A Panel Partial Break Model for Forecasting Stock Returns in the Presence of Parameter Instability

Nuttapat Bumrungrat* and Thaisiri Watewai

Department of Banking and Finance, Faculty of Commerce and Accountancy,
Chulalongkorn University, Bangkok 10330, Thailand
*Corresponding author, E-mail: bumnuttapat@gmail.com

Abstract

Predicting stock returns is one of the most fascinating problems in finance. However, predicting stock returns is also one of the most challenging problems due to the instability and noisy nature of stock returns. One of the promising directions to handle both problems is to use a panel break model. Recently, a panel common break model has been proposed and has been shown to generate superior predictive performance. However, the model assumes that every parameter breaks simultaneously, which is not aligned with empirical data. In this article, we propose a novel panel break model that addresses the main limitation of the common break model while still retaining its main advantage by allowing each type of parameter to break separately. Moreover, our model allows correlated breaks between each parameter type through a common hidden time-varying break probability. We evaluated its performance on the top 100 largest US stocks from January 2002 to December 2021. The results show that our model provides improved performance when stocks experience a series of extreme returns, as our model is quite sensitive to data. On one hand, this can be helpful for faster detection of high-impact breaks during crises. On the other hand, it can result in too many false detections. Further model restrictions and the use of more data may further improve the model's performance.

Keywords: *Bayesian Panel Break Model, Stock Return Prediction, Structural Break, Parameter Instability*

1. Introduction

Stock return prediction has long been a topic of interest. However, many predictive models suffer from poor out-of-sample performances. Model instability is one of the main problems with such models. A series of models have been proposed to overcome this problem. One difficulty is detecting breaks in the model parameters in a timely and accurate manner. In this paper, we propose a novel predictive model that improves on the existing models.

Predicting stock returns is one of the most fascinating problems in finance. For practitioners, predicted stock returns could serve as inputs for making diverse financial decisions. The problem is also highly relevant for academics. A deeper understanding of the nature of stock return predictability could guide researchers to discover a more realistic market equilibrium model, which could lead to a more effective test for market efficiency.

However, predicting stock returns is also one of the most challenging problems. Although a large body of literature suggests that a variety of financial and economic variables could be used to predict stock returns ex-post, Welch and Goyal (2008) show that forecasting stock returns based on a variety of the earlier-proposed financial and economic variables using a constant parameter predictive regression model fails to reliably deliver superior out-of-sample performance relative to the simple historical average benchmark in terms of the mean squared prediction error.

The model's poor out-of-sample performance could be attributed to instability in the predictive model. Various studies have documented evidence of parameter instability in a predictive regression for a wide range of predictors (Farmer et al., 2023; Georgiev et al., 2018; Lettau and Van Nieuwerburgh, 2008; Pettenuzzo and Timmermann, 2011; Pitarakis, 2017; Rossi, 2021; Smith and Timmermann, 2022; Tu and Xie, 2023; Zhu et al., 2022).

Building on the growing evidence of instability, many researchers have proposed a variety of approaches to formally account for instability to, hopefully, find a successful predictive model (Dangl and Halling, 2012; Farmer et al., 2023; Guidolin and Timmermann, 2007; Henkel et al., 2011; Johannes et al., 2014; Smith and Timmermann, 2021; Smith and Timmermann, 2022; Tu and Xie, 2023). However, most of the early models utilize only time-series information, making it quite hard to timely and precisely detect breaks in the noisy environment of stock markets.

One promising way to overcome this problem is to use a panel break model to exploit both time-series and cross-sectional information simultaneously. Smith and Timmermann (2021) introduce a panel break model, which assumes that a break can simultaneously affect every parameter of every stock in the sample. However, their result also shows that a break in any parameters in the model tends to force the model to identify that all of the other parameters are also hit by the break on that date, resulting in detecting the excessive number of breaks. This suggests that the assumption that a break simultaneously hits every parameter might be too strict.

The main contribution of this article is to propose a novel panel break model that addresses the main limitation of the common break model while still retaining its main advantage. We achieve this by partially relaxing the common break assumption of Smith and Timmermann (2021). Specifically, we assume that each type of parameter can break separately. However, we still assume that parameters of the same type for every stock break simultaneously. Moreover, our model allows breaks in each type of parameter to be dependent on each other by assuming that every type of parameter shares the same hidden time-varying break probability, allowing information regarding break dates between each type of parameter to be utilized.

2. Objectives

The objectives of this study are

1) To propose a novel model that addresses the main limitation of the Smith and Timmermann (2021) model,

2) To evaluate the proposed model's out-of-sample predictive performance on large US stocks.

3. Materials and Methods

3.1 Model

We propose a Bayesian panel change-point model that relaxes the common break assumption across all parameters of Smith and Timmermann (2021) while still retaining the benefit of assuming that a break hits all stocks simultaneously to utilize both cross-sectional and time-series information. In this model, we assume that every type of regression parameter shares the same time-varying break probability. Parameters of the same type from every stock are assumed to always break together, while parameters from different types are assumed to break separately. Furthermore, we follow Smith and Timmermann (2021) by introducing the observable common factor f_t to model cross-sectional dependency among individual stocks. This allows us to parsimoniously account for cross-sectional covariance instead of relying on a highly parameterized approach that directly assumes a full covariance matrix. To be precise, we assume that the excess return of stock i at time t, or $r_{i,t}$, follows the following equation:

$$r_{i,t} = \theta_{i,0,t} + \sum_{j=1}^{p} \theta_{i,j,t} x_{j,t-1} + \theta_{i,f,t} f_t + \sqrt{\theta_{i,\sigma,t}} \epsilon_{i,t}$$

where $x_{j,t-1}$ is the j^{th} predictor at time $t-1$, $\epsilon_{i,t} \sim N(0,1)$ is the independent error term of stock i at time t, and $\theta_{i,j,t}$ is the type-j regression parameter of stock i at time t.

Following Maheu and Song (2018), breaks in type-j parameters are parameterized using the regime duration $d_{j,t}$. The time when the regime duration is reset to one, $d_{j,t} = 1$, corresponds to the time when type-j parameters break and, thus, the new type-j parameters are drawn from the type-specific distribution F_j, while the time when the regime duration is increased by one, $d_{j,t} = d_{j,t-1} + 1$, corresponds to the time when type-j parameters do not break and, hence, the values of the parameters remain the same for another period. More specifically, the type-j regression parameter of stock i evolves as follows:

$$\theta_{i,j,t} = \begin{cases} \sim F_j, & if\ \ d_{j,t} = 1 \\ \theta_{i,j,t-1}, & if\ \ d_{j,t} = d_{j,t-1} + 1 \end{cases}.$$

At time t, each type-j parameter shares the common *hidden* break probability p_{s_t}, where the state variable s_t is either 1 with probability p^* or 2 with probability $1 - p^*$. By assuming that the break probability p_{s_t} explicitly depends on the hidden state s_t, we allow correlated breaks across parameter types.

To complete the Bayesian model, we specify the following prior distributions:

$$F_j = N\left(\underline{u}_j, \underline{v}_j\right), \underline{u}_j \in R, \underline{v}_j > 0, \quad for \quad j \in 0,1, \dots, p, f, \underline{F_\sigma} = IG\left(\underline{a}_1, \underline{b}_1\right), \underline{a}_1, \underline{b}_1 > 0, p_1, p_2 \\ \sim Beta\left(\underline{a}_2, \underline{b}_2\right), \underline{a}_2, \underline{b}_2 > 0, p^* \sim Beta\left(\underline{a}_3, \underline{b}_3\right), \underline{a}_3, \underline{b}_3 > 0,$$

where the underlined variables denote the hyper-parameters of the prior distributions.

To summarize, the parameters to be chosen by the authors are $\{\underline{u}_j, \underline{v}_j : j \in \{0,1, \dots, p, f, \sigma\}\}$ and $\{\underline{a}_j, \underline{b}_j : j \in \{1,2,3\}\}$, and the parameters and hidden variables to be estimated are

$$\Theta \coloneqq \bigcup_{j \in \{0,1,\dots,p,f,\sigma\}} \Theta_j, D \coloneqq \bigcup_{j \in \{0,1,\dots,p,f,\sigma\}} D_j,$$

$$S \coloneqq \{s_t : t \in [2, T]\},$$

$$\mathrm{P} \coloneqq \{p_1, p_2, p^*\}$$

where $\Theta_j \coloneqq \{\theta_{i,j,t} : i \in \bigcup_\tau I_\tau, t \in B_j\}$ is the set of distinct type-j parameters from every regime and every existing stock, $B_j \coloneqq \{\tau \leq T : d_{j,\tau} = 1\}$ is the set of break dates of type-j parameters, $D_j \coloneqq \{d_{j,t} : t \in [2, T]\}$, T is the time the model is estimated, and I_t is the set of stocks that can be traded at time t.

3.2 Estimation

The model can be estimated using a Gibbs sampler, which is a Markov chain Monte Carlo (MCMC) technique. Each set of parameters or unobserved variables is sequentially drawn from their conditional distribution, given all the other parameters and variables. The process is repeated to obtain a sufficiently large sample converging to their posterior distribution. Here, the predictors and the observable common factor are treated as exogenous variables and are implicitly given in every expression below. The Gibbs sampler consists of the following steps:

1) For $j \in \{0, 1, \dots, p, f, \sigma\}$, sample D_j and Θ_j from $P\left(Y_T, \Theta_{\setminus j}, D_{\setminus j}, S, \mathrm{P}\right)$, where $D_{\setminus j} \coloneqq D \setminus D_j, \Theta_{\setminus j} \coloneqq \Theta \setminus \Theta_j$, and Y_T is the set of available excess returns up to time T. The sampling method from Maheu and Song (2018) can be employed. This step can be done by first sampling each $d_{j,t}$ from $P\left(Y_T, \Theta_{\setminus j}, D_{\setminus j}, S, \mathrm{P}, D_j^{t+1}\right)$, where

Proceedings of RSU International Research Conference (RSUCON-2024)

$$D_j^{t+1} := \{d_{j,t+1}, d_{j,t+2}, \dots, d_{j,T}\},$$

$$P\left(Y_T, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ, D_j^{t+1}\right) = \{P\left(Y_t, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right) \; if \; d_{j,t+1} = 1 \qquad 1_{\{k=d_{j,t+1}-1\}} \; if \; d_{j,t+1} \geq 2 \,,$$

and $\mathbb{1}$ denotes the indicator function. $P\left(Y_t, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right)$ can be calculated forward in time. Given that $P\left(Y_{t-1}, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right)$ is already known,
$P\left(Y_t, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right)$ can be calculated in two steps as follows:

$$P\left(Y_t, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right) \propto P\left(d_{j,t} = k, Y_{t-1}, \Theta_{\setminus j}, D_{\setminus j}\right) \times P\left(Y_{t-1}, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right),$$

$$P\left(Y_{t-1}, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right) = \sum_{l} \quad P\left(d_{j,t-1} = l, s_t, p_{s_t}\right) P\left(Y_{t-1}, \Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right),$$

where $P\left(\Theta_{\setminus j}, D_{\setminus j}, S, Ƥ\right) = 1$.

For $j \in \{0,1,\dots,p,f\}$, $P\left(d_{j,t} = k, Y_{t-1}, \Theta_{\setminus j}, D_{\setminus j}\right) =$

$$\prod_i \quad \frac{1}{\sqrt{2\pi\theta_{i,\sigma,t} v_{i,j,t}\left(\frac{x_{j,t}^2}{\theta_{i,\sigma,t}} + \frac{1}{v_{i,j,t}}\right)}} exp\left\{-\frac{1}{2}\left(\frac{\left(r_{i,t} - \mu_{i,\setminus j,t}\right)^2}{\theta_{i,\sigma,t}} + \frac{u_{i,j,t}^2}{v_{i,j,t}} - \frac{\left(\frac{\tilde{x}_{j,t}\left(r_{i,t} - \mu_{i,\setminus j,t}\right)}{\theta_{i,\sigma,t}} + \frac{u_{i,j,t}}{v_{i,j,t}}\right)^2}{\frac{x_{j,t}^2}{\theta_{i,\sigma,t}} + \frac{1}{v_{i,j,t}}}\right)\right\},$$

$$P\left(d_{\sigma,t} = k, Y_{t-1}, \Theta_{\setminus\sigma}, D_{\setminus\sigma}\right) = \prod_i \quad \frac{1}{\sqrt{2\pi}} \frac{b_{i,t}^{a_{i,t}}}{\Gamma\left(a_{i,t}\right)} \frac{\Gamma\left(a_{i,t} + \frac{1}{2}\right)}{\left(\frac{1}{2}\left(r_{i,t} - \mu_{i,t}\right)^2 + b_{i,t}\right)^{a_{i,t}+\frac{1}{2}}},$$

$$u_{i,j,t} := \frac{\sum_{\tau \in Ŧ_{i,j,t}} \quad \frac{\tilde{x}_{j,\tau}\left(r_{i,\tau} - \mu_{i,\setminus j,\tau}\right)}{\theta_{i,\sigma,\tau}} + \frac{\underline{u}_j}{\underline{v}_j}}{\sum_{\tau \in Ŧ_{i,j,t}} \quad \frac{\tilde{x}_{j,\tau}^2}{\theta_{i,\sigma,\tau}} + \frac{1}{\underline{v}_j}},$$

$$v_{i,j,t} := \frac{1}{\sum_{\tau \in Ŧ_{i,j,t}} \quad \frac{\tilde{x}_{j,\tau}^2}{\theta_{i,\sigma,\tau}} + \frac{1}{\underline{v}_j}},$$

$$a_{i,t} := \frac{\left|Ŧ_{i,\sigma,t}\right|}{2} + \underline{a}_1,$$

$$b_{i,t} := \frac{1}{2} \sum_{\tau \in Ŧ_{i,\sigma,t}} \quad \left(r_{i,\tau} - \mu_{i,\tau}\right)^2 + \underline{b}_1,$$

$$\tilde{x}_{j,t} := \{x_{j,t-1} \; if \quad j \neq f \qquad f_t \; if \quad j = f \quad ,$$

$$\mu_{i,\backslash j,t} \coloneqq \mu_{i,t} - \theta_{i,j,t}\tilde{x}_{j,t}, \mu_{i,t} \coloneqq \theta_{i,0,t} + \sum_{j=1}^{p} \theta_{i,j,t}x_{j,t-1} + \theta_{i,f,t}f_t,$$

$x_{0,t}$ is equal to one for any t, $r_t \coloneqq \{r_{i,t}: i \in I_t\}$, and $Ŧ_{i,j,t} \coloneqq [t+1-d_{j,t}, t-1] \cap \{\tau \leq t-1: i \in I_\tau\}$ is the set of the time periods from the latest break of type-j parameters to time $t - 1$ when stock i is tradable.

Then, sample Θ_j from $P(Y_T, \Theta_{\backslash j}, D)$, where for $j \in \{0,1,\dots,p,f\}$, $P(Y_T, \Theta_{\backslash j}, D) =$

$$\prod_i \prod_{t \in B_j} N\left(\frac{\sum_{\tau \in Ŧ_{i,j,t}} \frac{\tilde{x}_{j,\tau}(r_{i,\tau} - \mu_{i,\backslash j,\tau})}{\theta_{i,\sigma,\tau}} + \frac{\underline{u}_j}{\underline{v}_j}}{\sum_{\tau \in Ŧ_{i,j,t}} \frac{\tilde{x}^2_{j,\tau}}{\theta_{i,\sigma,\tau}} + \frac{1}{\underline{v}_j}}, \frac{1}{\sum_{\tau \in Ŧ_{i,j,t}} \frac{\tilde{x}^2_{j,\tau}}{\theta_{i,\sigma,\tau}} + \frac{1}{\underline{v}_j}}\right),$$

$$P(Y_T, \Theta_{\backslash \sigma}, D) = \prod_i \prod_{t \in B_\sigma} IG\left(\frac{|Ŧ_{i,\sigma,t}|}{2} + \underline{a}_1, \frac{1}{2}\sum_{\tau \in Ŧ_{i,\sigma,t}} (r_{i,\tau} - \mu_{i,\tau})^2 + \underline{b}_1\right),$$

and $Ŧ_{i,j,t} \coloneqq \{\tau \leq T: \tau + 1 - d_{j,\tau} = t\} \cap \{\tau \leq T: i \in I_\tau\}$ is the set of the time periods when stock i is tradable and uses the same type-j parameter as that at time t.

2) Sample S from $P(Y_T, \Theta, D, \mathrm{P})$. This can be done by sampling each s_t from the distribution that is proportional to $\left(p_{s_t}\right)^{|\{j: d_{j,t}=1\}|}\left(1 - p_{s_t}\right)^{|\{j: d_{j,t}\neq 1\}|}(p^*)^{2-s_t}(1-p^*)^{1-(2-s_t)}$.
3) Sample p_1 and p_2 from $P(Y_T, \Theta, D, S, p^*)$ by sampling each p_k from $Beta\left(\sum_{\{t: s_{t+1}=k\}} |\{j: d_{j,t+1} = 1\}| + \underline{a}_2, \sum_{\{t: s_{t+1}=k\}} |\{j: d_{j,t+1} \neq 1\}| + \underline{b}_2\right)$.
4) Sample p^* from $P(Y_T, \Theta, D, S, p_1, p_2) = Beta\left(|\{t: s_{t+1} = 1\}| + \underline{a}_3, |\{t: s_{t+1} = 2\}| + \underline{b}_3\right)$.

After repeating the above steps a sufficiently large number of times, the resulting sample of $(\Theta, D, S, \mathrm{P})^{(n)}$ after the burn-in period can be used to approximate the posterior distribution$P(Y_T)$. Derivations for each Gibbs sampler step are available upon request.

3.3 Prior elicitation

The hyper-parameters for our empirical application that have their counterparts in Smith and Timmermann (2021) are adapted from them. The other hyper-parameters are chosen to represent non-informative priors. The adapted hyper-parameters are as follows:

$$\underline{u}_j = 0 \quad for \quad j \in \{0,1,\dots,p,f\},$$

$$\underline{v}_0 = 0.05^2 \times \hat{\sigma}_f^2,$$

$$\underline{v}_j = \frac{0.04^2}{\hat{\sigma}_j^2}\hat{\sigma}_f^2 \quad for \quad j \in \{1,\dots,p,f\},$$

$$\underline{a}_1 = 2,$$

$$\underline{b}_1 = 0.0049,$$

where $\hat{\sigma}_f^2$ and $\hat{\sigma}_j^2$ are the sample variance of the excess return of the market and that of the j^{th} predictor using the available data at the time the model is estimated, respectively. The other parameters, including $\underline{a}_2, \underline{b}_2, \underline{a}_3$, and $\underline{b}_3$, are chosen to be equal to one.

3.4 Evaluation

The benchmark model is the equal-weighted combined forecast used by Smith and Timmermann (2021) using the four single-predictor common break models. This model will be referred to as the cb-avg.

The proposed model is evaluated based on its out-of-sample predictive ability. The out-of-sample period is from January 2002 to December 2021. During this period, the model is re-estimated monthly using the available data of the US stocks listed on the NYSE, AMEX, or NASDAQ that are the top 100 largest stocks by market capitalization at the time from January 1982 to the estimation date. Then, the estimated parameters will be used to form predictions of the excess returns of the corresponding stocks in the next month. Excess returns are computed as returns less the risk-free rate. Specifically, at each time t, the excess return prediction for stock i is generated as follows:

$$\hat{r}_{i,t+1} = \hat{\theta}_{i,0,t} + \sum_{j=1}^{p} \hat{\theta}_{i,j,t} x_{j,t} + \hat{\theta}_{i,f,t} \underline{f_t},$$

where $\hat{r}_{i,t+1}$ is the predicted excess return of stock i at time $t+1$, $\underline{f_t}$ is the average excess return of the common factor using the data up to time t, and $\hat{\theta}_{i,j,t}$ is the average of the type-j parameter of stock i at time t from the MCMC sample after the burn-in period.

When required, the predicted covariance matrix $\hat{C}_{t+1} = [\hat{c}_{t+1,i,i'}]$ is generated as follows:

$$\hat{c}_{t+1,i,i'} = \hat{\theta}_{i,f,t} \hat{\theta}_{i',f,t} \hat{\sigma}_{f,t}^2 + 1_{\{i=i'\}} \hat{\theta}_{i,\sigma,t},$$

where $\hat{c}_{t+1,i,i'}$ is the element on the i^{th} row and the i'^{th} column of $\hat{C}_{t+1}$ and $\hat{\sigma}_{f,t}^2$ is the sample variance of f computed at time t. The covariance matrix for the cb-avg model is the average of the covariance matrices from the four single-predictor common break models, where each covariance matrix of each common break model is calculated using the above equation.

The model will be statistically and economically evaluated. First, the statistical performance will be visually assessed using the modified cumulative sum of squared error difference (MCSSED):

$$MCSSED_t = \sum_{\tau=1}^{t} \sum_{i} \left[\left(r_{i,\tau} - \hat{r}_{Bmk,i,\tau}\right)^2 - \left(r_{i,\tau} - \hat{r}_{Alt,i,\tau}\right)^2 \right],$$

where $MCSSED_t$ is the modified cumulative sum of squared error difference at time t, and $r_{Bmk,i,\tau}$ and $r_{Alt,i,\tau}$ are the predicted excess return of stock i at time τ from the benchmark model and the evaluated model, respectively. The MCSSED is extended from the cumulative sum of squared error difference (CSSED) proposed in Welch and Goyal (2008) so that it can handle the case where multiple stocks are evaluated while still retaining the key feature of the CSSED. The MCSSED places an equal weight on every stock. This is suitable if each stock in the investment universe is relatively equally investable, which holds true in our application since our investment universe contains only large market capitalization stocks. The rising value of the MCSSED implies that the evaluated model is outperforming the benchmark model in that particular period, while the declining value implies that the evaluated model is underperforming the benchmark model in that period. This visual tool can reveal whether the relative performance of the evaluated model is stable throughout the entire sample or dominated by a specific set of observations. Moreover, it can uncover which periods the evaluated model outperforms and underperforms the benchmark model. It also facilitates answering how changing starting or ending dates affects the overall squared prediction error, which can be done by shifting the horizontal axis to the desired position. Then, the statistical performance will be formally evaluated using the test statistic of Clark and West (2007) that allows for comparing nested models (the CW statistic hereafter). To evaluate the economic performance of the model, we follow earlier studies in the stock return predictability literature by using the utility gain of a small, myopic mean-variance investor with a risk-averse coefficient equal to 3 (Campbell and Thompson, 2008; Dangl and Halling, 2012; Rapach et al., 2010; Smith and Timmermann, 2021). This utility gain can also be interpreted as the certainty equivalent return (CER). Hence, the difference in the utility gain between using the evaluated model and the benchmark model to guide portfolio decisions can be interpreted as the maximum cost that the investor is willing to pay to access the information in the evaluated model.

3.5 Data

We use monthly excess returns from January 1982 to December 2021 of the US stocks listed on the NYSE, AMEX, or NASDAQ that are the top 100 largest stocks by market capitalization in some months during the period from January 2002 to December 2021. The return data are obtained from the CRSP database, while the risk-free rate data are obtained from the updated data of Welch and Goyal (2008) from Amit Goyal’s website (Goyal, n.d.). Following Smith and Timmermann (2021), we use the monthly predictors from Welch and Goyal (2008). The predictors include the aggregate dividend-price ratio (dp), the Treasury-bill rate (tbl), the term spread (tms), and the default spread (dfy). All the predictors are constructed as in Welch and Goyal (2008). To be more precise, dp is the log of 1-year moving sums of dividends paid on the S&P 500 index minus the log of the index's prices; tbl is the 3-month Treasury Bill rates; tms is the yields on long-term government bonds minus the Treasury-bill rates; and dfy is the yields on BAA-rated corporate bonds minus the yields on AAA-rated corporate bonds. The observable common factor is the value-weighted market portfolio’s monthly excess return. The predictor and the market portfolio data are obtained from the updated data on Amit Goyal’s website. Welch and Goyal (2008) used the original dataset, and the website provides its updated version. The datasets consist of the market portfolio’s returns, the US risk-free rate, and the data to construct the predictors.

4. Results and Discussion

4.1 Results

Figure 1 shows the MCSSED of the proposed model relative to the cb-avg model over the out-of-sample period. As we can see, the plot has a general downward trend, suggesting that the proposed model consistently underperforms the cb-avg model for most of the out-of-sample period except only during mid-2009, when the plot has an upward trend. This result seems to suggest that our model tends to perform better during a crisis period. In Section 4.2, we investigate this observation in more detail.

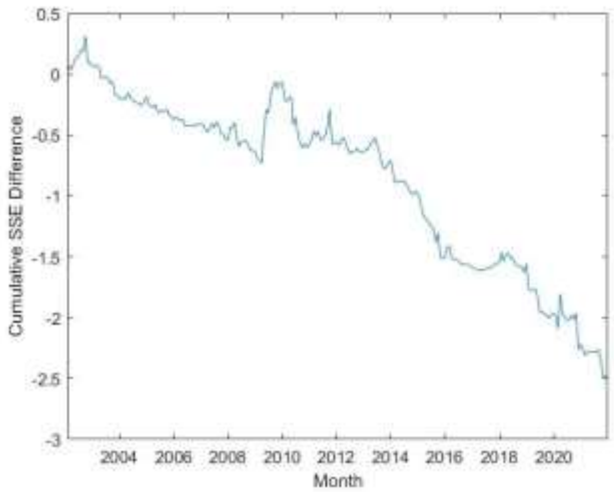

Figure 1: The MCSSED of the proposed model relative to the cb-avg model

Next, the statistical performances are formally evaluated using the CW statistic. To lower the chance of having spurious results, we follow Smith and Timmermann (2021) by using only stocks with at least 60 data points during the out-of-sample period to evaluate results. This includes a total of 130 stocks. Table 1 shows the number of stocks that have values of the test statistic within each interval, with the percentage of the total number of stocks shown in parentheses below. Each interval can be interpreted as the proposed model significantly underperforms ($t \leq -1.64$), insignificantly underperforms ($-1.64 < t \leq 0$), insignificantly outperforms ($0 < t \leq 1.64$), and significantly outperforms ($t > 1.64$) a benchmark. From the table, it appears that the proposed model outperforms the cb-avg model for only around 29% of the stocks with no statistically significant result, while the benchmark outperforms our model by around 71%, and only 13% of the stocks are statistically significant. Consistent with the result from Figure 1, the benchmark delivers superior overall performance. Note, however, that these small proportions of significant results suggest that more data might be needed.

Table 1: Distribution of the CW statistics

Benchmark	$t \leq -1.64$	$-1.64 < t \leq 0$	$0 < t \leq 1.64$	$t > 1.64$
cb-avg	17 (13.08%)	75 (57.69%)	38 (29.23%)	0 (0.00%)

Next, the economic performance is evaluated. The proposed model faces a sizeable utility loss against the cb-avg model at -2.49% per annum. This result is expected since the benchmark model provides better overall prediction accuracy in both the time dimension (MCSSED) and the stock dimension (CW statistics).

4.2 Discussion

To explain why our model only performed better in 2009, we performed a detailed analysis by splitting the predictions into six subgroups. Each subgroup consists of the predictions of every stock when the previous values of the excess returns of the corresponding stock are less than or equal to -10%, from -10% to -5%, from -5% to 0%, from 0% to 5%, from 5% to 10%, and greater than 10%, respectively. Figure 2 shows the comparison of the average predictions from both models in each subgroup. It appears that our model tends to predict larger excess returns in magnitude, as it tends to generate lower (higher) predictions after having observed negative (positive) excess returns. This implies that our model's predictions are more sensitive to data than the predictions from the benchmark. On the one hand, our model might predict less accurately when there are consecutive small returns. On the other hand, our model seems to be better at predicting a series of extreme value excess returns that exhibit a consistent directional trend. This can be confirmed from Figure 3, which shows the alternative MCSSED, which cumulates the sum squared error differences of the months in which there are at most a certain number of stocks exhibiting extreme value excess returns (±10%) in that month and the previous month with a consistent direction. It appears that the more stocks showing consistent directional extreme returns, the better our model performs, while the fewer stocks showing the behavior, the worse the model performs. Since months with a relatively large number of stocks exhibiting the said behavior only concentrate in 2009, this explains why our model performed better during 2009 but worse during the other periods.

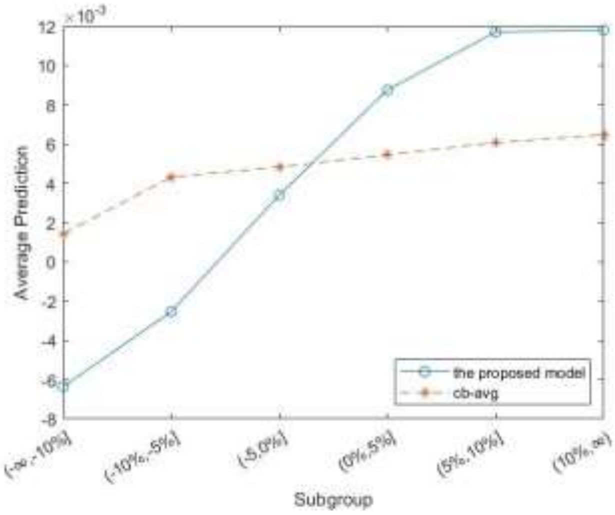

Figure 2: The average predictions from both models in each subgroup

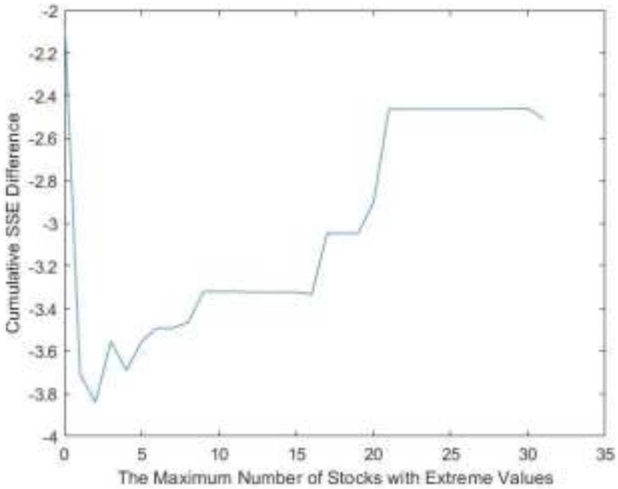

Figure 3: The alternative MCSSED

Next, Figure 4 displays the averaged break rates of type-1 parameters, which are equal to the regression coefficient of the aggregate dividend-price ratio (dp) predictor, for each estimation date. The bluer the color, the higher the chance of having a break on that corresponding date, while the lighter the color, the lesser the chance of having a break. The y-axis is the estimation date, and the x-axis is the date of the corresponding break probabilities. The model is deemed to detect a break on a certain date if it assigns a noticeable break probability on that date. Thus, a long-blue-vertical stripe indicates that a detected break on a certain date is supported by subsequent data, while an isolated blue dot or short-lived stripe indicates that the detected break is not supported by subsequent data. A break on each date t can be detected for the first time when it is included in the estimation period. Those dates correspond to the ones on the leftmost edge of the upper black triangle (diagonal line). The figure shows that breaks are typically quickly detected once the dates are included in the estimation sample, as there are many dark blue colors appearing close to the diagonal line. However, many of these breaks seem quite short-lived, indicating that they are false alarms. Such behavior may lead to predictions that are too sensitive to new data points, which can explain our model's overall underperformance. Nevertheless, such a quick detection can be very useful at the start of a crisis when breaks are expected to happen to many parameters. This observation supports our analysis above. Similar results are found for the other parameter types, except for type-f parameters, where no break is noticeable (Bumrungrat, 2024).

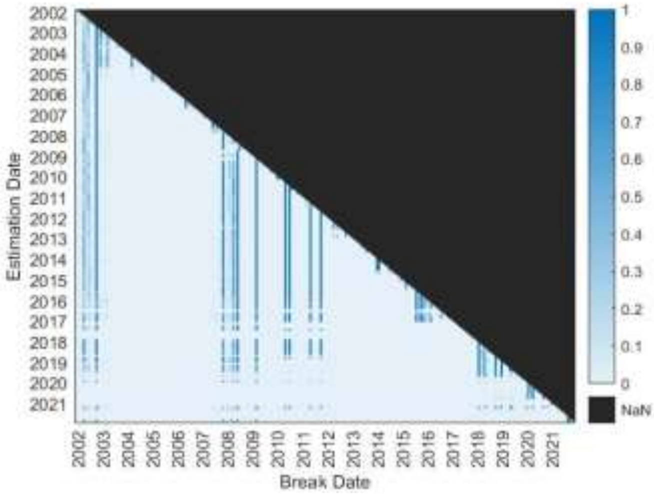

Figure 4: The smoothed break probabilities of type-1 parameters

While Smith and Timmermann (2021) force all the parameters to break simultaneously, we relax their assumption of common breaks and allow for partial breaks. The benefit of our relaxation over the model of Smith and Timmermann (2021) is faster break detection, which can be very useful at the start of a crisis. However, it causes an overfitting problem and leads to predictions that are overly sensitive to new data points, and hence lowers prediction accuracy. Since crises do not happen frequently, our model provides poorer performance over the full-sample period. One possible way to further improve the model is to use more data to fit the model. However, the current estimation method is computationally costly, making it expensive to use more data. A balance between our model and the model of Smith and Timmermann (2021) may improve the model's performance. One may assume no breaks for certain types of model parameters or assume that a subset of parameter types breaks together. We leave this for future work.

5. Conclusion

In this study, we propose the novel panel break model, addressing the limitations of the common break model of Smith and Timmermann (2021). Our model allows each type of parameter to break separately while still being able to utilize information from all stocks to quickly detect breaks.

We evaluate the model's performance by performing expanding-window out-of-sample predictions on the top 100 largest US stocks by market capitalization during the period from January 2002 to December 2021. The results show that our model improves the prediction results when stocks experience a series of extreme returns. This improved performance was clearly noticeable during the 2009 crisis. However, for the other periods, a large number of the detected breaks are quite short-lived and can be deemed false alarms. This signals overfitting. To further improve the model, we suggest using more data and adding some restrictions to the model, such as assuming no breaks for certain model parameter types or allowing common breaks for a subset of parameter types.

6. Acknowledgements

The authors would like to thank Chulalongkorn Business School for providing excellent research facilities.

7. References

Bumrungrat, N. (2024). *A panel partial break model for forecasting stock returns with parameter instability* [Unpublished master's thesis]. Chulalongkorn University.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, *21*(4), 1509–1531. https://doi.org/10.1093/rfs/hhm055

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291–311. https://doi.org/10.1016/j.jeconom.2006.05.023

Dangl, T., & Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*, *106*(1), 157–181. https://doi.org/10.1016/j.jfineco.2012.04.003

Farmer, L. E., Schmidt, L., & Timmermann, A. (2023). Pockets of predictability. *The Journal of Finance*, *78*(3), 1279–1341. https://doi.org/10.1111/jofi.13229

Georgiev, I., Harvey, D. I., Leybourne, S. J., & Taylor, R. (2018). Testing for parameter instability in predictive regression models. *Journal of Econometrics*, *204*(1), 101–118. https://doi.org/10.1016/j.jeconom.2018.01.005

Goyal, A. (n.d.). *Publications*. Retrieved June 13, 2023, from https://sites.google.com/view/agoyal145

Guidolin, M., & Timmermann, A. (2007). Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, *31*(11), 3503–3544. https://doi.org/10.1016/j.jedc.2006.12.004

Henkel, S. J., Martin, J. S., & Nardari, F. (2011). Time-varying short-horizon predictability☆. *Journal of Financial Economics*, *99*(3), 560–580. https://doi.org/10.1016/j.jfineco.2010.09.008

Johannes, M., Korteweg, A. G., & Polson, N. G. (2014). Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance*, *69*(2), 611–644. https://doi.org/10.1111/jofi.12121

Lettau, M., & Van Nieuwerburgh, S. (2008). Reconciling the return predictability evidence. *The Review of Financial Studies*, *21*(4), 1607–1652. https://doi.org/10.1093/rfs/hhm074

Maheu, J. M., & Song, Y. (2018). An efficient Bayesian approach to multiple structural change in multivariate time series. *Journal of Applied Econometrics*, *33*(2), 251–270. https://doi.org/10.1002/jae.2606

Pettenuzzo, D., & Timmermann, A. (2011). Predictability of stock returns and asset allocation under structural breaks. *Journal of Econometrics*, *164*(1), 60–78. https://doi.org/10.1016/j.jeconom.2011.02.019

Pitarakis, J. (2017). A simple approach for diagnosing instabilities in predictive regressions. *Oxford Bulletin of Economics and Statistics*, *79*(5), 851–874. https://doi.org/10.1111/obes.12184

Rapach, D. E., Strauss, J., & Zhou, G. (2010). Out-of-Sample Equity Premium Prediction: combination forecasts and links to the real economy. *The Review of Financial Studies*, *23*(2), 821–862. https://doi.org/10.1093/rfs/hhp063

Rossi, B. (2021). Forecasting in the presence of instabilities: How we know whether models predict well and how to improve them. *Journal of Economic Literature*, *59*(4), 1135–1190. https://doi.org/10.1257/jel.20201479

Smith, S. C., & Timmermann, A. (2021). Break risk. *The Review of Financial Studies*, *34*(4), 2045–2100. https://doi.org/10.1093/rfs/hhaa084

Smith, S. C., & Timmermann, A. (2022). Have risk premia vanished? *Journal of Financial Economics*, *145*(2), 553–576. https://doi.org/10.1016/j.jfineco.2021.08.019

Tu, Y., & Xie, X. (2023). Penetrating sporadic return predictability. *Journal of Econometrics*, *237*(1), 105509. https://doi.org/10.1016/j.jeconom.2023.105509

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. https://doi.org/10.1093/rfs/hhm014

Zhu, F., Liu, M., Ling, S., & Cai, Z. (2022). Testing for structural change of predictive regression model to threshold predictive regression model. *Journal of Business & Economic Statistics*, *41*(1), 228–240. https://doi.org/10.1080/07350015.2021.2008406