

De Franco, Carmine; Nicolle, Johann; Pham, Huyên

Article

Dealing with drift uncertainty: A Bayesian learning approach

Risks

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: De Franco, Carmine; Nicolle, Johann; Pham, Huyên (2019) : Dealing with drift uncertainty: A Bayesian learning approach, *Risks*, ISSN 2227-9091, MDPI, Basel, Vol. 7, Iss. 1, pp. 1-18,
<https://doi.org/10.3390/risks7010005>

This Version is available at:

<https://hdl.handle.net/10419/257843>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Article

Dealing with Drift Uncertainty: A Bayesian Learning Approach

Carmine De Franco ¹, Johann Nicolle ^{1,2} and Huy en Pham ^{2,*}

¹ OSSIAM, 75017 Paris, France; carmine.de-franco@ossiam.com

² LPSM, Universit  Paris Diderot, 75013 Paris, France; johann.nicolle@ossiam.com

* Correspondence: pham@math.univ-paris-diderot.fr; Tel.: +33-157-279-106

Received: 20 November 2018; Accepted: 4 January 2019; Published: 9 January 2019



Abstract: One of the main challenges investors have to face is model uncertainty. Typically, the dynamic of the assets is modeled using two parameters: the drift vector and the covariance matrix, which are both uncertain. Since the variance/covariance parameter is assumed to be estimated with a certain level of confidence, we focus on drift uncertainty in this paper. Building on filtering techniques and learning methods, we use a Bayesian learning approach to solve the Markowitz problem and provide a simple and practical procedure to implement optimal strategy. To illustrate the value added of using the optimal Bayesian learning strategy, we compare it with an optimal nonlearning strategy that keeps the drift constant at all times. In order to emphasize the prevalence of the Bayesian learning strategy above the nonlearning one in different situations, we experiment three different investment universes: indices of various asset classes, currencies and smart beta strategies.

Keywords: Bayesian learning; Markowitz problem; optimal portfolio; portfolio selection

1. Introduction

The seminal work by [Markowitz \(1952\)](#) initiated modern portfolio theory and provided a solution to the portfolio selection problem according to a mean-variance criterion. The mean-variance optimal portfolio is the one that maximizes its expected return at a given level of risk, measured by its variance, or conversely the one that minimizes its risk at a given level of expected return. This one period model has known multiple extensions, among them the discrete time multi-period model by [Samuelson \(1969\)](#) and the continuous time model by [Merton \(1969\)](#).

Initially, the parameters of these models have been considered known and constant, especially the parameters that drive the behavior of the assets: drifts and covariances. It not only oversimplifies the reality but also it raises the question of estimating these parameters. The most basic and widely used method consists of estimating drifts and covariances from past data and fix them once and for all. Estimating volatility in this way appears to give relatively good results in practice, while estimating the drift seems to be more difficult or even impossible, see [Merton \(1980\)](#). Moreover, optimal portfolios are very sensitive to the level of expected returns, as shown in [Best and Grauer \(1991\)](#), and a wrong estimation can result in very suboptimal portfolios a posteriori. See also [Elliott et al. \(1998\)](#) and [Siegel and Woodgate \(2007\)](#) for more details about these issues.

This explains the development of the literature on parameters uncertainty in portfolio analysis, see for instance [Barry \(1974\)](#) and [Klein and Bawa \(1976\)](#), and especially on Bayesian statistics, see ([Aguilar and West 2000](#); [Avramov and Zhou 2010](#); [Bodnar et al. 2017](#); [Frost and Savarino 1986](#)). Nonetheless, these models remain static and cannot benefit from the flow of information which results in a nonadaptive strategy, unable to process the most recent information conveyed by the assets market prices. It is one of the main reasons why literature on filtering and learning techniques in a partial information framework has developed, see ([Cvitani  et al. \(2006\)](#); [Rogers \(2001\)](#); [Lakner 1995, 1998](#)).

The Bayesian learning approach consists of modeling the uncertainty of a set of parameters by a prior distribution, representing the beliefs of the investor on the potential values of the parameters, which is updated with incoming information, for instance assets market prices. In particular, in our companion paper [De Franco et al. \(2018\)](#), we have solved the multidimensional Markowitz problem in the case of an uncertain drift using Bayesian learning with a Gaussian prior and have provided a brief analysis of sensitivities of the optimal solution w.r.t. the different parameters.

In this paper, we adapt the results in [De Franco et al. \(2018\)](#) to implement the strategy in practice. Indeed, the solution provided is in continuous time and the amounts invested in the assets are unconstrained. Of course, in reality, trading is discrete and amounts invested in assets are limited, so we discretize the optimal solution of the continuous Markowitz problem, turn amounts into proportions and cap the optimal proportions to be invested in order to fit the portfolio management constraints. Our purpose is to show the prevalence of the Bayesian learning strategy above the nonlearning one which considers the drift constant. To do so, we illustrate our point by confronting both strategies to different datasets representing different investment universes. First, we use a panel of major indices in four different asset classes: corporate bonds, sovereign bonds, commodities and equities completed with cash. Then, we consider bank accounts in foreign banks that pay the local interest rate but are valued in EUR, in order to study the performance of both strategies with respect to foreign exchange rates. Finally, we implement both strategies in an investment universe composed of smart beta strategies. Moreover, using the first dataset, we provide a sensitivity analysis of both strategies to various parameters: the uncertainty in the model, the impact of the leverage, the review frequency and the rebalancing frequency. We do not show this analysis for the two other datasets since we would find similar results.

The paper is organized as follows. Section 2 details the model and the discretized optimal strategies while Section 3 depicts the market data and the workflow. Section 4 shows the results in the case of the first dataset, and Section 5 is about the sensitivity analysis of the Bayesian learning and the nonlearning strategies applied to this dataset. Finally, Section 6 deals with foreign exchange rates and Section 7 with smart beta strategies.

2. The Framework

We consider a financial market consisting of one risk-free asset, whose return is denoted by r^f , and n risky assets whose returns r_t are modeled by

$$\begin{cases} r_t &= \mathbf{B} + \sigma \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_t &\sim N(0, I_n). \end{cases} \quad (1)$$

We shall assume that the random vectors $\boldsymbol{\xi}_t$ are independent for all t . Model (1) includes major linear models available in the financial literature, such as CAPM ([Lintner 1965](#); [Sharpe 1964](#)), discrete-time Black-Scholes ([Black and Scholes 1973](#)) or Fama-French models ([Fama and French 1993, 2015, 2016](#)). $\mathbf{B} \in \mathbb{R}^n$ is the vector of the expected returns of the risky assets, while $\boldsymbol{\Sigma} := \sigma \sigma' \in \mathbb{R}^{n \times n}$ is the covariance matrix of the risky assets. We assume that σ^{-1} exists. We denote by \mathcal{A} the set of admissible investment strategies. An admissible strategy $\mathbf{w} = (\mathbf{w}_t)_t \in \mathcal{A}$ represents the fraction of wealth invested in the assets at any time t . Recalling from the self-financing condition that,

$$X_{s+1} = X_s \left(1 + \mathbf{w}'_s \mathbf{r}_{s+1} + (1 - \mathbf{w}'_s \mathbf{1}) r^f \right), \quad s = 0, \dots, T-1,$$

we write the wealth at maturity $T > 0$ as

$$X_T^{\mathbf{w}} = X_0 \prod_{s=0}^{T-1} \left(1 + \mathbf{w}'_s \mathbf{r}_{s+1} + (1 - \mathbf{w}'_s \mathbf{1}) r^f \right).$$

We consider an investor who is aiming to solve the Markowitz problem:

$$\begin{cases} \max_{\mathbf{w} \in \mathcal{A}} \mathbb{E} [X_T^{\mathbf{w}}] \\ \text{Var} (X_T^{\mathbf{w}}) \leq \vartheta, \end{cases} \tag{2}$$

where $\vartheta > 0$ is the risk tolerance for the investor.

The initial version of the Markowitz problem (Markowitz 1952), which was stated for a single period, has been widely studied and solutions in the multi-period framework (such ours) in both discrete and continuous time have been provided (see e.g., Karatzas et al. (1987); Merton 1969, 1975; Samuelson (1969) among others). The common assumptions in previous works are that both expected return and volatility coefficients (\mathbf{B} and $\mathbf{\Sigma}$ in our framework) are known. In practice, these parameters are not directly observable and must be estimated from the data or input by the investor at inception. In both cases, biased parameters can significantly affect ex-post performance of the optimal strategy. Although the parameter $\mathbf{\Sigma}$ can be estimated from the data with some degree of confidence, the estimation of \mathbf{B} turns out to be quite difficult, if not impossible. Because the optimal strategy strongly depends on both \mathbf{B} and $\mathbf{\Sigma}$, a wrong estimation could significantly affect the optimal strategy.

To get closer to reality and account for model uncertainty, we assume reasonably that the investor has an a priori view on the risky assets and their expected returns, but she is uncertain about how good her forecast is. Introducing uncertainty into the problem brings it closer to the real-life situation, where not only does the investor not know the parameters of the model, but is also forced to admit that her estimates are uncertain. More precisely, the investor does not observe \mathbf{B} and only assumes that $\mathbf{B} \sim \mu$, where μ is a probability distribution in \mathbb{R}^n centered at \mathbf{b}_0 ($\mathbb{E} [\mathbf{B}] = \mathbf{b}_0$). The parameter \mathbf{b}_0 is the vector of returns the investor is expecting, while μ translates her uncertainty about it.

Remark 1. When $\mu = \delta_{\mathbf{b}_0}$ —the Dirac distribution at \mathbf{b}_0 —the investor has no uncertainty about her forecast.

Remark 2 (Discretization). *Implementing optimal strategies in practice leads to continuous adjustments to the optimal solutions in time and makes them suboptimal. Here, we discretize the continuous optimal solutions and controls, and cap them. To propose a solution to Problem (2), we suggest confronting the discretized continuous optimal Bayesian learning and nonlearning solutions since the investor can observe assets prices nearly continuously but trades in discrete time.*

The paper De Franco et al. (2018) solved the continuous-time version of problem (1) for a large class of distributions μ with $\mathbb{E}[|B|^2] < \infty$ in a Bayesian framework using dynamic programming techniques. We report here the discretized version of the results in the Gaussian case $\mu = N(\mathbf{b}_0, \mathbf{\Sigma}_0)$: Let $\mu = N(\mathbf{b}_0, \mathbf{\Sigma}_0)$ with $\mathbf{\Sigma}_0 > 0$. Then, the discretized version of the continuous-time optimal solution of the Bayesian–Markowitz problem is given by (the Bayesian learning case, BL)

$$\mathbf{w}_s^{BL} = \alpha (s, X_s, \hat{\mathbf{B}}_s) / X_s \tag{3}$$

where X_s is the wealth process at time s and

- $\hat{\mathbf{B}}_s = (\mathbf{\Sigma}_0^{-1} + s\mathbf{\Sigma}^{-1}) (\mathbf{\Sigma}_0^{-1}\mathbf{b}_0 + (\sigma')^{-1} \mathbf{Y}_s)$,
- \mathbf{Y}_s is defined as $[\sigma \mathbf{Y}_s]^i = \prod_{u=0}^{s-1} (1 + r_u^i) + \frac{s}{2} \mathbf{\Sigma}_{ii}$,
- $\alpha (s, x, \mathbf{b}) = \left(x_0 - x + \sqrt{\vartheta} \frac{e^{R(0, \mathbf{b}_0)}}{\sqrt{e^{R(0, \mathbf{b}_0)} - 1}} \right) (\mathbf{\Sigma}^{-1} \mathbf{b} - (\boldsymbol{\psi}(s) \sigma^{-1})' \nabla_{\mathbf{b}} R(s, \mathbf{b}))$,
- $\boldsymbol{\psi}(s) = \mathbf{\Sigma}_0 (\mathbf{\Sigma} + s\mathbf{\Sigma}_0)^{-1} \sigma$,
- R is the unique solution to the following semi-linear parabolic PDE

$$\begin{cases} 0 = -\partial_t R - \frac{1}{2} \text{tr}(\boldsymbol{\psi} \boldsymbol{\psi}' \mathcal{D}_b^2 R) + 2(\boldsymbol{\psi} \boldsymbol{\sigma}^{-1} b)' \nabla_b R - \frac{1}{2} |\boldsymbol{\psi}' \nabla_b R|^2 - |\boldsymbol{\sigma}^{-1} b|^2, \\ 0 = R(T, b), \end{cases}$$

which in the case of a Gaussian prior proves to be of the form $R(s, \mathbf{b}) = \mathbf{b}' \mathbf{M}(s) \mathbf{b} + r(s)$, where \mathbf{M} is the solution to a multidimensional Riccati equation, for which the solution can be found explicitly, and r is the solution to a first-order linear differential equation depending on $\boldsymbol{\psi}$ and \mathbf{M} , which can also be explicitly calculated. See De Franco et al. (2018) for further details.

The process $\hat{\mathbf{B}}$ is the conditional expectation of \mathbf{B} given the current observation of the assets returns, which are given by \mathbf{Y} . The matrix $\boldsymbol{\Sigma}_0$ represents the uncertainty around \mathbf{b}_0 . As time goes by, we observe more returns which in turns improves our knowledge of \mathbf{B} , as one can see from the fact that when $s \rightarrow T$, more weight is put on $\boldsymbol{\Sigma}^{-1}$ in the definition of $\hat{\mathbf{B}}$. The matrix valued function $\boldsymbol{\psi}$ is linked to the conditional covariance of \mathbf{B} given \mathbf{Y} .

Remark 3. When there is no uncertainty ($\mu = \delta_{b_0}$), or when the investor directly inputs her estimate \mathbf{b}_0 , the structure of the discretized version of the optimal continuous solution is simplified as follows (the nonlearning case, NL):

$$\mathbf{w}_s^{NL} = \alpha^{NL}(s, X_s) / X_s \tag{4}$$

where

- $\alpha^{NL}(s, x) = \left(x_0 - x + \sqrt{\vartheta} \frac{e^{R(0, b_0)}}{\sqrt{e^{R(0, b_0)} - 1}} \right) \boldsymbol{\Sigma}^{-1} \mathbf{b}_0,$
- $R(s, b) = \mathbf{b}' \boldsymbol{\Sigma}^{-1} \mathbf{b} (T - s).$

The main differences between the Bayesian learning and the nonlearning strategies arise from

- the market risk premium R in the leverage coefficient,
- the correction term $(\boldsymbol{\psi}(s) \boldsymbol{\sigma}^{-1})' \nabla_b R(s, \mathbf{b})$ which is zero for the nonlearning strategy.

As time goes by, we observe realized returns and we learn more about \mathbf{B} . Indeed, with uncertainty, the Bayesian learning strategy is updated with $\hat{\mathbf{B}}$, which is the conditional expectation of \mathbf{B} given the current observation (new knowledge). The trade $\boldsymbol{\Sigma}^{-1} \hat{\mathbf{B}}$ is modified with the corrective term $(\boldsymbol{\psi}(s) \boldsymbol{\sigma}^{-1})' \nabla_b R(s, \hat{\mathbf{B}})$.

Both \mathbf{w}^{BL} and \mathbf{w}^{NL} are unconstrained. In order to provide a realistic analysis on both strategies, we consider capped versions of them.

Definition 1. The capping operator for an investment strategy \mathbf{w} with a leverage $l > 0$ is

$$c(\mathbf{w}, l) := \mathbf{w} 1_{|\mathbf{w}|_1 \leq l} + \frac{l}{|\mathbf{w}|_1} \mathbf{w} 1_{|\mathbf{w}|_1 > l},$$

where $|\mathbf{w}|_1 = \sum_{i=1}^n |\mathbf{w}_i|$.

In Section 3, we will implement both the Bayesian learning and the nonlearning strategies in the context of asset allocation with real market data, to get insights on the effect of learning and its value added (value of information).

3. Market Data

We consider four asset classes across different regions, each of which is represented by a well-known market index detailed in Table 1, and the EONIA rate as the risk-free rate for an investor whose base currency is the Euro.

Table 1. Major market indices relative to each currency and asset class.

Bloomberg Ticker	Name	Currency	Asset Class
SPDYCITR Index	S&P GSCI Dynamic Roll TR	USD	Commodity
GOLDLNPM Index	LBMA Gold Price PM	USD	Commodity
LECPTR EU Index	Bloomberg Barclays Euro Aggregate Corporate TR Index	EUR	Corporate Bonds
IBOXHY Index	iBoxx USD Liquid High Yield Index	USD	Corporate Bonds
IBOXIG Index	iBoxx USD Liquid Investment Grade Index	USD	Corporate Bonds
IBOXXMJA Index	EUR Corporate Liquid High Yield	EUR	Corporate Bonds
SPTR500N Index	S&P 500 Net TR Index	USD	Equity
SX5T Index	Eurostoxx 50 Net TR Index	EUR	Equity
TUKXG Index	FTSE 100 TR Index	GBP	Equity
NDUEEGF Index	MSCI Emerging Net TR Index	USD	Equity
SPTPXN Index	S&P Topix 150 NR	JPY	Equity
LUATTRUU Index	Bloomberg Barclays US Treasury TR Unhedged Index	USD	Sovereign Bonds
LEATTREU Index	Bloomberg Barclays EurAgg Treasury TR Unhedged Index	EUR	Sovereign Bonds
JPEICORE Index	JP Morgan EMBI Global Core Index	USD	Sovereign Bonds
FTFIBGT Index	FTSE Actuaries UK Conv. GILTs All Stocks TR Index	GBP	Sovereign Bonds
EONIA Index	EMMI EURO Overnight Index average	EUR	Cash

Non-Euro-denominated indices are hedged against the Euro simply by implementing a monthly rolled hedging overlay with one-month forward contracts. We collected data from December 1998, except for the EUR Corporate Liquid High Yield for which we only obtained data from January 2006. Prices were sourced from Bloomberg, while currency spot and forward rates came from Datastream and both refer to the 4 p.m. London fixing. Our choice of market indices is motivated by their popularity among investors, their liquidity, and the wide range of financial products available in the market that give exposures to these indices (such as listed futures and ETFs). While limited in number, they provide well-diversified exposures to major global asset classes. Therefore, this suits the underlying premises of the Markowitz problem which is less suited for asset allocation in presence of many underlyings.

The testing period is from January 2000 to June 2018. To implement both *BL* and *NL*, we iteratively followed the workflow outlined below and Table 2 collects the parameters (in bold in the workflow) used for our test.

Work Flow 1.

- Let $T > 0$ and consider the time-frame $\{t_0 < t_1 < \dots < t_n = T\}$.
- We call t_0 a Review date since at this date we estimate all the parameters and calculate the function R that define \mathbf{w}^{BL} in (3) and \mathbf{w}^{NL} in (4). To ensure a realistic implementation of both strategies, all data-based estimations are performed with data available before the Review date t_0 and ending on $t_0 - \mathbf{Lag}$. The first Review date is 21 January 2000.
- We call t_k a Rebalancing date if, at this date, we updated the portfolio weights according to (3) and (4). To limit turnover and transaction costs, the subsets of Rebalancing dates in $[0, T]$ is relatively small. We assume that both Review and Rebalancing dates are Fridays, so that the number of Rebalancing dates in the $[0, T]$ period is given by the frequency **Freq**.
- \mathbf{b}_0 is estimated as the sample mean over the past r^w days ending on $t_0 - \mathbf{Lag}$ or over the maximum data available with a minimum of 30 data points. Σ is estimated as the sample covariance matrix over s^w days ending on $t_0 - \mathbf{Lag}$ or over the maximum data available with a minimum of 30 data points.
- We consider a parametric function for ϑ as follows:

$$\vartheta = \vartheta(\mathbf{d}) := T (\mathbf{d} \times X_{t_0})^2 / (1.96^2), \quad \mathbf{d} \in (0, 1).$$

The motivation behind our choice comes from Problem (2). Indeed, we expect that

$$X_T \succeq \mathbb{E}[X_T] - 1.96 \times \sqrt{\frac{\text{Var}(X_T)}{T}} \geq \mathbb{E}[X_T] - 1.96 \times \sqrt{\frac{\vartheta}{T}},$$

and we set ϑ so that within our confidence interval, the gap to the expected value of X_T is a fraction of the initial wealth

$$1.96 \times \sqrt{\frac{\vartheta}{T}} \sim \mathbf{d} \times X_{t_0}.$$

- The uncertainty around the estimate \mathbf{b}_0 is measured by Σ_0 . We assume Σ_0 to be diagonal. Therefore, each diagonal entry measures the degree of confidence we have on the relative entry of \mathbf{b}_0 . Each diagonal entry is modeled as follows:

$$\Sigma_0^{ii} = \Sigma_0^{ii}(\mathbf{unc}) := \mathbf{unc} \times \left(\frac{pct_{95}^i - pct_5^i}{2} \right)^2,$$

where pct_{95}^i (pct_5^i) is the 95% (5%) quantile of the empirical distribution of the return r^i . To calculate these quantiles, we consider the time series of the returns r^i spanning for r^q days and ending on $t_0 - \mathbf{Lag}$. Up to the parameter \mathbf{unc} , the square root of Σ_0^{ii} is half of the segment, centered at the median, that contains 90% of the empirical distribution over r^q days.

- Strategies \mathbf{w}^{BL} and \mathbf{w}^{NL} are considered in their capped versions according to Definition 1 with a maximum leverage \mathbf{L} .
- At time T , we rerun the workflow by setting $t_0 = T$ and update all parameters according to the new data available.

It should be noted that the choice of diagonal Σ_0 does not imply that assets are uncorrelated. The matrix Σ_0 represents the uncertainty the investor has around her initial prior \mathbf{b}_0 . While assets are correlated, we assume that the investor makes her initial expectations on \mathbf{B} and she only includes her anticipated error due to uncertainty on her expectations. To simplify, we do not include off-diagonal terms (uncertainty on pairwise cross-expectations), because this allows for a simple, one-parameter modeling of uncertainty in Σ_0 .

Our choice of Σ_0^{ii} reflects the assumption that investors anticipate the assets returns to be equal to $\mathbf{b}_0 + \text{error}$, and the error, which is unknown and represents their uncertainty around the expectation \mathbf{b}_0 , depends on the width of the empirical distribution of returns. In practice, if we decide to proxy expected returns with the average of past returns (which of course is known to be a very poor indicator of future returns), then the width of the empirical distribution, here measured by the difference between the 95% and 5% quantiles, serves as a proxy for uncertainty. The fine-tuning parameter \mathbf{unc} is a simple way to increase the uncertainty in the initial guess \mathbf{b}_0 .

Table 2. Parameters used in the implementation of *BL* and *NL* as defined in the workflow.

Parameter	Value
X_0	100
T	3 months
Lag	1 day
$Freq$	Monthly, 3rd Friday of each month.
r^w	750 days
s^w	125 days
d	10%
\mathbf{unc}	100
r^q	125
L	200%

The workflow above and Table 2 represent the *Base Case*. In Section 4, we report the results for both the Bayesian learning (*BL*) and the nonlearning (*NL*) strategies over the last 18 years. In Section 5, we will assess the sensitivity of both strategies to changes in the parameters of the *Base Case* to get a better grasp of the value added of learning in the context of portfolio construction.

4. The Base Case Result

We implemented the workflow to build the wealth processes X^{BL} and X^{NL} as in (2) related to the *BL* and *NL* strategies in the Base Case framework (Table 2). For the sake of simplicity, we refer to *BL* as either the strategy weights w^{BL} or the associated wealth process X^{BL} and we do the same for *NL*. Table 3 collects some long term statistics for both strategies. Although our choice of statistics is limited and many other interesting properties could have been easily derived, such as skew, kurtosis, VaR, CVaR or turnover (as e.g., in Scaillet (2004)), we prefer to show the statistics that are usually considered at any initial due diligence for investment strategies, such as annualized performance, volatility, maximum drawdown, Sharpe ratios and information ratios. The main reason behind our choice is to provide evidence on the effect of learning in the improvement of long-term performance.

Table 3. Statistics for the *BL* and *NL* strategies. The Sharpe Ratio is calculated as the ratio between annualized performance and annualized volatility. In this case, the difference is given in relative terms. Data from January 2000 to June 2018.

Statistics	<i>BL</i>	<i>NL</i>	Difference
Ann. Performance	5.96%	3.96%	2.00%
Ann. Volatility	5.51%	4.92%	0.59%
Max. Drawdown	−8.51%	−11.2%	2.70%
Sharpe ratio	1.08	0.80	34.26%
Information ratio	0.55	-	-

Over the period from January 2000 to June 2018, *BL* delivered an annualized performance of 5.96% while *NL* reached 3.96%. Incorporating uncertainty and learning from the data yielded an annualized 2% excess return. In terms of risk metrics, annualized volatility is slightly higher for *BL* (0.59% difference) but also shows a better maximum drawdown figure (−8.51% for *BL* versus −11.20% for *NL*). Finally, the Sharpe ratio is 1.08 for *BL* versus 0.80 for *NL*, or a 34.26% improvement in relative terms. The information ratio of *BL* over *NL* is about 0.55, meaning that the performance of *BL* comes from a better ability to process incoming information.

Figure 1a shows the historical levels of the wealth processes *BL* and *NL*, while Figure 1b provides the relative strength of *BL* over *NL* as well as the growth line. An increasing relative strength index signals outperformance of *BL* over *NL*, while a decreasing index shows underperformance.

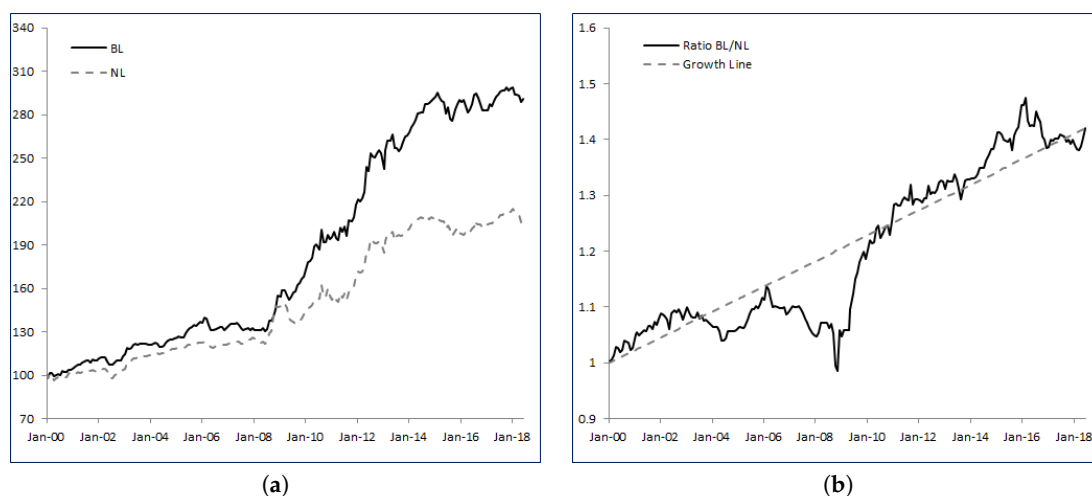


Figure 1. Historical values of the portfolios calculated with both w^{BL} and w^{NL} in the Base Case (a) and their ratio (b).

Looking at Figure 1b, *BL* mostly delivers a more robust and more regular performance than *NL*. Integrating uncertainty in the drift, learning from the data and adjusting the strategy accordingly clearly

adds value over time. Specifically, we identify three distinct periods: from inception until February 2006 the ratio BL/NL has an upward trend at moderate pace, reaching 1.13 (or equivalently 13% cumulated outperformance); then BL underperforms NL until October 2008; and finally, from October 2008 to June 2018, BL strongly dominates NL with 40% of cumulated outperformance.

5. Sensitivity Analysis

In the following paragraphs, we will stress-test the BL strategy by measuring the effects of the main parameters on long term results. Our stress-test methodology consists in fixing all but one of the parameters detailed in Table 2 and varying the remaining parameter across plausible values.

5.1. Impact of Uncertainty

Here, we study the effect of the uncertainty parameter unc in the strategy. Higher values of unc signal a higher volatility of the estimate b_0 , hence a higher uncertainty on the estimate of the expected returns. We study the BL strategy for $unc \in \{10, 50, 100, 200, 300\}$, where the value $unc = 100$ corresponds to the Base Case detailed in Section 4. Table 4 shows the results.

Table 4. Statistics for BL strategies with different values of unc and NL as defined in the workflow. The Sharpe ratio improvement is relative to NL .

	BL					NL	
	unc	10	50	100	200		300
Ann. Performance		4.59%	5.76%	5.96%	6.41%	6.51%	3.96%
Ann. Volatility		5.25%	5.49%	5.51%	5.62%	5.76%	4.92%
Max. Drawdown		−10.21%	−8.56%	−8.51%	−8.27%	−8.58%	−11.2%
Sharpe ratio		0.87	1.05	1.08	1.14	1.13	0.8
Sharpe ratio impr.		8.64%	30.45%	34.26%	41.72%	40.54%	-
Information ratio		0.34	0.58	0.55	0.57	0.54	-

Intuitively, among the BL strategies detailed in Table 4, the strategy with $unc = 10$ is the closest to NL and this is confirmed by the performance and volatility figures. Indeed, with such a low level of uncertainty, we are confident in our initial estimate b_0 . As we increase unc , the confidence we place in b_0 decreases.

Very interestingly, as unc increases, the excess return of the corresponding BL strategy over NL increases without significantly increasing the risk. Therefore, Sharpe ratios increase with unc from 0.87 for $unc = 10$ to 1.13 for $unc = 300$. More striking is the relative increase in Sharpe ratios with respect to NL —for $unc = 10$ we only have an increase of 8.64% and for $unc = 300$ the relative increase is 40.54%. Finally, the Information ratio rapidly stabilizes around 0.5.

Figure 2a,b graphically renders the increase in both annualized performances, excess returns and Sharpe ratios with respect to NL as a function of unc .

Figure 3 shows the historical levels of the strategy BL with $unc = 10$ (minimum uncertainty), $unc = 300$ (high uncertainty) and $unc = 100$ (Base Case) compared to NL . The BL strategies share the same profile, but the higher the unc parameter, the higher the excess return we find in the long run.

Clearly, the standard solution to the Markowitz problem suffers from the poor estimate of b_0 , while the BL strategy is able to adjust and react to new observable data. Moreover, the higher the uncertainty, the better BL behaves compared to NL .

Unreported tests showed us that with this particular dataset, we can increase the unc parameter even further. At some point though, around 10,000, the BL strategy underperforms NL . Indeed, when uncertainty is extremely high, the matrix Σ_0 that controls the a priori knowledge we have on B is simply uninformative because we are allowing B to span too vast a region of potential values. Therefore, the learning process, over a relatively short period of time of three months, is slow and does not add value.

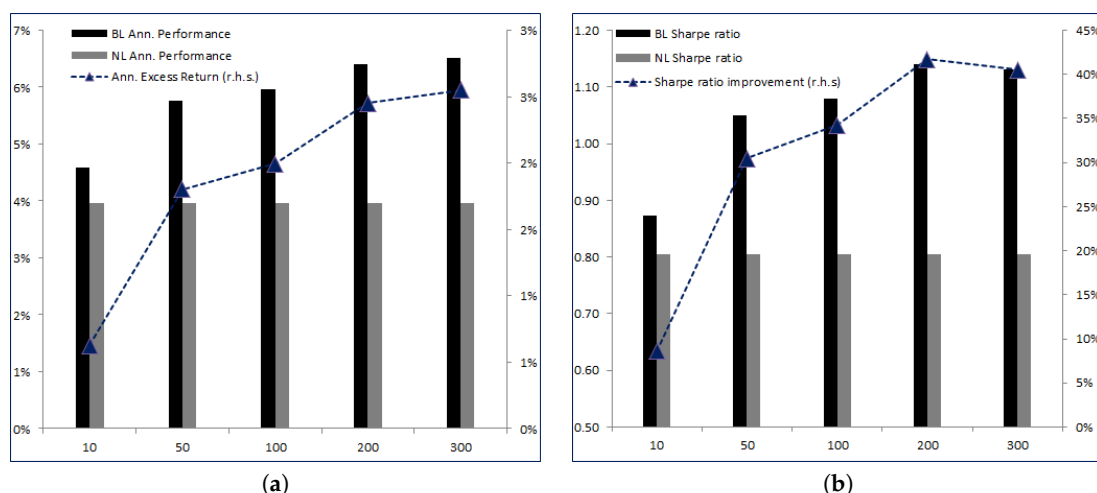


Figure 2. Annualized performances and excess returns of BL as a function of *unc* (a); Sharpe ratios and relative improvement (b).

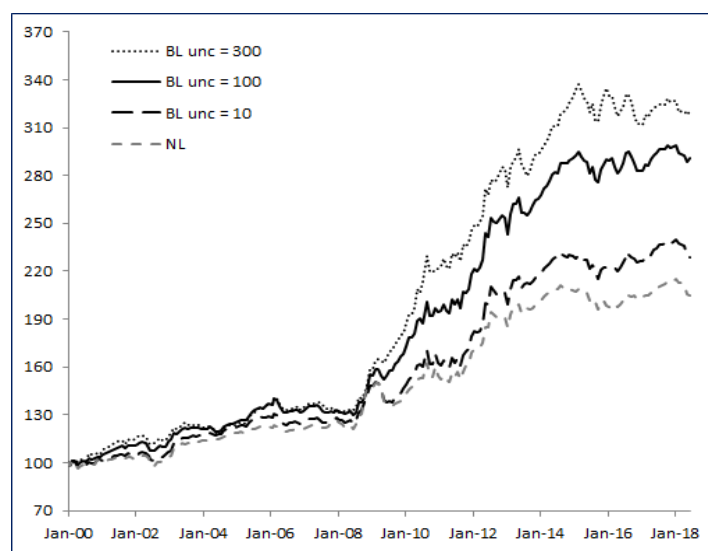


Figure 3. Historical values of the BL portfolios calculated with *unc* = 10, 100, 300 and NL.

5.2. Impact of Leverage

We now look at the maximum leverage parameter L in both strategies. Higher values of L make both $c(\mathbf{w}^{BL}, L)$ and $c(\mathbf{w}^{NL}, L)$ closer to the unconstrained weights. We test $L \in \{100\%, 150\%, 200\%, 250\%, 300\%\}$, where the value $L = 200\%$ corresponds to the Base Case detailed in Section 4. Table 5 collects the results, while Figure 4a,b gives a graphical overview of the impact on long term statistics.

Table 5. Statistics for BL and NL strategies with different levels of maximum leverage L . The Sharpe ratio improvement is calculated as the relative difference between BL and NL Sharpe ratios for the same level of maximum leverage.

L	100%		150%		200%		250%		300%	
	BL	NL	BL	NL	BL	NL	BL	NL	BL	NL
Ann. Performance	3.78%	2.48%	4.89%	3.23%	5.96%	3.96%	7.06%	4.39%	7.96%	5.00%
Ann. Volatility	2.78%	2.48%	4.12%	3.64%	5.51%	4.92%	6.85%	6.09%	8.11%	7.23%
Max. Drawdown	-3.96%	-5.59%	-5.92%	-8.43%	-8.51%	-11.2%	-12.49%	-13.01%	-16.33%	-14.07%
Sharpe ratio	1.36	1	1.19	0.89	1.08	0.8	1.03	0.72	0.98	0.69
Sharpe ratio impr.	36.36%	-	33.96%	-	34.26%	-	42.89%	-	41.88%	-
Information Ratio	0.69	-	0.65	-	0.55	-	0.58	-	0.53	-

As the maximum leverage L increases, the corresponding excess return of the BL strategy over NL increases. For $L = 100\%$, the excess return is 1.30% and goes up to 2.96% for $L = 300\%$.

Clearly, and as expected, the excess return increases with the leverage and on average we observe that adding 100% of leverage brings 0.80% in extra excess return. When we look at the Sharpe ratios, we see that BL always outperforms NL but the relation is more complex than the previous one. Both BL and NL Sharpe ratios decrease with L . Indeed, as L increases, we observe higher performances but also higher volatilities, and the volatility grows faster than the performance. This is obvious, because when L increases the strategy becomes more sensitive to any error in the estimated parameters, which is reflected in higher volatility. As $L \rightarrow \infty$, the strategies become unconstrained and bring all the instability that is well known to go alongside with price-based strategies. On a relative basis, the improvement in the Sharpe ratios first decreases as L reaches 200% and then increases.

To conclude, it appears that the impact of the maximum leverage L is as expected—it brings extra performance at the cost of higher risk. Because we know real data do not need to follow model (1), unconstrained BL and NL (as any price-based strategies) tend to amplify any model error, making the leverage constraint a wise feature to consider.

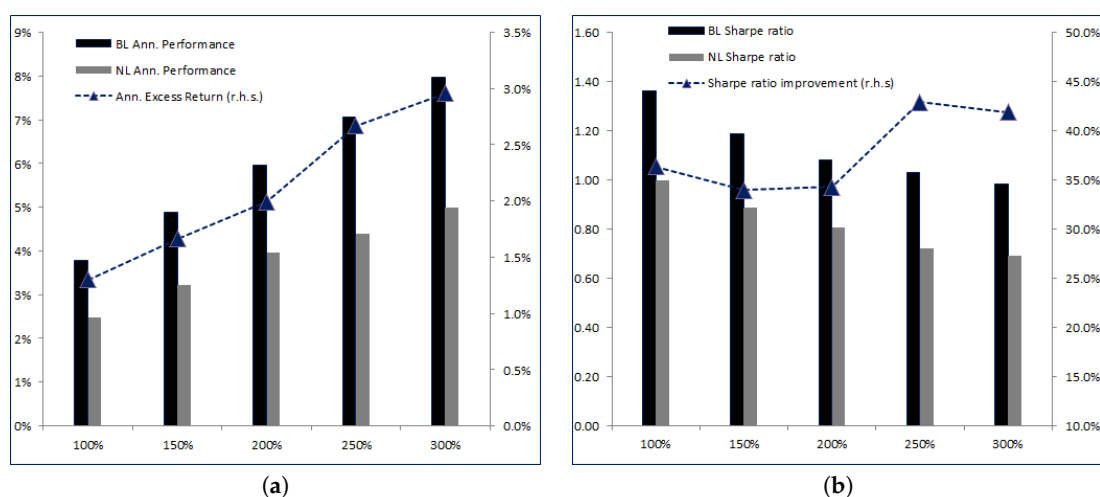


Figure 4. Annualized performances and excess returns of BL as a function of L (a); Sharpe ratios and relative improvement (b).

5.3. Impact of the Review Frequency

Probably one of the most important features of the learning effect on the portfolio is the frequency at which we recalibrate the parameters of the model. We have chosen a three-month frequency in the Base Case. In other words, every three months we compute new estimates of b_0 and Σ and input a new matrix Σ_0 , driving the uncertainty. We expected that the lower the frequency, the better BL would compare to NL . The main reason behind this thesis is that the strategy NL will be stuck with parameters b_0 and Σ for a long period of time, during which we know that they will most likely become obsolete as the market evolves and new information is processed. As investors know very well, forecasting is difficult and outdated forecasts are badly suited for portfolio construction¹. On the other side, BL can adapt over time because it embeds uncertainty about b_0 .

¹ *It's tough to make predictions, especially about the future.* The quote, which clearly applies in the context of portfolio construction, has been reported by many different people. According to <https://quoteinvestigator.com/2013/10/20/no-predict/>, it first appeared in Danish documents, but well known personalities such as Neils Bohr, Mark Kac, Stanislaw M. Ulam, and, probably misattributed, Mark Twain and Yogi Berra.

Let us consider the frequency parameter $Freq$ in both strategies varying from three to twelve months, $Freq \in \{3 M, 6 M, 9 M, 12 M\}$, where the value $Freq = 3 M$ corresponds to the Base Case detailed in Section 4. Table 6 shows the results.

Table 6. Statistics for BL and NL strategies with different levels of the review frequency $Freq$. The Sharpe ratio improvement is calculated as the relative difference between BL and NL Sharpe ratios for the same review frequency.

$Freq$	3 M		6 M		9 M		12 M	
	BL	NL	BL	NL	BL	NL	BL	NL
Ann. Performance	5.96%	3.96%	4.74%	3.25%	4.28%	2.41%	4.75%	2.21%
Ann. Volatility	5.51%	4.92%	5.5%	5.35%	5.71%	5.45%	5.36%	4.5%
Max. Drawdown	−8.51%	−11.2%	−11.62%	−13.71%	−12.44%	−16.15%	−8.9%	−16.05%
Sharpe ratio	1.08	0.8	0.86	0.61	0.75	0.44	0.89	0.49
Sharpe ratio impr.	34.26%	-	41.74%	-	69.38%	-	80.4%	-
Information Ratio	0.55	-	0.33	-	0.35	-	0.57	-

As we lower the review frequency, we see the performances of both BL and NL decreasing. Nevertheless, the excess return of BL over NL actually increases at frequencies lower than 6-month. The fact that it is not monotonic is related to the timing effect. The best metric to assess the value added of the learning feature remains the relative improvement in Sharpe ratios, as shown in Figure 5b. Here we see that as the frequency lowers, more risk-adjusted value added comes from the learning effect: at 3 months, the relative improvement in Sharpe ratio between BL and NL is 34.26%. At 6 months it goes up to 41.47%, then 69.38% at 9 months and finally 80.4% relative improvement at the 12-month frequency. Clearly, BL outperforms NL if we do not review the parameters of the model quite often, and this is clearly attributable to the fact that BL can adjust over time according to the data observed and the adjustments it can make on the a priori distribution of B .

For the sake of simplicity, we do not report the results here, but it is possible to increase the efficiency of BL compared to NL if we simultaneously modify the frequency $Freq$ and the uncertainty parameter unc in Σ_0 . Over long investment horizons, the investor would definitely have more uncertainty on her a priori estimate b_0 , therefore it makes perfect sense to consider $Freq = 12$ coupled with $unc = 1600$ or higher².

² We suggest 1600 simply as a rule of thumb: the ratio of frequency between 3 M and 12 M is 4 so that we pass a $4^2 = 16$ factor in unc . Of course, we do not expect this parameter to be the best choice since, empirically, uncertainty grows more than linearly with time, but even this, provides slightly higher Sharpe ratio improvements.

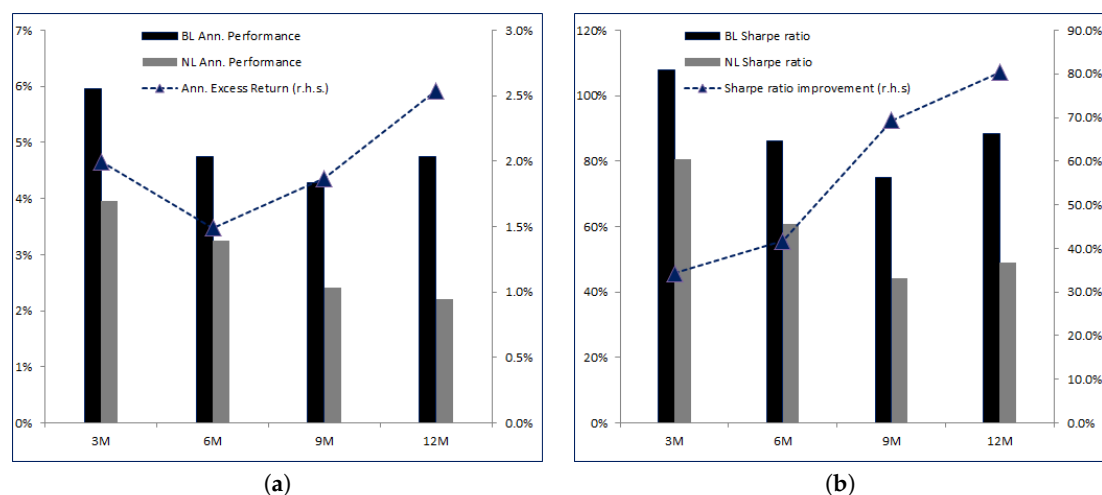


Figure 5. Annualized performances and excess returns of *BL* as a function of *L* (a); Sharpe ratios and relative improvement (b).

5.4. Impact of Rebalancing Frequency

We conclude this section with an overview on the impact of the rebalancing frequency on both *BL* and *NL*. This exercise is only theoretical, since within the Base Case we only performed three trades (one each month before the next review). A lower rebalancing frequency of trading would mean that we do not fully exploit the power of learning by adjusting the portfolio. Higher rebalancing frequency carries a turnover (and transaction cost) issue, that would make any benefit only hypothetical. Furthermore, when we rebalance too often, we embark significant amount of noise coming from daily, short term, price movements. Usually, practitioners consider higher rebalancing frequencies only for a small portion of the portfolio, i.e., they only rebalance a fraction x of their portfolio while keeping the remaining $1 - x$ unchanged and they roll over. Because this goes beyond the scope of this paper, we limit ourselves to monthly versus biweekly rebalancing frequency. Table 7 collects summary statistics for *BL* and *NL* when the optimal weights in the Base Case are implemented on both monthly and biweekly frequencies.

Table 7. Statistics for *BL* and *NL* strategies with different rebalancing frequencies *Reb Freq*, monthly versus biweekly. The Sharpe ratio improvement is calculated as the relative difference between *BL* and *NL* Sharpe ratios for the same rebalancing frequency.

<i>Reb Freq</i>	Monthly		Bi-Weekly	
	<i>BL</i>	<i>NL</i>	<i>BL</i>	<i>NL</i>
Ann. Performance	5.96%	3.96%	5.42%	2.43%
Ann. Volatility	5.51%	4.92%	5.18%	4.93%
Max. Drawdown	−8.51%	−11.2%	−13.62%	−14.08%
Sharpe ratio	1.08	0.8	1.05	0.49
Sharpe ratio impr.	34.26%	-	112.01%	-
Information Ratio	0.55	-	0.77	-

As we go from the monthly to the biweekly rebalancing frequency, we observe a small decrease in annualized performance of *BL*, while the drop is more significant for *NL*. Although these numbers should not be taken as very informative due to the timing effect, there is clearly not a strong incentive for *BL* to rebalance more often, since performance goes down slightly, as does volatility. In the end, the Sharpe ratio seems fairly stable. On the other hand, *NL* experiences a large drop in performance as well as an important increase in its maximum drawdown. Indeed, when we look at the structure of the strategy *NL* in (4), we see that the trade $\Sigma^{-1}\mathbf{b}_0$ does not change at higher frequency. So, *NL* will implement the same trade (up to the maximum leverage). If, for example, the *NL* strategy was

successful in the first two weeks, at the next rebalancing it will most likely reverse this successful trade because of the $x_{t_0} - x_t$ part of \mathbf{w}^{NL} . Conversely, if the *NL* strategy was unsuccessful over a two-week period with a trade, at the next rebalancing it will increase this trade for the same reason. Therefore, *NL* tends to have a reversal feature at short horizon/high frequency.

As far as *BL* is concerned, because of uncertainty, the strategy can accommodate for larger deviation of the process *B* relatively to the forecast \mathbf{b}_0 , so that it does not necessarily have to reverse a successful trade or leverage an unsuccessful one. Indeed, when we look at (3), a successful trade will lower the leverage, but this can be compensated by the corrective term in the trade, which depends on ψ and $\nabla_b R$, so that in theory *BL* does not systematically have a reversal feature.

Finally, given our dataset and the fact that the momentum premium (see for example [Asness et al. \(2013\)](#); [Blitz and Van Vliet \(2008\)](#); [Carhart \(1997\)](#); [Geczy and Samonov \(2017\)](#)) exists in the multi-asset framework, it is not surprising that *NL* experiences a drop in performance while *BL* is almost unaffected.

This is confirmed by the significant Sharpe ratio improvement: from a +34.26% in the Base Case, we improve the final Sharpe ratio at the biweekly frequency by 112.01% (mainly driven by the drop in *NL* Sharpe ratio). Furthermore, we see how *BL* is able to extract more information (or *alpha*) from the market because the Information ratio increases from 0.55 with the monthly rebalancing to 0.77 (a 40% increase) with the biweekly rebalancing.

6. Investing in Foreign Currencies

In our second example, we consider investing in different currencies: the Australian Dollar (AUD), the Canadian Dollar (CAD), the Euro (EUR), the British Pound (GBP), the Japanese Yen (JPY) and the U.S. Dollar (USD). Usually set by the central banks, the local risk-free interest rate, such as the federal funds rate in the U.S., together with the foreign exchange rate of currencies versus the Euro are the sources of performance for the investor. Therefore, the underlying assets available to the investor are bank accounts in foreign banks that pay the local interest rate but are valued in EUR. Details are collected in Table 8.

Table 8. Currency and their reference rates.

Currency	Rate	Source
AUD	Thomson Reuters Australian Dollar Overnight Deposit	Thomson Reuters
CAD	Canada Money Market Overnight	Bank of Canada
EUR	Eonia (Euro OverNight Index Average)	European Banking Federation
GBP	United Kingdom Sonia	Wholesale Markets Brokers' Association
JPY	Japan Uncollateralized Overnight	Bank of Japan
USD	United States Federal Funds Effective Rate	Federal Reserve, United States

The workflow detailed in Section 3 was implemented with the parameters listed in Table 9.

Table 9. Parameters used in the implementation of *BL* and *NL* as defined in the workflow for the foreign currency strategy.

Parameter	Value
X_0	100
T	12 months
Lag	1 day
$Freq$	Weekly, Friday
r^w	60 days
s^w	250 days
d	10%
unc	100
r^q	30
L	100%

With respect to Table 2, here we consider a longer investment horizon (one year), we rebalance more often (weekly) because of reduced transaction costs in highly liquid foreign currencies, we consider short windows for estimating the drift parameter and we do not allow leverage.

The result of the *BL* and *NL* strategies from January 2002 to June 2018 are reported in Table 10, while the historical performance is shown in Figure 6.

Table 10. Statistics for the *BL* and *NL* strategies. The Sharpe Ratio is calculated as the ratio between annualized performance and annualized volatility. For this case, the difference is given in relative terms. Data from January 2002 to June 2018.

	<i>BL</i>	<i>NL</i>	Difference
Ann. Performance	4.00%	2.83%	1.17%
Ann. Volatility	5.97%	4.93%	1.04%
Max. Drawdown	−14.17%	−10.51%	−3.66%
Sharpe ratio	0.67	0.57	17.5%
Information Ratio	0.42	-	-

Over the period used, *BL* outperformed *NL* by 1.17% annualized. This is quite impressive given the low yields of developed countries' currencies (mainly EUR and JPY) that can usually be observed in the market. Therefore, the effect of learning clearly brings value to the investor by adding extra performance, although this comes at slightly higher risk (roughly 1% more volatile and 3.66% larger maximum drawdown). Nevertheless, the improvement in the Sharpe ratio is consistent. Figure 6b shows the relative strength of *BL* over *NL* and the growth line. We can see that, except for a few months in the second half of 2008 where *BL* strongly underperformed, after that time it regularly outperformed *NL*.

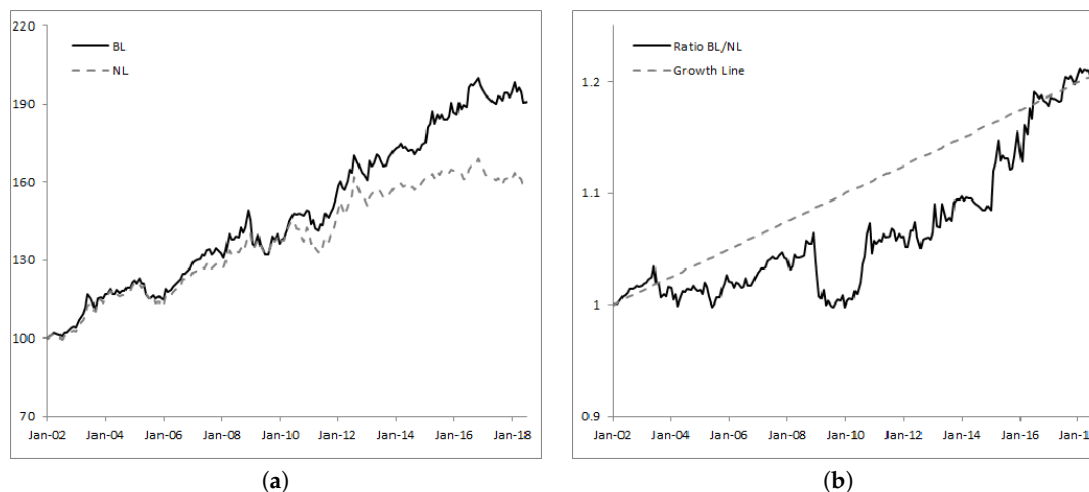


Figure 6. Historical values of the portfolios calculated with both w^{BL} and w^{NL} under the Base Case (a) and their ratio (b).

7. Investing in Factor Strategies

In the last few years, investors have embraced alternative strategies that target specific, well-established equity factors (such as size, value, volatility or momentum). These strategies offer an efficient and direct exposure to the main driving factors of equity markets and allow for an optimal allocation across factors. The main challenge is to invest in the right factor at the right time, as their performance usually shows cyclical patterns. Table 11 contains more details on the strategies we considered. All of them are based on the U.S. large capitalization equity market.

Table 11. Factor-based strategies and cash. For the volatility factor, we chained two strategies on 18 December 2009. Data in USD.

Factor	Name	Source
Dividend	S&P 500 Dividend Aristocrats Net Total Return Index	S&P
Growth	MSCI USA Growth Net Total Return Index	MSCI
Momentum	MSCI USA Momentum USD Net Total Return Index	MSCI
Quality	MSCI USA Quality Net Total Return Index	MSCI
Size	MSCI USA Small Cap Net Total Return Index	MSCI
Value	Shiller Barclays CAPE US Sector Value Net TR Index	Barclays
Volatility	Ossiam US Minimum Variance Index Net Return/ ESG US Minimum Variance Index	S&P, Solactive
Cash	United States Federal Funds Effective Rate	Federal Reserve, United States

We follow the workflow of Section 3 with the parameters listed in Table 12 to derive both *BL* and *NL* strategies.

Table 12. Parameters used in the implementation of the *BL* and *NL* factor rotation strategies as defined in the workflow.

Parameter	Value
X_0	100
T	12 months
Lag	1 day
$Freq$	Monthly, 3rd Friday of each month.
r^w	90 days
s^w	250 days
d	10%
unc	100
r^q	60
L	100%

The results from January 2002 to June 2018 are shown in Table 13. In this example as well, the *BL* strategy has significantly outperformed the *NL* strategy. Again, the learning effect clearly brings value to the investor. On an annualized basis, *BL* improves the performance by a bit more than 2%, with a lower maximum drawdown (−8.24% versus −10.43% for *NL*) but a slightly higher annualized volatility (3.61% for *BL* versus 2.26% for *NL*).

Table 13. Statistics for the *BL* and *NL* strategies. The Sharpe Ratio is calculated as the ratio between annualized performance and annualized volatility. For this case, the difference is given in relative terms. Data from January 2002 to June 2018.

	<i>BL</i>	<i>NL</i>	Difference
Ann. Performance	3.1%	1.03%	2.07%
Ann. Volatility	3.61%	2.26%	1.34%
Max. Drawdown	−8.24%	−10.43%	2.2%
Sharpe ratio	0.86	0.45	91.11%
Information Ratio	0.55	-	-

The improvement in the Sharpe ratio is particularly high—0.86 for the *BL* factor rotation strategy versus 0.45 for the *NL* strategy (an improvement of 91.11%). Figure 7 shows the historical performance (Figure 7a) and the strength ratio (Figure 7b) of the *BL* strategy over *NL*. Clearly, learning over time produces regular outperformance, as shown by the regular upwardly increasing ratio between *BL* and *NL*.

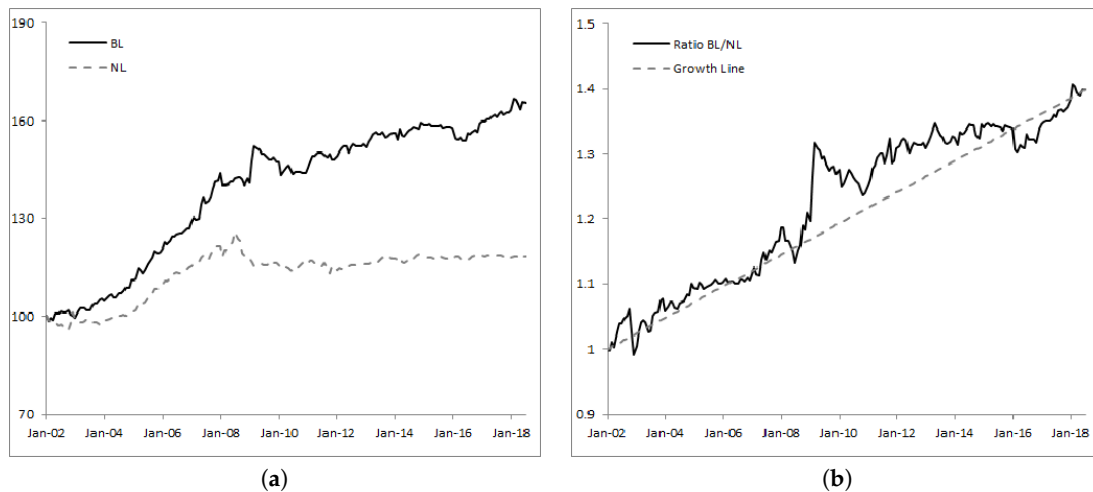


Figure 7. Historical values of the portfolios calculated with both w^{BL} and w^{NL} under the Base Case (a) and their ratio (b).

Author Contributions: C.D.F. designed the experiments; C.D.F. and J.N. performed the experiments; C.D.F. and J.N. analyzed the data; C.D.F., J.N. and H.P. contributed to the mathematical analysis; H.P. designed the theoretical framework of this study.

Funding: This research received no external funding.

Acknowledgments: This work is issued from a CIFRE collaboration between OSSIAM and LPSM.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CAPM	Capital Asset Pricing Model
AUD	Australian Dollar
BL	Bayesian learning (strategy)
CAD	Canadian Dollar
CAPE	Cyclically Adjusted Price-to-Earnings
Conv.	Conventional
EONIA	Euro OverNight Index Average
ETF	Exchange-Traded Fund
FTSE	Financial Times Stock Exchange
EMBI	Emerging Markets Bond Index
EMMI	European Money Markets Institute
EUR	Euro
EurAgg	Euro Aggregate
GBP	Great British Pound
GSCI	Goldman Sachs Commodities Index
JPY	Japanese Yen
LBMA	London Bullion Market Association
MSCI	Morgan Stanley Capital International
NL	nonlearning (strategy)
NR	Net Return
S&P	Standard and Poors
Topix	Tokyo Stock Price Index
TR	Total Return

References

- Aguilar, Omar, and Mike West. 2000. Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics* 18: 338–57.
- Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen. 2013. Value and momentum everywhere. *Journal of Finance* 68: 929–86. [CrossRef]
- Avramov, Doron, and Guofu Zhou. 2010. Bayesian portfolio analysis. *Annual Review of Financial Economics* 2: 25–47. [CrossRef]
- Barry, Christopher B. 1974. Portfolio analysis under uncertain means, variances, and covariances. *The Journal of Finance* 29: 515–22. [CrossRef]
- Best, Michael J., and Robert R. Grauer. 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: 315–42. [CrossRef]
- Black, Fisher, and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81: 637–54. [CrossRef]
- Blitz, David, and Pim Van Vliet. 2008. Global tactical cross-asset allocation: Applying value and momentum across asset classes. *Journal of Portfolio Management* 35: 23–28. [CrossRef]
- Bodnar, Taras, Stepan Mazur, and Yarema Okhrin. 2017. Bayesian estimation of the global minimum variance portfolio. *European Journal of Operational Research* 256: 292–307. [CrossRef]
- Carhart, Mark M. 1997. On persistence in mutual fund performance. *The Journal of Finance* 52: 57–82. [CrossRef]
- Cvitanić, Jakša, Ali Lazrak, Lionel Martellini, and Fernando Zapatero. 2006. Dynamic portfolio choice with parameter uncertainty and the economic value of analysts' recommendations. *The Review of Financial Studies* 19: 1113–56. [CrossRef]
- De Franco, Carmine, Johann Nicolle, and Huyèn Pham. 2018. Bayesian Learning for Markowitz Portfolio Selection Problem. Working paper, Cornell University, Ithaca, NY, USA, arXiv:1811.06893.
- Elliott, Robert J., William C. Hunter, and Barbara M. Jamieson. 1998. Drift and volatility estimation in discrete time. *Journal of Economic Dynamics and Control* 22: 209–18. [CrossRef]
- Fama, Eugene F., and Kenneth R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56. [CrossRef]
- Fama, Eugene F., and Kenneth R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116: 1–22. [CrossRef]
- Fama, Eugene F., and Kenneth R. French. 2016. Dissecting anomalies with a five-factor model. *The Review of Financial Studies* 29: 69–103. [CrossRef]
- Frost, Peter A., and James E. Savarino. 1986. An empirical bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis* 21: 293–305. [CrossRef]
- Geczy, Christopher, and Mikhail Samonov. 2017. Two Centuries of Multi-Asset Momentum (Equities, Bonds, Currencies, Commodities, Sectors and Stocks). Available online: <http://dx.doi.org/10.2139/ssrn.2607730> (accessed on 20 May 2015).
- Karatzas, Ioannis, John P. Lehoczky, and Steven E. Shreve. 1987. Optimal portfolio and consumption decisions for a "small investor" on a finite horizon. *SIAM Journal on Control and Optimization* 25: 1557–86. [CrossRef]
- Klein, Roger W., and Vijay S. Bawa. 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3: 215–31. [CrossRef]
- Lakner, Peter. 1995. Utility maximization with partial information. *Stochastic Processes and Their Applications* 56: 247–73. [CrossRef]
- Lakner, Peter. 1998. Optimal trading strategy for an investor: The case of partial information. *Stochastic Processes and Their Applications* 76: 77–97. [CrossRef]
- Lintner, John. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47: 13–37. [CrossRef]
- Markowitz, Harry. 1952. Portfolio Selection. *The Journal of Finance* 7: 77–91.
- Merton, Robert C. 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics* 51: 247–57. [CrossRef]
- Merton, Robert C. 1975. Optimum consumption and portfolio rules in a continuous-time model. In *Stochastic Optimization Models in Finance*. New York: Elsevier, pp. 621–61.

- Merton, Robert C. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8: 323–61. [[CrossRef](#)]
- Rogers, Leonard. C. G. 2001. The relaxed investor and parameter uncertainty. *Finance and Stochastics* 5: 131–54. [[CrossRef](#)]
- Samuelson, Paul. 1969. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economics and Statistics* 51: 239–46. [[CrossRef](#)]
- Scaillet, Olivier. 2004. Nonparametric estimation and sensitivity analysis of expected shortfall. *Mathematical Finance* 14: 115–29. [[CrossRef](#)]
- Sharpe, William F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19: 425–42.
- Siegel, Andrew F., and Artemiza Woodgate. 2007. Performance of portfolios optimized with estimation error. *Management Science* 53: 1005–15. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).