Potential losses from incorporating return predictability into portfolio allocation

Dragon Yongjun Tang
School of Economics and Finance, The University of Hong Kong, Hong Kong

Abstract
The extant literature demonstrates the importance of stock return predictability for portfolio allocation. The usefulness of incorporating return predictability into portfolio decisions is most evident for Bayesian investors who build their portfolios based on their prior beliefs. I show that the magnitude of economic significance of stock return predictability largely depends on the choice of prior beliefs. An investor could suffer substantial utility loss when he delegates portfolio management to a manager with a different belief about stock return predictability. The consideration of Bayesian prior robustness in portfolio analysis can be as important as return predictability itself.

Keywords
Bayesian robustness, portfolio selection, return predictability

“… a good Bayesian will beat a non-Bayesian, who will do better than a bad Bayesian.”

(Granger, 1986, p. 16)

Introduction
Bayesian methods have been extensively applied to portfolio analysis, as reviewed by Avramov and Zhou (2010). The advantage of the Bayesian approach is its facilitation of incorporating sound prior views into observed data (rather than relying purely on observed data). The flip side is that the prior belief is subjective. As pointed out by Fabozzi et al (2010), we need to better understand robust portfolio rules in real-world applications. This article explores the effect of Bayesian

Corresponding author:
Dragon Yongjun Tang, School of Economics and Finance, The University of Hong Kong, Pokfulam Road, Hong Kong.
Email: yjtang@hku.hk
robustness in the context of optimal portfolio allocation with stock return predictability building upon Kandel and Stambaugh (1996).

Stock return predictability has long been of great interest to finance academics and practitioners. For example, Dou et al. (2012) carry out a comprehensive study of stock return predictability in Australia. A recent string of research examines the implications of return predictability to optimal portfolio selection. Kandel and Stambaugh (1996) explicitly take into account parameter uncertainty using Bayesian methods and demonstrate that investors’ utility can be significantly improved even with statistically weak return predictability. However, Kandel and Stambaugh’s (1996) predictions are not fully supported by empirical studies using historical financial data. Pesaran and Timmermann (1995) find that investors cannot always exploit stock return predictability in historical data. Moreover, Hand and Tiwari (2006) argue that investors cannot benefit from data-based predictability in real time. Our study aims to bridge the gap between the theory and the data on the usefulness of return predictability when investors’ prior beliefs are uncertain.

Bayesian investors generally are uncertain with both the likelihood function of the data-generating process (model uncertainty) and prior distribution. Such concern is warranted as Gray (2008) shows that the economic significance of stock return predictability is sample specific and not robust using Australian data. This study focuses on prior robustness in financial applications. The importance of model uncertainty is analyzed by, among others, Tu and Zhou (2004). Brown (2011) stresses the important role of prior beliefs in causing the global financial crisis 2007–2008. Investors and financial analysts have heterogeneous prior beliefs which affect their portfolio selection. Lettau and Van Nieuwerburgh (2008) imply that investors are ex ante very uncertain regarding future return predictability. Therefore, prior beliefs on return predictability can be very different across investors and time. Previous Bayesian financial studies recognize the impact of prior belief on economic conclusions. We do a thorough robustness analysis on the economic significance of stock return predictability in a select prior space using contamination techniques from Bayesian statistics.

We show that the magnitude of economic significance of stock return predictability varies substantially over a small prior space. The conclusion drawn by Kandel and Stambaugh (1996) (hereafter referred to as KS) is quite optimistic. Their results often represent the most significant results over a set of possible priors. The most conservative investors may find that ignoring stock return predictability results in utility losses less than five basis points per month, and may be smaller than the transaction costs induced by portfolio rebalancing. Moreover, we show that a portfolio manager who makes investment decisions solely based on his own prior beliefs can cause a large utility loss (1% per month) for his investors who hold different prior beliefs. The economic significance of prior robustness is often much larger than the economic significance of return predictability, especially during market booms. Caution needs to be exercised when we make inferences using Bayesian methods in financial studies.

**Incorporating return predictability into portfolio allocation**

The usefulness of the Bayesian approach to asset allocation is probably best illustrated by KS, who consider a Bayesian investor’s asset allocation problem with return predictability information. For brevity, we only reproduce the key steps here. Readers interested in the details are referred to the original paper or the review by Avramov and Zhou (2010). The investor has constant relative risk aversion (CRRA) utility function with coefficient $A$

$$v(W) = \begin{cases} 
\frac{1}{1-A} W^{1-A} & \text{for } A > 0 \text{ and } A \neq 1 \\
\ln W & \text{for } A = 1
\end{cases}$$  

(1)
Provided with historical data $\Phi_T$, the investor needs to make a one-period portfolio allocation decision with one risky asset (stock) and one risk-free asset

$$\max_w E[v(W_{T+1})] = \int v(W_{T+1}) p(r_{T+1} \mid \Phi_T) dr_{T+1},$$

s.t. $W_{T+1} = W_T[w \exp\{r_{T+1} + r_{fT+1}\} + (1-w) \exp\{r_{fT+1}\}]$

where $p(r_{T+1} \mid \Phi_T)$ is the predictive probability distribution function (pdf), assuming known parameters, for the (uncertain) time $T+1$ continuously compounded excess returns on the stock in the portfolio $r_{T+1}$, $r_{fT+1}$ is the risk-free rate.

KS assume that the investor has certain belief about the data generating processes of the stock return $r_{T+1}$ and a vector of $N$ predictors $x_T$. Specifically, they use the following VAR(1) model for the predictive regression over sample period $T$:

$$Y = XB + U, \quad \text{vec}(U) \sim N(0, \Sigma \otimes I)$$

where the prediction error $U$ has a ‘matrix-variate normal’ (MN) distribution in $\Sigma$. The dependent variable $y_{T+1}$ is $(r_{T+1}, x_{T+1})$ so $Y$ is a $T \times (N+1)$ matrix. The independent variable is $(1, x_T)$ and $X$ is a $T \times (N+1)$ matrix. Here $B$ is a $(N+1) \times (N+1)$ regression coefficient. This specification is used to account for the dynamics of $x_T$ and minimize the potential bias in predictive regressions. The MN assumption gives the following likelihood function:

$$l(\Phi_T \mid B, \Sigma) \propto |\Sigma|^{-T/2} \exp\left[-\frac{1}{2} tr(Y - XB)'(Y - XB)\Sigma^{-1}\right]$$

In order to obtain the predictive pdf $p(y_{T+1} \mid \Phi_T)$ and the first component $p(r_{T+1} \mid \Phi_T)$, the prior distributions for the parameters $B$ and $\Sigma$ need to be specified.

KS employ both a diffuse prior and a no-predictability informative prior. In this study, we focus on robustness with respect to the diffuse prior for two reasons. First, diffuse priors are more robust than informative priors. Robustness findings regarding diffuse priors are likely to hold for informative priors as well. Second, KS’s results using diffuse priors are stronger than the results from informative priors. The diffuse prior used by KS is the Jeffreys prior

$$\pi(B) = \text{const}, \quad \pi(\Sigma) \propto |\Sigma|^{-(N+2)/2}$$

With these distributions, $p(r_{T+1} \mid \Phi_T)$ follows a Student’s $t$ distribution, when $T > 2N + 1$,

$$p(r_{T+1} \mid \Phi_T) = \frac{\Gamma((v + 1)/2)}{\Gamma(1/2)\Gamma(v/2)} \left(\frac{1}{v - 2}\right)^{1/2} \left[1 + \left(\frac{1}{v - 2}\right)\left(\frac{r_{T+1} - \mu_T}{\sigma_T^2}\right)^2\right]^{-(v+1)/2}$$

where the distribution parameters $(\mu_T, \sigma_T, v)$ are calculated from the following sample statistics:

$$\mu_T = \bar{r} + \hat{\beta}'(y_T - \bar{y}), \quad \hat{\beta}'(y_T - \bar{y}) = \hat{\delta} \sqrt{R^2}\hat{\sigma}_r$$

$$\sigma_T^2 = \frac{T}{T - 2(N+1)} \left(1 - R^2\right) \left[1 + \frac{1}{T} (1 + q)\right] \hat{\sigma}_r^2$$
where \( R^2 \) is the in-sample fitting statistics of regression (4), \( \delta \) represents the number of standard deviations the predicted value \( r_{T+1} \) is away from the sample mean \( \mu_T \). The closed-form predictive pdf \( p(r_{T+1} | \Phi_T) \) enables KS to numerically solve investor’s optimal asset allocation problem given investor’s risk aversion \( A \) and risk-free rate \( r_f \). For varying \( \delta \), the economic significance of predictability is characterized by certainty equivalent gains over ignoring predictability (focus on sample means and treat \( \delta = 0 \)), in contrast to statistical significance \( R^2 \) and corresponding \( p \)-value.

Let \( w^0 \) be the optimal portfolio rule accounting for predictability, and \( w^a \) be the suboptimal portfolio rule ignoring predictability, the economic significance of predictability is measured by gains in certainty equivalent returns

\[
\Delta CER = CER(w^0, \pi_0, \Phi_T) - CER(w^a, \pi_a, \Phi_T)
\]

where \( CER(w, \pi, \Phi_T) \) is the certainty equivalent return for the investor with prior \( \pi \), data \( \Phi_T \), and portfolio allocation \( w \). We reproduce KS’s key results (their Panel A of Table I on p. 405) in Table 1 in the rows denoted KS. The mean monthly return is assumed to be \( \bar{r} = 0.49\% \) and standard deviation of \( \sigma_r = 5.6\% \). The monthly risk-free rate is taken to be \( r_f = 0.235\% \). Also assume the observed data consists of \( T = 804 \) monthly observations and \( N = 25 \) predictive variables. Risk aversion is taken to be \( A = 1, 2, 5 \). Short selling and borrowing are restricted so that \( w \in [0, 0.99] \). In Table 1, the statistical significance is measured by un-adjusted \( R^2 \) and corresponding \( p \)-value. As argued by KS, stock return predictability can be economically significant, generating an improved \( CER \) of about 3% per annum (28.5 or 24.0 basis points per month) when predicted return is one standard deviation away from sample mean, while it is statistically weak with a \( R^2 \) of 2.5% and \( p \)-value of 0.75 for an investor with risk aversion \( A = 2 \).

KS conduct a robustness check by assuming a different prior which is derived from a hypothetical sample (\( \text{\`a la} \) empirical Bayesian approach). However, the effectiveness of this robustness check is limited. Basic Bayesian analysis tells us that splitting the sample does not affect the belief updating process. In order to investigate to what extent KS’s conclusions are robust with respect to prior choice, we need to consider alternative priors.

The Bayesian prior robustness

Berger (1994) provides guidelines on choosing an appropriate class of priors \( \Gamma \) for robustness analysis and suggests the \( \varepsilon \)-contamination prior class \( \Gamma = \{(1 - \varepsilon)\pi_0 + \varepsilon q : q \in Q\} \) for its generality and tractability. Hence, we investigate the robustness of KS’s results over all priors in \( \Gamma \) using their diffuse prior as the default prior \( \pi_0 \). Next, we need to choose the contaminations \( Q \) for parameters \((B, \Sigma)\).

The straightforward choice for \( Q \) is all distributions, but this choice loses all tractability. One intuitive and plausible choice for \( Q \), as suggested by Sivaganesan and Berger (1989), is \( Q_{SU} \) (all symmetric unimodal distributions with the same mode as that of \( \pi_0 \)). This contamination generates the prior space: \( \Gamma_{SU} = \{(1 - \varepsilon)\pi_0 + \varepsilon q : q \in Q_{SU}\} \). A tractable and general element of \( Q_{SU} \) is the following prior:

\[
\pi(B | \Sigma) \propto MN(B_0, \Sigma, M_0^{-1}), \quad \pi(\Sigma) \propto IW(S_0, v_0, N)
\]
Table 1. Bayesian prior robustness analysis of Kandel and Stambaugh (1996). The robustness analysis of Kandel and Stambaugh’s results with diffuse prior (KS Panel A of Table I on p. 405). KS’s results are reproduced in rows denoted ‘KS’ for the certainty equivalent gains ($\Delta CER$) of accounting for predictability when predicted return is $\delta$ sample standard deviations away from the sample average with in-sample predictive regression $R^2$ and corresponding $p$-value for investor risk aversion parameters $A = 1, 2, 5$. Range is the minimum and maximum of $\Delta CER$ spanning over the prior contamination space. $Ratio^{max}$ is the maximum of $(CER(w', \pi, \Phi_T^{'}) - CER(w^0, \pi, \Phi_T^{'0}))^2 / V^\pi$. Assume the observed historical data contains $T = 804$ months and $N = 25$ predictive variables.

<table>
<thead>
<tr>
<th>$R^2$</th>
<th>$p$-value</th>
<th>A</th>
<th>$\Delta CER$ (basis points) for $\delta$ equal to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$-1$</td>
</tr>
<tr>
<td>0.025</td>
<td>0.75</td>
<td>1</td>
<td>KS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$Ratio^{max}$</td>
</tr>
<tr>
<td>0.025</td>
<td>0.75</td>
<td>2</td>
<td>KS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$Ratio^{max}$</td>
</tr>
<tr>
<td>0.025</td>
<td>0.75</td>
<td>5</td>
<td>KS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$Ratio^{max}$</td>
</tr>
<tr>
<td>0.055</td>
<td>0.01</td>
<td>1</td>
<td>KS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$Ratio^{max}$</td>
</tr>
<tr>
<td>0.055</td>
<td>0.01</td>
<td>2</td>
<td>KS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$Ratio^{max}$</td>
</tr>
<tr>
<td>0.055</td>
<td>0.01</td>
<td>5</td>
<td>KS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$Ratio^{max}$</td>
</tr>
</tbody>
</table>

where ‘IW’ is the ‘inverted Wishart’ distribution:

$$
\pi(\Sigma) = k \left| S_0 \right| ^{\nu_0 / 2} \left| \Sigma \right| ^{-\left(\nu_0 + N + 1\right) / 2} \exp\left[ -\frac{1}{2} tr(\Sigma^{-1} S_0) \right] 
$$

(14)

We choose the parameters as follows $B_0 = c_1 \hat{B}$, $\hat{B} = (X'X)^{-1} X' Y$, $M_0 = c_2 X' X$, $S_0 = c_3 S$, $S = (Y - XB)' (Y - XB) \nu_0 = T - 2(N + 1)$, where $\hat{B}$ is in-sample estimation coefficients, $c_1, c_2, c_3$ are hyper-parameters for the distributions. In particular, let $c_1 = 1$, $c_2 = 0$, these priors evolve into Zellner’s $g$ -prior, which is midway between the informative and diffuse priors. As $c_2$ varies, we get a class of Zellner’s $g$ -priors, therefore we define our $\varepsilon$-contamination class as $Q_N = \{ \pi(B \mid \Sigma) \pi(\Sigma) \mid c_2 \in \mathbb{R}^+ \}$ which is a subset of $Q_{SU}$. Note that $Q_N$ is a rather small prior space. Therefore, our analysis provides a lower bound for prior robustness.

We can derive the predictive pdf $p_q(r_{T+1} \mid \Phi_T)$ with prior $q \in Q_N$, which is also a univariate Student’s $t$ distribution:
\[ p_q(r_{T+1} \mid \Phi_T) = \frac{\Gamma[(v+1)/2]}{\Gamma(1/2)\Gamma(v/2)} \left( \frac{1}{(v-2)\sigma_T^2} \right)^{1/2} \left[ 1 + \left( \frac{1}{v-2} \right) \left( \frac{r_{T+1} - \mu_T}{\sigma_T} \right)^2 \right]^{-\frac{(v+1)/2}{2}} \]  

(15)

with the sample distribution parameters except the variance and degree of freedom, which are given below:

\[ \sigma_T^2 = \frac{T}{2T-N-3}(1-R^2) \left[ 1 + c_v + \frac{1}{T}(1+q) \right] \hat{\sigma_T}^2 \]  

(16)

\[ v = 2T-N-1 \]  

(17)

Other parameters are the same as in Equations (5)–(8). After we obtain the posterior under \( \pi_0 \) and \( q \), we can easily calculate the combined posterior under \( \pi \) according to Berger (1985, p. 206) by weighing the two priors:

\[ p_\pi(r_{T+1} \mid \Phi_T) = (1-\varepsilon^*)p_{\pi_0}(r_{T+1} \mid \Phi_T) + \varepsilon^* p_q(r_{T+1} \mid \Phi_T), \quad \varepsilon^* = \frac{m(\Phi_T \mid q)}{m(\Phi_T \mid \pi)} \]  

(18)

where \( m(\Phi_T \mid \pi) \) denotes the marginal likelihood of data \( \Phi_T \) under prior \( \pi \), and the weight \( \varepsilon^* \) is a constant given \( \varepsilon, q, \) and \( \pi_0 \). Note that \( p_\pi(r_{T+1} \mid \Phi_T) \) is well defined even with non-informative priors because the priors only need to be specified in the correct proportion.

**Measuring the effect of Bayesian robustness**

We first graphically illustrate *local robustness* (changing \( \varepsilon^* \) while fixing \( q \) or \( c_2 \)) in Figure 1 and *global robustness* (changing \( c_2 \) while fixing \( \varepsilon^* \)) in Figure 2 of KS’s results. Figures 1 and 2 plot the optimal portfolio rules \( w^* \) and \( \Delta CER \) for different combinations of \( A \) and \( \delta \). In Figure 1, we fix \( c_2 = 3 \) so that both the default prior of KS and the contamination \( q \) have the same variance. Spanning \( \varepsilon^* \) over \([0,1]\), both optimal asset allocation \( w^* \) and gains in certainty equivalent returns \( \Delta CER \) are relatively stable, which indicates that KS’s results are locally robust. In Figure 2, we fix \( \varepsilon^* = 0.2 \), meaning that there is 20% probability that the default prior is misspecified. Spanning \( c_2 \) over \( R^+ \) (the results converge well after \( c_2 = 100 \)), we observe pronounced variations in \( \Delta CER \) (as well as in \( w^* \), not reported). For \( R^2 = 2.5\%, A = 5 \) and \( \delta = -1 \), \( \Delta CER \) varies from 12.8 basis points to 2.2 basis points. The variation is more significant for lower risk aversion \( A \). Hence, KS’s results are less robust in a global sense.

We need a metric to measure the robustness of inference using Bayesian methods. For the portfolio selection problem we consider in this paper, two meaningful robustness measures can be applied. The first measure is the *range* of the \( \Delta CER \)

\[ \text{Range} = \left[ \inf_{\pi \in \Gamma} \Delta CER(w^*, \pi, \Phi_T), \sup_{\pi \in \Gamma} \Delta CER(w^*, \pi, \Phi_T) \right] \]  

(19)

The second measure is a relative sensitivity measure or *ratio* measure:

\[ \text{Ratio}^{\max} = \sup_{\pi \in \Gamma} \frac{(\Delta CER(w^*, \pi, \Phi_T) - \Delta CER(w^0, \pi_0, \Phi_T))^2}{V^\pi} \]  

(20)
where $V^\pi$ is the posterior variance of $\Delta CER(w^*, \pi, \Phi_T)$. (In our analysis, the minimum of the ratio is essentially zero, therefore we only report the maximum.) If $\text{Range}$ is wide or $\text{Ratio}_{\text{max}}$ is large, we say the results are not Bayesian robust. Berger (1994) shows that both measures have acceptably desired asymptotic properties.

Figure 1. Local robustness. Plot of the gains in certainty equivalent returns $\Delta CER$ (in basis points) against contamination $\varepsilon^*$ for $R^2 = 2.5\%, 5.5\%$ and $A = 1, 2, 5$ (the six scenarios considered by Kandel and Stambaugh (1996)). Here $c_2 = 3$ is chosen so that the mixed and the default predictive probability distribution function have the same variance.
We report the two robustness measures for KS’s analysis in Table 1. The results reveal that the economic significance of return predictability varies substantially across different prior choices. For the Range measure, we see that the range is wide and KS’s results are often close to the maximum. For $R^2 = 5.5\%$, $A = 2$ and $\delta = -1$, the range is $[13.0, 61.3]$ while KS’s result is 55.8. The ratio

![Figure 2](image_url)

**Figure 2.** Global robustness. Plot of the gains in certainty equivalent returns $\Delta CER$ (in basis points) against contamination $c_2$ for $R^2 = 2.5\%, 5.5\%$ and $A = 1, 2, 5$ (the six scenarios considered by Kandel and Stambaugh (1996)). Here $c^* = 0.2$, meaning there is a 20% probability that the base prior is misspecified.

We report the two robustness measures for KS’s analysis in Table 1. The results reveal that the economic significance of return predictability varies substantially across different prior choices. For the Range measure, we see that the range is wide and KS’s results are often close to the maximum. For $R^2 = 5.5\%$, $A = 2$ and $\delta = -1$, the range is $[13.0, 61.3]$ while KS’s result is 55.8. The ratio
measure can be better understood as comparable of the $t$-statistic in frequentist statistics. The maximum ratio is at the level of 10 (comparable with a $t$-statistic of approximately 3.2), which indicates substantial non-robustness. More importantly, the magnitude of the minimum $\Delta CER$ is often small in an economic sense. For example, when $A = 2$, and when predicted value is one standard deviation away from the sample mean, the utility gain from considering return predictability can be as small as 6 basis points per month, which is smaller than stock and exchange traded fund (ETF) bid–ask spreads.

The economic significance for delegated portfolio management

In this section we explore the economic implications of Bayesian robustness. If a fund manager has a different prior belief from fund investors, the portfolio chosen by fund manager will be viewed as sub-optimal by fund investors. Investors managing their own portfolios can also choose a suboptimal portfolio if their priors are misspecified (e.g. their prior beliefs will be different regarding different data sets).

If the manager or investor ignores the robustness issue and treats $\pi_0$ as if it were the only possibility, he will employ a suboptimal portfolio rule $w^0$ and obtain $\Delta CER(w^0, \pi_0, \Phi_T)$. But the prior can be any one in the class $\pi \in \Gamma$, the corresponding optimal portfolio rule is $w^*$ and resulting $\Delta CER(w^*, \pi, \Phi_T)$. Along the same line of KS, we use

$$Cost^{\text{max}} = \sup_{\pi \in \Gamma} \left[ \Delta CER(w^*, \pi, \Phi_T) - \Delta CER(w^0, \pi, \Phi_T) \right]$$

(21)

to measure the economic significance of robustness, that is, the utility cost the investor would suffer if he ignores this issue. We use the maximum rather than the mean, minimum or others to be consistent with investor’s uncertainty aversion in the Knightian model uncertainty literature. Note that the minimum of utility cost is almost always zero.

We report the economic significance of ignoring Bayesian robustness in Table 2 for the scenarios considered by KS. The utility cost is astounding, often exceeding 1% per month. The utility

<table>
<thead>
<tr>
<th>$R^2$</th>
<th>$p$-value</th>
<th>$A$</th>
<th>$Cost^{\text{max}}$ for $\delta$ equal to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$-1$</td>
</tr>
<tr>
<td>0.025</td>
<td>0.75</td>
<td>1</td>
<td>25.0</td>
</tr>
<tr>
<td>0.025</td>
<td>0.75</td>
<td>2</td>
<td>12.3</td>
</tr>
<tr>
<td>0.025</td>
<td>0.75</td>
<td>5</td>
<td>4.9</td>
</tr>
<tr>
<td>0.055</td>
<td>0.01</td>
<td>1</td>
<td>10.8</td>
</tr>
<tr>
<td>0.055</td>
<td>0.01</td>
<td>2</td>
<td>5.4</td>
</tr>
<tr>
<td>0.055</td>
<td>0.01</td>
<td>5</td>
<td>2.1</td>
</tr>
</tbody>
</table>
cost is higher during market booms than market downturns (most likely due to borrowing constraints). Furthermore, the economic significance of Bayesian robustness can be much larger than the economic significance of return predictability. For $R^2 = 2.5\%$, $A = 5$ and $\delta = 0.5$, ignoring return predictability has a cost of 2.9 basis points per month (shown in row seven of Table 1) while ignoring prior robustness can have a cost as large as 36.5 basis points per month. Our finding indicates that a portfolio manager could cost his investors substantially if he ignores prior robustness.

**Summary and concluding remarks**

Our main objective is to call for caution when applying Bayesian methods and making inferences solely based on Bayesian results in portfolio management. We show that, in the context of portfolio allocation with asset return predictability, Bayesian results are not robust with respect to prior selection. The economic significance of stock return predictability may well be within the bid–ask spread bounds for some prior choices. Our finding is consistent with Gray (2008) using Australian data and explains why in real data investors benefit little from data-based predictability. We also show that ignoring Bayesian robustness can have important economic consequences. We conclude that discretion needs to be taken when one applies Bayesian methods to portfolio allocation aiming to take advantage of return predictability.

Some may refute the relevance of Bayesian robustness by arguing that investors can form informative priors using financial theories. However, data-implied prior beliefs are often unrealistically restrictive. Financial theories alone cannot resolve prior selection issue. Therefore, we advocate alternative approaches for portfolio allocation such as the shrinkage approach of Wang (2005), the multi-prior approach of Garlappi et al. (2007) and the combination forecast of Rapach et al (2010), Kong et al (2011), and Tu and Zhou (2011).

**Acknowledgements**

I thank an anonymous referee as well as Lucy Ackert, David Chapman, Ron Kaniel, Gabriel Ramirez, Garry Twite, and Zhenyu Wang for helpful comments and discussions.

**Funding**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Notes**

1. Harvey et al (2008) illustrate how a minor change in the prior distribution and estimation can affect the results in an investment contest. Tu and Zhou (2010) and Wachter and Warusawitharana (2009) show that portfolio allocation outcomes can be sensitive to prior choice.
2. Informative priors are often employed in financial applications. Tu and Zhou (2010) demonstrate ways to incorporate economic objectives into informative priors.
3. Hong et al (2008) show that financial advisors can mislead investors into over-allocating to risky stocks.

**References**


Date of acceptance of final transcript: 28 August 2012.

Accepted by Associate Editor, Garry Twite (Finance).