# Dynamic portfolio allocation by parametric portfolio policies

## Master thesis Quantitative Finance


Luka Kranjcev

(569864)

Erasmus University Rotterdam [*]

Erasmus School of Economics


Supervisor: dr. Rasmus Lönn

Second assessor: dr. Gustavo Freire

November 16, 2022


**Abstract**

Over the last few decades, several competing approaches to portfolio selection have been proposed. In this thesis, I reexamine the parametric portfolio policy approach of Brandt and Santa-Clara (2006) which enables market timing of portfolios and a simple implementation of dynamic portfolios. I extend the original study by applying several methods aimed at reducing parameter estimation error and combine it with cross-sectional information. An empirical implementation of this approach and its extension is presented in different investment settings. I find that gains from including conditioning information (i.e. market timing) are difficult to realize out-of-sample regardless of asset allocation model or asset set.


[*]The content of this thesis is the sole responsibility of the author and does not reflect the view of the supervisor, second assessor, Erasmus School of Economics or Erasmus University.

# Contents

# 1 Introduction

The problem of portfolio selection is of great importance both in financial research and asset management industry. Quantitative approach to portfolio selection relies on mathematical models and statistical analysis to make investment decisions. The most famous framework in this context is the mean-variance analysis (Markowitz (1952)) which gives optimal portfolio weights as a function of assets' expected returns, variances and covariances for an investor that trades off the mean and variance of portfolio returns. Most other portfolio management approaches also derive optimal portfolio choices as functions of moments of financial returns. It is well known that such traditional implementations of portfolio allocations suffer from serious shortcomings (Jobson and Korkie (1980), Best and Grauer (1991), Aït-Sahalia and Brandt (2001)).

In a single period setting these issues usually stem from large parameter uncertainty in expected returns and covariance matrices due to large amounts of noise in financial returns data. This means that the resulting portfolio weights can be far away from the theoretically optimal portfolio weights, leading to sub-optimal investments. The importance of this problem, particularly in terms of economic gains and losses, was demonstrated already in Jobson and Korkie (1980). Nowadays, most implementations of mean-variance analysis include some sort of remedy to deal with parameter uncertainty issues.

In multiple period allocation problems, there are additional issues resulting from empirically observed return predictability, which makes the problem extremely complex and difficult to solve. This is usually done by means of dynamic programming which involves solving the Bellman equation (e.g. Brandt et al. (2005)), and alternatively by stochastic optimal control (e.g. Chellathurai and Draviam (2007)) or reinforcement learning (e.g. Hens and Wöhrmann (2007)). Neither of these methods is easy to implement. A more recent approach to portfolio allocation, introduced by Brandt and Santa-Clara (2006) and Brandt et al. (2009) avoids portfolio weights estimation through assets expected returns and covariances, and instead directly establishes a model for portfolio weights.

A particularly interesting area of portfolio management is a long-horizon dynamic portfolio allocation. This is because many investors face long term investment decisions which allow rebalancing before the terminal period. Bodnar et al. (2015) notes that much less has been done in that area compared to the single-period portfolio choice, while being of greater importance practically (e.g. investing for retirement). In addition, it is known already since Samuelson (1969) that if returns are not independent across time, the optimal long-term allocation differs from the short-term allocation.

Despite this, as Brandt (2010) notes, only a few financial institutions implement a dynamic instead of a static, single period solution. One of the reasons for this is that its use is prevented by difficult and cumbersome implementation based on approximations and numerical techniques. Furthermore, long-term allocations include a hedge component which is subject to additional estimation error, such that out-of-sample performance might not be superior to that of repeated myopic allocations (Diris et al. (2015)). On the other hand, many studies (e.g. Barberis (2000)) show that gains from utilizing return predictability are significant. Hence, this application is particularly relevant from practitioners' point of view.

To avoid the need of having to work with complex optimization problems which require the use of numerical techniques, different papers have proposed various approximations or continuous time setting to obtain more straightforward solutions. One of such approximations is the idea of directly modeling portfolio weights by circumventing the use of expected returns, covariances and other moments of returns. Brandt and Santa-Clara (2006) are the first to consider directly modeling portfolio weights in a dynamic setting. The basic idea behind the method is that optimal portfolio weights are directly parameterized as a function of state variables. Furthermore, they show that by using some simple linear algebra rules, the solution can be formulated in terms of a standard, static Markowitz solution. The obvious advantage of this approach and thus practical relevance is that we do not have to resort to dynamic optimization to derive optimal portfolio allocations but can use simple statistics such as OLS. The disadvantage is that the solution is only an approximation which deteriorates with investment horizon and rebalancing frequency. However, Brandt and Santa-Clara (2006) show that for horizons up to 5 years, it still offers satisfactory performance.

The framework also allows straightforward inference on the coefficients of the weight function due to Britten-Jones (1999), who shows that the standard OLS regression inference can be applied to portfolio weights. This enables us to say which state variables or characteristics are important for the optimal portfolio policy by implicitly affecting return moments. While it is known that for example dividend-price ratio affects expected returns, it could be that it also affects covariances in an offsetting way such that optimal portfolio weights do not change. Hence, considering the problem in this framework also offers an answer to the question of which economic variables actually change the composition of optimal portfolios.

Another important advantage of this approach is that the resulting optimal portfolio weights take into account higher order moments of returns and their effect on the distribution of portfolio returns and ultimately expected utility. This is particularly appealing if investors do not have quadratic utility, in which case higher order moments count. With a large number of assets, the traditional approach of modeling return moments as functions of characteristics is infeasible since too many parameters need to be estimated. Hence, the traditional approach is restricted to mean-variance preferences which is not desirable considering investors' probable utility functions and empirical characteristics of asset returns.

Hence, in principle this approach is very appealing on a number of grounds. However, this approach to portfolio allocation, while relatively new, has received surprisingly little attention in the literature. This thesis intends to fill the gap in the literature by reexamining and extending the method. It also provides a thorough out-of-sample evaluation which is an important consideration as one of the intentions of Brandt and Santa-Clara (2006) was to introduce it in the real world portfolio management. Furthermore, since this approach leads to a static Markowitz solution, many of the tools developed to improve the original solution by reducing parameter uncertainty are examined. These include shrinkage, portfolio restrictions, prior information, etc. Furthermore, the basic performance in the dynamic setting has not been yet evaluated, since Brandt and Santa-Clara (2006) evaluate its performance on its own, without comparing it to the theoretically optimal solution based on the Bellman equation solution. Finally, there is even some value in replicating the original results since often important results in quantitative finance have proven not to be robust.

I find that many appealing results from Brandt and Santa-Clara (2006) no longer hold when evaluated out-of-sample. In particular, portfolio policies which use conditioning information and perform well in-sample, often perform equally or even worse when assessed out of sample. Careful choice of estimators and state variables is needed in order to improve on unconditional portfolio policies. This is the case for both, single-period and multi-period investment problems. To deal with parameter uncertainty I consider moment shrinkage, weights shrinkage, portfolio constraints and the Black-Litterman model. As with the standard assets such as stocks or industry portfolios, neither of these techniques proves superior when applied to the managed portfolios of parametric portfolio policies. I also find that in the dynamic setting the approximation which the method provides is not inferior to the traditional exact solutions. However, the multi-period solution can quickly involve a large number of unknown parameters in which case this approach is unappealing despite its relative ease of dealing with parameter uncertainty. I also find that investment performance can be improved by applying the method to other sets of assets such as industry portfolios or combining it with cross-sectional information.

The rest of the thesis is organized as follows. Section 2 summarizes the literature on parameter uncertainty, the use of conditioning information and dynamics in portfolio allocation. In Section 3, the mechanics of Brandt and Santa-Clara (2006) method are discussed and the proposed extensions. Section 4 describes the data and its sources. The methods are implemented in an empirical study in Section 5. Finally, Section 6 concludes.

## 2 Literature

Much research has been done on the single period portfolio choice starting with the seminal paper of Markowitz (1952) which laid the theoretical foundation of the portfolio choice problem. The basic formulation of the problem is based on a number of simplifying assumptions such as mean-variance preferences and frictionless markets. Thus, a large part of the literature focuses on relaxing these assumptions to arrive at more realistic investment setups (see Brandt (2010) for a short summary). The second focus stems from the fact that the inputs to the optimization problem, expected returns and covariance matrix, are not known in practice. Using naive sample analogues has proven to result in portfolios that perform suboptimally due to large uncertainty in input estimates (Jobson and Korkie, 1980). Hence, the econometrics aspect of portfolio choice deals with the question of how to optimally use historic data for portfolio allocation.

To deal with large uncertainty in optimal weights resulting from imprecisely estimated input parameters, several techniques have been proposed in the literature (Kolm et al. (2014)). Some of the most popular ones include shrinkage, portfolio restrictions, factor models and mixed estimation. The idea of shrinkage estimation comes from the surprising result that with three or more normal random variables, the joint mean squared error can always be reduced by combining sample means with any constant (James and Stein, (1961)). Jorion (1986) showed that this result also holds in the context of portfolio choice when applied to expected returns with the objective of maximizing utility. Portfolio constraints on the other hand are direct constraints on portfolio weights. Extreme weights are likely the result of severe estimation errors due to the nature of optimization as shown in Michaud (1989), therefore eliminating extreme positions should improve portfolio performance. Economic theory can also help improve the inputs of optimization, in particular the covariance matrix. Modeling financial markets using factor structure where factors are justified as risk sources greatly reduces the number of elements of covariance matrix that need estimation, resulting in more stable estimates. Lastly, the Black-Litterman model combines both the statistical and economic insights using the framework of Bayesian mixed estimation to determine the moments of returns. All of the above techniques are in practice often used side by side since there is no single approach which is superior in all applications (DeMiguel et al. (2009)).

Recent econometric advances also take advantage of the empirical finding of time-varying return distributions, i.e. predictability and conditional hetersokedasticity (Fama and French (1988), Engle (1982)). The use of conditioning information in portfolio formation which captures changing economic conditions dates back to Hansen and Richard (1987). Initially it was applied to model return moments as a function of state variables and using the conditional moments to form portfolios (e.g. Ferson and Siegel (2001)). This is straightforward for expected returns, however, with N assets the covariance matrix requires estimation of N(N+1)/2 functions, where each function includes K unknown parameters corresponding to K state variables. This makes the approach susceptible to the original issue of parameter uncertainty. Brandt and Santa-Clara (2006) proposed an alternative which circumvents the need of modeling return distributions. Namely, they assume that portfolio weights can be modeled as (linear) functions of state variables. The so-called parametric portfolio policy (PPP) approach has the advantage of being much more parsimonious with only N functions (or N×K parameters) to be estimated which results in considerable estimation and computation efficiency. This is especially convenient when considering more realistic utility functions involving higher moments of returns, since PPP implicitly accounts for the effect of state variables on all relevant moments of asset returns.

While the single-period portfolio choice is appealing due to its tractability, most real-world investment problems involve long horizons with the possibility of rebalancing such as pension investing. The multi-period counterpart received much less attention until the breakthrough by Samuelson (1969) and Merton (1969, 1973) who derived the conditions in discrete and continuous time under which myopic and dynamic optimal allocations are equivalent. Most importantly, this happens if asset returns are independent and identically distributed. However, while theoretically the problem is not difficult to formulate, its empirical implementation is much harder due to the complexity of the problem and the lack of closed form solutions. In addition, it was thought until late 1980s that asset returns, in particular stock returns, were unpredictable (Cochrane, 2011) in which case the single-period solution to the long-horizon problem is optimal. Since then there has been ample literature showing that many economic variables related to business cycle such as price-dividend ration, term spread, etc. are able to predict financial asset returns, with the fit increasing in the time horizon (e.g. Campbell and Shiller (1988), Campbell and Viceira (2002)). However, more recent studies such as Welch and Goyal (2008) have argued that this finding is much less robust when examined out-of-sample. Using kernel regressions, Farmer et al. (2022) find some evidence of out-of-sample predictability which is, however, restricted to certain subsamples of up to two years ("pockets") which represent only 15% of the whole sample.

The ability to predict returns means that changes in the investment opportunity set can be anticipated which creates intertemporal hedging demands. Hence, long-term solution at any point in time can be decomposed in the myopic solution and hedging demand. There have been several studies showing that these hedging demands tend to be large and significant (Barberis (2000), Larsen and Munk (2012)). For example, Basak and Chabakauri (2010) show that the percentage hedging demand can range from 18% to 84% in different economic settings. In general, it is found that in asset allocation problems, long-term portfolios allocate more to stocks compared to the single-period due to the mean-reversion in stocks.

Numerical solutions to dynamic portfolio allocation problems can be obtained by stochastic dynamic programming, stochastic optimal control or reinforcement learning (e.g. Cvitanic et al. (2003), Cong and Oosterlee (2017)). While the general problems require the use of numerical techniques, analytical solutions have been developed in continuous-time setting and under various return process assumptions (Lim and Zhou (2002), Basak and Chabakauri (2010)). Closed form solutions in the more difficult case of discrete time have been obtained by considering various approximations (e.g. Leippold et al. (2004), Bodnar et al. (2015)). Importantly, Brandt and Santa-Clara (2006) derive an approximate solution to the multi-period portfolio problem by ignoring compounding and forming the so called timing portfolios which replicate dynamic allocations using a static setting. Consequently, the optimal dynamic solution is obtained from the standard single-period Markowitz solution. This implies that all the above mentioned tools of dealing with parameter uncertainty can be used to address the otherwise difficult issue in dynamic portfolio choice.

This methodology has been implemented in de Roon et al. (2010) who consider the effects of a lockup period characteristic of many hedge funds. Similarly, Plazzi et al. (2010) use identical setup as Brandt and Santa-Clara (2006) but also add commercial real estate as available asset for investment. As such, these studies only use the method to obtain empirical results without examining or extending the method itself.

# 3 Methodology

## 3.1 Parametric portfolio policies

The core of this study relies on the parametric portfolio policy proposed in Brandt and Santa-Clara (2006). They assume that optimal portfolio weights are linear in state variables summarizing the state of the economy and derive the resulting closed form solutions. While it is shown that implementation is as simple as the standard Markowitz solution, it requires expanding the set of assets to also include managed portfolios. In particular, their portfolio optimization technique essentially relies on two ideas, the conditional and timing portfolios.

### 3.1.1 Unconditional and conditional portfolios

Working in the expected utility framework and assuming mean-variance preferences over next period's wealth, analytical solutions can be derived for portfolio weights in the standard way by maximizing expected utility subject to the budget constraint:

$$\max E_t \left( W_{t+1} - \frac{b_t}{2} W_{t+1}^2 \right) \qquad s.t. \quad W_{t+1} = W_t \left( R_t^f + r_{t+1}^p \right), \tag{1}$$

where $W_{t+1}$ is next period's wealth, $b_t$ is a sufficiently small coefficient such that utility does not decrease with wealth, $R_t^f$ is the gross risk-free rate and $r_{t+1}^p$ is the excess portfolio return.

Substituting the budget constraint in the maximization and ignoring the constants the problem can be written as a trade-off between portfolio's expected return and variance:

$$\max E_t \left( x_t' r_{t+1} - \frac{\gamma}{2} x_t' r_{t+1} r_{t+1}' x_t \right), \tag{2}$$

where $r_{t+1}$ is an $N \times 1$ vector of asset returns, $x_t$ is an $N \times 1$ vector of weights and $\gamma$ is a coefficient summarizing the investor's risk aversion.

When returns are assumed to be independent and identically distributed (i.i.d.), portfolio weights are constant over time, which means that the conditional expectation is equal to the unconditional expectation. This results in the familiar Markowitz solution where population moments have been replaced with the sample averages:

$$x = \frac{1}{\gamma} \left[ \sum_{t=1}^{T-1} r_{t+1} r_{t+1}' \right]^{-1} \left[ \sum_{t=1}^{T-1} r_{t+1} \right]. \tag{3}$$

Given that the i.i.d. assumption is unlikely, Brandt and Santa-Clara (2006) assume that portfolio weights are linear in a vector of K state variables $z_t$ through an N x K matrix of constant coefficients $\Theta$:

$$x_t \;=\; \Theta z_t. \tag{4}$$

This leads to the so-called conditional portfolios which make investment in each asset proportional to the level of the conditioning variables while simultaneously maintaining the optimization simplicity of the Markowitz framework. Analytical solutions can be derived for portfolio weights of assets in the augmented set by first replacing $x_t$ in equation (2) by the functional form in (4):

$$\max_t \left[(\Theta z_t)' \, r_{t+1} - \frac{\gamma}{2} (\Theta z_t)' \, r_{t+1} r_{t+1}' \, (\Theta z_t)\right]. \tag{5}$$

Using the following result from linear algebra allows separation of vector of parameters from the data:

$$(\Theta z_t)' \, r_{t+1} = z_t' \Theta' r_{t+1} = vec\,(\Theta)' \, (z_t \otimes r_{t+1})\,, \tag{6}$$

where $vec(\Theta)$ stacks the columns of matrix $\Theta$ into a vector and $\otimes$ denotes a Kronecker product. These variables constitute a transformed vector of weights and returns so that the familiar optimization problem is obtained:

$$\max E_t \, \left(\tilde{x}' \tilde{r}_{t+1} - \frac{\gamma}{2} \tilde{x}' \tilde{r}_{t+1} \tilde{r}_{t+1}' \tilde{x}\right), \tag{7}$$

where $\tilde{x} = vec(\Theta)$ and $\tilde{r}_{t+1} = z_t \otimes r_{t+1}$. Hence, the solution gives the optimal weights of managed portfolios, which are equivalent to the basis assets whose returns are scaled by the state variables:

$$\widetilde{x} = \frac{1}{\gamma} \left[\sum_{t=0}^{T} \left(z_t z_t'\right) \otimes \left(r_{t+1} r_{t+1}'\right)\right]^{-1} \left[\sum_{t=0}^{T} \; z_t \otimes r_{t+1}\right]. \tag{8}$$

Inference on the efficient weights on managed portfolios in (8) can be done by first noting that (8) is equivalent to the basic solution in (3) with transformed returns $\tilde{r}_{t+1}$. And then using the results from Britten-Jones (1999) who derives the sampling distribution of estimates of efficient portfolio weights. In this particular case, it can be shown that optimal portfolio weights are equal to the scaled (by $1/\gamma$) OLS coefficients obtained in a regression of a vector of ones on asset returns

(Brandt and Santa-Clara (2006)). Hence, the standard inference that applies to OLS coefficients is used to test hypotheses about portfolio weights. In general, if $\hat{\beta}$ is a vector of estimated regression coefficients, $\sigma^2$ is the variance of the error term and $X$ is the matrix of regressors, then $Var(\hat{\beta}) = \sigma^2(X'X)^{-1}$. In our case with $X = \tilde{r}$ and the errors $\iota_T - \tilde{r}\tilde{x}$, the covariance matrix is given by:

$$Var(\tilde{x}) \;=\; \frac{1}{\gamma^2}\frac{1}{T-(N\times K)}(\iota_T - \tilde{r}\tilde{x})'(\iota_T - \tilde{r}\tilde{x})(\tilde{r}'\tilde{r})^{-1}. \tag{9}$$

#### 3.1.2 Timing portfolios

On the other hand, the idea of timing portfolio enables a multi-period portfolio problem to be solved as a single period problem. However, it comes at the cost of being an approximation which has been examined in a simulation study in Brandt and Santa-Clara (2006). The formulas below give a solution for a two-period problem which can be easily generalized to any H-period problem. Starting from the mean-variance objective:

$$\max E_t \left( r^p_{t\to t+2} - \frac{\gamma}{2}(r^p_{t\to t+2})^2 \right), \tag{10}$$

the portfolio return $r^p_{t\to t+2}$ now represents the excess return of the two-period investment strategy calculated as:

$$r^p_{t\to t+2} = \left(R^f_t + x'_t r_{t+1}\right)\left(R^f_{t+1} + x'_{t+1} r_{t+2}\right) - R^f_t R^f_{t+1}, \tag{11}$$

which can also be written in the following way:

$$r^p_{t\to t+2} = x'_t\left(R^f_{t+1} r_{t+1}\right) + x'_{t+1}\left(R^f_t r_{t+2}\right) + \left(x'_t r_{t+1}\right)\left(x'_{t+1} r_{t+2}\right). \tag{12}$$

The familiar Markowitz solution can be obtained by ignoring the last term which represents compounding of returns and is thus much smaller compared to the first two terms. This causes the solution to be an approximation which could be satisfactory for short horizon and large magnitudes of returns. The solution is then:

$$\widetilde{x} \;=\; \frac{1}{\gamma}\left[\sum_{t=1}^{T-2} \widetilde{r}_{t\to t+2}\widetilde{r}'_{t\to t+2}\right]^{-1}\left[\sum_{t=1}^{T-2} \widetilde{r}_{t\to t+2}\right], \tag{13}$$

where $\widetilde{r}_{t\to t+2} \;=\; \left[R^f_{t+1} r_{t+1}, R^f_t r_{t+2}\right]$.

This can be easily combined with conditioning portfolios by expanding the asset set to include returns scaled by conditioning variables. In particular, $r_{t+1}$ and $r_{t+2}$ are replaced by $z_t \otimes r_{t+1}$ and $z_{t+1} \otimes r_{t+2}$, where $z_t$ includes a constant and variables which are thought to influence return distributions.

### 3.2 Dynamic programming

The benchmark which represents the optimal dynamic allocation is usually obtained through stochastic dynamic programming which involves solving the Bellman equation. The alternative solution techniques found in the literature are stochastic optimal control and reinforcement learning which are used less often. Several methods have been proposed to solve the Bellman equation, which differ in how the conditional expectation of utility of wealth is approximated (Brandt et al. (2005)). In the empirical study below the simulation approach is chosen. This approach is based on estimating a return model from which simulations of return and state variables can be obtained. The advantage of this approach is that it can handle complex objectives and constraints together with the fact that conditional expectation can be approximated arbitrarily close by increasing the number of simulations from the model. However, the drawback is that the return model might be misspecified or parameters estimated with significant error. Van Binsbergen and Brandt (2007) and more recently, Denault and Simonato (2017) present two similar methods (value function vs portfolio weights recursion) of approximating the Bellman equation using simulations and regressions.

In general, the Bellman equation $V_t$ can then be solved by breaking down the multi-period problem into simpler single-period problems and using backward recursion:

$$\begin{aligned} V_t\left(W_t, z_t\right) &= \max_{\{x_s\}_{s=t}^{T-1}} E_t\left[u\left(W_T\right)\right] = \max_{x_t} E_t\left\{\max_{\{x_s\}_{s=t+1}^{T-1}} E_{t+1}\left[u\left(W_T\right)\right]\right\} = \\ &= \max_{x_t} E_t\left\{V_{t+1}\left[W_t\left(x_t' r_{t+1} + R_t^f\right), z_{t+1}\right]\right\}. \end{aligned} \tag{14}$$

The algorithm in Van Binsbergen and Brandt (2007) proceeds in five steps to find the optimal weights:

1. Create a grid of portfolio weights and simulate N sample paths with length T of asset returns and state variables. The simulation model used can be for example a vector autoregressive model (VAR) of (log) asset returns and state variables which is often used in the portfolio choice literature (Campbell and Viceira, 2002).

2. For a given combination of weights from the grid, determine the conditional portfolio return moments. This is done by first calculating portfolio return of each sample path associated

with that combination of weights. Then all (squared) portfolio returns are regressed on a constant and a set of state variables. Finally, the conditional return mean (variance) of each sample path can be obtained as a fitted value from that regression.

3. Approximate the expected utility by a Taylor series expansion using conditional return moments (this is not necessary with mean variance preferences). The expected utility of each sample path can then be obtained by plugging in the state variables.

4. Repeat 2. and 3. for all combinations of weights and for each simulation path select those that maximize the expected utility.

5. Proceed recursively backward to the beginning time period by repeating the above steps. The difference between value function and portfolio weights recursion is whether expected utility or actual utility in used to evaluate the Bellman equation.

## 3.3 Extensions

As shown above, solutions to optimal conditional and dynamic portfolios can be expressed in the form of Markowitz solution. Therefore, many of the refinements developed to improve empirical implementation of the Markowitz approach can be applied. To avoid extreme portfolio weights, a number of techniques from literature can be considered. These include shrinkage, portfolio constraints, economic models restrictions and inclusion of beliefs through Bayesian analysis. While Brandt and Santa-Clara (2006) argue that weights obtained by the proposed portfolio policy are much more stable than the ones resulting from traditional approach based on expected returns and covariances, their empirical study still yields large variation in weights which frequently exceed 400%. Expanding the asset set beyond the three assets in their study could make the problem even worse, and thus these techniques are fair to consider.

### 3.3.1 Shrinkage

Shrinkage techniques can be applied to inputs of the optimization which traditionally are expected returns and covariance matrix. The general idea is to multiply the sample moments with a coefficient smaller than 1 (i.e. shrink) and combine them with some other information to obtain new moment estimates:

$$\hat{\theta}^{sh} = (1-\rho)\hat{\theta} + \rho\theta^{target}. \tag{15}$$

This is often found to result in smaller total loss resulting from estimation. One of the first applications of shrinkage estimation in portfolio optimization was proposed by Jorion (1986) who shrinks the vector of asset means towards the mean of the minimum variance portfolio by means of Bayesian analysis. The so called Bayes-Stein approach is still widely used in empirical studies (e.g. Bessler et al., 2017). More recently, DeMiguel et al. (2013) conduct a comprehensive study of shrinkage estimators in the context of asset allocation by considering different calibration criteria and approaches for computing the shrinkage intensity. They find that calculating the shrinkage intensity by minimizing expected quadratic loss rather than through empirical-Bayes approach proposed by Jorion (1986) results in a shrinkage vector of means and portfolios that outperform the ones based on Jorion (1986). In addition, assuming only iid returns, they show that when minimizing expected quadratic loss of mean vector:

$$\min_{\alpha} E\Big[\|\mu_{sh} - \mu\|_2^2\Big], \tag{16}$$

the optimal shrinkage intensity has a closed form expression:

$$\alpha_\mu = \frac{E\Big(\|\mu_{sp} - \mu\|_2^2\Big)}{E\Big(\|\mu_{sp} - \mu\|_2^2\Big) + \|\mu_{tg} - \mu\|_2^2} = \frac{(N/T)\overline{\sigma^2}}{(N/T)\overline{\sigma^2} + \|\mu_{tg} - \mu\|_2^2}, \tag{17}$$

where $\overline{\sigma^2} = trace(\Sigma)/N$, $\mu$ is the population mean vector, $\mu_{sp}$ is the sample mean vector and $\mu_{tg}$ is the target mean vector. In practice $\mu_{sp}$ is used as a proxy for $\mu$.

The alternative to shrinking the moments of asset returns is shrinking the portfolio weights directly:

$$w_{sh} = (1 - \alpha)w_{sp} + \alpha w_{tg}. \tag{18}$$

This might be particularly suitable for the parametric portfolio policy where the resulting weights contain not only original asset weights but also the weights on managed portfolios containing state variables. The latter can be shrunk to zero implying that state variables have no predictive power in portfolio choice. On the other hand, using insights from economic theory, the weights on basis weights can be shrunk towards model implied weights such as those from the CAPM or some other factor model. Alternatively, the prior on can be based on the rules often used in practice. For example, financial planners often advise to split wealth among the three assets in certain proportion that might depend on the age (e.g. 70% into stocks, 20% into bonds, 10% into cash). Hence, using economic information from the state variables, these proportions can be tilted to exploit or hedge against changing investment opportunities. Another option, used in my empirical study is the 1/N portfolio which has been shown to perform very well in practice, beating many sophisticated portfolio allocation strategies (DeMiguel et al. (2009)). To determine the shrinkage intensity several criteria have been proposed in the literature (e.g. DeMiguel et al. (2013)). These include expected utility/Sharpe ratio maximization or expected quadratic loss/variance minimization. However, as opposed to moment shrinkage, to obtain closed form expressions one must assume iid normal returns which is not appealing from the practical perspective. Hence, in the empirical section a simple grid search is used to find the optimal shrinkage intensity. In particular, for a given estimation subsample in a rolling or expanding window approach, I use the value from a grid of values between 0 and 1 which maximizes the in-sample Sharpe ratio.

### 3.3.2 Portfolio restrictions

Portfolio restrictions are even more straightforward solution since they can be implemented by simply introducing no short-selling or no borrowing constraints in the optimization problem. These are often requirements for many institutional investors therefore examining their effect on performance is of interest on its own. In addition, it has been shown that many types of norm constraints on portfolio weights lead to the same portfolio allocation as shrinking the covariance matrix (e.g. Jagannathan and Ma (2003), DeMiguel et al. (2009)). Furthermore, Fan et al. (2012) presented a theoretical result that imposing an L1 norm constraint on portfolio weights creates an upper bound on the estimation risk.

### 3.3.3 Black-Litterman model

The above mentioned techniques have been found to be relatively successful in alleviating the parameter uncertainty issue, however their effectiveness depends on the application of interest. In a number of applications, the Black-Litterman model proved to be a more effective alternative (e.g. Bessler et al. (2014). The Black-Litterman model is similar to the Bayes-Stein shrinkage estimator where the target now comes from the investor's views and the shrinkage scalar parameter is replaced with a matrix. The possibility to include subjective opinions and their reliability in a quantitative model has made this model very popular in practice, despite not receiving much attention in the academic literature. While in theory investor's views can come from any source, it is most convenient in this setting to use historic data to form the views. This avoids the hardship of obtaining the views and enables straightforward comparison with other portfolio strategies. Such sample based approach to the Black-Litterman model was proposed by Bessler et al. (2014).

The other important element of this model is the reference portfolio $w*$ which is usually derived by assuming that the financial market is in equilibrium. It is also possible to use portfolios that perform well empirically such as the 1/N portfolio or the minimum variance portfolio. Assuming that investors mean-variance optimize, the reference portfolio corresponds to the following implied mean vector:

$$\mu_I = \gamma \Sigma w^*. \tag{19}$$

Combining it with the views $\mu_V$ results in the Black-Litterman model implied mean vector and covariance matrix:

$$\hat{\mu}_{BL} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\mu_I + P'\Omega^{-1}\mu_V], \tag{20}$$

$$\Sigma_{BL} = \Sigma + [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}. \tag{21}$$

where $P$ is a binary matrix used to form the views (an identity if views are sample estimates), $\Omega$ and $\tau$ are measures of uncertainty of the views and implied return estimates, respectively.
An important practical issue concerns the choice of $\tau$ and $\Omega$. In the literature, the value for $\tau$ is usually between 0.025 and 0.300 (Idzorek (2007)). The reliability of the views, $\Omega$, can be either proportional to $\Sigma$ by factor $1/c$ or time-varying which is based on historical forecast errors of the views (Bessler et al. (2014)). The former, proposed by Meucci (2010), sets a constant $\Omega$:

$$\Omega = \left( P \left( \frac{1}{c} \Sigma \right) P' \right), \tag{22}$$

while the latter employs a rolling window approach to calculate forecast errors and uses the covariance matrix of these errors as $\Omega$. I compare the performance of the two approaches in the empirical study.

Table 1 summarizes the asset allocation models considered above which are used in the empirical study below.

Table 1: Asset allocation models

| # | Name | Description |
|---|---|---|
| 0 | 1/N | Benchmark allocation which uses equal weights for all assets |
| 1 | Mean variance | Sample-based mean-variance allocation |
| 2 | Mean shrinkage | Shrinks the sample means towards a target with shrinkage intensity which minimizes expected quadratic loss |
| 3 | Weights shrinkage | Shrinks the sample mean-variance weights towards a target with shrinkage intensity which maximizes Sharpe ratio in the estimation sample |
| 4 | Portfolio restrictions | Imposes bounds on the sample mean-variance weights |
| 5 | Black-Litterman | Black-Litterman model with sample based views and implied mean vector based on 1/N allocation. Reliability of views is either fixed or varying over time |
| 6 | Bellman | Allocation obtained using dynamic programming |

This table lists the models used for portfolio selection in the empirical study. Models numbered 1-5 can be used with parametric portfolio policies of Brandt and Santa-Clara (2006) while models 0 and 6 represent separate approaches to asset allocation.

#### 3.3.4 Objective functions and parameterizations

Another extension concerns the type of investor's objective function. The optimal mean-variance allocation might not be optimal for investors with constant relative risk aversion utility functions which are empirically found to be appealing and more realistic than mean-variance utility.
The objective as in equation (5) can be easily generalized to other utility functions in the form of:

$$\max_{t} \left( u \left( R_t^f + (\Theta z_t)' \, r_{t+1} \right) \right). \tag{23}$$

In addition, the basic linear specification of the parametric portfolio policy (4) might not be suitable. Nonlinearities can be introduced by including as state variables for instance polynomials of the state variables.
However, the above proposed modifications such as different objective functions or functional forms mean that numerical optimization is required to derive optimal solutions. Since the problem is still inherently static over the augmented set, optimization remains relatively simple and can be done with standard algorithms such as the Newton's method. Alternatively, higher order approximation of utility function based on Taylor series expansion can be considered which avoids the need to use numerical optimization. For example, Brandt et al. (2005) proposes a fourth-order approximation of expected utility around risk-free growth of wealth which accounts for skewness and kurtosis of returns:

$$
\begin{aligned}
E_t[u(W_{t+1})] \;\approx\; & E_t\Big[u(W_tR_t^f) + u'(W_tR_t^f)(W_tx_t'r_{t+1}) + \frac{1}{2}u''(W_tR_t^f)(W_tx_t'r_{t+1})^2 + \\
& \frac{1}{6}u'''(W_tR_t^f)(W_tx_t'r_{t+1})^3 + \frac{1}{24}u''''(W_tR_t^f)(W_tx_t'r_{t+1})^4\Big].
\end{aligned}
\tag{24}
$$

In this case, the first order conditions result only in an implicit solution for the weights:

$$
\begin{aligned}
x_t \;\approx\; & -\Big(E_t[u''(W_tR_t^f)(r_{t+1}r_{t+1})]W_t^2\Big)^{-1} \times \Big(E_t[u'(W_tR_t^f)(r_{t+1})]W_t \\
& + \frac{1}{2}E_t[u'''(W_tR_t^f)(x_t'r_{t+1})^2 r_{t+1}]W_t^3 \;+\; \frac{1}{6}E_t[u''''(W_tR_t^f)(x_t'r_{t+1})^3 r_{t+1}]W_t^4\Big).
\end{aligned}
\tag{25}
$$

The optimal solution can then be obtained by starting with an initial guess for $x_t$ and iterating with the newly obtained solutions until convergence.

#### 3.3.5 Cross-sectional information

To exploit also well-known differences in expected returns in the cross-section of stocks, several options are available. The factor idea which aims to improve and replace the value-weighted stock market index by leveraging the well-known differences in cross-sectional risk premia, can be implemented in several ways. For example, it can be assumed that the three Fama-French factors are available to invest in, which requires no estimation. Alternatively, statistical factors can be considered, which can be obtained by the principal component analysis. However, investing in factors rather than stock market index could result in very high turnover, which is why it is sensible to include transaction costs in the analysis. Brandt et al. (2009) shows that parametric portfolio strategy can easily accommodate high dimensionality when considering a cross-section of stocks in the single period case. As opposed to time-series parametric policy where the variables are common for all assets but coefficients on those variables are different, cross-sectional parametric policy assumes common coefficients on firm-specific variables:

$$
x_{i,t} = \overline{x}_{i,t} + \frac{1}{N_t}\theta' c_{i,t} \tag{26}
$$

where $\overline{x}_{i,t}$ are benchmark portfolio weights, $\theta$ is a vector of common coefficients and $c_{i,t}$ is a vector of characteristics of firm $i$ at time $t$. A candidate set for characteristics includes variables which have been found to explain the cross-section of average stock returns, such as size and book-to-market ratio (Fama and French (1992)).

I propose a way of combining the two dimensions by using the framework of Brandt and Santa-Clara

(2006) and expanding the state variable set to include firm characteristics. In particular, taking the same parametric form as in (4), $z_t$ can be expanded to include all firm specific characteristics. This results in a new asset space which also includes managed portfolios formed on characteristics. Similarly, $\Theta$ can be modified to include parameters on characteristics. These can be assumed to be identical for all firms as in Brandt et al. (2009). This can be illustrated in a simple example with one state variable and one firm characteristic:

$$x_t = \Theta z_t = \begin{bmatrix} \theta_{11} & \theta_{1c} & 0 \\ \theta_{21} & 0 & \theta_{2c} \end{bmatrix} \begin{bmatrix} z_{1,t} \\ c_{1,t} \\ c_{2,t} \end{bmatrix}. \tag{27}$$

The solution that satisfies the restrictions on $\Theta$, $\theta_{1c} = \theta_{2c}$ and zeros can then be obtained by using the restricted least squares (RLS). The calculation of RLS estimates requires the restrictions to be formulated in the matrix form, which is in this example expressed in the following way:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \theta_{11} \\ \theta_{21} \\ \theta_{1c} \\ 0 \\ 0 \\ \theta_{2c} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{28}$$

However, as it can be seen, the dimension of $\Theta$ increases very quickly when new firm characteristics or assets are added which might make this approach infeasible with many stocks and state variables.

### 3.3.6 Performance and transaction cost measures

All of the above methods are implemented and evaluated with real world data in an empirical study. There are numerous performance measures used in the literature to evaluate constructed portfolios. The simplest ones calculate the moments of portfolio returns such as the mean and variance. A well-known Sharpe ratio compares the average portfolio return to the risk-free rate and adjusts it by its risk which is its standard deviation:

$$SR = \frac{E(r) - R_f}{\sigma}. \tag{29}$$

Since the investor is assumed to optimize the trade-off between portfolio mean and variance, Sharpe ratio is the appropriate measure of performance in this case. In addition, Sharpe ratios for different investment horizons can be compared by first calculating returns of portfolios for a common investment horizon. For example, when using monthly returns, annualized Sharpe ratios are calculated as monthly Sharpe ratios multiplied by $\sqrt{12}$.

Performance of different strategies can be also evaluated using the equalization fee which is defined as the percentage fee that an investor would be willing pay to use a superior allocation instead of an inferior one. It is calculated as the difference between certainty equivalents of the two strategies.

The turnover quantifies the average trading from one period to the next required to implement a specific portfolio strategy. With $N$ assets and $T$ time periods where the first $M$ periods are used for estimation, the turnover of strategy $i$ is defined as:

$$Turnover_i = \frac{1}{T-M}\sum_{t=1}^{T-M}\sum_{j=1}^{N}(|x_{i,j,t+1} - x_{i,j,t+}|). \tag{30}$$

where $x_{i,j,t+}$ denotes the portfolio weight $t+1$ after return realizations at time $t$ and before rebalancing at time $t+1$.

Turnover can be used to estimate transaction costs which are directly connected to the trading activity of a portfolio through broker commissions and bid-ask spreads. A popular choice is proportional transaction costs where transaction costs are estimated as a certain percentage of the turnover rate. In the empirical analysis below I assume this to be 0.5%.

# 4 Data

The purpose of the first part of the empirical study is to replicate the key results from Brandt and Santa-Clara (2006). In line with their data, the sample period is from January 1945 to December 2000 with return data at a monthly frequency. CRSP value-weighted stock market index, which is comprised of stocks listed on the NYSE, NYSE MKT, NASDAQ and Arca stock exchanges, is used as a proxy for stocks. The index of long-term US government bonds (US) serves as a proxy for bonds and the index of US Treasury bills as a proxy for cash/risk-free rate. The data for the latter two, known as Ibbotson Stocks, Bonds, Bills, and Inflation (SBBI) data, come from Ibbotson Associates (now part of Morningstar).

The state variables used in the study are dividend yield, default spread, detrended short-term interest rate and term spread. These variables have been found to be important return predictors in previous empirical studies (Ferson and Siegel (2001), Brandt and Santa-Clara (2006)). Dividend yield is defined as the log of the ratio of last 12 months of dividends and the current value of securities in CRSP stock market index. Dividends in each month are obtained by taking the difference between returns including dividends and excluding dividends, and multiplying it by the total value of the securities in the index at the beginning of the month. Detrended short-term interest rate is constructed as the difference between the current Treasury bill rate and 12-month moving average. Interest data in Brandt and Santa-Clara (2006) comes from the DRI/Citibase database which has since become part of IHS Markit, and is thus not freely available anymore. Therefore, this data is obtained from a variety of other sources. Default spread is the difference between yields on Moody's Baa and Aaa rated corporate bonds, which is obtained from Federal Reserve Economic Data (FRED) database. Term spread is the difference between yields on 10-year and 1-year US Treasury bonds. The data from 1953 onward comes from FRED, while the data prior to 1953 comes from the European Central Bank database for 10-year yields and from Capital Markets Data for 1-year yields.

The cross-sectional element of the study requires the use of assets beyond indexes of stocks and bonds, and cash. Using insights from asset pricing theory, the study is extended by using risk factors which are associated with higher expected returns. For this purpose I use Fama and French's market, size and value factor. Since size and value factors are zero-net investment portfolios, 6 portfolios sorted on size and value are instead used in the optimization. This data comes from Kenneth Frenchs' data library. The second set of assets which is applied to the framework of Brandt et al. (2009) is comprised of value-weighted returns on 5 industry portfolios which also comes from

Kenneth Frenchs' data library. The corresponding firm/industry characteristics are the average value-weighted firm size and the book-to-market ratio. These two characteristics are found to be important predictors of stock returns, i.e. they are thought to proxy for undiversifiable risk (e.g. Fama and French, 1992).

# 5 Empirical analysis

## 5.1 Single-period portfolio policies

The first part of the empirical application of parametric portfolio policies is the investment problem of a myopic investor with mean-variance preferences, coefficient of risk aversion $\gamma = 5$, and a monthly or annual holding period. Table 2 shows the in-sample mean-variance optimization results for two risky assets, stock and bond index, and a risk-free asset proxied by Treasury bill index. Conditional portfolios using the formulation in equation (4) are formed using four state variables, dividend yield (D/P), default spread (Default), detrended short-term interest rate (Tbill) and term spread (Term). Unconditional results show that the optimal investment in stocks is 77.2% and 60.7% for monthly and annual investments, respectively. The difference can be explained by the fact that, as opposed to monthly frequency, at annual frequency stocks have a slight negative autocorrelation which increases their volatility, thereby making them less attractive.[1] The investor allocates almost no wealth to long-term bonds which can be explained by the unattractive risk-return profile of long-term bonds. Their annual mean return is around 1% and volatility around 9%. In contrast, stocks have the annual average return of around 8% with still relatively low volatility at 15%. Moreover, positive correlation with stocks at monthly and annual frequency means that holding bonds does not offer significant diversification benefits. Since the weights must sum to 1, the investment in the risk-free rate is 22.8% and 39.3%, respectively.

The second set of results in Table 2 shows the parameter estimates for a linear conditional portfolio policy. All conditioning variables are standardized to have mean zero and standard deviation one which eases the interpretation of coefficient estimates. In particular, while the allocation to stocks and bonds varies each time period depending on state variable realizations, the estimates of a constant give the average allocation. At a monthly frequency bond allocation stays the same while the average stock allocation increases to 84.5%. For an annual holding period, stock allocation slightly decreases while the bond allocation turns to be large and negative at -49.4%. The estimate is, however, not statistically significant as seen from the standard error of 0.408 in parentheses.

[1] It should be noted that negative autocorrelation makes stocks more attractive for long-term investors since the variability of multi-period returns decreases.

This is also the case for coefficients on the state variables which have large standard errors and are thus in general not statistically significant at the 5% level. However, the coefficient estimates tend to have signs which are in line with economic theory and previous empirical studies. For example, Fama and French (1989) find that the equity premium on US stocks has a positive relation with the slope of the yield curve, which is in line with the positive TBill coefficient. However, the results in Table 1 relate state variables to portfolio weights rather than just the expected returns. The offseting effect on the variance could explain why most variables appear to be insignificant predictors. Despite that, the hypothesis that all state variable parameters are zero is rejected with a p-value of 0. Moreover, it seems that conditioning variables add valuable information for portfolio performance which is evident from the Sharpe ratios. The increase from 0.563 to 0.851 for monthly portfolios and from 0.518 to 0.918 for annual portfolios seems rather large. Furthermore, the equalization fee of approximately 4% for both holding periods tells us that investors would be willing to pay 4% to use the conditional rather than the unconditional strategy.

Table 2: Single-Period Portfolio Policies

| | State variable | Monthly | | Annual | |
|---|---|---|---|---|---|
| | | Unconditional | Conditional | Unconditional | Conditional |
| Stock | Const | 0.772 (0.187) | 0.845 (0.192) | 0.607 (0.143) | 0.571 (0.213) |
| | Term | | 0.496 (0.214) | | 0.437 (0.256) |
| | Default | | -0.147 (0.195) | | -0.121 (0.205) |
| | D/P | | 0.331 (0.171) | | -0.146 (0.157) |
| | TBill | | -0.231 (0.213) | | 0.485 (0.435) |
| Bond | Const | -0.003 (0.326) | -0.014 (0.414) | 0.066 (0.280) | -0.494 (0.408) |
| | Term | | 0.553 (0.336) | | 0.133 (0.417) |
| | Default | | 0.215 (0.349) | | 0.106 (0.339) |
| | D/P | | 0.079 (0.457) | | -0.499 (0.345) |
| | TBill | | 0.579 (0.316) | | -0.544 (0.481) |
| p-value | | | 0.000 | | 0.000 |
| Mean excess return | | 0.062 | 0.137 | 0.043 | 0.092 |
| SD return | | 0.110 | 0.161 | 0.083 | 0.101 |
| Sharpe ratio | | 0.563 | 0.851 | 0.518 | 0.918 |
| Equalization fee | | 0.041 | | 0.042 | |

This table shows the coefficient estimates of single-period (monthly and annual) parametric policies with and without conditioning information. Coefficients are reported for standardized state variables. The standard errors which are calculated using equation (9) are reported in parentheses. Below are reported the p-values of conditional policies which correspond to an F-test of the joint significance of conditioning variables. The next three rows report in-sample return statistics which are annualized for monthly data. Lastly, equalization fee is the fee an investor would be willing to pay to use the conditional rather than the unconditional allocation.

## 5.2 Single-period out-of-sample performance

As seen in Table 2 there seems to be an economically significant difference between unconditional and conditional portfolio policies. This result, but found to be even stronger, was also established in the original study by Brandt and Santa-Clara (2006). However, it is well known that many important in-sample results derived from portfolio strategies (e.g. DeMiguel et al. (2009)) do not hold when evaluated out-of-sample due to severe parameter estimation errors. Table 3 presents an out-of-sample analysis of single-period parametric portfolio policies based on a monthly holding period. The out-of-sample results for the annual holding period are reported in the Appendix since the pattern and the quantities are very similar to the monthly results. The table compares the Sharpe ratios of different portfolio policies which are obtained using sample estimators and a number of other estimators and constraints that attempt to address parameter uncertainty. In addition, the results are compared to a naive 1/N allocation which has proven to be a hard to beat allocation in many settings (DeMiguel et al. (2009)).

The first estimation approach, called Mean variance, uses sample moments to form portfolios and thus ignores parameter uncertainty. This approach was also used to obtain in-sample results in Table 2. The Moment shrinkage approach shrinks the mean vector of assets/managed portfolios returns towards a target where shrinkage intensity is calculated by minimizing expected quadratic loss, as explained in Section 3.3.1. As a target mean vector, the first option considered is the grand mean (g-mean) of returns of all basis assets and conditional portfolios. The second option is the states mean (s-mean) where basis asset and conditional portfolios' mean returns are each shrunk towards their own means. Since the conditioning variables are standardized to have zero mean, it follows that returns on conditional portfolios also have zero mean which is thus the shrinkage target. [2] As opposed to mean vector shrinkage, Weights shrinkage shrinks the Mean variance (sample) weights towards some target weights. Here I consider 1/N weights for basis assets and zeros for the weights on conditional portfolios. The latter implies that state variables have no effect on the allocation of stocks and bonds. The next approach Portfolio restrictions imposes hard bounds on the asset/conditional portfolios weights. For basis assets, no short-selling and borrowing constraints are imposed, limiting the weights to be between 0 and 1. For conditional portfolios coefficient bounds of -0.5 and 0.5 are chosen. Since all state variables are standardized, if a normal distribution is assumed, it is very unlikely that any state variable shifts the allocation in stocks or bonds by more than 100%. [3] Finally, the Black-Litterman model is applied where the views are chosen to be the sample means and the reference portfolio is 1/N for basis assets and zeros for the conditional portfolios. Two variations of the model are considered, one where the variability of the views is time-invariant (fixed), and one where it is time-varying (varying).

The estimation is based on the two most common approaches in the literature, rolling (RW) and expanding (EW) estimation windows. In the rolling window case, the first 120 months (10 years)

[2] This justifies choosing separate targets as basis assets (stocks and bonds) have mean returns different from zero.

[3] To impose bounds of 0 and 1 on the final weights in basis assets, the maximum and minimum values of state variables must be taken into account when setting state variable coefficient bounds.

of data are used to estimate portfolio policy parameters which are then used to construct portfolio weights and evaluate portfolio return in the first out-of-sample month. That month is then included in the sample while the first month is dropped from estimation, proceeding until the last period. Expanding window approach proceeds in the same way but without dropping the oldest data point in each step.

Table 3: Sharpe Ratios Single-Period Monthly

| Strategy | Unconditional EW | Unconditional RW | Conditional EW | Conditional RW |
|---|---|---|---|---|
| Mean variance | 0.356 | 0.340 | 0.359 | 0.253 |
| Moment shrinkage (g-mean) | 0.301 | 0.299 | 0.381 | 0.294 |
| Moment shrinkage (s-mean) | / | / | 0.426 | 0.276 |
| Weights shrinkage | 0.425 | 0.448 | 0.474 | 0.419 |
| Portfolio restrictions | 0.389 | 0.483 | 0.578 | 0.475 |
| Black-Litterman (fixed) | 0.416 | 0.442 | 0.471 | 0.419 |
| Black-Litterman (varying) | 0.467 | 0.499 | 0.437 | 0.412 |
| 1/N | 0.405 | 0.405 | 0.405 | 0.405 |

This table shows the out-of-sample Sharpe ratios of the single-period (monthly) parametric portfolio policies using different estimators and restrictions. The Sharpe ratios are annualized. The results are reported for both expanding window and rolling window estimation approaches, both based on 120 months of data. Optimal shrinkage intensities for means are obtained using analytic expression in (16), while those for weights shrinkage are found using grid search. Uncertainty parameters used in the Black-Litterman model are also obtained using grid search.

Focusing first on the unconditional portfolios, I find that the Black-Litterman approach with varying reliability of the views proves to be superior both in the EW and RW case with Sharpe ratios of 0.467 and 0.499, respectively. All other approaches with the exception of Mean variance and Moment shrinkage tend to improve the performance relative to the 1/N allocation which has a Sharpe ratio of 0.405. Surprisingly, the Mean variance (sample) approach still performs relatively well which can be attributed to the fact that only two risky assets are considered. Relative to the in-sample results, there is a noticeable decline of around 0.20 from the in-sample Sharpe ratio of 0.563. However, as shown in the table, much of the out-of-sample performance can be salvaged using a number of different estimation error remedies.

The results also show that in the unconditional case the RW Sharpe ratios are comparable to the approach with EW estimation. The similarity of Sharpe ratios suggests that there is no significant trade-off between capturing the changes in the data generating process achieved by rolling window approach and the reduced parameter uncertainty when using more data. This is however not the case when conditioning information is included in the formation of portfolios. With two risky assets and four state variables the number of first moments increases from 2 to 10 and the number of second moments increases from 4 to 100. This imposes significant parameter uncertainty which is reflected in low Sharpe ratios of the naive mean variance and mean vector shrinkage approaches. Specifically, the Mean variance approach results in a Sharpe ratio as low as 0.253 which implies that including additional information actually hurts the performance. A comparison with the in-sample results in Table 2 reveals a considerable reduction of a rather high Sharpe ratio of 0.851. Furthermore, most approaches do not improve the mid 0.4 Sharpe ratio of the unconditional case, with an exception of EW Portfolio restrictions where the Sharpe ratio is 0.578. From these results it can be concluded that timing the market with economic states information might not be possible, and it might actually hurt the portfolio performance if sufficient amount of data or estimation procedure are not used.

In addition, it is expected that conditional portfolio policies require more drastic portfolio changes since the portfolio weights depend on state variables realizations which are different each period. This results in transaction costs which could severely decrease the returns on conditional portfolios. Table 4 shows the per-period turnover of each portfolio strategy from Table 3. The turnovers are expressed relative to the turnover of the 1/N strategy which only requires rebalancing due to assets' different return realizations.

Table 4: Turnovers Single-Period Monthly

| Strategy | Unconditional EW | Unconditional RW | Conditional EW | Conditional RW |
|---|---|---|---|---|
| Mean variance | 4.967 | 11.958 | 112.014 | 149.523 |
| Moment shrinkage (g-mean) | 5.392 | 11.317 | 159.946 | 194.253 |
| Moment shrinkage (s-mean) | / | / | 131.814 | 187.682 |
| Weights shrinkage | 1.492 | 3.825 | 24.643 | 22.429 |
| Portfolio restrictions | 1.934 | 4.128 | 44.549 | 48.672 |
| Black-Litterman (fixed) | 1.015 | 2.436 | 29.155 | 31.882 |
| Black-Litterman (varying) | 2.638 | 4.866 | 21.867 | 25.317 |
| 1/N | 0.012 | 0.012 | 0.012 | 0.012 |

This table reports the corresponding turnovers of portfolio strategies considered in Table 3. The last row reports the average monthly turnover of the 1/N allocation. All other strategies show the average monthly turnovers relative to the turnover of the 1/N allocation.

Comparing the two estimation alternatives, it is no surprise that expanding window and unconditional portfolio policies result in much lower turnovers. The issue of extreme portfolio weights of Mean variance strategy is well reflected in its turnovers. The conditional RW turnover of 149.523*0.012 implies an average rebalancing of 179% each period. The issue of excessive trading is somewhat improved when other approaches are used, however, it still persists in the conditional case. For example, the lowest turnover associated with the varying Black-Litterman model, results in the average rebalancing of 26% per month.

To provide a rough estimate of portfolio performance net of transaction costs, I use an ad hoc approach which adjusts the mean of asset returns downwards by 0.5% times the average portfolio turnover. For example, if the average monthly turnover was 50%, the monthly mean return would be 0.25% lower. For the 1/N allocation with a small rebalancing of 1.2% per month, the adjusted Sharpe ratio is 0.398. In the unconditional case, the highest adjusted Sharpe ratio of 0.455 is obtained using portfolio restrictions with RW estimation. Hence, with modest rebalancing it is still possible to beat the 1/N allocation. This is no longer the case for conditional portfolio policies which involve much more aggressive rebalancing. The highest adjusted Sharpe ratio still results from portfolio restrictions with EW estimation, however it decreases from 0.578 to 0.394. This is consistent with many studies which show that high levels of trading are strongly negatively correlated with investment performance (e.g. Barber and Odean (2000)). Hence, the conclusion of Brandt and Santa-Clara (2006) that conditional and unconditional return distributions are very different would be very difficult to capitalize on in practice.

### 5.3 Multi-period portfolio policies

For the multi-period problem I consider one year investment horizon with the possibility of monthly portfolio rebalancing. Since there are two assets and twelve timing portfolios for each, this results in 24 parameters to be estimated. For the multi-period conditional case the issue of dimensionality quickly exacerbates with the number of state variables. Hence, due to convenience when reporting the results I use one state variable, resulting in 48 unknown parameters. Second, as explained below, the out-of-sample performance deteriorates with the inclusion of additional variables. The choice of state variable is based on the in-sample explanatory ability of each variable. Detrended short-term interest rate (Tbill) and term spread (Term) result in Sharpe ratios of 0.664 and 0.659, respectively. In contrast, dividend yield (D/P) and default spread (Spread) give Sharpe ratios of 0.589 and 0.581. It is not surprising that the variables which are thought to be stronger bond rather than stock return predictors are more helpful since it is well-known that at short horizons bonds are more predictable than stocks (e.g. Baltussen et al. (2021)). Thus, Tbill is the state variable used to form conditional multi-period portfolios.

Table 5 shows the in-sample estimation results. For convenience only months 1,4,8 and 12 are reported where the values for intermediate months follow the pattern in the reported months. In the unconditional case, the pattern of allocation to stocks is decreasing as the end of the horizon approaches. In particular, the stock holding decreases from 62.7% in the first month to 47.7% in the last month. Conversely, the holding of bonds sharply increases from -70.3% to 47.3%. Relative to the single-period case in Table 2, the change in the stock allocation from 77.2% is not large. However, long-term bonds play a much bigger role in the multi-period portfolio as opposed to the single-period where the allocation was around 0%. This suggests that there are considerable horizon effects when investing dynamically rather than myopically. However, the standard errors show that the weights estimates for bonds are less certain and for some months not statistically significantly different from 0. Turning to the conditional case, it can be noticed that the average allocations to stocks and bonds are almost the same as in the unconditional case. This implies that the horizon effects come from the serial correlation structure of stock and bond returns rather than their relation with the state variable. Nevertheless, there is a small increase in Sharpe ratio from 0.576 to 0.664 when conditioning information is used. The equalization fee is rather small at 0.8%.

Table 5: Multi-Period Portfolio Policies

| Asset | Month | State variable | Monthly Unconditional | | Conditional | |
|---|---|---|---|---|---|---|
| Stock | 1 | Cnst | 0.6269 | (0.180) | 0.592 | (0.190) |
| | | TBill | | | -0.316 | (0.182) |
| | 4 | Cnst | 0.6164 | (0.182) | 0.563 | (0.192) |
| | | Tbill | | | -0.192 | (0.195) |
| | 8 | Cnst | 0.6478 | (0.184) | 0.646 | (0.194) |
| | | TBill | | | -0.114 | (0.196) |
| | 12 | Cnst | 0.4774 | (0.182) | 0.493 | (0.188) |
| | | TBill | | | -0.255 | (0.184) |
| Bond | 1 | Cnst | -0.7028 | (0.315) | -0.717 | (0.328) |
| | | TBill | | | 0.005 | (0.221) |
| | 4 | Cnst | -0.3485 | (0.316) | -0.331 | (0.333) |
| | | TBill | | | 0.016 | (0.228) |
| | 8 | Cnst | 0.0272 | (0.312) | -0.007 | (0.332) |
| | | TBill | | | 0.044 | (0.230) |
| | 12 | Cnst | 0.4725 | (0.309) | 0.529 | (0.327) |
| | | TBill | | | -0.005 | (0.218) |
| p-value | | | 0.000 | | 0.000 | |
| Mean excess return | | | 0.050 | | 0.058 | |
| SD return | | | 0.088 | | 0.088 | |
| Sharpe ratio | | | 0.576 | | 0.664 | |
| Equalization fee | | | 0.008 | | | |

This table shows the coefficient estimates of multi-period parametric policy with a 1-year holding period and monthly rebalancing. For convenience, only the results for months 1,4,8 and 12 are shown. The standard errors which are calculated using equation (9) are reported in parentheses. Below are reported the p-values of conditional policies which correspond to an F-test of the joint significance of conditioning variables. The next three rows report in-sample return statistics which are annualized for monthly data. Lastly, equalization fee is the fee an investor would be willing to pay to use the conditional rather than the unconditional allocation.

As explained in Methodology (section 3.1.2), the parametric portfolio policy of Brandt and Santa-Clara (2006) is an approximation since the timing portfolios disregard compounding of returns. This might not be a significant issue for monthly returns where the magnitudes of returns are smaller. Comparing the results with the theoretically optimal solution obtained by solving the Bellman equation can shed some light on the extent of approximation issue. Equations (30) and (31) report the OLS estimates of a restricted VAR for log returns with Tbill as the state variable. This specification is a popular choice in the asset allocation literature (e.g. Campbell and Viceira

(2002), Brandt and Santa-Clara (2006)).

$$\begin{bmatrix} ln(1+r^s_{t+1}) \\ ln(1+r^b_{t+1}) \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} 0.0059 \\ 0.0007 \\ 0.0007 \end{bmatrix} + \begin{bmatrix} -0.0046 \\ -0.0006 \\ 0.7112 \end{bmatrix} \times z_t + \begin{bmatrix} \epsilon^s_{t+1} \\ \epsilon^b_{t+1} \\ \epsilon^z_{t+1} \end{bmatrix} \tag{31}$$

$$\begin{bmatrix} \epsilon^s_{t+1} \\ \epsilon^b_{t+1} \\ \epsilon^z_{t+1} \end{bmatrix} \sim MVN \begin{bmatrix} 0, \begin{bmatrix} 0.0018 & 0.0003 & -0.0008 \\ 0.0003 & 0.0006 & 0.0003 \\ -0.0008 & 0.0003 & 0.4933 \end{bmatrix} \end{bmatrix} \tag{32}$$

Coefficient estimates are in line with the two studies mentioned above which find a strong persistence in the state variable and a weak relation with the log return processes.

The estimated model is used to simulate sample paths of returns and the state variable. Then using the simulations and regressions, the Bellman equation is solved using the method of Van Binsbergen and Brandt (2007) described in section 3.2. The associated in-sample Sharpe ratio is 0.679 which implies that there is an improvement of 0.015 over the conditional parametric policy. This is roughly in line with the result from Brandt and Santa-Clara (2006), Table I, who examined in a simulation study how accurate are approximation solutions using timing/conditional portfolios. They showed that with the same investment setting and term spread as the state variable, the Sharpe ratio difference between timing/conditional portfolios approximation and exact solution is 0.0064.

### 5.4 Multi-period out-of-sample performance

As in the single-period setting, it is crucial to evaluate out-of-sample performance of the methods to determine their robustness and usefulness in practice. While in the single period case parameterizing the portfolio weights instead of moments of returns as functions of state variables leads to a significant reduction in dimensionality this is not always the case in the multiperiod problems. For example one year ahead portfolio allocation with monthly rebalancing of only two risky assets and one state variable results in 48 parameters to be estimated. If the number of state variables is increased to four as in the single period case, there are 120 unknown parameters. If the rolling window estimation approach was used it could quickly result in a problem where N>T and the associated non-invertibility issues. This would greatly increase the complexity of the problem and statistical estimation error which would make this solution approach inferior to traditional approaches based on the Bellman equation formulation and numerical solutions. In fact, I observe that out-of-sample performance is highest with only one state variable which quickly deteriorates as more variables are added.

Table 6 shows the results of a similar out-of-sample performance as in Table 3 but now considering a multi-period problem with a one year horizon and monthly rebalancing. It is well-known that shrinkage greatly improves and is essential for single-period portfolio optimization, but it is less clear how helpful it is for dynamic strategies, in particular for timing portfolios. As mentioned above, the rolling window approach is less suitable in the dynamic case since it can lead to significant estimation errors. In fact, the Sharpe ratios are often much lower when the RW approach is used, therefore Table 6 reports only the EW approach.

Table 6: Sharpe Ratios Multi-Period

| Strategy | Unconditional EW | Conditional EW |
|---|---|---|
| Mean variance | 0.337 | 0.298 |
| Moment shrinkage (g-mean) | 0.277 | 0.308 |
| Moment shrinkage (s-mean) | / | 0.322 |
| Weights shrinkage | 0.450 | 0.453 |
| Portfolio restrictions | 0.409 | 0.409 |
| Black-Litterman (fixed) | 0.448 | 0.437 |
| Black-Litterman (varying) | 0.471 | 0.429 |
| Bellman equation | / | 0.488 |
| 1/N | 0.405 | 0.405 |

This table shows the out-of-sample Sharpe ratios of the multi-period parametric portfolio policies from Table 5 using different estimators and restrictions. The results are reported for expanding window estimation approach with 120 months of data used for initial estimation. Optimal shrinkage intensities for means are obtained using analytic expression in (16), while those for weights shrinkage are found using grid search. Uncertainty parameters used in the Black-Litterman model are also obtained using grid search.

Similar to the single-period case, there is a significant reduction in the Sharpe ratios of Mean variance approach when evaluated out-of-sample. This is especially true for the conditional case where adding the state variable Tbill leads to a decrease in portfolio performance. In fact, all strategies in the conditional case underperform relative to the unconditional case. Similar to the single-period problem, Weights shrinkage and the Black-Litterman model seem to be the most successful approaches to deal with estimation error. However, unlike in the single-period they are unable to improve upon the unconditional case. This can be in part explained by the fact that in-sample performance was roughly the same with and without conditioning information. However, the Bellman equation solution gives the Sharpe ratio of 0.488 which improves upon the Black-Litterman model's 0.471 in the unconditional case. This suggests that conditional information does add some value when investing dynamically in the short-term. As opposed to the in-sample result when timing/conditional portfolios performed almost as well as the Bellman equation solution, the difference is bigger out-of-sample. A possible explanation for this is that the former requires a much larger number of parameter estimates which leads to more adverse effect of parameter parameter uncertainty.

Table 7: Turnovers Multi-Period

| Strategy | Unconditional EW | Conditional EW |
|---|---|---|
| Mean variance | 17.668 | 26.841 |
| Moment shrinkage (g-mean) | 17.035 | 23.679 |
| Moment shrinkage (s-mean) | / | 21.563 |
| Weights shrinkage | 7.067 | 8.052 |
| Portfolio restrictions | 5.658 | 8.751 |
| Black-Litterman (fixed) | 7.146 | 4.273 |
| Black-Litterman (varying) | 7.921 | 5.198 |
| Bellman equation | / | 18.552 |
| 1/N | 0.012 | 0.012 |

This table reports the corresponding turnovers of portfolio strategies considered in Table 6. The last row reports the average monthly turnover of the 1/N allocation. All other strategies show the average monthly turnovers relative to the 1/N allocation.

Table 7 examines the turnover of the multi-period portfolio policies. It can be noticed that in contrast to the single-period, the unconditional policies now lead to bigger turnovers relative to the 1/N allocation. This can be explained by the dynamic weights on stocks and bonds, which depend on the month in the investment horizon. Unsurprisingly, Mean variance approach which is subject to the largest portfolio weights variance, leads to the biggest turnover. Hence, shrinkage and other remedies for parameter uncertainty, not only improve gross performance but also reduce trading costs. Finally, the conditional information generally increases the turnover but not significantly. This is in line with the finding from Table 5 that including a state variable does not change much the portfolio allocation.

### 5.5 Myopic vs dynamic portfolio allocations

This section answers the question of how big is the loss from investing myopically instead of dynamically. The focus is again on the one year holding period with monthly rebalancing, but now the Sharpe ratio is evaluated based on compounding of monthly returns. This means that the Sharpe ratio for the myopic investor with monthly investment horizon is evaluated using the one year objective with monthly rebalancing. Table 8 compares the annual Sharpe ratio performance evaluated out-of-sample using the expanding window estimation.

Table 8: Myopic vs Dynamic allocation Sharpe Ratios

| Strategy | Uncond. Myopic | Uncond. Dynamic | Cond. Myopic | Cond. Dynamic |
|---|---|---|---|---|
| Mean variance | 0.319 | 0.342 | 0.324 | 0.312 |
| Moment shrinkage (g-mean) | 0.288 | 0.293 | 0.329 | 0.316 |
| Moment shrinkage (s-mean) | / | / | 0.352 | 0.338 |
| Weights shrinkage | 0.401 | 0.455 | 0.468 | 0.458 |
| Portfolio restrictions | 0.369 | 0.396 | 0.489 | 0.427 |
| Black-Litterman (fixed) | 0.402 | 0.447 | 0.462 | 0.426 |
| Black-Litterman (varying) | 0.417 | 0.461 | 0.454 | 0.431 |
| 1/N | 0.390 | 0.390 | 0.390 | 0.390 |

This table reports the corresponding turnovers of portfolio strategies considered in Table 6. The last row reports the average monthly turnover of the 1/N allocation. All other strategies show the average monthly turnovers relative to the 1/N allocation.

Focusing on the first two columns, it is clear that there is a noticeable difference between investing myopically and dynamically. The Sharpe ratio can be improved by as much as 0.054 in the case of weights shrinkage when accounting for the serial correlation structure of returns. In addition, all strategies benefit from dynamic investing. However, surprisingly this is no longer the case when conditioning information is used. For all strategies accounting for horizon effects is detrimental for out-of-sample portfolio performance. The likely explanation is the estimation loss resulting from a significant increase of dimensionality, from 10 parameters in myopic case to 48 parameters in dynamic case. Despite the use of different estimators such as shrinkage, the horizon effects are too small to outweigh the additional estimation error loss.

## 5.6 Timing and cross-sectional information

So far, the conditional portfolio allocations focused on market timing. This means that conditioning information was used to determine the state of the business/market cycle and the response of different assets' returns to the state, as summarized in a matrix of coefficients $\Theta$. This is most appropriate when investing in different asset classes since it is known that risk premia on asset classes change depending on the state of the economy. However, when investing in individual components of a particular asset class (e.g. stocks), individual characteristics, such as firm size, also matter for the distribution of returns. In this section, I consider portfolio selection problem with 5 industry portfolios where one might expect that both, market timing and cross-sectional (i.e. characteristics) information are important for optimal allocations. For convenience, only one state variable and one characteristic is used. The most informative state variable leading to the highest Sharpe ratios is default spread. The Sharpe ratios of other state variables are reported in the Appendix, Table 13. As a portfolio characteristic, the average value-weighted ratio of book value to market value (B/M) is used. [4] Health industry portfolio has the lowest average BM ratio

[4] Using the average firm size results in similar portfolio performance

at 0.31 while Other industries have the highest BM ratio at 1.1. This suggests there is a lot of variation in BM ratio across industry portfolios, which could potentially be exploited in portfolio allocation.

Table 9: Market Timing and Cross-Section

| | State variable/ Characteristic | Market timing | | Cross-section | Market timing & Cross-section |
|---|---|---|---|---|---|
| | | Unconditional | Conditional | | |
| Cnsmr | Const | 0.449 | 0.378 | 0.449 | 0.197 |
| | Default | | 0.450 | | 0.199 |
| | B/M | | | -0.054 | 0.174 |
| Manuf | Const | 1.040 | 1.118 | 1.040 | 0.549 |
| | Default | | 0.223 | | 0.101 |
| | B/M | | | -0.054 | 0.174 |
| HiTech | Const | 0.218 | 0.178 | 0.218 | 0.031 |
| | Default | | 0.469 | | 0.242 |
| | B/M | | | -0.054 | 0.174 |
| Hlth | Const | 0.555 | 0.450 | 0.555 | 0.318 |
| | Default | | -0.694 | | -0.334 |
| | B/M | | | -0.054 | 0.174 |
| Other | Const | -0.789 | -0.704 | -0.789 | -0.539 |
| | Default | | 0.100 | | 0.047 |
| | B/M | | | -0.054 | 0.174 |
| Mean excess return | | 0.152 | 0.236 | 0.183 | 0.273 |
| SD return | | 0.181 | 0.227 | 0.208 | 0.250 |
| Sharpe ratio | | 0.841 | 1.039 | 0.882 | 1.092 |
| Equalization fee | | 0.047 | 0.010 | 0.042 | |

This table shows the coefficient estimates of single-period (monthly) parametric policies using different information sets. The first column is the mean-variance result based only on return data. The next two columns also include state variables and cross-sectional characteristics, respectively, while the last column combines both. Below are three rows reporting in-sample return statistics which are annualized for monthly data. Lastly, equalization fee is the fee an investor would be willing to pay to use both timing and cross-sectional information instead of either of them or none.

Table 9 shows the estimated coefficients of four different single-period portfolio policies which use either market timing information or characteristics, both, or neither. The first two columns show the same two strategies as considered before in Table 2, but now only with state variable Default. The next column shows the cross-sectional strategy that uses only characteristic B/M. As briefly described in section 3.3.4, it assumes equal coefficients for all securities. Here I choose mean-variance weights as the benchmark portfolio weights. The last column shows the approach illustrated in equation (26) in section 3.3.4 which combines both types of information in the framework of Brand and Santa-Clara (2006). However, it requires estimation by restricted least squares rather than OLS, as shown in section 3.3.5, since the coefficients on the characteristics are assumed to be equal with the absence of cross effects. While the interpretation of coefficients would require an understanding of the five industries, it can be noticed that including variable Default does not change average portfolio weights significantly. However, the Sharpe ratio increases by almost 0.20 which suggests variation in Default is important in predicting industry returns. On the other hand, including book-to-market ratio is not as helpful as suggested by Sharpe ratio increase of 0.041. This is also reflected in the final allocation which is uses Default and B/M. The Sharpe ratio increases only marginally relative to the market timing case. The performance is also evaluated using equalization fee with respect to the last strategy. Most importantly, an investor is prepared to pay a fee of 4.7% to use market timing and cross-sectional information as opposed to the simple mean-variance allocation.

### 5.7 Factor timing

The previous section examined how important are state variables compared to cross-sectional characteristics. The results suggest that variables such as default spread are more informative about portfolio weights and consequently portfolio performance than characteristics such as book-to-market ratio. In addition, the Sharpe ratio increases only marginally when a characteristic is added to the model. This could be due to the incorrect linear specification of parametric portfolio policy or it could be that the predictable variation in expected returns is offset by variances and covariances.

Alternatively, average characteristics used on industry portfolios could have weaker predictive power since the firms within portfolios possibly contain large variation in a particular characteristic. This issue could be fixed by considering portfolios which are sorted on characteristics, the so called portfolios sorts (Fama and French, (1993)). I consider 6 portfolios sorted on B/M ratio and size which can replicate the famous value and size factors by taking long and short positions in these portfolios. In addition, a value-weighted stock market index is used as a market factor. Examining their performance answers the question of how much influence variation in risk premia have on investment performance. In particular, performance is evaluated relative to cases where broad asset classes (e.g. stocks and long-term bonds) or industry portfolios are available for investment. In addition, the framework of Brand and Santa-Clara (2006) allows for an easy test of whether factor timing is possible.

Table 10: Single-Period Sharpe Ratios using Factors

| Strategy | Unconditional EW | Unconditional RW | Conditional EW | Conditional RW |
|---|---|---|---|---|
| Mean variance (in sample) | 1.072 | 1.072 | 1.493 | 1.493 |
| Mean variance | 0.859 | 0.819 | 0.862 | 0.798 |
| Moment shrinkage (g-mean) | 0.665 | 0.714 | 0.891 | 0.727 |
| Moment shrinkage (s-mean) | / | / | 0.968 | 0.738 |
| Weights shrinkage | 0.889 | 0.873 | 0.913 | 0.860 |
| Portfolio restrictions | 0.593 | 0.576 | 0.584 | 0.515 |
| Black-Litterman (fixed) | 0.820 | 0.846 | 0.908 | 0.859 |
| Black-Litterman (varying) | 0.906 | 0.851 | 0.871 | 0.881 |
| 1/N | 0.534 | 0.534 | 0.534 | 0.534 |

This table shows the out-of-sample Sharpe ratios of the single-period parametric portfolio policies applied to monthly returns on size/value portfolio sorts and stock market index. It shows the results for different estimators and restrictions using both expanding window and rolling window estimation approaches, both based on 120 months of data. The Sharpe ratios are annualized. Optimal shrinkage intensities for means are obtained using analytic expression in (16), while those for weights shrinkage are found using grid search. Uncertainty parameters used in the Black-Litterman model are also obtained using grid search.

Table 10 reports the in-sample and out-of-sample performance of factor investing. The first thing that can be noticed is that the in-sample Sharpe ratios are much higher compared to when stock and bond index or industry portfolios were available for investment. This is not surprising given the vast literature documenting significant risk premia associated with these factors. However, the conditional parametric policy also suggests that factor returns are heavily influenced by economic state information since the Sharpe ratio increases by more than 0.4 relative to the unconditional case. This in-sample result, however, no longer holds when out-of-sample performance is examined. In general, factor timing does not result in superior investment performance while it is likely to result in considerably larger transaction costs. Nevertheless, out-of-sample performance remains high when compared to the unconditional in-sample result. In particular, Mean variance approach performs surprisingly well which suggests that there is not much estimation error. Using the same dataset of 3 factors but a different sample period, this was also documented in DeMiguel et al. (2009), Table 3. The low performance of 1/N strategy and portfolio restrictions suggests that being able to take short positions, used to form factors, is necessary for good performance with this asset set. This confirms the observation that the performance of different tools for dealing with parameter uncertainty crucially depends on the application of interest.

# 6 Conclusion

The purpose of this thesis was to revisit the parametric portfolio policy approach to portfolio selection by Brandt and Santa-Clara (2006) and provide a thorough out-of-sample evaluation. The basic method was complemented with a number of techniques to alleviate parameter uncertainty issues which have been successful in more traditional approaches to asset allocation. These are moment shrinkage, weights shrinkage, portfolio constraints and the Black-Litterman model. I find that these techniques also improve the sample mean-variance allocations of parametric portfolio policies. However, the addition of conditioning information of parametric portfolio policies does not lead to large performance gains out-of-sample as is documented in-sample. This finding is robust across different asset spaces, investment horizons and state variables. In fact, market timing in multi-period portfolio allocations tends to decrease the performance relative to repeated myopic allocations. This is possibly due to dimensionality issues of dynamic portfolio policies and/or weak predictive power of state variables. Therefore, further research could examine other methods aimed at reducing estimation error or develop new methods which are specifically designed for parametric portfolios. Alternatively, the approach could be applied with the recent findings on return predictability in mind (Farmer et al. (2022)), which suggest that there exist only short-periods of return predictability.

# 7 Bibliography

## References


Aït-sahali, Y., & Brandt, M. W. (2001). Variable selection for portfolio choice. *The Journal of Finance*, 56(4), 1297-1351.

Baltussen, G., Martens, M., Penninga, O. (2021). Predicting Bond Returns: 70 Years of International Evidence. *Financial Analysts Journal*, 77(3), 133-155.

Barber, B. M., Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance*, 55(2), 773-806.

Barberis, N. (2000). Investing for the Long Run when Returns Are Predictable. *The Journal of Finance*, 55(1), 225-264.

Barberis, N. (2000). Investing for the long run when returns are predictable. *The Journal of Finance, 55(1)*, 225-264.

Basak, S., & Chabakauri, G. (2010). Dynamic mean-variance asset allocation. *The Review of Financial Studies*, 23(8), 2970-3016.

Bellman, R. (1958). Dynamic programming and stochastic control processes. *Information and control*, 1(3), 228-239.

Bessler, W., Opfer, H., & Wolff, D. (2017). Multi-asset portfolio optimization and out-of-sample performance: an evaluation of Black–Litterman, mean-variance, and naïve diversification approaches. *The European Journal of Finance*, 23(1), 1-30.

Best, M. J., & Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The review of financial studies*, 4(2), 315-342.

Bodnar, T., Parolya, N., & Schmid, W. (2015). A closed-form solution of the multi-period portfolio choice problem for a quadratic utility function. *Annals of Operations Research*, 229(1), 121-158.

Brandt, M. W., Goyal, A., Santa-Clara, P., & Stroud, J. R. (2005). A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *The Review of Financial Studies*, 18(3), 831-873.

Brandt, M. W. (2010). Portfolio choice problems. In *Handbook of financial econometrics: Tools and techniques* (pp. 269-336). North-Holland.

Brandt, M. W., & Santa-Clara, P. (2006). Dynamic portfolio selection by augmenting the asset space. *The journal of Finance*, 61(5), 2187-2217.

Brandt, M. W., Santa-Clara, P., & Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies*, 22(9), 3411-3447.

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance*, 54(2), 655-671.

Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3), 195-228.

Chellathurai, T., & Draviam, T. (2007). Dynamic portfolio selection with fixed and/or proportional transaction costs using non-singular stochastic optimal control theory. *Journal of Economic Dynamics and Control*, 31(7), 2168-2195.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of finance*, 66(4), 1047-1108.

Cong, F., & Oosterlee, C. W. (2017). Accurate and robust numerical methods for the dynamic portfolio management problem. *Computational Economics*, 49(3), 433-458.

Cooper, M. J., Gulen, H., & Vassalou, M. (2001). Investing in size and book-to-market portfolios using information about the macroeconomy: some new trading rules. *Available at SSRN 265181.*

Cvitanić, J., Goukasian, L., & Zapatero, F. (2003). Monte Carlo computation of optimal portfolios in complete markets. *Journal of Economic Dynamics and Control*, 27(6), 971-986.

DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management science*, 55(5), 798-812.

DeMiguel, V., Martin-Utrera, A., & Nogales, F. J. (2013). Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking  Finance*, 37(8), 3018-3034.

de Roon, F., Guo, J., & ter Horst, J. (2010). Being locked up hurts. In *EFA 2009 Bergen Meetings Paper*.

Denault, M., & Simonato, J. G. (2017). Dynamic portfolio choices by simulation-and-regression: Revisiting the issue of value function vs portfolio weight recursions. *Computers  Operations Research*,79, 174-189.

Diris, B., Palm, F., & Schotman, P. (2015). Long-term strategic asset allocation: an out-of-sample evaluation. *Management Science*, 61(9), 2185-2202.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987-1007.

Fan, J., Zhang, J., & Yu, K. (2012). Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association*, 107(498), 592-606.

Fama, E. F., French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427-465.

Fama, E. F., French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56.

Farmer, L., Schmidt, L., & Timmermann, A. (2022). Pockets of predictability. *Journal of Finance*, forthcoming.

Ferson, W. E., & Siegel, A. F. (2001). The efficient use of conditioning information in portfolios. *The Journal of Finance*, 56(3), 967-982.

Hens, T., & Wöhrmann, P. (2007). Strategic asset allocation and market timing: a reinforcement learning approach. *Computational Economics*, 29(3), 369-381.

Kolm, P. N., Tütüncü, R., Fabozzi, F. J. (2014). 60 Years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2), 356-371.

Larsen, L. S., & Munk, C. (2012). The costs of suboptimal dynamic asset allocation: General results and applications to interest rate risk, stock volatility risk, and growth/value tilts. *Journal of Economic Dynamics and Control*, 36(2), 266-293.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4), 1651-1683.

Jobson, J. D., & Korkie, B. (1980). Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association*, 75(371), 544-554.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative analysis*, 21(3), 279-292.

Idzorek, T. (2007). A step-by-step guide to the Black-Litterman model: Incorporating user-specified confidence levels. *In Forecasting expected returns in the financial markets*, 17-38.

Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5), 603-621.

Lim, A. E., & Zhou, X. Y. (2002). Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, 27(1), 101-120.

Markowitz, H.M. (1952). Portfolio selection. *Journal of Finance*, 7, 77–91.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, 247-257.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867-887. Meucci, A. (2010). The Black-Litterman approach: Original model and extensions. *The Encyclopedia of Quantitative Finance*, Wiley.

Pástor, Ľ. (2000). Portfolio selection and asset pricing models. *The Journal of Finance*, 55(1), 179-223.

Plazzi, A., Torous, W., & Valkanov, R. (2010). Expected returns and expected growth in rents of commercial real estate. *The Review of Financial Studies*, 23(9), 3469-3519.

Samuelson, P. (1969). Lifetime Portfolio Selection By Dynamic Stochastic Programming. *The Review of Economics and Statistics*, 51(3), 239-246.

Van Binsbergen, J. H., & Brandt, M. W. (2007). Solving dynamic portfolio choice problems by recursing on optimized portfolio weights or on the value function?. *Computational Economics*, 29(3), 355-367.

Welch, I., Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4), 1455-1508.

# 8 Appendix

Table 11: Sharpe Ratios Single-Period Annual

| Strategy | Unconditional EW | Unconditional RW | Conditional EW | Conditional RW |
|---|---|---|---|---|
| Mean variance | 0.322 | 0.205 | 0.285 | 0.418 |
| Moment shrinkage (g-mean) | 0.213 | 0.118 | 0.245 | 0.017 |
| Moment shrinkage (s-mean) | / | / | 0.211 | 0.049 |
| Weights shrinkage | 0.414 | 0.381 | 0.429 | 0.538 |
| Portfolio restrictions | 0.379 | 0.292 | 0.566 | 0.521 |
| Black-Litterman (fixed) | 0.409 | 0.374 | 0.428 | 0.448 |
| Black-Litterman (varying) | 0.411 | 0.382 | 0.395 | 0.405 |
| 1/N | 0.399 | 0.399 | 0.399 | 0.399 |

This table shows the out-of-sample Sharpe ratios of the single-period parametric portfolio policies using different estimators and restrictions. The results are reported for both expanding window and rolling window estimation approaches, both based on 10 years of data. Optimal shrinkage intensities for means are obtained using analytic expression in (16), while those for weights shrinkage are found using grid search. Uncertainty parameters used in the Black-Litterman model are also obtained using grid search.

Table 12: Turnovers Single-Period Annual

| Strategy | Unconditional EW | Unconditional RW | Conditional EW | Conditional RW |
|---|---|---|---|---|
| Mean variance | 11.579 | 23.459 | 259.53 | 582.859 |
| Moment shrinkage (g-mean) | 15.816 | 24.109 | 363.035 | 1789.113 |
| Moment shrinkage (s-mean) | / | / | 351.238 | 1121.810 |
| Weights shrinkage | 3.474 | 7.507 | 57.097 | 104.915 |
| Portfolio restrictions | 5.581 | 16.058 | 112.126 | 124.614 |
| Black-Litterman (fixed) | 4.946 | 3.511 | 205.661 | 40.036 |
| Black-Litterman (varying) | 4.165 | 2.847 | 196.97 | 38.991 |
| 1/N | 0.083 | 0.083 | 0.083 | 0.083 |

This table reports the corresponding turnovers of portfolio strategies considered in Table 10. The last row reports the average annual turnover of the 1/N allocation. All other strategies show the average annual turnovers relative to the average monthly turnover of 1/N allocation from Table 3.

Table 13: Sharpe Ratios using different state variables

| | Market timing | Market timing & Cross-section |
|---|---|---|
| Term | 0.995 | 1.016 |
| Default | 1.039 | 1.092 |
| D/P | 1.031 | 1.090 |
| TBill | 0.972 | 0.988 |

This table reports the annualized Sharpe ratios of different conditional portfolio policies using 5 industry portfolios considered in section 5.6. Market timing strategies show results for 4 different state variables considered separately while Market timing&Cross-section strategy also includes B/M as a characteristic.