an intertemporal hedge to changes in investment opportunities. Consequently, common definitions for tactical asset allocation strategies as myopic short-term strategies seem to be empirically justified.

1 Introduction

Stock market predictability is now taken almost as a feature of the data. Indeed, it has been called a “new fact in finance” by Cochrane (1999). This paper examines the implications of this return predictability for investors seeking to make sensible portfolio allocation decisions. Based on Barberis (2000), we approach this question from the perspective of horizon effects. Given the evidence for stock market predictability, should long-term investors allocate their wealth differently from short-term investors? The classic work of Samuelson (1969) and Merton (1969, 1971, 1973) find that when asset returns are independently and identically distributed (i.i.d.), investors with power utility who rebalance
their portfolio optimally should choose the same asset allocation, regardless of their planning horizon. When a tactical asset allocation strategy is motivated by return predictability, however, the investor’s planning horizon may no longer be irrelevant. The extent to which the planning horizon does play a role then demonstrates how return predictability affects portfolio choice. In addition, the results may also clarify the common but controversial advice that investors with long planning horizons should allocate more heavily to stocks.\(^1\)

Of course, on the theoretical side, already Merton (1973) notes that time variation in expected excess returns can potentially introduce horizon effects. Extending the work of Barberis (2000), our contribution is therefore primarily an empirical one. Based on actual historical data on asset returns and predictive variables, we critically review how stock market predictability affects optimal portfolio choice by computing optimal asset allocations for both static buy-and-hold and optimally rebalancing investors.

The investors considered here are assumed to find the historical evidence of return predictability useful in assessing investment opportunities. However, observing only a finite sample, they do not know the true parameter values. As a result, part of the risk they rationally perceive arises due to parameter uncertainty. It is thus an important aspect of our analysis that we account for the fact that the true extent of return predictability is highly uncertain. This is of particular concern in this context because the evidence of time variation in expected excess returns is often weak (see Rey (2003a,b) for a thorough discussion regarding this). While some investors might react to this weak statistical evidence by completely discarding the notion that returns are predictable, others might instead even ignore this uncertainty about the true predictive power of the prevailing predictive variables and assume that the regression parameters are known precisely. Both of these views are probably flawed. Instead, when constructing optimal portfolios, we explicitly account for the uncertainty about the regression parameters, i.e., for estimation risk.

Based on Barberis (2000), we approach the portfolio choice problem in a discrete-time setting for an investor with power utility over terminal wealth. This corresponds to the definition for tactical asset allocation proposed in Rey (2004). There are two assets: a short-term risk-free asset and a stock market index. The investor is assumed to use a VAR model to predict excess equity returns, where the state vector in the VAR includes asset returns and a number of predictive variables. Thus, by varying the number of these predictive variables in the state vector, we can easily compare the optimal allocation of investors who take return predictability into account to that of investors who ignore it.

We follow Barberis (2000) and incorporate parameter uncertainty with a Bayesian approach. Given the data, the uncertainty about the regression parameters is summarized by the posterior distribution of the parameters. Rather than constructing the distribution of future excess returns conditional on fixed parameter estimates, the Bayesian approach of Barberis (2000) integrates over

---

\(^1\)See Samuelson (1994), Bodie (1995), Siegel (1998), and particularly Campbell and Viceira (2002) for recent contributions to this debate.
the uncertainty in the parameters captured by the posterior distribution. The resulting predictive distribution of future returns is then solely based on observed data, and no longer on any fixed parameter values. Finally, the effect of parameter uncertainty on the portfolio allocation problem is revealed by comparing the optimal solutions conditioned on fixed parameters and where we integrate over the posterior distribution.

We do not restrict the analysis to the dividend-price ratio as predictive variable. Instead, the empirical analysis for the Swiss stock market includes a number of predictive variables found important in previous studies of return predictability, as well as their stochastically detrended counterparts. As in Barberis (2000), our first set of results relates to the case where parameter uncertainty is ignored. To be more precisely, we first assume that investors allocate their portfolio taking the parameters as fixed at their estimated values and that they either solve a myopic buy-and-hold portfolio choice problem or a dynamic problem with optimal rebalancing.

In the buy-and-hold case, we find that there is enough return predictability to make short-term investors time the market. However, optimal weights for long-term investors are generally not very sensitive to the initial value of the predictive variable. There is no horizon effect. The only exception is when we consider the dividend-price ratio as predictive variable and restrict the analysis to the time period from 1975 to 1988. Here we find that return predictability leads to strong horizon effects. Investors with a planning horizon of, say, ten years allocate significantly more to stocks than investors with, say, a one-year horizon. Thus, it seems that time variation in expected excess returns induces mean reversion in realized returns, thereby lowering the variance of cumulative returns over long horizons. As a result, stocks appear less risky to long-horizon investors and leads them to allocate more to equities than would investors with shorter planning horizons.

In the case of the dividend-price ratio, we also find strong horizon effects when we assume that investors solve a dynamic portfolio choice problem and therefore rebalance optimally at regular intervals. The results here must be interpreted differently, however. Investors only hold substantially more in equities at longer planning horizons when they are more risk averse than log-utility investors. In this case, the additional stock holdings of investors with a long planning horizon are the intertemporal hedging demands first described in Merton (1973). When expected excess returns are parameterized by the dividend-price ratio, shocks to expected excess returns are reliably negatively correlated with shocks to realized returns. Therefore, when investors choose to hedge these changes in the investment opportunity set, they do so by increasing their holdings of stocks. Recently, however, from 1989 to 2002, any evidence of return predictability by the dividend-price ratio, and therefore intertemporal hedging demands, seems to have completely disappeared. Similarly, for the remaining and stochastically detrended variables parameterizing expected excess returns, the stock market does not seem to provide a good intertemporal hedge to changes in investment opportunities.

So far, we assumed that investors ignore parameter uncertainty. Yet, while
incorporating parameter uncertainty does not seem to significantly change the asset allocations of short-term buy-and-hold investors, both static buy-and-hold and dynamic investors with long planning horizons may reduce their allocations to equities when they account for estimation risk. This is particularly true in the case of the dividend-price ratio as predictive variable and when we restrict the analysis to the early subsample. After incorporating estimation risk, horizon effects are still present, but much less prominent. Indeed, uncertainty about the regression parameters can be large enough to reverse the direction of the results. Investors with a long planning horizon who ignore it run the risk to overallocate to stocks by a sizeable amount.

Our analysis critically reviews the empirical work on portfolio choice in the presence of time-varying expected excess returns. Besides Barberis (2000), the two most important contributions are from Brennan, Schwartz, and Lagnado (1997) and Campbell and Viceira (1999, 2002). While Brennan, Schwartz, and Lagnado (1997) work in continuous time and analyze a dynamic programming portfolio choice problem for a small number of assets and predictive variables, Campbell and Viceira (1999, 2002) develop an analytical approximation to the more general problem of deriving both optimal consumption and portfolio rules for an infinite-horizon investor with Epstein-Zin utility. Both papers, however, ignore the effect of parameter uncertainty on optimal portfolio choice. Among the first papers that acknowledge the issue of parameter uncertainty are Bawa and Klein (1976) and Bawa, Brown, and Klein (1979). Their analysis is restricted to the context of i.i.d. returns, however. In contrast, Kandel and Stambaugh (1996) demonstrate the importance of parameter uncertainty when returns are predictable. Using a Bayesian framework similar to Barberis (2000), they show that for a myopic short-horizon investor, the optimal allocation to stocks can be sensitive to the current values of predictive variables, even though the statistical evidence for this return predictability may be weak. We follow Barberis (2000) and examine a wider range of planning horizons, including the intertemporal portfolio choice problem.

In contrast to Barberis (2000), however, we suggest that when a tactical asset allocation strategy is motivated by return predictability (which is a sine qua non of tactical asset allocation; Brennan, Schwartz, and Lagnado, 1997), the consideration of intertemporal hedging demands may be neglected. In general, dynamic investors who follow a dynamic rebalancing strategy do not hold more stocks and do not time the market more aggressively than myopic buy-and-hold investors. Consequently, common definitions of tactical asset allocation strategies as myopic short-term buy-and-hold strategies are therefore empirically justified (see Rey, 2004).

We proceed as follows. Section 2 reviews the asset allocation framework for myopic buy-and-hold investors originally developed in Barberis (2000) and also presents the respective empirical results. Section 3 turns to the intertemporal problem of optimal rebalancing and contrasts the results with those in the myopic buy-and-hold case. Section 4 concludes. A summary of the Swiss stock market data and the numerical procedures is provided in Appendix A and B, respectively.
2 Asset Allocation Framework for a Buy-and-Hold Investor

This section reviews Barberis’ (2000) framework for investigating how return predictability and estimation risk affect the optimal portfolio choice of buy-and-hold investors.

First of all, however, it is necessary to specify the exact choices that investors with a long planning horizon are allowed to make. We distinguish between two different ways of formulating the portfolio choice problem.

The first possibility is a buy-and-hold strategy. In this case, investors with, say, a ten-year planning horizon choose an allocation at the beginning of the first year, and do neither adjust nor rebalance their portfolio again until the ten years are over. The second possibility of formulating the portfolio choice problem is a dynamic rebalancing strategy. Following this strategy, investors split their planning horizon of ten years into several investment horizons (rebalancing intervals) of arbitrary length, say, one year. They then choose an allocation at the beginning of the first year, knowing that at the start of every new year, they will optimally adjust their portfolio using the new information available at that time. In other words, in contrast to the buy-and-hold strategy, the planning and investment horizon are not the same for dynamic investors. It is important to distinguish between “dynamic rebalancing” and “myopic rebalancing”. Following a myopic rebalancing strategy, investors choose an allocation at the beginning of the first year, knowing that they will reset (i.e., effectively rebalance) their portfolio to that same allocation at the start of every new year. While this is myopic in the sense that investors do not use any of the new information they have once a year has passed, it is the optimal intertemporal strategy when returns are assumed to be i.i.d. In contrast, dynamic rebalancing means that investors reoptimize and thus adjust their portfolios using the new information at each time, likely resulting in optimal portfolio weights varying over time.

2.1 The Myopic Portfolio Choice Problem

Assume that a buy-and-hold investor is at time $T$ and has a planning horizon of $T$ months. He may choose from the following two assets: a short-term risk-free interest rate and a stock market index. For simplicity, the continuously compounded real monthly risk-free rate is assumed to be constant, $r_f$.

If initial wealth, $W_T$, is normalized to $W_T = 1$, and $\omega$ is the allocation to stocks, then end-of-horizon wealth, $W_{T+\hat{T}}$, is trivially given by

$$W_{T+\hat{T}} = (1 - \omega) \exp\left(r_f \hat{T}\right) + \omega \exp\left(r_f \hat{T} + e_{T+1} + ... + e_{T+\hat{T}}\right), \tag{1}$$

where $e_t$ denotes the continuously compounded excess return over month $t$.

As in Barberis (2000), we further assume that the investor’s preferences over terminal wealth are described by constant relative risk aversion (CRRA) power
utility functions

\[ v(W) = \frac{W^{1-\zeta} - 1}{1-\zeta}. \]  (2)

Given the cumulative log excess return over \( \bar{T} \) periods by

\[ e_{T-T+\bar{T}} \equiv e_{T+1} + e_{T+2} + ... + e_{T+\bar{T}}, \]

the buy-and-hold investor maximizes expected utility given by

\[ \max_{\omega} E_T \left\{ \frac{(1-\omega) \exp \left( r_f \hat{T} \right) + \omega \exp \left( r_f \hat{T} + e_{T+\bar{T}} \right) }{1-\zeta} \right\}^{1-\zeta}. \]  (3)

Of course, maximizing expected utility solely over final wealth corresponds to the definition for tactical asset allocation proposed in Rey (2004).

One of the crucial questions addressed in this paper is the issue of which distribution the investor should use in calculating the conditional expectation in equation (3). Depending on whether he recognizes return predictability and/or accounts for estimation risk, the distribution may be very different. In either case of predictability and no predictability, the conditional distribution of cumulative excess returns, \( e_{T-T+\bar{T}} \), is, in general, given by \( p \left( e_{T-T+\bar{T}} \mid \theta, z \right) \), where \( z \) is the data observed by the investor up until the start of his planning horizon and \( \theta \) denotes the set of regression parameters. Given an estimate \( \hat{\theta} \) of the parameter values, the investor then maximizes expected utility according to (Barberis, 2000, eq. 7)

\[ \max_{\omega} \int v \left( W_{T+\bar{T}} \right) p \left( e_{T-T+\bar{T}} \mid \hat{\theta}, z \right) de_{T-T+\bar{T}}. \]  (4)

However, the problem with this approach is that it ignores the fact that the true values of the parameters in \( \theta \) are not known precisely. Given only a limited sample of data, \( z \), there may be substantial uncertainty about \( \theta \) and it thus seems reasonable to account for this estimation risk, particularly so for long-horizon investors.

To take estimation risk into account, Barberis (2000) proposes a Bayesian approach of a posterior distribution, \( p \left( \theta \mid z \right) \). This posterior distribution summarizes the uncertainty about the regression parameters conditional only on the data observed so far. The predictive distribution of long-horizon returns is then obtained by integrating over this posterior distribution. Barberis (2000, eq. 8) shows that the predictive distribution is conditioned only on the sample observed, and no longer on any fixed parameter estimate:

\[ p \left( e_{T-T+\bar{T}} \mid z \right) = \int p \left( e_{T-T+\bar{T}} \mid \theta, z \right) p \left( \theta \mid z \right) d\theta. \]  (5)

The resulting portfolio choice problem for the buy-and-hold investor to solve is then (Barberis, 2000, eq. 9)

\[ \max_{\omega} \int v \left( W_{T+\bar{T}} \right) p \left( e_{T-T+\bar{T}} \mid z \right) de_{T-T+\bar{T}}. \]  (6)
As described in Appendix B, these integrals are evaluated numerically by simulation. In our examples, the conditional distribution $p\left(e_{T^{+}\rightarrow T}|\hat{\theta},z\right)$ is Normal. Therefore, the integral in equation (4) is approximated by generating a large sample of independent draws from this Normal distribution, and averaging $v(W_{T^{+}\rightarrow T})$ over all the draws. In the case of equation (6), Barberis (2000, eq. 10) shows that it is helpful to rewrite the problem as

$$
\max_{\omega} \int v(W_{T^{+}\rightarrow T}) p\left(e_{T^{+}\rightarrow T},\theta|z\right) de_{T^{+}\rightarrow T}d\theta = \\
\max_{\omega} \int v(W_{T^{+}\rightarrow T}) p\left(e_{T^{+}\rightarrow T}|\theta,z\right) p(\theta|z) de_{T^{+}\rightarrow T}d\theta. \tag{7}
$$

Therefore, the integral can be evaluated by sampling from the joint distribution $p\left(e_{T^{+}\rightarrow T},\theta|z\right)$, and then averaging $v(W_{T^{+}\rightarrow T})$ over those draws. To be more precise, we sample from the joint distribution by first sampling from the posterior $p(\theta|z)$ and then from the conditional $p\left(e_{T^{+}\rightarrow T}|\theta,z\right)$.

By comparing the solution to problem (4), which ignores parameter uncertainty, with the solution to problem (6), which takes it into account, we may easily see how estimation risk affects optimal portfolio choice.

Rather than incorporate both predictability and estimation risk at once, we introduce them one at a time. That is, we start out in the next section by considering the special case of i.i.d. returns and look at how parameter uncertainty alone affects portfolio allocation. Only then we move to the more general case which additionally allows for predictability in returns.

### 2.2 Case I: i.i.d. Returns

As a base case, suppose that stock index returns are unpredictable and follow a random walk with normally distributed increments,

$$e_t = \mu + \xi_t, \tag{8}$$

where $\xi_t \sim \text{i.i.d. } N(0,\sigma^2)$.

The investor has two choices of distribution for maximizing expected utility. When he incorporates parameter uncertainty, he may use the predictive distribution of excess returns, $p\left(e_{T^{+}\rightarrow T}|\hat{\theta},\hat{\sigma}^2,z\right)$, where the data is summarized in $e = [e_1 \cdots e_T]^\prime$. Alternatively, when he ignores parameter uncertainty, he may calculate the expectation in equation (4) over the distribution of excess returns conditional on fixed parameter values, $p\left(e_{T^{+}\rightarrow T}|\hat{\mu},\hat{\sigma}^2,\hat{\theta}\right)$.

#### 2.2.1 Ignoring Estimation Risk

When estimation risk is ignored, a buy-and-hold investor maximizes expected utility as stated in equation (4), using the return distribution conditional on the
parameter values fixed at their estimated values, \( p(e_{T \rightarrow T^+} | \hat{\mu}, \hat{\sigma}^2, e) \). Conditional on \( \mu \) and \( \sigma^2 \) and with normally distributed increments, we have
\[
e_{T+1} = \mu + \xi_{T+1}, \ldots, e_{T^+} = \mu + \xi_{T^+},
\]
such that cumulative excess returns \( e_{T \rightarrow T^+} \) are normally distributed with mean \( \hat{T}\mu \) and variance \( \hat{T}\sigma^2 \). Since we assume that the investor treats the estimated mean \( \hat{\mu} \) and variance \( \hat{\sigma}^2 \) as the true values of \( \mu \) and \( \sigma^2 \), we compute the optimal allocation to stocks, \( \omega \), by drawing a large sample from a Normal distribution with mean \( \hat{T}\mu \) and variance \( \hat{T}\sigma^2 \).

The numerical procedure, described in Appendix B, is then repeated for planning horizons, \( \hat{T} \), ranging from one year to ten years in one-year increments (as in Barberis, 2000), and for several values of risk aversion coefficients, \( \zeta \).

### 2.2.2 Accounting for Parameter Uncertainty

As pointed out in equation (7), there are two steps to sampling from the predictive distribution of long-horizon returns, \( p(e_{T \rightarrow T^+} | e) \). To produce a large sample from the predictive distribution, we generate a large sample from the posterior distribution of the parameters \( p(\mu, \sigma^2 | e) \), and then, for each of the \( (\mu, \sigma^2) \) pairs drawn, sample once from the distribution of long-horizon returns conditional on past data and the parameters \( p(e_{T \rightarrow T^+} | \mu, \sigma^2, e) \), a Normal distribution.

To construct the posterior distribution \( p(\mu, \sigma^2 | e) \), we require a prior. As in Barberis (2000, p. 235), we use a conservative uninformative prior,
\[
p(\mu, \sigma^2) \propto \frac{1}{\sigma^2}.
\]

Zellner (1971) and Barberis (2000) show that the posterior is then given by
\[
\sigma^2 | e \sim IG \left( \frac{T-1}{2}, \frac{1}{2} \sum_{t=1}^{T} (e_t - \bar{e})^2 \right)
\]
\[
\mu | \sigma^2, e \sim N \left( \bar{e}, \frac{\sigma^2}{T} \right),
\]
where
\[
\bar{e} = \frac{1}{T} \sum_{t=1}^{T} e_t.
\]
To get an accurate representation of the posterior distribution \( p(\mu, \sigma^2 | e) \), we first sample from the marginal \( p(\sigma^2 | e) \), an Inverse Gamma (IG) distribution,

---

\(^2\)As indicated in Barberis (2000, footnote 10), another reasonable approach would be to use a more informative prior that puts zero weight on negative values of \( \mu \), reflecting the observation in Merton (1980) that the expected market risk premium should be positive; see also Boudoukh, Richardson, and Smith (1993).
and then, given the $\sigma^2$ draw, from the conditional $p(\mu \mid \sigma^2, e)$, a Normal distribution. To obtain a large sample from the predictive distribution $p(e_{T-T+\hat{T}} \mid e)$, we sample one point from the Normal distribution with mean $\hat{T}\mu$ and variance $T\sigma^2$ for each draw of $\mu$ and $\sigma^2$ from the posterior $p(\mu, \sigma^2 \mid e)$.

The numerical procedure is again repeated for planning horizons ranging from one year to 10 years in one-year increments, and for several values of risk aversion coefficients, $\zeta$.  

2.3 Case II: Return Predictability

To see whether predictability in stock market returns has any effect on portfolio choice, we compare the allocation of an investor who recognizes return predictability to that of an investor who is blind to it, i.e., who assumes that excess returns are unpredictable i.i.d. as in the previous section. In the case of return predictability, we model excess returns on the stock index using a VAR framework similar to that in Campbell (1991), Hodrick (1992), and particularly Barberis (2000, eq. 2). It takes the form

$$z_t = a + Bx_{t-1} + \xi_t,$$

with $z'_t = [e_t, x'_t]$, $x_t = [x_{1,t} \cdots x_{n,t}]'$, and $\xi_t \sim \text{i.i.d. } N(0, \Sigma)$. $e_t$, the first component of $z_t$, denotes the continuously compounded excess return over month $t$, as before. The remaining components of $z_t$, $x_t$, consist of the $n$ predictive variables suspected relevant for predicting excess returns, such as the dividend-price ratio.

The problem faced at time $T$ by a buy-and-hold investor with a planning horizon of $\hat{T}$ months is given by equation (4), which ignores parameter uncertainty, and equation (6), which takes parameter uncertainty into account, respectively. An investor who ignores estimation risk uses the distribution of future excess returns conditional on both past data and fixed parameter values, $p(e_{T-T+\hat{T}} \mid \theta, z)$. In contrast, the investor who takes parameter uncertainty into account samples from the predictive distribution, conditional only on past data and not on the parameters, $p(e_{T-T+\hat{T}} \mid z)$.  

2.3.1 Ignoring Estimation Risk

When estimation risk is ignored, a buy-and-hold investor maximizes expected utility as stated in equation (4), using the return distribution conditional on the parameter values fixed at their estimated values. Once the set of predictive variables is specified, a standard procedure is to estimate the VAR parameters $\theta = (a, B, \Sigma)$, and then iterate the model forward with the parameters fixed at their estimated values. This generates a distribution of future excess returns conditional on the set of estimated parameter values.

Note that since $z_t = a + Bx_{t-1} + \xi_t$, we may write $z_t = a + B_0 z_{t-1} + \xi_t$, with $B_0 = [0 \ B]$, where $0$ is an $(n+1, 1)$ vector of zeros. Therefore (Barberis,
2000, eq. 17),

\[
\begin{align*}
    z_{T+1} &= a + B_0 z_T + \xi_{T+1} \\
    z_{T+2} &= a + B_0 a + B_0^2 z_T + \xi_{T+2} + B_0 \xi_{T+1} \\
    &\vdots \\
    z_{T+\hat{T}} &= a + B_0 a + B_0^2 a + \ldots + B_0^{\hat{T}-1} \\
    &+ B_0^2 z_T \\
    &+ \xi_{T+\hat{T}} + B_0 \xi_{T+\hat{T}-1} + B_0^2 \xi_{T+\hat{T}-2} + \ldots + B_0^{\hat{T}-2} \xi_{T+2} + B_0^{\hat{T}-1} \xi_{T+1}.
\end{align*}
\]

Conditional on \(a, B\) and \(\Sigma\), \(Z_{T\rightarrow T+\hat{T}} \equiv z_{T+1} + z_{T+2} + \ldots + z_{T+\hat{T}}\) is normally distributed with mean vector (Barberis, 2000, eq. 18)

\[
\mu_{T\rightarrow T+\hat{T}} = \hat{T} a + \left(\hat{T} - 1\right) B_0 a + \left(\hat{T} - 2\right) B_0^2 a + \ldots + B_0^{\hat{T}-1} a \\
+ \left(B_0 + B_0^2 + \ldots + B_0^{\hat{T}-1}\right) z_T,
\]

and variance matrix given by (Barberis, 2000, eq. 19)

\[
\begin{align*}
\Sigma_{T\rightarrow T+\hat{T}} &= \Sigma \\
&+ \left(I + B_0\right) \Sigma \left(I + B_0\right)' \\
&+ \left(I + B_0 + B_0^2\right) \Sigma \left(I + B_0 + B_0^2\right)' \\
&\quad \vdots \\
&+ \left(I + B_0 + B_0^2 + \ldots + B_0^{\hat{T}-1}\right) \Sigma \left(I + B_0 + B_0^2 + \ldots + B_0^{\hat{T}-1}\right)'.
\end{align*}
\]

When parameter uncertainty is ignored, the investor treats the estimated parameter values as the true values of \(a, B\) and \(\Sigma\), and we can therefore compute the optimal allocation to stocks, \(\omega\), by drawing a large sample from a Normal distribution with mean vector \(\mu_{T\rightarrow T+\hat{T}}\) and variance matrix \(\Sigma_{T\rightarrow T+\hat{T}}\).

Of course, the investor’s distribution of future excess returns depends on the value of the vector of predictive variables at the beginning of the planning horizon, \(x_T\). The initial value of the vector \(x_T\) enters equation (10) through \(z_T\).

As in the case of i.i.d. returns, the numerical procedure, described in Appendix B, is then repeated for planning horizons ranging from one year to ten years in one-year increments, for several values of risk aversion coefficients, \(\zeta\), and for several initial values of the predictive variables, \(x_T\).

### 2.3.2 Accounting for Parameter Uncertainty

The problem with the approach described in the previous section is that it ignores the fact that the true values of \(a, B\) and \(\Sigma\) are not known precisely.
There may be substantial uncertainty about the regression coefficients and the variance matrix. For a long-horizon investor in particular, it is important to take estimation risk into account. The procedure for sampling from the predictive distribution is similar to the case with i.i.d. returns. To obtain a large sample from the predictive distribution of cumulative excess returns, conditional only on past data, we first generate a large sample from the posterior distribution of the parameters, \( p(\mathbf{a}, \mathbf{B}, \Sigma \mid \mathbf{z}) \), and then, for each of the set of the parameter values drawn, sample once from the distribution of returns conditional on past data and the parameters, a Normal distribution.

To compute the posterior distribution \( p(\mathbf{a}, \mathbf{B}, \Sigma \mid \mathbf{z}) \), Barberis (2000, eq. 15 and 16) rewrites the model as

\[
\begin{bmatrix}
\mathbf{z}_2' \\
\vdots \\
\mathbf{z}_T'
\end{bmatrix} =
\begin{bmatrix}
1 & \mathbf{x}_1' \\
\vdots & \vdots \\
1 & \mathbf{x}_{T-1}'
\end{bmatrix}
\begin{bmatrix}
\mathbf{a}' \\
\mathbf{B}'
\end{bmatrix} +
\begin{bmatrix}
\xi_2' \\
\vdots \\
\xi_T'
\end{bmatrix},
\]

(12)
or

\[
\mathbf{Z} = \mathbf{X}\mathbf{C} + \mathbf{E},
\]

where \( \mathbf{Z} \) is a \((T-1, n+1)\) matrix with the vectors \( \mathbf{z}_2, \ldots, \mathbf{z}_T \) as rows; \( \mathbf{X} \) is a \((T-1, n+1)\) matrix with the vectors \([1 \mathbf{x}_1] \cdots [1 \mathbf{x}_{T-1}]\) as rows, and \( \mathbf{E} \) is a \((T-1, n+1)\) matrix with the vectors \( \xi_2, \ldots, \xi_T \) as rows. \( \mathbf{C} \) is a \((n+1, n+1)\) matrix with top row \( \mathbf{a}' \) and the matrix \( \mathbf{B}' \) below that.

The Bayesian analysis of a multivariate regression model in the traditional case with exogenous regressors is discussed in Zellner (1971). As briefly discussed in Barberis (2000, p. 240), his analysis carries over directly to our dynamic regression framework with endogenous regressors. A standard uninformative (diffuse) prior is (Hamilton, 1994; also Stambaugh, 1999)

\[
p(\mathbf{C}, \Sigma) \propto \det(\Sigma)^{-(n+2)/2}.
\]

The posterior \( p(\mathbf{C}, \Sigma^{-1} \mid \mathbf{z}) \) is then given by

\[
\Sigma^{-1} \mid \mathbf{z} \sim \text{Wish}(T - n - 2, \mathbf{S}^{-1}),
\]

where

\[
\mathbf{S} = (\mathbf{Z} - \mathbf{X}\hat{\mathbf{C}})' (\mathbf{Z} - \mathbf{X}\hat{\mathbf{C}}),
\]

and

\[
\text{vec}(\mathbf{C}) \mid \Sigma, \mathbf{z} \sim N\left(\text{vec}(\hat{\mathbf{C}}), \Sigma \otimes (\mathbf{X}'\mathbf{X})^{-1}\right),
\]

with \( \hat{\mathbf{C}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Z} \).

We sample from the posterior distribution by first drawing from the marginal \( p(\Sigma^{-1} \mid \mathbf{z}) \), a Wishart (Wish) distribution, and then, given the \( \Sigma^{-1} \) draw, from the conditional \( p(\mathbf{C} \mid \Sigma, \mathbf{z}) \), a multivariate Normal distribution. Repeating this many times gives an accurate representation of the posterior distribution. Second, for each draw of \( \mathbf{a}, \mathbf{B} \) and \( \Sigma \) from the posterior \( p(\mathbf{a}, \mathbf{B}, \Sigma \mid \mathbf{z}) \), we sample from the Normal distribution with mean vector \( \mathbf{\mu}_{T \rightarrow T + \hat{T}} \) and variance matrix \( \Sigma_{T \rightarrow T + \hat{T}} \). This gives a large sample of the predictive distribution.
\[ p \left( c_{T+\tau} \mid x \right) \], which we can use to compute the optimal allocation, \( \omega \), when taking both return predictability and estimation risk into account.

The numerical procedure is again repeated for planning horizons ranging from one year to ten years in one-year increments, for several values of risk aversion coefficients, \( \zeta \), and for several initial values of the predictive variables, \( x_T \).

2.4 Empirical Results

This section demonstrates the impact of return predictability and parameter uncertainty on portfolio choice of buy-and-hold investors, using data from the Swiss stock market over several time periods. To do this, we compute optimal allocations to stocks using the four different choices for the distribution of future excess returns introduced above. These distributions differ in whether they take return predictability and estimation risk into account. As in Barberis (2000), we do these calculations for planning horizons \( \hat{T} \) ranging from one year to ten years in one-year increments, for several values of risk aversion coefficients, \( \zeta \), and, in the case of return predictability, for several initial values of the predictive variables, \( x_T \).

Rather than showing the empirical results for return predictability and estimation risk at once, we will introduce them one at a time. That is, we again start out in the next section by considering the special case of i.i.d. returns and look at how parameter uncertainty alone affects portfolio allocation. We then move to the more general case and additionally allow for predictability in excess returns.

2.4.1 The Data

Our investment universe consists of monthly observations on continuously compounded excess stock market returns over January 1975 through December 2002 (336 observations). The stock market data is summarized in Appendix A. In our numerical work, we fix the continuously compounded real monthly Swiss interbank rate, \( r_f \), at the arithmetic average over the sample period under consideration.

In deciding which predictive variables to include, attention was given to those variables found important in previous studies of return predictability. Of course, there is a natural concern about return predictability uncovered through collective “data-snooping” by a series of researchers (Lo and MacKinlay, 1990; Foster, Smith, and Whaley, 1997; Ferson, Sarkissian, and Simin, 2003, 2004). However, most of this research is based on U.S. data and, to our knowledge, there is no study for the Swiss stock market that uses data covering the period starting in 1975 and that includes the recent bear market. We consider each of the following seven predictive variables:

(i) Dividend-price ratio, log (DPR),
(ii) Earnings-price ratio, log (EPR),
(iii) Term spread (TERM),
The dividend-price ratio/earnings-price ratio is measured as the sum of dividends/earnings paid on the index over the previous year, divided by the current level of the index. The term spread is the difference between the (log) nominal yield on long-term government bonds provided by IMF and the (log) nominal three-month Swiss interbank rate. In the same way as Goyal and Santa-Clara (2003), we compute the monthly realized variance of the real stock market returns using within-month daily return data for each month as

\[
\text{Var}_{\text{Market}}^t = \sum_{d=1}^{D_t} r_{m,d}^2 + 2 \sum_{d=2}^{D_t} r_{m,d} r_{m,d-1},
\]

where \(D_t\) is the number of days in month \(t\) and \(r_{m,d}\) is the continuously compounded real stock market return on day \(d\). The second term on the right-hand side adjusts for the autocorrelation in daily returns using the approach proposed by French, Schwert, and Stambaugh (1987). The U.S. TED spread is calculated as the difference between (log) three-month Eurodollar rates and (log) three-month Treasury Bill rates, provided by the Federal Reserve Board of Governors. Finally, the U.S. default risk spread is formed as the difference in annualized (log) yields of Moody’s Baa and Aaa rated bonds.

Again, monthly data are used throughout, spanning 336/337 months from December 1974 to November/December 2002.

Motivated by the recent contributions of Ferson, Sarkissian, and Simín (2003, 2004), a second subset includes the same seven predictive variables, but now transformed in the following simple way. We transform the predictive variables by subtracting off a trailing moving average of its own past values,

\[
x_{t-1}^* = x_{t-1} - \frac{1}{12} \sum_{\tau=1}^{12} x_{t-1-\tau}.
\]

In words, we subtract a backward one-year moving average of past values from the prevailing value of the predictive variable to get a “stochastically detrended” time series that is equivalent to a triangularly weighted moving average of past changes in the predictive variable, where the weights decline as one moves back in time. Accordingly, the detrended time series is stationary if changes in the predictive variable are stationary. While this stochastic detrending method has already been used by Campbell (1991) and Hodrick (1992), only recently Ferson, Sarkissian, and Simín (2003, 2004) show that this is the most practically useful insurance against spurious regression bias (and therefore data mining). Since most of the above predictive variables are either manifestly non-stationary (realized stock market volatility is the exception), or, if not, their behavior is close enough to unit-root non-stationarity for small-sample statistics to be affected, it is interesting to compare the characteristics of the two data subsets.

---

3 A full list of references is provided in Rey (2003a,b).
Table 1:

Parameter estimates for an i.i.d. model of stock returns.
The results in this table are based on the model given in equation (8). The table gives estimates for the (annualized) mean and the (annualized) standard deviation of continuously compounded monthly excess returns over three different time periods. The top row uses data from January 1975 to December 2002. Estimates for two subsamples are showed below. The first subsample uses data from January 1975 to December 1988; the second subsample is based on the time period from January 1989 to December 2002. The table includes the average values of the continuously compounded real monthly Swiss interbank rates over the respective time periods.

2.4.2 Case I: i.i.d. Returns

When estimation risk is ignored, buy-and-hold investors maximize expected utility using the return distribution conditional on the parameter values fixed at their estimated values. When excess returns are assumed i.i.d., we thus need an estimate of the realized mean excess return, \( \hat{\mu} \), and the standard deviation of the realized risk premium, \( \hat{\sigma} \). Table 1 presents these estimates for three different time periods. The top row uses the full sample of monthly data on stock index returns from January 1975 to December 2002. Estimates for two subsamples are showed below. The first subsample uses data from January 1975 to December 1988, covering the first half of the total time period, the second subsample is based on data from January 1989 to December 2002, covering the second half of the full sample. Table 1 also includes the average values of the continuously compounded real monthly Swiss interbank rate over the respective time periods.

When investors use the full sample from 1975 to 2002, for example, they are confronted with a mean of monthly excess returns of 6.02% p.a. and with a volatility of 16.14% p.a. Over the first time period from 1975 to 1988, the average excess return was higher with 6.65% p.a., while the volatility was lower at 14.75% p.a. In contrast, stocks were somewhat riskier over the more recent subsample and did not perform equally well; the average risk premium was only 5.39% p.a. with a volatility of 17.45% p.a. Consequently, we should not be surprised when the optimal allocation to stocks is much lower using parameter estimates based on the second, more recent data sample.

Figure 1 shows the optimal percentage 100\( \omega \) percent allocated to the stock market index, plotted against the planning horizon in years. The graph on the left shows optimal allocations to stocks for investors with a risk aversion coefficient of \( \zeta = 5 \), the one on the right is based on \( \zeta = 10 \). In both cases,
The investor follows a buy-and-hold strategy, uses an i.i.d. model for asset returns, and has power utility over terminal wealth. The solid lines correspond to the case where the investor ignores parameter uncertainty, the dotted lines to the case where he accounts for it. Both graphs use data from 1975 to 2002. \( \zeta \) denotes the coefficient of relative risk aversion.

investors use all data from 1975 to 2002. The solid lines show the allocation conditional on fixed parameter values, and the dotted lines show the allocation when investors account for parameter uncertainty.

The solid line is horizontal for both levels of risk aversion. Thus, investors who ignore the uncertainty about the mean and volatility of excess returns allocate the same amount to stocks, regardless of their planning horizon. However, when parameter uncertainty is explicitly incorporated into the investor’s decision-making framework, Figure 1 also shows that the allocation to stocks falls as the planning horizon increases. Parameter uncertainty can thus introduce (negative) horizon effects even within the context of an i.i.d. model for returns.

The magnitude of the negative horizon effect induced by parameter uncertainty is quite substantial. At the ten-year horizon and with \( \zeta = 5 \), the allocation to stocks is \( \omega = 56\% \) for investors who ignore parameter uncertainty, and \( \omega = 44\% \) for investors who account for it. If investors are more risk averse with \( \zeta = 10 \), the reduction in the allocation to stocks increases even further expressed as a percentage.

According to Barberis (2000, p. 238), these results can be explained as follows. When parameter uncertainty is ignored, both the mean and the variance grow linearly with the planning horizon, \( T \). However, when parameter uncertainty is taken into account, the distribution of long-horizon returns faces an additional source of uncertainty. Parameter uncertainty thus makes the variance of the distribution of cumulative excess returns increase faster than linearly with the planning horizon. Indeed, in the presence of parameter uncertainty and from the viewpoint of the investors, returns are no longer i.i.d., but rather positively serially correlated. Positive serial correlation makes stocks look riskier in the long run, and investors therefore reduce the amount they allocate to equities.
Figure 2: Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses an i.i.d. model for asset returns, and has power utility over terminal wealth. The solid lines correspond to the case where the investor ignores parameter uncertainty, the dotted lines to the case where he accounts for it. The top row graphs use data from 1975 to 1988, the lower two use data from 1989 to 2002. $\zeta$ denotes the coefficient of relative risk aversion.
As demonstrated in Figure 2, investors reduce their allocation to equities even further when their investment decisions are based on one of the subsamples from 1975 to 1988 or 1989 to 2002, respectively. This reflects the greater impact of the higher parameter uncertainty faced by investors who use only a short data sample (168 instead of 336 monthly observations). Because of the lower mean and higher volatility in the more recent subsample (see Table 1), it is not surprising that Figure 2 also shows that the optimal allocation to stocks is significantly lower for this time period.

2.4.3 Case II: Return Predictability

Now that the impact of parameter uncertainty alone has been illustrated, return predictability can be introduced as well. We return to the VAR framework of equation (9) and consider the special case of \( n = 1 \), i.e., we restrict the vector \( z_t \) to only two components: the excess stock index return, \( e_t \), and a single predictive variable, \( x_t \equiv x_{1,t} \), which potentially captures an important component of the variation in expected excess returns. We can thus write the predictive regression model as

\[
\begin{align*}
  e_t &= \alpha + \beta x_{t-1} + \xi_t \\
  x_t &= \gamma + \delta x_{t-1} + \eta_t,
\end{align*}
\]

(13)

where

\[
\begin{pmatrix}
  \xi_t \\
  \eta_t
\end{pmatrix}
\sim N\left(0, \begin{pmatrix}
  \sigma_\xi^2 & \sigma_{\xi\eta} \\
  \sigma_{\xi\eta} & \sigma_\eta^2
\end{pmatrix}\right).
\]

We start out with the dividend-price ratio and discuss the results in detail. Only then we move to the analysis of the other predictive variables. We still fix the continuously compounded real monthly Swiss interbank rate, \( r_f \), at the arithmetic average over the sample period under consideration.

**Dividend-Price Ratio** When estimation risk is ignored, buy-and-hold investors maximize expected utility using the return distribution conditional on the parameter values fixed at their estimated values. Given the above VAR framework, they thus need estimates for \( \alpha, \beta, \) and \( \Sigma \). Table 2 presents these estimates for three different time periods when the predictive variable is the (log) dividend-price ratio.

The top two rows of Table 2 use the full sample of monthly data on stock index returns and the dividend-price ratio. Estimates for the two subsamples are obtained using data from 1975 to 1988 and from 1989 to 2002. The coefficient on the dividend-price ratio is not significantly different from zero, and the \( R^2 \) is very low, estimated at 0.18%. However, there is more evidence of return predictability in the first subsample from 1975 to 1988. Over this period, the results show the often-documented predictive power of the dividend-price ratio for stock returns. The coefficient on the dividend-price ratio is significant at the 5% significance level, and the \( R^2 \) is 2.73%. The dividend-price ratio is highly persistent and the variance matrix shows a strong negative correlation between
Table 2: 
Dividend-price ratio. Parameter estimates for a VAR model of stock returns.
The results in this table are based on the model given in equation (13), where the predictive 
variable is the (log) dividend-price ratio. The table gives estimates of the respective regression 
coefficients \( a \) and \( B \), the coefficients of determination, and the variance matrix over three 
different time periods. The top two rows use data from January 1975 to December 2002. 
Estimates for the two subsamples are showed below. The first subsample uses data from 
January 1975 to December 1988; the second subsample is based on the time period from 
January 1989 to December 2002. The figures above the diagonal in the variance matrices 
are correlations. */**/*** indicate p-values less than 0.1/0.05/0.01 (using the standardized 
normal distribution).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>Variance matrix</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>B</td>
<td>RSq.</td>
<td></td>
</tr>
<tr>
<td>1975:01—2002:12</td>
<td>0.0280</td>
<td>0.0060</td>
<td>0.18%</td>
<td>0.0022</td>
</tr>
<tr>
<td></td>
<td>-0.0653</td>
<td>0.9836</td>
<td>97.45%</td>
<td>0.0027</td>
</tr>
<tr>
<td>1975:01—1988:12</td>
<td>0.1638 **</td>
<td>0.0444</td>
<td>2.73%</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td>-0.1738</td>
<td>0.9523</td>
<td>92.40%</td>
<td>0.0019</td>
</tr>
<tr>
<td>1989:01—2002:12</td>
<td>0.0000</td>
<td>-0.0011</td>
<td>0.00%</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>-0.1108</td>
<td>0.9731</td>
<td>95.15%</td>
<td>0.0035</td>
</tr>
</tbody>
</table>
innovations in excess returns and the dividend-price ratio, estimated here at \(-0.9574\). On the other hand, any evidence of stock market predictability seems to have disappeared over the more recent subsample. First, the coefficient on the dividend-price ratio is far from being statistically significant at any reasonable significance level. Furthermore and even worse, the sign is wrong. Finally, the coefficient of determination is a bleak 0.00%.

Recall that the distribution of future excess returns depends on the value of the dividend-price ratio at the beginning of the planning horizon. The initial value of the dividend-price ratio enters equation (10) through \(z_T\). For now, however, we abstract from this effect and set the initial value of the dividend-price ratio to its mean over the respective sample period, and investigate how the optimal allocation to equities changes with the planning horizon for this fixed initial value of the dividend-price ratio. Later, we look at how the results are affected when the initial value of the dividend-price ratio varies over the state space.

Within each graph of Figure 3, the lines show the percentage 100\(\omega\)% percent allocated to stocks plotted against the planning horizon ranging from one to ten years. The four lines on each graph correspond to the four predictive distributions that the investors could use to predict future excess returns. The graphs on the left correspond to a risk aversion level of \(\zeta = 5\), the graphs on the right to \(\zeta = 10\). While the top row graphs are based on computations using the full data sample from 1975 to 2002, the lower graphs are based on data from 1975 to 1988, where the evidence of return predictability is much higher. The fine lines in the graphs represent the cases where investors ignore return predictability. In this case, of course, the model for excess returns reduces to the i.i.d. model discussed above and the lines are exactly the same as those in Figure 1 and 2, respectively. In contrast, the bold lines correspond to the cases where the dividend-price ratio is included in the analysis. Focus on the lower graph on the right, which presents the results most clearly. It is easy to see that when investors ignore parameter uncertainty (the solid lines), the optimal allocation to stocks for long-horizon investors is both higher than for short-horizon investors and than for investors who ignore return predictability. When investors take the uncertainty about the parameters into account (the dashed lines), however, the horizon effect is no longer clear. The allocation to equities is still higher compared to investors who ignore return predictability and account for estimation risk, but the positive horizon effect is restricted to, say, the first three years. For longer planning horizons, the allocation to stocks reduces more and more and even falls below the case of i.i.d. returns.

Following Barberis (2000, p. 243), these results can be explained as follows. Expected utility and hence optimal allocations to stocks are determined by the predictive distribution of future excess returns. In the case where estimation risk is ignored, the role of both the conditional mean and variance of cumulative log excess returns, as given in equations (10) and (11), respectively, is best illustrated using the following approximation developed in Campbell and Viceira (2002). With power utility and normally distributed log returns, they show that
Figure 3: Dividend-price ratio. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the (log) dividend-price ratio. The solid lines correspond to the case where the investor ignores parameter uncertainty, the dashed and dotted lines to the case where he accounts for it. Bold lines correspond to the cases where the investor takes return predictability into account, fine lines to the cases where he ignores it (i.i.d. returns). The top row graphs use data from 1975 to 2002, the lower two use data from 1975 to 1988. ζ denotes the coefficient of relative risk aversion.
the optimal allocation to stocks can be approximated by

$$\omega \approx \frac{1}{\zeta} \frac{E_T \left(e_{T+T+\hat{T}}\right) + \frac{1}{2} Var_T \left(e_{T+T+\hat{T}}\right)}{Var_T \left(e_{T+T+\hat{T}}\right)}$$

(14)

Accordingly, the optimal allocation to stocks depends on the ratio between expected cumulative log excess returns with the addition of one-half the conditional variance (to convert from log returns to simple returns that are ultimately of concern to investors), and the conditional variance. It has already been argued that when equity returns are modeled as i.i.d., the conditional mean and variance of cumulative log returns grow linearly over time, leading to identical allocations to stocks, regardless of the planning horizon. However, when investors take return predictability into account, this is no longer true. In this case, the conditional variance of cumulative log excess returns may grow slower than linearly with the planning horizon, making stocks look relatively less risky at longer horizons, which in turn leads to higher allocations to stocks in the optimal portfolio.

Barberis (2000, eq. 22 and 23) demonstrates this point as follows. Given the regression model in equation (13), the conditional variances of one- and two-period cumulative excess returns are given by

$$Var_T (e_{T+T+1}) = \sigma_\xi^2,$$

$$Var_T (e_{T+T+2}) = 2\sigma_\xi^2 + \beta^2 \sigma_\eta^2 + 2\beta \sigma_\xi \sigma_\eta.$$  (15)

For the parameter values estimated from the data (1975–1988, see Table 2), we have $\beta^2 \sigma_\eta^2 + 2\beta \sigma_\xi \sigma_\eta < 0$. This implies that the conditional variance of two-period returns is less than twice the conditional variance of one-period returns. In other words, the conditional variance ratio, $VR_T (\cdot)$, generally defined as

$$VR_T (\hat{T}) = \frac{Var_T (e_{T+T+\hat{T}})}{Var_T (e_{T+T+1})},$$  (16)

is below one, $VR_T (2) < 1$. This generalizes to any planning horizon; the ratio is less than one for all horizons $\hat{T}$. Consequently, when investors take the predictive power of the dividend-price ratio into account, conditional variances grow slower than linearly with the planning horizon, lowering the perceived risk of stocks in the long run. As a result, given expected cumulative excess returns, the optimal allocation to stocks increases.

The intuition behind this effect is perfectly pointed out in Barberis (2000, p. 245). In brief, since $\sigma_\xi \sigma_\eta < 0$ and $\beta > 0$, an unexpected fall of the dividend-price ratio is likely to be accompanied by a contemporaneous positive shock to realized returns. Given that the dividend-price ratio is now lower, however, excess returns as well are expected to be lower in the future. This pattern generates a component of negative serial correlation in realized returns. With a growing planning horizon, this perceived negative serial correlation slows the evolution of
the conditional variance of cumulative returns. Indeed, the idea that time variation in expected excess returns may induce mean reversion in realized returns has a strong economic intuition. If expected excess returns suddenly increase, it is only reasonable to assume that realized returns suffer a contemporaneous negative shock: after all, the discount rate for discounting future cash flows has suddenly increased (see, e.g., Campbell, Lo, and MacKinlay, 1997, Ch. 7). Negative shocks to current realized returns, which are followed by the higher excess returns predicted for the future, are the source of this perceived mean reversion.

Still, horizon effects can be present even without mean reversion in realized returns. Put it differently, return predictability may be sufficient to make stocks more attractive at long horizons, without being strong enough to induce mean reversion, i.e., an unconditional variance ratio less than one for \( \hat{T} > 1 \). Barberis (2000, eq. 24 and 25) notes that

\[
\begin{align*}
\text{Cov} (e_t, e_{t+1}) &= \frac{\beta^2 \delta \sigma_q^2}{1 - \delta^2} + \beta \sigma_q \xi,
\text{Cov} (e_t, e_{t+i}) &= \delta^i - 1 \text{Cov} (e_t, e_{t+1}),
\end{align*}
\]

\( i > 1 \). Thus, we may choose the regression parameters so that returns are serially uncorrelated at all lags and yet, by equation (15), the two-period conditional variance is still less than twice the one-period conditional variance. Although there is no mean reversion in realized excess returns in this case, a two-period investor would potentially allocate more to stocks than a one-period investor. Furthermore, note that the unconditional variance can be expressed as

\[
\text{Var} \left( e_{T-\hat{T}} \right) = \frac{\text{Var}_T \left( e_{T-\hat{T}} \right) \left( 1 - R^2 (e_{T-\hat{T}}) \right)}{1 - R^2 (e_{T-\hat{T}})},
\]

where \( R^2 (e_{T-\hat{T}}) \) denotes the coefficient of determination of a so-called long-horizon regression of the form

\[
e_{t-t+K-1} = \alpha (K) + \beta (K) x_{t-1} + \xi_{t-t+K-1}.
\]

Moreover, it is easy to show that the unconditional variance ratio, \( VR (\cdot) \), is given by

\[
VR \left( \hat{T} \right) = VR_T \left( \hat{T} \right) \frac{1 - R^2 (e_{T-\hat{T}})}{1 - R^2 (e_{T-\hat{T}})},
\]

where \( R^2 (e_{T-1}) \) corresponds to the coefficient of determination of the predictive regression in equation (13). In words, the unconditional variance ratio is always greater than the conditional variance ratio if the long-horizon \( R^2 \) is higher than the explanatory power of the corresponding short-horizon predictive regression. The difference between the two ratios can be substantial, since long-horizon \( R^2 \)s can be as large as 40% (see, e.g., Campbell, 1991). Consequently, empirical
results based on mean reversion, i.e., the unconditional variance ratio, may
understate the risk-reduction that is relevant for long-term investors; it may be
the case that the unconditional variance ratio equals one (as in the case of i.i.d.
returns), but the conditional variance ratio is below one.

Figure 3 also shows that incorporating parameter uncertainty may substan-
tially reduce the size of the horizon effect. For $\zeta = 10$ and the period from 1975
to 1988, ignoring parameter uncertainty may lead to an overallocation to stocks
of more than 60% at the ten-year planning horizon.

When investors take parameter uncertainty into account, they acknowledge
that they are not only uncertain about the average equity premium (as in the
i.i.d. case), but also about the true predictive power of the dividend-price ratio
and whether the dividend-price ratio really does induce (conditional) negative
serial correlations in realized returns. Consequently, with respect to their asset
allocation, investors are more cautious about stocks and allocate less to them.\(^4\)

Overall, thus, when investors ignore parameter uncertainty, return predictabil-
ity makes stocks look less risky in the long run. On the other hand, incorporating
parameter uncertainty makes them look more risky. These two effects go in the
opposite directions. Portfolio decisions are therefore not necessarily monotonic
as a function of the planning horizon. The dashed lines in Figure 3 show that
the allocation to stocks rises only for horizons up to about three years, for longer
planning horizons, estimation risk increases more and more and makes stocks
look less attractive.

So far we have focused on just one effect of including the dividend-price ra-
tio as a predictive variable in the VAR, namely, that the dividend-price ratio
reduces the conditional variance of predicted long-horizon cumulative returns.
Obviously, however, conditioning on the dividend-price ratio has another im-
\footnotetext{\(^4\)As discussed in Barberis (2000, footnote 16), when investors take parameter uncertainty
into account, they acknowledge both that the predictive power of the dividend-price ratio
may be weaker than the parameter estimates suggest or that it may be stronger. In the
first case they would allocate less to stocks at long horizons, in second case, however, they
would increase their equity holdings. In sum, thus, because investors are risk averse and hence
dislike the mean-preserving spread that accounting for estimation risk adds to the distribution
of future excess returns, they invest less at long horizons.}

It is only evident that these values do not only depend on the estimated para-

\begin{align*}
E_T (e_{T\rightarrow T+1}) &= \alpha + \beta x_T, \\
E_T (e_{T\rightarrow T+2}) &= 2\alpha + \beta \gamma + \beta (1 + \delta) x_T. 
\end{align*}
Figure 4: 
Dividend-price ratio. Optimal allocation to stocks plotted against the planning horizon in years. 
The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the (log) dividend-price ratio. The graphs on the left ignore parameter uncertainty, those on the right account for it. The five lines within each graph correspond to different initial values of the dividend-price ratio: 2.23% (dashed), 2.52% (dash dot), 2.77% (solid), 3.05% (dash dot dot), and 3.45% (dotted). Parameter values are estimated over the 1975 to 1988 sample period. ζ denotes the coefficient of relative risk aversion.
We now demonstrate how different initial values of the dividend-price ratio effect optimal allocation to stocks. Figure 4 presents the results for the sample period from 1975 to 1988, where the effect of return predictability is the strongest. The graphs on the left illustrate the optimal allocations when parameter uncertainty is ignored; the graphs on the right incorporate it. Within each graph, we plot optimal allocations to stocks as a function of the planning horizon and five different initial values of the dividend-price ratio. We take the interval ranging from the maximum and minimum value of the dividend-price ratio over the respective sample period, and discretize this range with 25 equally spaced grid points, \( x_T^j \) for \( j = 1, \ldots, 25 \). In particular, the five values we use are \( x_T^6 = 2.23\% \), \( x_T^{10} = 2.52\% \), \( x_T^{13} = 2.77\% \) (close to the mean value used above), \( x_T^{16} = 3.05\% \), and \( x_T^{20} = 3.45\% \).

The graphs on the left side of Figure 4 show that for all initial values of the dividend-price ratio, the optimal allocation to stocks rises with the planning horizon. Recall that investors expect higher future excess returns when the initial value of the dividend-price ratio is higher. Consequently, for any fixed horizon, the optimal allocation to equities is higher for higher values of the dividend-price ratio. In addition, the optimal allocation for stocks in the long run is just as sensitive to the initial value of the dividend-price ratio as the optimal allocation to stocks in the short run: the allocation lines do not seem to converge. Note, however, that expected future returns converge at longer horizons. For a one-year planning horizon, \( E_T (e_{T-T+12}) = -4.06\% \) p.a. for \( x_T^6 = 2.23\% \), and \( E_T (e_{T-T+12}) = +14.07\% \) p.a. for a high initial value of the dividend-price ratio, \( x_T^{20} = 3.45\% \). At the ten-year horizon, this difference shrinks considerably from 18.13\% to 4.07\%, with \( E_T (e_{T-T+120}) = +1.23\% \) p.a. for \( x_T^6 \), and \( E_T (e_{T-T+120}) = +5.30\% \) p.a. for \( x_T^{20} \). At the same time, however, the conditional variance shrinks from \( Var_T (e_{T-T+12}) = 1.37\% \) p.a. at the one-year horizon to \( Var_T (e_{T-T+120}) = 0.35\% \) p.a. at the ten-year horizon. This slow-down of the evolution of the conditional variance is very strong. Indeed, even in the case of a high initial dividend-price ratio and, accordingly, decreasing expected future returns over long horizons, it leads to higher allocations to stocks in the long run. In our case, buy-and-hold investors allocate all of their wealth to equities when the planning horizon is long enough (approximately thirty years), irrespective of the initial value of the dividend-price ratio.

However, this pattern significantly changes when investors take parameter uncertainty into account. The graphs on the right-hand side of Figure 4 show that for low initial values of the dividend-price ratio, the optimal allocation to stocks rises with the planning horizon. For high initial dividend-price ratios, however, the allocation to equities falls in the long run. Obviously, thus, when investors take parameter uncertainty into account, the allocation lines converge: the resulting optimal allocation to stocks in the long run is less sensitive to the initial value of the dividend-price ratio than the allocation to stocks in the short run, and much less sensitive than the allocation to stocks of investors with a long planning horizon who ignore parameter uncertainty. Of course, if the true predictive power of the dividend-price ratio is uncertain, the allocation to risky assets should be less sensitive to the initial value of the dividend-price ratio,
Figure 5: Dividend-price ratio. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the (log) dividend-price ratio. The graph on the left ignores parameter uncertainty, the one on the right accounts for it. The five lines within each graph correspond to different initial values of the dividend-price ratio: 1.22% (dashed), 1.48% (dash dot), 1.71% (solid), 1.98% (dash dot dot), and 2.41% (dotted). Parameter values are estimated over the 1989 to 2002 sample period. $\zeta$ denotes the coefficient of relative risk aversion.

Barberis (2000, p. 250), partly based on Kandel and Stambaugh (1996), highlight another surprising fact. For a given planning horizon and risk aversion level, the optimal allocation to stocks is not necessarily increasing in the initial value of the dividend-price ratio. If the initial value of the dividend-price ratio is $x_{T}^{20} = 3.45\%$ rather than $x_{T}^{16} = 3.05\%$, for example, the predictive distribution has a higher posterior mean. This should lead to a higher allocation to equities. Because, in addition, the variance of the predictive distribution is not sensitive to the initial value of the dividend-price ratio, this cannot explain the non-monotonicity result in Figure 4. In fact it is the third moment of the predictive distribution, skewness, that is important here. While for low initial values of the dividend-price ratio, incorporating parameter uncertainty generates positive skewness in the predictive distribution, it generates negative skewness for high initial values of the dividend-price ratio. In the case of power utility, this negative skewness makes stock less attractive, and makes optimal stock holdings non-monotonic in the initial value of the dividend-price ratio.

The true predictive power of the dividend-price ratio may change over time. Investors thus may prefer to estimate the relationship over the more recent subsample from 1989 to 2002. However, as indicated in Table 2, any evidence for stock market predictability using the dividend-price ratio seems to have disappeared recently. Nevertheless, Figure 5 repeats the calculation of Figure 4 for the case where investors use the more recent subsample in making their portfolio decisions. Not surprisingly, however, the resulting allocations to stocks show neither a horizon effect nor are they sensitive to the initial value of the
Dividend-price ratio. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the (log) dividend-price ratio. The state space is the interval ranging from the maximum and minimum value of the dividend-price ratio over the respective sample period, discretized with 25 equally spaced grid points. The graphs on the left ignore parameter uncertainty, those on the right account for it. The top row graphs use data from 1975 to 1988, the lower two use data from 1989 to 2002. $\zeta$ denotes the coefficient of relative risk aversion.

dividend-price ratio. Only when parameter uncertainty is incorporated, the recommended allocation to equities decreases with the planning horizon – in the same way as discussed in the case of i.i.d. returns.

Finally, Figure 6 shows optimal allocations to equities over the whole state space, i.e., for all initial values of the dividend-price ratio from $x_1^T$ to $x_{25}^T$. When investors account for parameter uncertainty, the non-monotonicity and lower level of the allocation to stocks is only obvious. For the period from 1989 to 2002, there is no horizon effect at all – the predictive ability of the dividend-price ratio has completely disappeared.

It is now important to discuss whether above results for the dividend-price ratio generalize to the other predictive variables suggested in previous studies of return predictability. In what follows, we repeat the same calculations for those variables as well and demonstrate the resulting allocations to equities. To save space, we restrict the analysis to the two subsamples from 1975 to 1988.
Table 3: Earnings-price ratio. Parameter estimates for a VAR model of stock returns.

The results in this table are based on the model given in equation (13), where the predictive variable is the earnings-price ratio. The table gives estimates of the respective regression coefficients $a$ and $B$, the coefficients of determination, and the variance matrix over three different time periods. The top two rows use data from January 1975 to December 2002. Estimates for the two subsamples are showed below. The first subsample uses data from January 1975 to December 1988; the second subsample is based on the time period from January 1989 to December 2002. The figures above the diagonal in the variance matrices are correlations. */**/*** indicate p-values less than 0.1/0.05/0.01 (using the standardized normal distribution).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a$</td>
<td>$B$</td>
<td>$\Sigma$</td>
</tr>
<tr>
<td></td>
<td>0.0195</td>
<td>0.0056</td>
<td>0.12%</td>
</tr>
<tr>
<td></td>
<td>-0.0638</td>
<td>0.9762</td>
<td>96.65%</td>
</tr>
<tr>
<td></td>
<td>0.0203</td>
<td>0.0995</td>
<td>8.28%</td>
</tr>
<tr>
<td></td>
<td>-0.3020</td>
<td>0.8757</td>
<td>83.28%</td>
</tr>
<tr>
<td></td>
<td>-0.0208</td>
<td>-0.0091</td>
<td>0.28%</td>
</tr>
<tr>
<td></td>
<td>-0.0675</td>
<td>0.9764</td>
<td>96.46%</td>
</tr>
<tr>
<td></td>
<td>0.0022</td>
<td>-0.7894</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td>0.0017</td>
<td>-0.7762</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>0.0025</td>
<td>-0.7951</td>
<td>0.0030</td>
</tr>
</tbody>
</table>

and 1989 to 2002, respectively.

**Earnings-Price Ratio** To begin with, we consider the earnings-price ratio as predictive variable. Estimates for $a$, $B$ and $\Sigma$ are given in Table 3. The slope coefficients are positive over the full and the first time period. Over the first sample period from 1975 to 1988, the coefficient is even statistically significantly different from zero, and the $R^2$ is very high with 8.28%. Over the more recent subsample, however, the slope coefficient has the wrong sign. The negative sign would indicate that a higher earnings-price ratio predicts lower future excess returns, contrary to intuition.

As in the case of the dividend-price ratio, the variance matrix shows that innovations to excess returns and earnings-price ratios are highly negatively correlated.

Figure 7 shows optimal allocations to stocks. For any fixed planning horizon and the period from 1975 to 1988, the optimal allocation to stocks is higher for higher initial values of the earnings-price ratio, as suggested by intuition. However, this relationship has completely changed over the more recent time period. In addition, despite the fact that $\sigma_{\xi \eta} < 0$ and $\beta > 0$ (the latter over the first subperiod), there is no positive horizon effect. Optimal allocations to stocks do not generally increase with the planning horizon.

29
Figure 7: Earnings-price ratio. Optimal allocation to stocks plotted against the planning horizon in years.
The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the earnings-price ratio. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The five lines within each graph correspond to different initial values of the earnings-price ratio: 12.15%/9.48% (dotted), 10.95%/7.75% (dash dot dot), 10.07%/6.66% (solid), 9.26%/5.72% (dash dot), and 8.27%/4.68% (dashed) for the first/second time period. $\zeta$ denotes the coefficient of relative risk aversion.
We discuss this result in the following case of the term spread as predictive variable.

**Term Spread** Consider now the case where the single predetermined variable, $x_t$, is the term spread. Estimates for $a$, $B$ and $\Sigma$ are given in Table 4.

The coefficients on the term spread are positive, establishing that a high term spread predicts high future excess returns. However, the coefficients are not significantly different from zero, and the $R^2$s are rather low, estimated at 0.57\% and 0.80\%, respectively. The variance matrix shows that innovations to excess returns and the term spread are not highly correlated. Over the period from 1975 to 1988, the correlation is even positive, estimated at 0.2037. In other words, in contrast to the dividend-price ratio, where $\sigma_{\xi y} < 0$, a decrease of the term spread may be accompanied by a contemporaneous negative shock to the excess return. Furthermore, since the term spread is lower, excess returns are predicted to be lower in the future, since $\beta > 0$. Therefore, the economic intuition behind the general idea that time variation in expected excess returns induces mean reversion in realized returns is not confirmed here. Realized returns no longer suffer a contemporaneous negative shock if there is a positive shock to the expected risk premium, despite the fact that the discount rate for discounting future cash flows has suddenly increased. Rather than mean reversion, this generates a component of positive serial correlation in returns which increases rather than decreases the variance of cumulative excess returns as the horizon.

### Table 4:
**Term spread. Parameter estimates for a VAR model of stock returns.**
The results in this table are based on the model given in equation (13), where the predictive variable is the term spread. The table gives estimates of the respective regression coefficients $a$ and $B$, the coefficients of determination, and the variance matrix over three different time periods. The top two rows use data from January 1975 to December 2002. Estimates for the two subsamples are showed below. The first subsample uses data from January 1975 to December 1988; the second subsample is based on the time period from January 1989 to December 2002. The figures above the diagonal in the variance matrices are correlations. */**/*** indicate $p$-values less than 0.1/0.05/0.01 (using the standardized normal distribution).

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>B</th>
<th>RSq.</th>
<th>Variance matrix</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975:01–2002:12</td>
<td>0.0038</td>
<td>0.2309</td>
<td>0.68%</td>
<td>0.0022 0.0974</td>
<td>336</td>
</tr>
<tr>
<td></td>
<td>0.0004</td>
<td>0.9494</td>
<td>90.40%</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>1975:01–1988:12</td>
<td>0.0044</td>
<td>0.1831</td>
<td>0.57%</td>
<td>0.0018 0.2037</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>0.0005</td>
<td>0.9231</td>
<td>85.81%</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>1989:01–2002:12</td>
<td>0.0033</td>
<td>0.2891</td>
<td>0.80%</td>
<td>0.0025 -0.0745</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>0.0002</td>
<td>0.9827</td>
<td>96.28%</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>

31
Figure 8: Term spread. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the term spread. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The five lines within each graph correspond to different initial values of the term spread: 2.30%/1.32% (dotted), 0.82%/0.38% (dash dot dot), -0.29%/-0.33% (solid), -1.39%/-1.03% (dash dot), and -2.87%/-1.98% (dashed) for the first/second time period. ζ denotes the coefficient of relative risk aversion.

Indeed, the conditional variance of cumulative excess returns grows faster than linearly with the planning horizon, making stocks look relatively more risky at longer horizons. Even over the more recent subsample, the conditional variance increases from \( \text{Var}_T (e_{T \rightarrow T+12}) = 3.02\% \) p.a. for the one-year horizon to \( \text{Var}_T (e_{T \rightarrow T+120}) = 3.96\% \) p.a. at the ten-year horizon. Although this increase is not huge, it potentially reduces the allocation to stocks in the optimal portfolio at longer horizons.

By its very nature as a predictive variable, the term spread also affects the conditional mean of the distribution of future returns. As depicted in Figure 8, for any fixed planning horizon, the optimal allocation to stocks is higher for higher values of the term spread since investors expect higher future excess returns. However, it is no longer true that the ten-year allocation is significantly
higher than the one-year allocation for any fixed initial value of the term spread. This is only true when the term spread is initially low. In this case and from 1989 to 2002, $E_T(e_{T-T+12}) = -1.87\%$ p.a. and $E_T(e_{T-T+120}) = +3.52\%$ p.a. for $x_T^6 = -1.98\%$. For a high initial value of the term spread, $x_T^{20} = 1.32\%$, $E_T(e_{T-T+12}) = +8.54\%$ p.a. and $E_T(e_{T-T+120}) = +8.34\%$ p.a. Overall, since the impact of the conditional variance is rather low, the convergence of the optimal allocation lines mainly mirrors the convergence of expected future excess returns. The optimal allocation to stocks of long-term investors is thus no longer as sensitive to the initial value of the term spread as the optimal portfolio choice of short-term investors.

The picture is not remarkably different when investors account for estimation risk. These results are shown in the graphs on the right-hand side of Figure 8. Basically, incorporating parameter uncertainty makes conditional variances grow more quickly as the horizon grows, tending to make stocks look more risky. The allocation to stocks is therefore lower than in the case where estimation risk is ignored. However, compared to the dividend-price ratio as predictive variable, the effect is quite modest and the allocation to stocks of long-horizon investors is only slightly less sensitive to the initial value of the term spread. Optimal allocations to equities already converge when investors ignore estimation risk.

Finally, Figure 9 shows optimal allocations to stocks over the whole interval of initial values of the term spread, i.e., for $x_T^1$ to $x_T^{25}$. In comparison to Figure 6, the impact of estimation risk is less evident. Clearly, even in the case of the term spread as predictive variable, the optimal allocation to equities is lower when parameter uncertainty is taken into account, but non-monotonic in either case of ignoring or accounting for parameter uncertainty.
Table 5: 
One-month Swiss interbank rate. Parameter estimates for a VAR model of stock returns.

The results in this table are based on the model given in equation (13), where the predictive variable is the one-month Swiss interbank rate. The table gives estimates of the respective regression coefficients $a$ and $B$, the coefficients of determination, and the variance matrix over three different time periods. The top two rows use data from January 1975 to December 2002. Estimates for the two subsamples are showed below. The first subsample uses data from January 1975 to December 1988; the second subsample is based on the time period from January 1989 to December 2002. The figures above the diagonal in the variance matrices are correlations. */**/*** indicate p-values less than 0.1/0.05/0.01 (using the standardized normal distribution).

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$B$</th>
<th>RSq.</th>
<th>Variance matrix</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975:01–2002:12</td>
<td>0.0085</td>
<td>-0.0926</td>
<td>0.23%</td>
<td>0.0022</td>
<td>-0.1179</td>
</tr>
<tr>
<td></td>
<td>0.0013</td>
<td>0.9587</td>
<td>93.09%</td>
<td></td>
<td>0.0000</td>
</tr>
<tr>
<td>1975:01–1988:12</td>
<td>0.0061</td>
<td>-0.0161</td>
<td>0.01%</td>
<td>0.0018</td>
<td>-0.1666</td>
</tr>
<tr>
<td></td>
<td>0.0032</td>
<td>0.9030</td>
<td>85.30%</td>
<td></td>
<td>0.0001</td>
</tr>
<tr>
<td>1989:01–2002:12</td>
<td>0.0101</td>
<td>-0.1432</td>
<td>0.55%</td>
<td>0.0025</td>
<td>-0.0493</td>
</tr>
<tr>
<td></td>
<td>-0.0001</td>
<td>0.9964</td>
<td>98.45%</td>
<td></td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Among others, Fama and French (1989) and Dahlquist and Harvey (2001) find that the form of the term structure is associated with the business cycle. In particular, the term spread is low around business-cycle peaks and high near business-cycle troughs. Given the general message that expected excess returns are lower when economic conditions are strong and higher when economic conditions are weak, our results can be easily interpreted. Investors should increase their allocation to equities when the term spread is high, i.e., near business-cycle troughs. Conversely, around business-cycle peaks, when the term structure is inverse and expected excess returns are low, investors should reduce their exposure to stocks.

**One-Month Swiss Interbank Rate** When the predictive variable is the nominal one-month Swiss interbank rate, estimates for $a$, $B$ and $\Sigma$ are given in Table 5. The coefficients on the interest rates are negative, establishing that high short rates predict low future excess returns. However, the coefficients are not significantly different from zero, and the $R^2$s are low, estimated only at 0.01% and 0.55%, respectively.

Clearly, short-term interest rates are highly persistent. The variance matrix shows that innovations to excess returns and short rates are not highly correlated. Despite the fact that the correlation between innovations to excess
returns and short rates is negative, the economic intuition that time variation in expected excess returns induces mean reversion in realized returns is hard to justify either. Realized returns do not suffer a contemporaneous positive shock if there is a negative shock to the expected risk premium. Indeed, over the more recent subsample, the conditional variance increases from 3.06% p.a. at the one-year horizon to 4.10% p.a. at the ten-year horizon, making stocks look relatively more risky at longer horizons. Although this increase is not huge, it potentially reduces the allocation to stocks in the optimal portfolio at longer horizons.

Figure 10 shows optimal allocations to stocks. For any fixed planning horizon, the optimal allocation to stocks is lower for higher initial interest rates since investors expect lower future excess returns. As in the case of the term spread as predictive variable, the impact of the conditional variance is rather low, the (slow) convergence of the optimal allocation lines mainly mirrors the convergence of expected future returns.

The picture is not remarkably different when investors account for estimation risk. The allocation to stocks is lower than in the case where estimation risk is ignored, and the convergence is more distinct at longer horizons. However, looking at the results for the first subsample, another important aspect of estimation risk needs to be stressed. As pointed out in equation (12), when investors account for estimation risk, the predictive distribution is conditioned on the first observation in the sample, \( z_1 \). The first observation, though, may be quite influential to the regression results. This seems to be the case here. When we drop the first observation, the regression coefficient is \( \beta = -0.1487 \) and the \( R^2 \) is estimated at 0.60%. In comparison to the case where parameter uncertainty is ignored and the allocation lines are merely horizontal, the effect of the initial value of the short rate is much more distinct when estimation risk is accounted for. Of course, this only demonstrates the sensitivity of the results to the underlying data.

Overall, the results are easily interpreted using the business-cycle argumentation developed above for the term spread. Note that the correlation between the short-term interest rate and the term spread is very high, estimated at \( -0.9080 \) and \( -0.9582 \), respectively. Since high short rates are likely accompanied by lower long-term rates, investors should decrease their allocation to equities when short rates are high. This presumably happens near business-cycle peaks. Conversely, around business-cycle troughs, when short rates are low and expected excess returns high, investors should increase their exposure to equities.

**Realized Stock Market Volatility** When we explore the link between realized stock market volatility and the market return, we regress realized excess returns on the lagged volatility measure. We use the lagged volatility as a proxy for the expectation of the current period’s volatility, which can be justified by the high persistence of the realized volatility series. Notice, however, that realized volatility is less persistent compared to the other predictive variables discussed
Figure 10:  
One-month Swiss interbank rate. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the one-month Swiss interbank rate. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The five lines within each graph correspond to different initial values of the one-month Swiss interbank rate: 7.96%/7.36% (dotted), 6.29%/5.93% (dash dot dot), 5.03%/4.87% (solid), 3.77%/3.80% (dash dot), and 2.10%/2.38% (dashed) for the first/second time period. $\zeta$ denotes the coefficient of relative risk aversion.
Table 6:
Realized stock market volatility. Parameter estimates for a VAR model of stock returns.
The results in this table are based on the model given in equation (13), where the predictive variable is the realized stock market volatility. The table gives estimates of the respective regression coefficients \( a \) and \( B \), the coefficients of determination, and the variance matrix over three different time periods. The top two rows use data from January 1975 to December 2002. Estimates for the two subsamples are showed below. The first subsample uses data from January 1975 to December 1988; the second subsample is based on the time period from January 1989 to December 2002. The figures above the diagonal in the variance matrices are correlations. */**/*** indicate p-values less than 0.1/0.05/0.01 (using the standardized normal distribution).

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>B</th>
<th>RSq.</th>
<th>Variance matrix</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975-01-2002:12</td>
<td>0.0113</td>
<td>-0.1636</td>
<td>0.73%</td>
<td>0.0022</td>
<td>-0.3848</td>
</tr>
<tr>
<td></td>
<td>0.0187</td>
<td>0.5137</td>
<td>26.38%</td>
<td>0.0004</td>
<td></td>
</tr>
<tr>
<td>1975-01-1988:12</td>
<td>0.0144</td>
<td>*-0.2719</td>
<td>1.87%</td>
<td>0.0018</td>
<td>-0.3623</td>
</tr>
<tr>
<td></td>
<td>0.0146</td>
<td>0.5506</td>
<td>30.30%</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>1989-01-2002:12</td>
<td>0.0087</td>
<td>-0.0961</td>
<td>0.24%</td>
<td>0.0025</td>
<td>-0.4056</td>
</tr>
<tr>
<td></td>
<td>0.0247</td>
<td>0.4404</td>
<td>19.44%</td>
<td>0.0005</td>
<td></td>
</tr>
</tbody>
</table>

so far.\(^5\) Estimates for \( a, B \) and \( \Sigma \) are given in Table 6. Both the regression coefficients and the correlation coefficients between innovations to excess returns and realized volatilities are negative.

Notice that other empirical studies present conflicting results on the sign of the regression coefficient. Our results are in line with Campbell (1987) and Glosten, Jagannathan, and Runkle (1993), who also find a negative relation, whereas French, Schwert, and Stambaugh (1987) and Campbell and Hentschel (1992) find a positive relation. In a recent contribution, Goyal and Santa-Clara (2003) find a significant positive relation between realized average stock variance (the arithmetic average of the monthly variance of each stock’s returns) and the return on the market.

Figure 11 shows that investors hold less of the risky asset when realized volatility is high. This provides some perspective on the so-called flight-to-quality phenomenon, which refers to investors moving capital from stock markets to government bond markets when the stock markets are more volatile than usual.

\(^5\)Goyal and Santa-Clara (2003) claim that the results are essentially the same if we replace the realized volatility series by the fitted values from an ARMA model. See also Schwert (1989).
Figure 11: Realized stock market volatility. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the realized stock market volatility. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The five lines within each graph correspond to different initial values of the realized stock market volatility: 68.30%/64.08% (dotted), 54.89%/51.78% (dash dot dot), 44.84%/42.55% (solid), 34.79%/33.33% (dash dot), and 21.39%/21.03% (dashed) for the first/second time period. $\zeta$ denotes the coefficient of relative risk aversion.
The results in this table are based on the model given in equation (13), where the predictive variable is the U.S. TED spread. The table gives estimates of the respective regression coefficients $\alpha$ and $\beta$, the coefficients of determination, and the variance matrix over three different time periods. The top two rows use data from January 1975 to December 2002. Estimates for the two subsamples are showed below. The first subsample uses data from January 1975 to December 1988; the second subsample is based on the time period from January 1989 to December 2002. The figures above the diagonal in the variance matrices are correlations. */**/*** indicate p-values less than 0.1/0.05/0.01 (using the standardized normal distribution).

**U.S. TED Spread** The U.S. TED spread is calculated as the difference between three-month Eurodollar rates and three-month Treasury Bill rates, provided by the Federal Reserve Board of Governors. It can be viewed as a political risk premium that reflects either actual or anticipated barriers to international investing. The yield differential widens when the risk of disruption in the global financial system increases. Hence, it is conceivable that a positive relationship between the TED spread and expected excess returns shows up. Indeed, Table 7 shows that the regression coefficients are positive, establishing that a high TED spread predicts high future excess returns.

Figure 12 shows optimal allocations to stocks. Over the recent subsample, the optimal allocation to equities is higher for a higher initial TED spread, since investors expect higher future excess returns. However, this is not true for the first subsample. While the allocation lines are merely flat when estimation risk is ignored, accounting for parameter uncertainty yields a fundamentally different result. A high initial TED spread leads to a low allocation to stocks, contrary to what we would expect based on the estimated VAR coefficients summarized in Table 7. Yet, the explanation is simple. As in the case of the one-month interest rate, this result can be explained by the fact that the predictive distribution is conditioned on the first observation in the sample, $z_1$, when estimation risk is taken into account. The first observation, though, is very influential for the regression results. When it is dropped, the regression

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>RSq.</th>
<th>Variance matrix</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975:01–2002:12</td>
<td>0.0026</td>
<td>0.6097</td>
<td>0.15%</td>
<td>0.0022 -0.2083</td>
<td>336</td>
</tr>
<tr>
<td></td>
<td>0.0003</td>
<td>0.9232</td>
<td>87.01%</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>1975:01–1988:12</td>
<td>0.0055</td>
<td>0.0120</td>
<td>0.00%</td>
<td>0.0018 -0.3140</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>0.0007</td>
<td>0.8814</td>
<td>79.78%</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>1989:01–2002:12</td>
<td>-0.0106</td>
<td><strong>6.7887</strong></td>
<td>2.40%</td>
<td>0.0025 -0.0674</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>0.0002</td>
<td>0.8758</td>
<td>79.10%</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: U.S. TED spread. Parameter estimates for a VAR model of stock returns.
Figure 12: U.S. TED spread. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the U.S. TED spread. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The five lines within each graph correspond to different initial values of the U.S. TED spread: 1.40%/0.46% (dotted), 1.14%/0.38% (dash dot dot), 0.95%/0.31% (solid), 0.75%/0.25% (dash dot), and 0.50%/0.16% (dashed) for the first/second time period. ζ denotes the coefficient of relative risk aversion.

Coefficient is no longer positive, but estimated at $\beta = -0.6518$. Of course, this negative relation between the TED spread and excess returns explains the allocation lines over the first time period. Again, the high sensitivity of the results to the underlying data is only obvious.

U.S. Default Risk Spread Keim and Stambaugh (1986) and Fama and French (1989) are among the first who apply measures of U.S. default spreads to explain time variation in expected U.S. stock returns. In times of a recession, investors will demand a higher return premium for investing in low-rated corporate bonds, implying a larger default risk spread. We would therefore expect a positive relation between differences in yields between Moody’s Baa and Aaa rated bonds and expected excess returns. Unfortunately, as indicated in Table 8, the sign of the coefficient on the default risk spread is wrong for the more
Table 8: U.S. default risk spread. Parameter estimates for a VAR model of stock returns.

The results in this table are based on the model given in equation (13), where the predictive variable is the U.S. default risk spread. The table gives estimates of the respective regression coefficients $a$ and $B$, the coefficients of determination, and the variance matrix over three different time periods. The top two rows use data from January 1975 to December 2002. Estimates for the two subsamples are showed below. The first subsample uses data from January 1975 to December 1988; the second subsample is based on the time period from January 1989 to December 2002. The figures above the diagonal in the variance matrices are correlations. */**/*** indicate p-values less than 0.1/0.05/0.01 (using the standardized normal distribution).

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$B$</th>
<th>RSq.</th>
<th>Variance matrix</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975:01–2002:12</td>
<td>0.0045</td>
<td>0.1094</td>
<td>0.00%</td>
<td>0.0022</td>
<td>336</td>
</tr>
<tr>
<td></td>
<td>0.0002</td>
<td>0.9606</td>
<td>92.81%</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>1975:01–1988:12</td>
<td>-0.0054</td>
<td>1.9847</td>
<td>0.60%</td>
<td>0.0018</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>0.0004</td>
<td>0.9337</td>
<td>87.25%</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>1989:01–2002:12</td>
<td>0.0319</td>
<td>*-8.2933</td>
<td>1.96%</td>
<td>0.0025</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.9595</td>
<td>90.48%</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>

**recent time period. Counter to intuition, the negative coefficient indicates that high default spreads predict low future excess returns. This is in fact hard to explain. Worse, it is in exactly this case where the coefficient is statistically significantly different from zero and the $R^2$ is relatively high.**

Apart from this result, however, Figure 13 does not reveal any new pattern of optimal allocations to stocks. Conditional variances slightly increase from the one-year to the ten-year horizon, making stocks look relatively more risky at longer horizons. For any fixed planning horizon, the optimal allocation to stocks is higher for high (low) initial values of the default risk spread in the period from 1975 to 1988 (1989 to 2002).

**Stochastically Detrended Predictive Variables** Before we turn to the dynamic portfolio choice problem, it is interesting to compare the above results with the results obtained by stochastically detrending the predictive variables. Since our results so far show the existence of a positive horizon effect only in the case of the dividend-price ratio as predictive variable (and over the first subperiod from 1975 to 1988), we restrict the following analysis to the stochastically detrended dividend-price ratio. The results of the other stochastically detrended predictive variables do not provide any new evidence.

Figure 14 shows the optimal allocations to stocks. Both graphs ignore pa-
Figure 13: U.S. default risk spread. Optimal allocation to stocks plotted against the planning horizon in years.

The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the U.S. default risk spread. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The five lines within each graph correspond to different initial values of the term spread: 0.87%/0.50% (dotted), 0.74%/0.44% (dash dot dot), 0.65%/0.40% (solid), 0.56%/0.35% (dash dot), and 0.43%/0.29% (dashed) for the first/second time period. ζ denotes the coefficient of relative risk aversion.
Figure 14:
Stochastically detrended dividend-price ratio. Optimal allocation to stocks plotted against the planning horizon in years.
The investor follows a buy-and-hold strategy, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the stochastically detrended dividend-price ratio. Both graphs ignore parameter uncertainty. Parameter values are estimated over the 1975 to 1988 (left) and the 1989 to 2002 sample period (right). The five lines within each graph correspond to different initial values of the stochastically detrended dividend-price ratio. ζ denotes the coefficient of relative risk aversion.

3 Asset Allocation Framework for a Dynamic Investor

So far, our analysis has covered the buy-and-hold portfolio choice problem. In this case, investors do neither adjust nor rebalance their portfolio until the end of their planning horizon. In effect, therefore, investors behave myopically. However, the long-term investors’ portfolio depends not only on their objective, but also on what they are allowed to do in each period. We now examine a dynamic rebalancing strategy, where investors optimally adjust their portfolio over the planning horizon. Specifically, we consider investors who are allowed to adjust their portfolio annually using the new information at the end of each year. Compared to myopic buy-and-hold investors with equal planning and investment horizons, dynamic investors distinguish between the planning and (the shorter) investment horizon. Given the length of the investment horizon (one year), we analyze how the optimal allocation depends on the planning horizon.
3.1 The Dynamic Portfolio Choice Problem

We use the same regression model as above, originally introduced in equations (9) and (13), with $\mathbf{z}_t = [x_t, x_t^2]$, where $x_t$ is a single predictive variable.

Investors who optimally adjust their portfolio at regular intervals are confronted with a dynamic programming problem. We follow Barberis (2000) and solve this problem by employing the standard technique of discretizing the state space and using backward induction.

3.1.1 Ignoring Estimation Risk

Again, assume that at time $T$, an investor has a planning horizon of $\hat{T}$ months. His planning horizon may then be divided into $K$ (rebalancing) intervals of equal length, $[t_0, t_1], [t_1, t_2], ..., [t_{K-1}, t_K]$, where the start and end of the investor’s planning horizon are $t_0 = T$ and $t_K = T + \hat{T}$, respectively. The investor is thus allowed to adjust his portfolio $K$ times over the planning horizon, at points $(t_0, t_1, ..., t_{K-1})$; his allocations to the stock index at times $(t_0, t_1, ..., t_{K-1})$ are given by $(\omega_0, \omega_1, ..., \omega_{K-1})$. The investor then optimizes (Barberis, 2000, eq. 26)

$$\max_{t_0=T} E_{t_0} \left( \frac{W^{1-\zeta}_{t_{K-1}}}{1-\zeta} \right),$$

where $\max_{t_0=T}$ means that he maximizes over all remaining decisions from time $t_0$ on. Of course,

$$W_{k+1} = W_k \left( (1 - \omega_k) \exp \left( \frac{\hat{T}}{K} \right) + \omega_k \exp \left( \frac{\hat{T}}{K} + e_{k-1} \right) \right),$$

and

$$e_{k-1} \equiv e_{t_k+1} + e_{t_k+2} + ... + e_{t_k+\hat{T}/K}$$

for $k = 0, ..., K - 1$. For the ease of exposition, we may write $W_k$ in place of $W_{t_k}$ for the investor’s wealth at time $t_k$. The cumulative excess return between rebalancing points $t_k$ and $t_{k+1}$ is denoted by $e_{k-1}$.

According to Barberis (2000, eq. 29), derived utility of wealth is defined as

$$J(W_k, x_k, t_k) = \max_{t_k} E_{t_k} \left( \frac{W^{1-\zeta}_{k}}{1-\zeta} \right),$$

where $\max_{t_k}$ means a maximization over all remaining portfolio choice decisions from time $t_k$ on. The so-called Bellman equation of optimality is then given by (Barberis, 2000, eq. 30)

$$J(W_k, x_k, t_k) = \max_{\omega_k} E_{t_k} \left( J(W_{k+1}, x_{k+1}, t_{k+1}) \right).$$

Barberis (2000, eq. 31) shows that derived utility may be written as

$$J(W_k, x_k, t_k) = \frac{W^{1-\zeta}_{k}}{1-\zeta} Q(x_k, t_k)$$

44
for $\zeta \neq 1$. The Bellman equation can thus be rewritten as (Barberis, 2000, eq. 33)

$$Q(x_k, t_k) = \max_{\omega_k} E_{t_k} \left\{ \left[ (1 - \omega_k) \exp \left( r_j \frac{\hat{T}}{K} \right) + \omega_k \exp \left( r_j \frac{\hat{T}}{K} + e_{k-1} \right) \right]^{1-\zeta} \right\} \times Q(x_{k+1}, t_{k+1}). \quad (27)$$

Parameter uncertainty is ignored in this section. Hence, the expectation in equation (27) is taken over the Normal distribution $p(e_{k-1}, x_{k+1} | \theta, x_k)$, conditioned on fixed parameter values.

The usual technique for solving a Bellman equation is to discretize the state space and then use backward induction. Slightly different to Barberis (2000), we take the interval ranging from the maximum and minimum value of the predictive variable over the respective sample period, and discretize this range with 25 equally spaced grid points, $x^j_T$ for $j = 1, ..., 25$. The exact numerical procedure is described in Appendix B.

### 3.1.2 Accounting for Parameter Uncertainty

In an intertemporal setting, the effects of parameter uncertainty are more involved. On the one hand, a first effect is analogous to the one investors face in the buy-and-hold portfolio choice problem. As before, when investors calculate the value function in equation (27), the conditional expectation should be taken over a predictive distribution that incorporates parameter uncertainty. On the other hand, parameter uncertainty may change over time as more data are received. With fresh data, investors may update their posterior distribution of the parameters. If investors anticipate this learning, it may affect their portfolio holdings.

Williams (1977) and Gennotte (1986) analyze the effect of a learning-based hedging demand theoretically. In our dynamic context, however, additionally incorporating dynamic learning is a formidable problem. In this case, the investment opportunity set would no longer be characterized by the predictive variable alone, but also by variables summarizing investors’ beliefs about the parameters $\theta = (a, B, \Sigma)$. These additional variables would dramatically increase the size of the state space, making the dynamic programming problem difficult to solve. To our knowledge, there is no suggestion as to how dynamic learning can be incorporated in a discrete time setting. The notable exception is Xia (2002), but she examines the effects of uncertainty about the return predictability in a continuous-time setting.

As in Barberis (2000), we thus simplify the problem and assume that although investors acknowledge that they are uncertain about the regression parameters, they ignore the impact on today’s optimal allocation of the fact that their beliefs about those parameters may change over time. As in the buy-and-hold case, these beliefs are summarized by the posterior distribution calculated...
conditional only on data up until $T$. Consequently, the investment opportunity set is still described by the prevailing predictive variable alone and investors still use equation (27) to calculate the value function. When investors account for parameter uncertainty, the expectation $E_{t_k}$ is taken over the predictive distribution $p(e_{k-1}, x_{k+1} \mid x_k)$ rather than over $p(e_{k-1}, x_{k+1} \mid \theta, x_k)$. Investors thus sample from the predictive distribution by taking a large sample from the posterior distribution $p(\theta \mid z)$ and then for each set of parameter values drawn, make a draw from $p(e_{k-1}, x_{k+1} \mid \theta, x_k)$, a Normal distribution.

Of course, this approach is possibly too simple, investors should at least recognize that the precision of the parameter estimates will improve over time. However, as suggested by Barberis (2000, footnote 20) and own empirical results not reported here, we might justify this by thinking of it in the context of a model with time-varying regression parameters. Indeed, in this reasonable case, it would be no longer true that the posterior distribution becomes tighter as more data are received; in fact, it may become even more dispersed.

### 3.2 Empirical Results

The aim of this section is to empirically examine the size of the intertemporal hedging demand. We use the same data from the Swiss stock market and the same predictive variables as above. As in Barberis (2000), we do the calculations for planning horizons ranging from one year to ten years in one-year increments, and consider investors who are allowed to adjust their portfolio annually using the new information at the end of each year. We start with the dividend-price ratio as the single predictive variable and discuss the results in detail. Only then we move to the analysis of the other predictive variables. We maintain the simplification that the continuously compounded real monthly Swiss interbank rate is a constant, $r_f$, per period.

**Dividend-Price Ratio and Earnings-Price Ratio**  The graphs on the left-hand (right-hand) side of Figure 15 present optimal allocations to stocks for planning horizons ranging from one to ten years and for the dividend-price ratio as predictive variable. It is assumed that the investors optimally adjust their portfolios once every year and ignore (account for) parameter uncertainty. The top row graphs use data from 1975 to 1988, the lower two use data from 1989 to 2002. Within each graph, each line corresponds to a different initial value of the dividend-price ratio.

Using data from 1975 to 1988, Figure 15 shows that the optimal allocation to stocks rises with the planning horizon. Although these results appear to be the same as in the buy-and-hold case (see Figures 4 and 5), they must be explained differently. In an intertemporal framework, the observed increase in the allocation to equities across planning horizons is due to the intertemporal hedging demand first described by Merton (1973). Note that expected excess returns are governed by the dividend-price ratio in our simple regression specification. As it changes over time, the investment opportunity set faced by the investors changes as well. Merton (1973) shows that investors that are
Figure 15:
Dividend-price ratio. Optimal allocation to stocks plotted against the planning horizon in years.
The investor rebalances optimally once a year, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the (log) dividend-price ratio. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The five lines within each graph correspond to different initial values of the dividend-price ratio: 2.23%/1.22% (dashed), 2.52%/1.48% (dash dot), 2.77%/1.71% (solid), 3.05%/1.98% (dash dot dot), and 3.45%/2.41% (dotted) for the first/second time period. ζ denotes the coefficient of relative risk aversion.
more risk averse than log-utility maximizers ($\zeta > 1$) may want to hedge these changes by investing in a way that gives them higher wealth precisely when investment opportunities are unattractive, i.e., when expected returns are low. Since shocks to expected excess returns are reliably negatively correlated with shocks to realized returns, holding more in stocks seems to be an ideal way of hedging against movements in expected returns, indeed. However, as indicated in Table 2, any evidence of stock market predictability using the dividend-price ratio seems to have disappeared recently. Indeed, the lower graph on the left of Figure 15 shows that the resulting allocations to stocks show neither a horizon effect nor are they sensitive to the initial value of the dividend-price ratio.

In accordance to Kim and Omberg (1996) and similar to Campbell and Viceira (2002), Figure 16 shows optimal allocations to stocks over the whole state space, i.e., for all initial values of the dividend-price ratio from $x_1^T$ to $x_{25}^T$. Again, over the first subsample, the intertemporal hedging demand is quite large, particularly for investors with a long planning horizon. Notice, however, that the intertemporal hedging demand seems to converge for long planning horizons. The difference between the resulting hedging demand for a five-year (dotted lines) and a one-year planning horizon is much larger than the difference between a ten-year and a five-year horizon hedging demand. On the other hand, however, investors who follow a dynamic rebalancing strategy do hardly time the market more aggressively than myopic buy-and-hold investors. The allocation lines in Figure 16 are largely parallel to one another.

Recently, after all, along with the evidence of return predictability, the intertemporal hedging demand has completely disappeared.

The graphs on the right-hand sides of Figures 15 and 16 demonstrate that when the uncertainty in the predictive power of the dividend-price ratio is taken into account, the allocation lines are flatter and the intertemporal hedging demands somewhat smaller. Of course, when investors acknowledge parameter uncertainty, they become more skeptical whether the investment opportunity set is really changing over time. As a consequence, they may not invest more heavily in stocks as a (intertemporal) hedge. Empirically, however, with respect to the magnitude of the intertemporal hedging demand, the effect of (static) estimation risk does not seem to be very important.

Once investors account for parameter uncertainty, Figures 15 and 16 also reveal the reduced sensitivity of the optimal allocation to the initial value of the dividend-price ratio. Changes in portfolio compositions thus occur more gradually over time. Therefore, the often cited results suggested in Brennan, Schwartz, and Lagnado (1997) and Campbell and Viceira (1999) may need to be interpreted with some caution. Since their dynamic asset allocation strategies ignore parameter uncertainty, the recommended allocations to stocks are probably too high and too sensitive to the variables parameterizing expected excess returns.

Finally, Figure 17 presents optimal allocations to stocks for the earnings-price ratio as predictive variable. In general, the picture is quite similar to the one arising for the dividend-price ratio. It seems that the intertemporal hedging demand is somewhat smaller, but the overall pattern corresponds to the case of
Figure 16: Dividend-price ratio. Optimal allocation to stocks plotted against the initial value of the dividend-price ratio.

The investor rebalances optimally once a year, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the (log) dividend-price ratio. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The state space is the interval ranging from the maximum and minimum value of the dividend-price ratio over the respective sample period, discretized with 25 equally spaced grid points. The ten lines within each graph correspond to different planning horizons, ranging from one year (dashed) to 10 year (solid). \( \zeta \) denotes the coefficient of relative risk aversion.
Figure 17: Earnings-price ratio. Optimal allocation to stocks plotted against the initial value of the earnings-price ratio.

The investor rebalances optimally once a year, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the (log) earnings-price ratio. The graphs on the left ignore parameter uncertainty, those on the right account for it. Parameter values are estimated over the 1975 to 1988 (top row graphs) and the 1989 to 2002 sample period. The state space is the interval ranging from the maximum and minimum value of the earnings-price ratio over the respective sample period, discretized with 25 equally spaced grid points. The ten lines within each graph correspond to different planning horizons, ranging from one year (dashed) to 10 year (solid). $\zeta$ denotes the coefficient of relative risk aversion.

Again, it is interesting to extend the analysis to the other predictive variables introduced above. In what follows, we repeat the same calculations for those variables as well and discuss the magnitude of the intertemporal hedging demand.

**U.S. TED Spread** Under the specification given in equation (13), available investment opportunities change over time along with the time variation in the predictive variable. In the case of the dividend-price ratio, for example, shocks to excess returns are negatively correlated with shocks to the dividend-price ratio. Moreover, when the dividend-price ratio falls, expected excess returns fall, too. Investors with $\zeta > 1$ may want to hedge these changes in the invest-
Figure 18: U.S. TED spread. Optimal allocation to stocks plotted against the initial value of the U.S. TED spread.

The investor rebalances optimally once a year, uses a VAR model which allows for return predictability, and has power utility over terminal wealth. The predictive variable is the U.S. TED spread. Parameter values are estimated over the 1975 to 1988 (graph on the left) and the 1989 to 2002 sample period. The state space is the interval ranging from the maximum and minimum value of the U.S. TED spread over the respective sample period, discretized with 25 equally spaced grid points. The ten lines within each graph correspond to different planning horizons, ranging from one year (dashed) to 10 year (solid). ζ denotes the coefficient of relative risk aversion.

Given the regression evidence in Tables 4 to 8, the only case where a similar picture evolves is when it is assumed that the U.S. TED spread governs expected excess returns. Table 7 shows that innovations to excess return and TED spreads are negatively correlated. Suppose that the TED spread falls unexpectedly. Since σξη < 0, this is likely to be accompanied by a contemporaneous positive shock to excess returns. However, since the TED spread is lower, excess returns are forecast to be lower in the future, since β > 0. This rise, followed by a fall in excess returns, generates not only a component of negative serial correlation in realized excess returns (which would be, loosely speaking, relevant for buy-and-hold investors), but also makes holding more in stocks an ideal way of hedging against movements in expected excess returns. However, as indicated in Table 7, the correlation between innovations to excess return and TED spreads is much lower (in absolute terms) compared to the case where the dividend-price ratio governs expected excess returns.

Indeed, Figure 18 shows that the resulting allocations to stocks do not show a horizon effect. The intertemporal hedging demand is zero for the U.S. TED spread as predictive variable.

For the sample period from 1975 to 1988, the correlation between innovations to excess return and TED spreads is, at least, −0.3140, but return predictability
is extremely weak with $R^2 = 0.00\%$. Consequently, the optimal allocations to stocks are merely flat over the respective state space. In contrast, over the recent subperiod, optimal allocations to stocks are lower when initial values of the TED spread are low, indicating significant return predictability. On the other hand, the correlation between innovations to excess return and TED spreads is only $-0.0674$, not enough to create an intertemporal hedging demand.

**Other Predictive and Stochastically Detrended Predictive Variables**

In general, Tables 4 to 8 show rather low (i.e., near-zero) correlations between innovations in excess returns and the respective predictive variables. Often, the correlations are even slightly positive. In this case, realized returns hardly suffer a contemporaneous negative shock if there is a positive shock to the expected risk premium, despite the fact that the discount rate for discounting future cash flows has suddenly increased. Consequently, the risky asset cannot be used to hedge the changes in the investment opportunity set and the intertemporal hedging component is zero as well. In other words, investors who are more risk averse than a log-utility maximizer would not like to hold more stocks (positive intertemporal hedging demand) since the bad returns will not likely be compensated by good investment opportunities.

This is true for the remaining variables parameterizing expected excess returns. Neither of the term spread, the one-month Swiss interbank rate, the U.S. default risk spread, and the realized stock market volatility give rise to a significant intertemporal hedging demand. The very same is true when the predictive variables are stochastically detrended, even including the stochastically detrended dividend-price ratio and the stochastically detrended earnings-price ratio.

Overall, thus, given our framework, stocks do not provide a good hedge to changes in the investment opportunity set. With the exception of the dividend-price ratio as predictive variable from 1975 to 1988, the effect of intertemporal hedging is not only empirically small, but rather non-existent. To save space, we do not report the results of the other predictive variables.

**Optimal Allocations to Stocks Over Time**

To highlight the effect of the intertemporal hedging demand, we may also track the optimal allocations to stocks over time. Figure 19 shows the optimal allocations to stocks of a myopic strategy (planning and investment horizon are one year) and of a strategy that optimally rebalances (planning horizon is ten year, investment horizon one year), when the predictive variables are the dividend-price ratio and the earnings-price ratio, respectively. The parameters are estimated over the full period from 1975 to 2002. The portfolios are adjusted at the beginning of each year, using the prevailing initial values of the predictive variables.

Figure 19 establishes that intertemporal investors invest more aggressively in equities. On average, their allocations to stocks are higher than for myopic buy-and-hold investors. When parameter uncertainty is taken into account, however, the difference, i.e., the intertemporal hedging demand, is much lower.
Figure 19: Optimal allocation to stocks plotted over time.

This figure shows optimal allocations to stocks over time. The predictive variables are the dividend-price ratio (the graphs on the top) and the earnings-price ratio. The parameters are estimated using the full sample period from 1975 to 2002. The graphs on the left ignore parameter uncertainty, the graphs on the right take it into account. $\zeta$ denotes the coefficient of relative risk aversion.
After all, even in the case of the dividend-price ratio and the earnings-price ratio, we may again conclude that the effect of intertemporal hedging is empirically small. The effect of a changing degree of risk aversion, for example, is much more influential to optimal allocations to stocks than the intertemporal hedging demand. In addition, it is likely that the effect of optimally rebalancing would diminish even further in an out-of-sample analysis (which would be computationally demanding, however).

4 Conclusion

This paper empirically examines the implications of time-varying excess returns for investors making sensible portfolio decisions. As in Barberis (2000), we use the sensitivity of the optimal portfolio allocation to the investor’s planning horizon as a way of thinking how return predictability and parameter uncertainty may effect optimal stock holdings of buy-and-hold and dynamically rebalancing investors.

When investors take return predictability into account, Barberis (2000) argues that, because time variation in expected excess returns induces mean reversion in realized returns, thus slowing the growth of conditional variances of long-horizon returns, buy-and-hold investors should invest substantially more in risky equities the longer their planning horizon. In much the same way, in a dynamic and intertemporal setting with optimal portfolio rebalancing, Barberis (2000) also claims that investors who are more risk averse than log-utility investors should allocate substantially more to equities the longer their planning horizon. In this case, the higher allocation to equities would provide the investors with an intertemporal hedge against changes in available investment opportunities.

However, our own empirical results based on Swiss stock market data from 1975 to 2002 shows that this is only true when expected excess returns are parameterized by the dividend-price ratio and the time period considered is from 1975 to 1988. For the remaining and stochastically detrended variables parameterizing expected excess returns, there are neither buy-and-hold horizon effects nor do dynamically rebalancing investors increase their equity holdings due to intertemporal hedging demands. In addition, recently, from 1989 to 2002, any evidence of stock market predictability by the dividend-price ratio seems to have disappeared anyway.

Moreover, when investors additionally take parameter uncertainty into account, our results suggest that any horizon effect largely disappears, even in the case of the dividend-price ratio and from 1975 to 1988. Thus, any investment advice that ignores parameter uncertainty may leads to stock holdings which are both too large and too sensitive to the predictive variable.

Still, although parameter uncertainty makes the optimal allocation to stocks somewhat less sensitive to the prevailing value of the predictive variable, the potential for short-term myopic tactical asset allocation strategies is not severely affected by estimation risk. Irrespective of whether short-term buy-and-hold
investors account for parameter uncertainty or not, there is enough predictability in returns to make them time the market. Finally, we conclude that when a tactical asset allocation strategy is motivated by return predictability, intertemporal hedging demands are generally small, if not non-existent. Dynamic investors who follow a dynamic rebalancing strategy do not generally hold more stocks and do not time the market more aggressively than myopic buy-and-hold investors. Common definitions for tactical asset allocation strategies as myopic buy-and-hold strategies seem therefore to be empirically justified.
A Description of Data

Appendix A describes the Swiss stock market data used throughout the paper.

A.1 The Stock Market Index

The empirical results in this paper are based on Swiss stock market data. Monthly data are used throughout, spanning 336 months from January 1975 to December 2002. We also investigate two subsamples of equal length (each 168 months). The first subsample uses data from January 1975 to December 1988, covering the first half of the total time period, the second subsample is based on data from January 1989 to December 2002, covering the second half of the full sample.

The Swiss stock market index is a value-weighted aggregate of the following industry sector indices:

- Airlines and Transportation (Datastream Mnemonic: AIRLNJBV),
- Finanicals (BANKSBV, INSURBV),
- Food (FOODSBV, BREWSB),
- Industrials (GENINBV),
- Pharma (PHARMBV),
- Retailers (MULTIBV, FDRETBV),
- Utilities (UTILSBV), and
- Other Businesses (OTHBUBV, LESURBV).

All return series include dividends (total returns). To obtain continuously compounded real returns, total returns are deflated using monthly rates of change in the Consumer Price Index (CPI), provided by the Swiss National Bank. Continuously compounded excess returns are less the prevailing one-month Swiss interbank rate (SWIBK1M) at the beginning of the month.

Figure 20 shows the nominal price index (January 1975 = 100) and the respective continuously compounded return series.

Table 9 presents summary statistics for the continuously compounded excess and real returns.

Mean Reversion in Stock Market Returns An interesting and important sidestep is to consider whether the stock market returns exhibit mean reversion over the respective sample periods. In particular, we calculate the $q$-period variance ratio statistic $VR(q)$ as

$$VR(q) \equiv \frac{Var(\epsilon_{t-q})}{qVar(\epsilon_t)} = 1 + 2 \sum_{j=1}^{q-1} \left( 1 - \frac{j}{q} \right) Corr(\epsilon_t, \epsilon_{t+j}),$$  

(28)
Figure 20:
The Swiss stock market.
The graph on the left shows the time series of the aggregated nominal price index (January 1975 = 100). The graph on the right plots the respective continuously compounded return series. Monthly data are used from January 1975 to December 2002.

Table 9:
Summary statistics of the Swiss stock market data.
The table presents summary statistics of continuously compounded real and excess monthly Swiss market stock returns. These include mean (ann.), median, maximum and minimum value, volatility (ann.), skewness, and kurtosis. The table includes the Jarque and Bera (1980) test of normality. Estimates are given for three different time periods: 1975 to 2002 (full sample), 1975 to 1988, and 1989 to 2002. */**/*** indicate p-values less than 0.1/0.05/0.01.
Figure 21: Mean reversion and variance ratio statistics.
The graphs show variance ratio statistics for time horizons varying from one to 60 months. A variance ratio below one indicates mean reversion, a ratio above one mean aversion. The graph on the left is based on data from 1975 to 1988. The graph on the right uses monthly data from 1989 to 2002. The fine lines plot variance ratios calculated based on the full sample (1975 to 2002). The dotted lines are based on the standard approach given in equation (29). Bold lines correspond to efficient and bias-adjusted variance ratio statistics (Lo and MacKinlay, 1988, 1989).

with \( e_{t-q} \equiv e_t + e_{t+1} + \ldots + e_{t+q-1} \) and \( \text{Corr}(e_t, e_{t+j}) \) is the \( j \)th-order autocorrelation coefficient of \( \{e_t\} \). If returns are positively autocorrelated, variances grow faster than linearly and the variance ratio is above one for \( q > 1 \), \( \text{VR}(q) > 1 \). Alternatively, in the presence of negative autocorrelation, the variance of the sum of one-month returns is smaller than the sum of the one-month return’s variances; hence \( \text{VR}(q) < 1 \), variances grow slower than linearly. This is the case of mean reversion. If returns are i.i.d., \( \text{VR}(q) = 1 \).

Figure 21 shows variance ratio statistics for time horizons, \( q \), varying from one to 60 months. Based on the data covering the first half of the total time period (1975 to 1988), continuously compounded excess returns seem to mean revert. The variance ratio statistic is well below one for long horizons. Recently, however, the mean reversion pattern has completely disappeared. Rather than mean reversion, the Swiss stock market exhibits strong mean aversion since 1989. However, if the estimates of the variance ratio statistics are based on overlapping \( q \)-period returns and corrected for the bias in the variance estimators (Lo and MacKinlay, 1988, 1989), the evidence of mean reversion over the first subsample vanishes. Over the recent subsample, the tendency of mean aversion is even amplified.

B Numerical Procedures
To complete the picture, Appendix B demonstrates how the integrals in Section 2 are evaluated numerically by simulation. We first describe the numerical procedure for the portfolio choice problem of a myopic buy-and-hold investor. Only then we move to the dynamic programming problem and describe the
discretization of the state space and the principle of backward induction.

The notation is based on Barberis (2000).

B.1 The Portfolio Choice Problem of a Buy-and-Hold Investor

How can problems (4) and (6) be solved? Given a risk aversion coefficient \( \zeta \), we calculate the integrals in those problems for \( \omega = 0, 0.01, 0.02, ..., 0.98, 0.99 \), and report the \( \omega \) that maximizes expected utility. As in Barberis (2000), we restrict the allocation to stocks to the interval \( 0 \leq \omega \leq 0.99 \), thus precluding short selling and buying on margin. We do not calculate expected utility for \( \omega = 1 \) because in this case the integral in equation (6) equals \( -\infty \). The problem is that when \( \omega = 1 \), wealth can be arbitrarily close to zero, but the left tail of the predictive distribution does not shrink fast enough to ensure that expected utility is bounded from below (Barberis, 2000, p. 232, footnote 7).

The integrals themselves are evaluated numerically by simulation. In general, we approximate the integral for expected utility by taking a sample

\[
\left( e^{(i)}_{T-T} \right)_{i=1}^{I}
\]

from one of the two possible distributions - \( p \left( e_{T-T+T} \mid \theta, z \right) \) when ignoring parameter uncertainty, the predictive distribution \( p \left( e_{T-T+T} \mid z \right) \) when accounting for estimation risk -, and then computing

\[
\frac{1}{I} \sum_{i=1}^{I} \left\{ (1 - \omega) \exp \left( r_f T \right) + \omega \exp \left( r_f \hat{T} + e^{(i)}_{T-T+T} \right) \right\}^{1-\zeta},
\]

where we take \( I = 10,000 \) throughout (Barberis, 2000, eq. 12). For each of the two cases where the investor either ignores or accounts for parameter uncertainty, we present the optimal allocations \( \omega \) that maximize equation (29) for a variety of risk aversion levels \( \zeta \) and planning horizons \( \hat{T} \).

Of course, an important issue is the accuracy of the numerical methods used to obtain the optimal portfolios. In an effort to ensure a very high degree of accuracy, Barberis (2000) uses a very large sample of \( I = 1,000,000 \) draws from the appropriate distribution when evaluating the integrals for expected utility. But his appendix also suggests that using \( I = 10,000 \) draws already provides a high degree of accuracy. Indeed, own checks on simulation errors (not reported) suggest that for the smaller sample size actually used to obtain our results, there does not appear to be any significant variation in the recommended portfolios.
B.2 The Dynamic Programming Problem

In the case of the dynamic portfolio choice problem, Barberis (2000, eq. 33) shows that the Bellman equation can be written as

\[ Q(x_k, t_k) = \max_{\omega_k} E_k \left\{ \left[ (1 - \omega_k) \exp \left( r_f \frac{\hat{T}}{K} \right) + \omega_k \exp \left( r_f \frac{\hat{T}}{K} + e_{k-1} \right) \right]^{1-\zeta} \right\} \times Q(x_{k+1}, t_{k+1}) \]

and indicated that the usual technique for solving a Bellman equation is to discretize the state space and then use backward induction. Following Barberis (2000), the aim of this section is to formalize this numerical technique.

In particular, we take the interval ranging from the maximum and minimum value of the predictive variable over the respective sample period, and discretize this range with 25 equally spaced grid points, which we write as

\[ x_j^{(k)} \quad j = 1, \ldots, 25. \]

Suppose that \( Q(x_{k+1}, t_{k+1}) \) is known for all \( x_{k+1} = x_j^{(k+1)}, j = 1, \ldots, 25 \). Clearly, this is true in the last period as \( Q(x_K, t_K) = 1, \forall x_K \). Then we can use equation (27) to obtain \( Q(x_j^{(k)}, t_k) \) and, in general, \( Q(x_j^{(k)}, t_k) \).

Specifically, for each \( x_j^{(k)}, j = 1, \ldots, 25 \), we draw a large sample

\[ (e_{t-1}^{(i)}, x_{t+1}^{(i)})_{i=1}^I \]

from the Normal distribution \( p(e_{t-1}, x_{t+1} | \theta, x_j^{(k)}) \), and set \( Q(x_j^{(k)}, t_k) \) equal to (Barberis, 2000, eq. 34)

\[ \max_{\omega_k} \frac{1}{I} \sum_{i=1}^I \left\{ (1 - \omega_k) \exp \left( r_f \frac{\hat{T}}{K} \right) + \omega_k \exp \left( r_f \frac{\hat{T}}{K} + e_{t-1}^{(i)} \right) \right\}^{1-\zeta} \times Q(x_j^{(k+1)}, t_{k+1}) \].

Of course, in general, we only know \( Q(x_{k+1}, t_{k+1}) \) for \( x_{k+1} = x_j^{(k+1)} \), so we approximate \( Q(x_{j+1}^{(k+1)}, t_{k+1}) \) by \( Q(x_j^{(k+1)}, t_{k+1}) \), where \( x_j^{(k+1)} \) is the closest element of the discretized state space to \( x_{j+1}^{(k+1)} \). This calculation gives \( Q(x_j^{(k)}, t_k) \) for all \( j = 1, \ldots, 25 \). Backward induction through all \( K \) rebalancing points eventually gives \( Q(x_0^{(k)}, t_0) \), and hence the optimal allocations \( \omega_0 \) (Barberis, 2000, p. 252).

We find that using a sample size of \( I = 20,000 \) from the distribution of cumulative excess returns provides a satisfying degree of accuracy. Another factor that affects the accuracy of the results is the number of grid points used to discretize the predictive variable. We use a range of 25 equally spaced grid points.
References


