Journal Pre-proof

Stock return predictability from mixed model perspective

Zhifeng Dai, Huan Zhu

PII: S0927-538X(19)30633-X
DOI: https://doi.org/10.1016/j.pacfin.2020.101267
Reference: PACFIN 101267

To appear in: *Pacific-Basin Finance Journal*

Received date: 23 October 2019
Revised date: 19 December 2019
Accepted date: 15 January 2020

Please cite this article as: Z. Dai and H. Zhu, Stock return predictability from mixed model perspective, *Pacific-Basin Finance Journal*(2019), https://doi.org/10.1016/j.pacfin.2020.101267

Stock return predictability from mixed model perspective

Zhifeng Dai* , Huan Zhu

Department of Statistics, Changsha University of Science and Technology, Hunan, 410114 China

Abstract: We find that mixing existing forecasting models can significantly improve prediction performance of stock returns. Empirical results suggest that the stock return forecasting by three proposed mixed models are more significant both in statistical and economic terms than the corresponding models in Campbell and Thompson (2008), Wang et al. (2018) and Zhang et al. (2019). This improvement of predictability is also remarkable when we employ the multivariate information to predict stock return. The prediction performance of mixed models is robust to a series of robustness test. Particularly, the three proposed mixed models obtain superior out-of-sample forecasting

* Corresponding author. E-mail : zhifengdai823@163.com.(Zhifeng Dai)

performance of stock return for business cycles, rolling window predictions and different out-of-sample periods.

Keywords: Mixed models; Stock return predictability; Out-of-sample forecast; Asset allocation.

JEL classification: C53；G11；G17

1. Introduction

Stock return prediction is of great significance to asset allocation, risk management and asset pricing. An influential research by Goyal and Welch (2008) indicates that it is difficult to find a predictor or a rational model to accurately forecast out-of-sample stock return. So far, numerous literatures have proposed predictors that could be used to predict stock returns, including interest rates (Ang and Bekaert, 2007; Fama and Schwert, 1977; Campbell, 1987), dividend ratios (Fama and French, 1988, 1989; Goyal and Welch, 2003; Lewellen, 2004), the consumption-wealth ratio (Lettau and Ludvigson, 2001), inflation (Campbell and Vuolteenaho, 2004), stock variances (Guo, 2006; Ludvigson and Ng, 2007), downside variance risk (Feunou et al. 2015; Kilic and Shaliastovich, 2018), the variance risk premium (Bollerslev et al. 2009; Bollerslev Bollerslev et al. 2014), economic policy uncertainty (Chen et al. 2017), investor sentiment (Huang et al. 2015), short interest index (Rapach et al. 2016), news-implied volatility (Manela and Moreira, 2017), technical indicators (Neely et al. 2014; Gao et al. 2018; Zhang et al. 2019), manager sentiment (Jiang et al. 2017), oil-related variables (Chiang and Hughen, 2017; Nonejad, 2018; Wang et al. 2019), and among others.

Additionally, a series of studies have also employed new models to improve the prediction performance for stock return. For example, Campbel and Thompson (2008) (CT hereafter) proposed non-negative economic constraint and coefficient economic constraint to make investor obtain appropriate forecasts. Pettenuzzo et al. (2014) proposed a reasonable bound in term of the conditional Sharpe ratio. Wang et al. (2018) (MoP hereafter) found the momentum of return predictability and proposed a strategy where the forecasting model are selected between the interested model and the benchmark model. Zhang et al. (2019) (ZY hereafter) used the famous three sigma rule to make the yield forecast within a reasonable range, and among others.

On the other hand, combination method also is a popular method to improve the prediction performance for stock return, for example, Rapach et al. (2010), Zhu and Zhu (2013), Lin et al. (2018), Zhang et al. (2018), and Bahrami et al. (2019). However, these literatures focus on taking predictors as a combination. Not much has been done to promote the prediction performance for stock return from a mixed model perspective. As we know, combination of predictors may lead to overfitting, so that they could hardly obtain more accurate stock return forecasts than existing models. For this consideration, the main purpose of our paper's is to develop simple but effective, mixed models to achieve superior out-of-sample predictability of stock return without constructing new forecasting models.

We choose existing models as the CT economic constraint, the ZY economic constraint and the MoP strategy, which have been well studied in Campbell and Thompson (2008), Wang et al. (2018), and Zhang et al. (2019). There are two reasons to explain this selection. Firstly, in real word, it is unlikely for a rational investor to adopt a negative return on investment or trust the forecast outlies, particularly large or small forecast. Hence, to solve those question, Campbell and Thompson (2008) and Zhang et al. (2019) suggested a non-negative economic constraint and non-outlier economic constraint, respectively, which shows realistic economic significance on the prediction of stock return.

Secondly, Goyal and Welch (2003,2008) emphasized the historical average is a very strict out-of-sample benchmark: if a predictor can forecast stock returns, the forecasts generated by this predictor are superior to the historical average benchmark. And Campbell and Thompson (2008) suggested that the historical average benchmark is beat by a few predictive regressions with weak restrictions as limiting sign of coefficients and individual forecasts. Those studies all show that the historical average is of great importance to the prediction of stock return. However, in the MoP strategy, Wang et al. found the momentum of predictability and designed a selection strategy between historical average and individual forecasts generated by interested model. Therefore, the MoP strategy could not only predict stock return, but integrate the prediction ability of historical average, which plays an important role on forecasting stock return.

Considering the common role of historic average, the CT constraint, MoP strategy and the ZY constraint, we propose the first mixed model, namely Generalized Non-outlier model (GN hereafter). Then, the second mixed model is developed by combining information of the CT and ZY economic constraint. This model is called as Non-outlier Positive model (NP hereafter). Finally, combining the ZY economic constraint with the MOP strategy, the third mixed model, Non-outlier Momentum model (NM hereafter), is presented.

We contribute to literatures in four ways. Firstly, we show that investors can use hybrid models to generate more accurate return forecasts in real time. Secondly, the mixed model not only plays a role in the economic recession, but also has a good predictive ability of economic expansion, which has made great progress in the prediction of stock returns. Most of existing models have merely a better predictability over recessions. (see, e.g., Rapach et al. (2010); Neely et al. (2014); Huang et al. (2015); Jiang et al. (2017); Zhang et al. (2019) and Wang et al. (2019)). Thirdly, the prediction ability of the mixed models is also stronger when we use rolling windows to predict stock returns. Finally, the mixed models applied to forecast stock return are unaffected by the out-of-sample periods, which fills a gap that the out-of-sample forecasting ability is largely associated with prediction period.

Using mixed models with each 14 macroeconomic variables supposed by Welch and Goyal (2008), we obtain out-of-sample forecasts of stock return from January 1948 to December 2017. According to the relevant literature, we employ the out-of-sample R^2 (R^2_{OoS}) to evaluate the performance of return prediction, that is, to study whether the out-of-sample prediction based on a given model outperforms the historical average benchmark. The CW statistic proposed by Clark and West (2007) is employed to test whether the stock return predictability is statistically significant.

The out-of-sample results show that the mixed models generate larger R^2_{OoS} than existing models. Naturally, the improvement of predictability is exhibited in those mixed models. Specifically, when applied to GN model, the R^2_{OoS} value of six predictors, EP, NTIS, LTR, DFY, DFR and INFL, improved from -0.585% to 0.132%, from -0.529% to 0.435%, from -0.677% to 0.182%, from -0.168% to 0.648% , from -0.429% to 0.32%, from -0.043% to 0.758%, respectively. Then, after implementing NP model, the R^2_{OoS} value of two predictors, EP and DE, increase from -0.808% to 0.689%, from -0.618% to 0.752%, respectively. In addition, there are 12, 11, 11, 12, 10 and 10 out of all predictors generate larger R^2_{OoS} than ZY constraint for look-back period from 1 to 12 months, respectively, when using the NM model to predict stock returns.

We evaluate economic value of stock return forecasts by the certainty equivalent return (CER) of mean variance investor in asset allocation between stocks and riskless assets. And we report the differences between CER from given models and the benchmark model following Campbell and Thompson (2008). The empirical results show that stock return forecasts from mixed models are all significant economically. Furthermore, the CER gains have been greatly improved when using hybrid models. For example, the CER values of NM (1) model applied to DP, DY, EP promote from-0.842% to 1.176%, from -1.055% to 0.942%, from -0.028% to 1.393%, respectively.

The improvement of predictability is also found in a few multivariate information methods. Those approaches include principal component analysis (PCA), average combination and five popular combination methods from Pettenuzzo et al. (2014). We find that the NM model applied to all combination methods generates more significant return predictability. The prediction ability is also improved after implementing GN and NP mixed models with larger R^2_{OoS} value.

To find the source of improvement, we link empirical analysis with business cycle. The empirical results show that all mixed models not only have better predictive effect in the recession period, but also have good prediction ability in the expansion period. Namely, the prediction performance of all mixed models is independent on business cycle. Notably, during recession period, the prediction accuracy of NTIS predictor applied to NP model, NM (1) model and NM (12) model is improved from -2.477% to 0.945%, from -0.410% to 0.361%, from -0.202% to 0.569%, respectively, and the relative models are ZY model, MoP (1) model, MoP (12) model, respectively.

To test the robustness of mixed models, we also employ another robustness analysis, including linking to business cycle, portfolio exercise considered transaction cost, alternative prediction approach, alternative out-of-sample period and portfolio exercise with alternative asset weight limitation. In summary, the mixed models are more efficient and stable. Particularly, superior out-of-sample forecasting performance is obtained by those three mixed models when employing rolling windows to produce out-of-sample forecasts. Additionally, when considering other five out-of-sample periods, 1979:01 to 2017:12, 1959:01-2017:12, 1969:01-2017:12, 1989:01-2017:12 and 1999:01-2017:12, we find that the prediction performance of existing models is sensitive to out-of-sample period, while our developed models still do a good job in different out-of-sample periods. This finding fills a gap in existing literatures.

The remainder of this paper is organized as follows: Section 2 provides our research data and summary statistics. Section 3 presents econometric methodology including existing models and mixed models. Section 4 reports out-of-sample prediction ability of all models. Section 5 investigates three extension analysis by combining multivariate information, portfolio performance and linking to business cycle. Section 6 presents the robustness test. The last Section gives the conclusion.

2. Data and summary statistics

In this paper, we use 14 macroeconomic variables[1] that originally used in Welch and Goyal (2008) as predictors to predict stock return, as those variables are

[1]Those 14 economic variables can be downloaded from homepage of Amit Goyal at http://www.hec.unil.ch/agoyal/

employed by a long strand of literature (see, e.g., Rapach et al. (2010); Christiansen et al. (2012); Neely et al. (2014); Dangl and Halling (2012); Wang et al. (2018) and Zhang et al. (2019)). For the sake of brevity, they are described as follows.

- Dividend-price ratio (log), DP: log of a 12-month moving sum of dividends paid on the S&P 500 index minus the log of stock prices (S&P 500 index).
- Dividend yield (log), DY: difference between the log of dividends and the log of lagged prices.
- Earnings-price ratio (log), EP: difference between the log of earnings on the S&P 500 index and the log of prices, where earnings is measured using a one-year moving sum.
- Dividend-payout ratio (log), DE: difference between the log of dividends and the log of earnings on the S&P 500 index.
- Stock return variance, SVAR: sum of squared daily returns on the S&P 500 index.
- Book-to-market ratio, BM: ratio of book value to market value for the Dow Jones Industrial Average.
- Net equity expansion, NTIS: ratio of 12-month moving sums of net issues by NYSE-listed stocks to total end- of-year market capitalization of NYSE stocks.
- Treasury bill rate, TBL: interest rate on a 3-month Treasury bill (secondary market).
- Long-term yield, LTY: long-term government bond yield.
- Long-term return, LTR: return on long-term government bonds.
- Term spread, TMS: difference between the long-term yield and the Treasury bill rate.
- Default yield spread, DFY: difference between BAA- and AAA-rated corporate bond yields.
- Default return spread, DFR: difference between the long-term corporate bond return and the long-term government bond return.
- Inflation, INFL: calculated from the Consumer Price Index (CPI) (all urban consumers) and used lagged for one month to account for the delay in the release of the CPI.

Insert Table 1 about here

Table1 reports summary statistics for the monthly excess stock return and 14 economic variables from January 1927 through December 2017. The excess stock return reaches 0.005 on average, together with a standard deviation of 0.0054 and produces a monthly Sharpe ratio of 0.094. In addition, the first-order autocorrelation of excess stock return is 0.0086, reflecting stock return is hardly forecasted or explained by its past. Although there is low autocorrelation as the excess stock return, most of those predictors shows high persistence. Overall, the summary statistics in Table 1 are basically consistent with the relative literature about the predictability of stock return. (see, e.g. Zhang et al. (2019)).

3. Econometric methodology

3.1. Univariate predictive regression model

We employ the conventional framework in a long strand of literature (see, e.g., Neely et al. (2014); Pettenuzzo et al. (2014); Rapach et al. (2016) and Zhang et al. (2019)) and use the following univariate regression model (original hereafter).

$$r_{t+1} = \alpha_i + \beta_i x_{i,t} + \varepsilon_{i,t+1}, \tag{1}$$

Where r_{t+1} denotes the excess stock return at month t+1; $x_{i,t}$ denotes the ith predictor of 14 macroeconomic variables at month t; and $\varepsilon_{i,t+1}$ denotes an error term assumed to follow an independent and identically normal distribution. α_i and β_i denote the parameter estimates being subject to the ordinary least squares (OLS) estimation.

Following Goyal and Welch (2008), Rapach et al. (2010) and Neely et al. (2014), the recursive estimation window is employed to generate out-of-sample forecasts of excess stock return. Particularly, the whole T observations (sample) of excess stock return and 14 predictors is divided into two portions. One is an in-sample portion which is consist of the first M observations, the other is out-of-sample portion which is consist of the last P observations (i.e., M+P=T). Then, the out-of-sample forecast at month M+1 can be obtained by

$$\hat{r}_{i,M+1} = \alpha_{i,M} + \beta_{i,M} x_{i,M}, \tag{2}$$

where $\hat{r}_{i,M+1}$ denotes the stock return forecast at month M+1; the parameters $\alpha_{i,M}$ and $\beta_{i,M}$ are the OLS estimation by regressing $\{r_t\}_{t=2}^{M}$ on a constant and $\{x_t\}_{t=1}^{M-1}$. The out-of-sample forecast at month M+2 can be obtained by

$$\hat{r}_{i,M+2} = \alpha_{i,M+1} + \beta_{i,M+1} x_{i,M+1}, \tag{3}$$

where the parameters $\alpha_{i,M+1}$ and $\beta_{i,M+1}$ are the OLS estimation by regressing $\{r_t\}_{t=2}^{M+1}$ on a constant and $\{x_t\}_{t=1}^{M}$. Going forward as this, when at the end of the sample period, a series of P=T-M stock return forecasts will be generated by this recursive window.

To evaluate precision of forecasts, we compare the stock return forecasts to historical average benchmark which is available by regressing returns on a constant ($r_{t+1} = \alpha_i + \varepsilon_{i,t+1}$). And the natural benchmark is given by

$$\bar{r}_{t+1} = \frac{1}{t}\sum_{i=1}^{t} r_i, \tag{4}$$

where $\bar{r}_{t+1}$ denotes the historical average forecast at month t+1; r_t denotes the actual value of stock return. Goyal and Welch (2003, 2008) manifest that (4) is a very strict out-of-sample benchmark: if the economic variables can predict stock returns, the forecasts $\hat{r}_{M+1}$ outperform the historical average benchmark.

3.2. **Existing economic model**

We consider three existing economic models as competing strategies relative to our developed models. Those existing models are descripted as follows.

3.2.1. **Economic constraint model**

Two existing economic constraint models are considered. The first is proposed by Campbell and Thompson (2008) (CT hereafter). They suggest that it is difficult for a rational investor to adopt a negative return on investment and set the forecast to zero whenever it is negative. The CT economic constrain is provided by

$$\hat{r}_{i,t+1}^{CT} = \max(0, \alpha_{i,t} + \beta_{i,t} x_{i,t}), \tag{5}$$

where $\hat{r}_{i,t+1}^{CT}$ denotes the CT forecast by the *i*-th predictor at month t+1.

The second model is proposed by Zhang et al. (2019) (ZY hereafter) where a rational investor will rule out the forecast outlies and select the general three-sigma rule to generate the return forecast within a rational range. The ZY economic constrain is given by

$$\hat{r}^{ZY}_{i,t+1} = \begin{cases} r_{i,t} + 3\sigma_t, if\ \hat{r}_{i,t+1} > r_{i,t} + 3\sigma_t \\ r_{i,t} - 3\sigma_t, if\ \hat{r}_{i,t+1} < r_{i,t} - 3\sigma_t \\ \hat{r}_{i,t+1}, otherwise \end{cases}, \tag{6}$$

where $\hat{r}^{ZYFY}_{i,t+1}$ denotes the ZY forecast with the *i*-th predictor at month t+1; $\hat{r}_{i,t+1}$ denotes the forecast obtained by original model; and σ_t denotes the standard deviation of the excess stock returns at month t.

3.2.2. **Momentum of return predictability**

Additionally, we consider the momentum of return predictability presented by Wang et al. (2018) (MoP hereafter). They fund momentum of predictability that if an interested model beats a benchmark model in the recent past period, it will usually outperform the benchmark in the current period. Therefore, the past predictability during period t can be calculated by

$$pp_t(k) = I(\sum_{j=t-k}^{t-1}(r_t - \hat{r}_t)^2 - \sum_{j=t-k}^{t-1}(r_t - \bar{r}_t)^2), \tag{7}$$

where k is the number of look-back period taking k=1,3,6,9,12 following Wang et al. (2018), and $I(\bullet)$ is an indicator function.

They integrate this momentum of predictability into a selection process, resulting in a series of more accurate forecasts. In this process, the prediction model switches between the interested model and the benchmark. In detail, when $pp_t(k)$ takes a value of 1, we employ the interested model to forecast future stock returns, and 0 otherwise, we adopt the history average. This selection process is given by

$$\begin{cases} \hat{r}_{t+1,MoP}(k) = \hat{r}_{t+1}, if\ PP_{t+1}(k) = 1 \\ \hat{r}_{t+1,MoP}(k) = \bar{r}_{t+1}, if\ PP_{t+1}(k) = 0 \end{cases} \tag{8}$$

3.3. **Mixed forecast model**

Although numerous studies have cited the above methods to predict the stock return, there is little evidence on the out-of-sample prediction by hybrid model. Given this, mixed models[2,3] based on the above methods are constructed. In this subsection,

---

[2] Besides those mixed models, there are other two mixed models which hardly outperform the existing models considered in this paper. One is that we consider both the CT constraint and the

we provide the definition of all mixed models.

3.3.1. **Generalized Non-outlier model**

Considering the role of historical average, CT and ZY constraint, we propose the first mixed model, namely Generalized Non-outlier model (GN hereafter). Different from the ZY economic constraint, we employ the median criterion to remove outliers initially. How to apply the GN model to forecast stock return is given as follows:

Step1: Considering the significance of history average and CT constraint, we conduct a maximum model, which is given by

$$\hat{r}^{1}_{i,t+1} = Max(0, \alpha_{i,t} + \beta_{i,t} x_{i,t}, \frac{1}{t}\sum_{j=1}^{t} r_j), \tag{9}$$

Step2: To remove outliers preliminarily, the median is considered as another criterion to control the rational range of stock return forecasts.

$$\hat{r}^{2}_{i,t+1} = Median(0, \alpha_{i,t} + \beta_{i,t} x_{i,t}, \frac{1}{t}\sum_{j=1}^{t} r_j), \tag{10}$$

where $\hat{r}^{2}_{i,t+1}$ denotes the forecasts generated by median criterion at month t+1;

Step3: We apply the ZY constraint to bound the return forecasts of $\hat{r}^{2}_{i,t+1}$ within a rational range.

$$\hat{r}^{GN}_{i,t+1} = \begin{cases} r_{i,t} + 3\sigma_t, & if\ \hat{r}^{2}_{i,t+1} > r_{i,t} + 3\sigma_t \\ r_{i,t} - 3\sigma_t, & if\ \hat{r}^{2}_{i,t+1} < r_{i,t} - 3\sigma_t \\ \hat{r}^{2}_{i,t+1}, & otherwise \end{cases}, \tag{11}$$

where $\hat{r}^{GN}_{i,t+1}$ denotes the GN forecast for month t+1.

3.3.2. **Non-outlier Positive model**

We propose the second mixed model, namely Non-outlier Positive model (NP hereafter), by integrating information of the CT constraint and the ZY constraint. In detail, we employ the CT constraint to rule out the negative forecasts, and then the ZY

MoP strategy, the other is combining information of all the existing models considered at the same time.

[3] We find that the out-of-sample performance of the mixed models is independent of the mixed way. Taking CZ model as an example, if we first use ZY constraint to eliminate outliers in return forecast, and then employ CT constraint to rule out negative forecasts, the ability of prediction will not change much.

constraint is used to remove outliers of the return forecasts. The NP model is given by

$$\hat{r}^{NP}_{i,t+1} = \begin{cases} r_{i,t} + 3\sigma_t, & if\ \hat{r}^{CT}_{i,t+1} > r_{i,t} + 3\sigma_t \\ r_{i,t} - 3\sigma_t, & if\ \hat{r}^{CT}_{i,t+1} < r_{i,t} - 3\sigma_t \\ \hat{r}^{CT}_{i,t+1}, & otherwise \end{cases}, \tag{12}$$

where $\hat{r}^{NP}_{i,t+1}$ denotes the NP forecast at month t+1;

3.3.3. **Non-outlier Momentum model**

The positive effect of the ZY constraint and the MoP strategy is combined to conduct the third mixed model, namely Non-outlier Momentum model (NM hereafter). Specifically, before the ZY constraint applied to make the return forecasts within a rational range, we perform the MoP strategy to select between the original model and benchmark. The NM model is given by

$$\hat{r}^{NM(k)}_{i,t+1} = \begin{cases} r_{i,t} + 3\sigma_t, & if\ \hat{r}^{MoP(k)}_{i,t+1} > r_{i,t} + 3\sigma_t \\ r_{i,t} - 3\sigma_t, & if\ \hat{r}^{MoP(k)}_{i,t+1} < r_{i,t} - 3\sigma_t \\ \hat{r}^{MoP(k)}_{i,t+1}, & otherwise \end{cases}, \tag{13}$$

where $\hat{r}^{NM(k)}_{i,t+1}$ denotes the NM forecast at month t+1. Being Similar to Wang et al. (2018), we also consider an average NM strategy (NM-AVG), which is given by

$$\hat{r}_{t+1,NM-AVG} = \frac{1}{N}\sum_{k=1}^{N} \hat{r}_{t+1,NM(k)} \tag{14}$$

where N denotes the total number of k-values.

3.4. **Forecast evaluation**

Following relative literatures (see, e.g., Ferreira and Santa-Clara (2011); Rapach et al. (2016); Jiang et al. (2017); Lin et al. (2018); Wang et al. (2018); Dai and Zhu (2019) and Zhang et al. (2019)) , we take a wide spread out-of-sample R^2 statistics (R^2_{OoS} hereafter) to evaluate prediction performance of given model. This statistic tests whether the out-of-sample prediction performance of given model outperforms the historical average benchmark. The R^2_{OoS} is computed by

$$R_{OoS}^2 = 1 - \frac{MSPE_{model}}{MSPE_{bench}}, \tag{15}$$

where $MSPE_{model} = \frac{1}{T-M}\sum_{t=M+1}^{T}(r_t - \hat{r}_t)^2$ and $MSPE_{bench} = \frac{1}{T-M}\sum_{t=M+1}^{T}(r_t - \bar{r}_t)^2$.

$MSPE_{bench}$ and $MSPE_{model}$ are the mean squared predictive errors (MSPE) of the benchmark model and the tested model, respectively; Different from in-sample R^2, the value of R_{OoS}^2 can be equal to negative. Intuitively, a positive R_{OoS}^2 indicates that the forecasts by tested model have lower MSPE than the benchmark, implying the greater accuracy of stock return predictability.

Furthermore, we use Clark and West (2007) statistics to further test whether the prediction model produced significant statistical improvements in MSFE. Mathematically, Clark and West (2007) statistics are defined as

$$f_{i,t} = (r_t - \bar{r}_t)^2 - (r_t - \hat{r}_t)^2 + (\bar{r}_t - \hat{r}_t)^2 . \tag{16}$$

Through regression $\{f_t\}_{M+1}^{T}$ on a constant, we can easily get the C-W statistics, that just is t-statistics for the constant. In addition, the p-value of one-sided test can be easily obtained from the standard normal distribution.

4. Out-of-sample forecasting results

In this section, we empirically analyze the out-of-sample prediction results of existing models and mixed models from January 1948 to December 2017.

Insert Table 2 about here

Table 2 reports out-of-sample prediction performance evaluated by R_{OoS}^2 in (16) and CW statistic. We find that it is difficult for the original model to beat the historical average benchmark, and for the existing models to significantly outperform the original counterpart. This finding is consistent with this literature as in Campbell and Thompson (2008), Zhang et al. (2019) and Wang et al. (2018).

After implementing the GN model to forecast stock return, 13 out of 14 predictors have results with larger R_{OoS}^2 than using ZY constraint. Furthermore, the accuracy of prediction of those four predictors, EP, DE, LTY, LTR, is improved significantly. In particular, the two predictors, EP and DE, the R_{OoS}^2 values increase from -0.808% to 0.689%, from -0.618% to 0.752%, respectively. Hence, we can conclude that the out-of-sample prediction performance of the GN model outperforms the ZY constraint.

Comparing the CT and the ZY constraint to the NP model, we find that 14 and 10 out of 14 predictors generate larger R^2_{OoS} by the NP model, respectively. Furthermore, 12 and 8 out of them are positive, respectively, which indicates outperformance correspond to historical average. Notably, when we contrast with the CT constraint and the NP model, the R^2_{OoS} values of those 12 predictors not only is non-negative, but also is improved significantly. It is obvious that the NP model applied to those 12 predictors outperforms the historical average benchmark. In addition, the R^2_{OoS} values of EP, NTIS, LTR, DFY, DFR, INFL improve from -0.585% to 0.132%, from -0.529% to 0.435%, from -0.677% to 0.182%, from -0.168% to 0.648% , from -0.429% to 0.32%, from -0.043% to 0.758%, respectively. Comparing the ZY constraint to the NP model, we find that improvement of prediction performance for four predictors, DY, EP, LTY, LTR, increases greatly. Particularly, the R^2_{OoS} value of EP improves from -0.808% to 0.132%, showing a better prediction ability. In summary, the NP model generates more accurate forecasts than the existing economic constraints considered in this paper.

When we compare out-of-sample the NM model to the MoP strategy, the R^2_{OoS} values of all predictors is higher using NM model. The same conclusion applies when we calculate the average. In addition, there are 5, 6, 6, 4 and 5 out of all indicators whose R^2_{OoS} is developed from negative to positive under different look-back period $k = 1,\ldots,12$ months, respectively. Therefore, we can conclude that the NM model generates lower forecasting error. And comparing the NM method to the ZY constraint, we find that 12, 11, 11, 12, 10 and 10 out of all predictors generate larger R^2_{OoS} for look-back period $k = 1,\ldots,12$ months, respectively. Consequently, we can draw a conclusion that the out-of-sample prediction of the NM model is stronger than both the MoP strategy and the ZY constraint model.

Overall, all mixed models improve significantly the prediction ability for stock return. And a better out-of-sample prediction power is exhibited in mixed models.

5. Extension analysis

We will provide extended analysis to further demonstrate that the prediction performance of mixed models is superior to existing models in this section. Firstly, we consider a mean-variance utility to further validate the economic value of stock return forecasts generated by the mixed models, which is presented in subsection 5.1. Then, to examine whether multivariate information contributes to forecast stock return, we employ a few combination approaches presented in subsection 5.2.

5.1. Portfolio performance

In the Section 4, we have manifested that the mixed models are associated with significant improvement of predictability to stock return over the existing models and the natural benchmark, statistically. A natural question is that whether the stock return forecasts form the mixed models are significant economically, which means how to evaluate the economic value of stock return considering risk aversion. To address this issue, an effective method is used, calculating a mean variance investor's certainty equivalent return (CER) in asset allocation between riskless assets and stocks.

Firstly, the investor could assign weight to stocks in the portfolio at month t, and the optimal weight is given by

$$w_{i,t} = \frac{1}{\gamma} \frac{\hat{r}_{i,t+1}}{\sigma^2_{i,t+1}}, \tag{17}$$

where γ denotes the risk aversion of investor, being equal to 3 following Rapach et al. (2010) and Wang et al. (2019); $\hat{r}_{i,t+1}$ and $\sigma^2_{i,t+1}$ are prediction of returns and variance, respectively. Following Campbell and Thompson (2008) and Neely et al. (2014), a five-year moving window of past monthly stock return is adopted to estimate the forecasts of variance. Naturally, the $1-w_t$ is assigned to the weight of risk-free assets. Different from the majority of existing literatures which limit stocks weight to lie in an interval at [0, 1.5] (e.g., Campbell and Thompson (2008); Rapach et al. (2010); Neely et al. (2014); Huang et al. (2015); Dai et al. (2019) and Jiang et al. (2019)), we make the weights of stocks within a range of -0.5 and 1.5 as the literature of Rapach et al. (2016) and Zhang et al. (2019). This is because it is difficult to promote the portfolio performance of the CT constraint and the MoP strategy when we make the weight lie in a range [0, 1.5].

Then, the portfolio returns at month t+1 is given by

$$R_{i,t+1} = \omega_i r_{t+1} + R_f \tag{18}$$

If a portfolio is constructed by Eq. (17) and (18), the investor obtains an average CER as

$$\mathrm{CER}_i = \mu_i - 0.5\gamma\sigma_i^2, \tag{19}$$

where μ_i and σ_i are the mean and variance of portfolio, respectively.

Insert Table 3 about here

Table 3 reports the portfolio performance of the existing models and the mixed models, respectively, evaluated by CER gains. We report the differences between CER of forecasts generated by our interested model and the natural benchmark. And the differences are multiplied by 1200 to get the annualized percent values. Comparing the CER gains of original model to the mixed model, we find that most of the CER value based on different mixed models and predictors is larger than relate forecasting models. Furthermore, after employing the mixed models, improvement of the CER value from different predictors increase significantly. Noted that all of the CER gains generated by the NP model and the NM (3) model are larger than the CT constraint and the MoP (3) model, respectively. The mixed models make the CER gains improved from negative to positive, such as the CER gains from NM(1) model taking DP, DY, EP as predictors promote from -0.842% to 1.176%, from -1.055% to 0.942%, from -0.028% to 1.393% , respectively.

In short, the results of portfolio performance show that the prediction of mixed models have economic significance.

5.2. **Multivariate results**

All prediction models, including existing models and mixed models, are extended on the basis of the original model. In those model, univariate predictive regression model is employed, namely, the stock return forecast is obtained by each single predictor. However, whether multivariate information can show predictability to stock return, it is a question. Fortunately, a long strand of studies has implied the role of multivariate information in predicting stock return (see, e.g., Rapach et al. (2010) ; Dangl and Halling (2012); Ludvigson and Ng (2007); Kelly and Pruitt (2013); Neely et al. (2014); Rapach et al. (2016); Zhu et al. (2013); Zhang et al. (2018); Wang et al. (2018); Wang et al. (2019)). Hence, it is necessary to broaden analysis to combining multivariate information of individual forecasts. In this subsection, multivariate information is considered to predict stock return based on the existing models and the mixed models.

Following Pettenuzzo et al. (2014), we primarily introduce five popular combination multivariate information methods.

The first approach uses a common factor been available by principal component analysis (PCA). The method is also called diffusion index (diffusion index hereafter) as in Ludvigson and Ng (2007); Kelly and Pruitt (2013); Neely et al. (2014). We take the common factor as predictor by univariate predictive regression model is given by

$$r_{t+1} = \alpha_{DI} + \beta_{DI} F_{DI,t} + \varepsilon_{t+1} \tag{20}$$

$$x_{i,t} = \lambda_t' F_{DI,t} + \varepsilon_{i,t+1} \tag{21}$$

Where $F_{DI,t}$ denotes q-vector of latent factors $(q \ll N)$, N denotes the total number of predictors, N=14; $F_{DI,t}$ can be obtained by principal components; λ_t denotes a q-vector of factor loadings; $\varepsilon_{i,t+1}$ denotes error term assumed to follow an independent and identically normal distribution. Noted that we follow Rapach and Zhou (2013) and Pettenuzzo et al. (2014) to employ the first principal component derived from 14 macroeconomic variables, parsimoniously. Because it is difficult to improve the ability of out-of-sample prediction by using multiple principal components.

The second approach is the mean combination forecast (mean hereafter). This method performs the average of the N forecasts, which is given by the following formula:

$$\hat{r}_{c,t+1} = \sum_{i=1}^{N} \omega_{i,t} \hat{r}_{i,t+1}, \tag{22}$$

where $\hat{r}_{c,t+1}$ denotes a combination forecasting for month t+1; $\hat{r}_{i,t+1}$ denotes the forecasts of i-th predictor at month t+1; $\omega_{i,t}$ denotes combination weight of the i-th forecast at month t. We employ the mean combination forecasts with the same weight, equaling $1/N$.

The third approach is median combination forecast (median hereafter), in which we simply take median value of individual forecasts as combination forecast.

The fourth approach is trimmed mean combination forecast (trimmed mean hereafter). Different from the mean approach, we take $\omega_{i,t}=1/(N-2)$ to avoid impact of the largest and smallest forecasting in $\{\hat{r}_{i,t+1}\}_{i=1}^{14}$.

We also consider discount mean squared prediction error combination approach (DMSPE hereafter). The weight in DMSPE method is given by

$$\omega_{i,t}=\phi_{i,t}^{-1}\sum\nolimits_{l=1}^{N}\phi_{i,t}^{-1}, \tag{23}$$

where

$$\phi_{i,t}=\sum\nolimits_{s=m+1}^{t}\theta^{t-s}(r_s-\hat{r}_{i,s})^2 \tag{24}$$

m denotes the all number of observations in sample, and θ is the discount factor. Following Rapach et al. (2013), Zhu and Zhu (2013), and Zhang et al. (2019), we take 1 and 0.9 as discount factor to generate two discount MSPE combination approaches, namely, DMSPE (1) and DMSPE (0.9), respectively.

In addition to these above combination approaches, we also employ average information of 14 predictors to forecast stock return (average hereafter), which is given by

$$r_{t+1}=\alpha+\beta\bar{x}_t+\varepsilon_{t+1}, \tag{25}$$

where

$$\bar{x}_t=\frac{1}{N}\sum_{j=1}^{N}x_{j,t}. \tag{26}$$

Intuitively, the methods of mean, median, trimmed mean, DMSPE (1) and DMSPE (0.9) calculate the weighted average of forecasts with different weights, while the average method and the PCA extract information from the predictor to forecast stock return.

Insert Table 4 about here

Table 4 reports the multivariate forecasting results of existing models and mixed models. Two important results can be obtained from the table. First, all the values of R_{OoS}^2 are positive among the different mixed models, and most of them are above 1%, manifesting that all mixed models significantly outperform the historical average benchmark. This important finding is consistent with numerous literatures on stock return forecast (see, e.g., Ludvigson and Ng, 2007; Rapach et al. (2013); Zhu and Zhu (2013); and Zhang et al. (2018)). Second and more importantly, the mixed models can generate larger R^2_{OoS} than most of the existing models. That is to say, through combining information from individual forecasts, the prediction accuracy of the hybrid model is improved. Unfortunately, the R^2_{OoS} value of most mixed models in different combinations is lower than the ZY economic constraint, which is because combining information shown above may be linked to over-fitting on stock return predictability.

In short, the mixed models of integrated information not only contribute to outperform the history average benchmark, but also produce more accurate forecasts than most existing models.

6. Robustness analysis

In this section, four robustness analysis are given to test the robustness of mixed models, including business cycle, portfolio exercise considered transaction cost, alternative prediction approach, alternative out-of-sample periods and alternative asset weight limitation.

6.1. **Business cycle**

Why can the mixed models improve the predictability of stock return? To answer this question, we provide a further analysis linked to business cycle in this part. This method has performed by a large body of related literatures (see, e.g., Cochrane (1999,2007); Rapach et al. (2010); Neely et al. (2014); Jiang et al. (2019); Ma et al. (2018) and Wang et al. (2018)), in which researchers suggest that the business cycle affects the trend of stock return, namely, the return forecasts is impacted by the alternation between expansion and recession periods. Following Rapach et al. (2010), Neely et al. (2014), and Zhang et al. (2019), the R^2_{OoS} linked to expansions and recessions ($R^2_{OS,\ \mathrm{EXP}}, R^2_{OS,\ \mathrm{REC}}$, respectively) is given by

$$R^2_{OS,c} = 1 - \frac{\sum_{t=m+1}^{T} (r_t - \hat{r}_{i,t})^2 I_t^c}{\sum_{t=m+1}^{T} (r_t - \bar{r}_{i,t})^2 I_t^c}, \text{for c=EXP, REC} \tag{27}$$

where c denotes the business cycles for expansions (EXP) or recessions (REC); I_t^{EXP} (I_t^{REC}) denotes an indicator variable that we take a value of 1 when month t belongs to an NBER expansions (recessions) period and zero otherwise[4]. We directly obtain the NBER-dated business cycle for expansions and recessions from the FRED Database.

Insert Table 5 and Table 6 about here

Table 5 and Table 6 report out-of-sample forecasting performance linked to business cycle based on the existing models and the mixed models, respectively. Two important findings emerge. First, most R_{OoS}^2 values produced by existing models in the recession period are larger than those in the expansion period when most of R_{OoS}^2 values are negative. We can draw a conclusion that the predictive ability of all the existing models applied to forecast stock return is concentrated over recessions. This finding is consistent with the relative literature as in Rapach et al. (2010), Neely et al. (2014), Huang et al. (2015), Jiang et al. (2019), Zhang et al. (2019), Wang et al. (2019), and among others. However, the prediction performance of mixed models focused not only on recessions but also on expansions. Particularly, most of R_{OoS}^2 values generated by mixed models are positive during recession and expansion periods, demonstrating that the prediction ability of mixed models is not affected by economic cycle. Furthermore, in the recession period, most of the mixed models applied to different predictors do a better job to forecast stock return than the existing models with larger R_{OoS}^2 value. Particularly, taking the predictor of NTIS as example, the prediction accuracy of NTIS predictor applied to NP model, NM (1) model and NM (12) model is improved from -2.477% to 0.945%, from -0.410% to 0.361%, from -0.202% to 0.569%, and the related existing models are ZY model, MoP (1) model, MoP (12) model, respectively.

In short, when we relate to business cycle, all mixed models not only have better predictive effect in the recession period, but also have good prediction ability in the expansion period, meaning the prediction performance of all mixed models is independent on business cycle.

[4]All month data will be included as either expansion or recession.

6.2. **Transaction cost**

In Subsection 5.1, we have provided extension analysis with portfolio exercise without transaction cost. However, in the real world, investors need to pay transaction cost when they trade stocks. And the transaction cost is very important for portfolio return. Given this, we present a portfolio analysis with transaction costs in this section to test the robustness of return forecasts for economic value. Following Neely et al. (2014) and Zhang et al. (2019), we also use the proportional transaction cost with a value of 50 basis points per transaction.

Insert Table 7 about here

Table 7 reports the CER gains of portfolio taking transaction costs into account. Two important observations can be drawn from the table. First, all models including existing models and mixed models have lower CER returns than non-transaction costs models because they are related to monthly inventory turnover. The more investors buy and sell stocks, the higher the transaction cost. This argument is consistent with related literature as in Neely et al. (2014) and Zhang et al. (2019). The second and more importantly, most of the CER gains generated by our mixed models are larger than the existing counterparts. Notably, all of the CER gains yielded by the NP model is higher than the CT constraint.

Overall, the CER gains are thus robust to portfolio performance considered transaction cost to our mixed models.

6.3. **Alternative prediction approach**

In this subsection, we use a rolling estimation window to generate stock return forecasts to test the robustness of out-of-sample prediction performance of mixed models. In detail, the first out-of-sample forecast with the *i*-th predictor of 14 economic variables is obtained by

$$\hat{r}_{i,M+1} = \alpha_{i,M} + \beta_{i,M} x_{i,M}, \tag{28}$$

where $\hat{r}_{i,M+1}$ denotes stock return forecasts at month M+1; $\alpha_{i,M}$ and $\beta_{i,M}$ are the OLS estimation obtained by regressing $\{r_t\}_{t=2}^{M}$ on a constant and $\{x_t\}_{t=1}^{M-1}$. The second out-of-sample forecast is obtained by

$$\hat{r}_{i,M+2} = \alpha_{i,M+1} + \beta_{i,M+1} x_{i,M+1}, \tag{29}$$

where $\alpha_{i,M+1}$ and $\beta_{i,M+1}$ denote the OLS estimation obtained by regressing $\{r_t\}_{t=3}^{M+1}$ on a constant and $\{x_t\}_{t=2}^{M}$. Going forward as this, when at the end of the sample period, a series of P=T-M stock return prediction will be generated by a rolling window method.

Insert Table 8 about here

Table 8 reports out-of-sample forecasting performance by a rolling estimation window approach. Different from the out-of-sample analysis in Section 4, we use a rolling estimation window to generate predictions to test the robustness of the mixed models. Firstly, the rolling prediction performance of all models is lower than that of recursive prediction, because stock return prediction is related to historical data dimension. This argument is consistent with related literature as in Wang et al. (2018). Then, most R_{OoS}^2 values of the existing models are negative, which indicates that the prediction performance of the existing models based on a rolling estimation prediction is inferior to the historical average benchmark. However, the mixed models based on a rolling prediction still yield positive R_{OoS}^2, suggesting that investors are willing to employ mixed models to predict stock return rather than existing models and historical average benchmark. This finding is consisted with the results based on a recursive window presented in section 4.

In short, though predictive performance based on a rolling estimation window method is inferior to a recursive window, stock return prediction yielded by the mixed models still shows more accurate than the existing model.

6.4. **Alternative out-of-sample period**

A large body of related literatures (e.g., Campbell and Thompson (2008); Rapach et al. (2016), Wang et al. (2019)) find that the out-of-sample forecasting ability is largely associated with the prediction period. The models of different sample periods have different ability to predict stock returns. Given this, in this subsection, we employ alternative two out-of-sample periods from 1969:01 to 2017:12 and from 1989:01 to 2017:12[5], respectively, to further test the robustness of out-of-sample prediction performance of our mixed models.

---

[5] we also consider other three different out-of-sample periods, namely, 1959:01-2017:12, 1969:01-2017:12, 1999:01-2017:12, respectively. All the results manifest that the mixed models

Insert Table 9 and Table 10 about here

Table 9 and Table 10 report out-of-sample forecasting performance considering alternative out-of-sample period based on existing models and mixed models, respectively. Firstly, compared with the out-of-sample result of the existing models and the mixed models, it is obvious that numerous existing models do a bad job with the negative R_{OoS}^2 value, while most of R_{OoS}^2 values generated by the mixed models are positive except for individual indicators regression, which shows that the mixed models have a more stable structure. Furthermore, the prediction ability of the mixed models is stronger over more recent out-of-sample period, which is consistent with Rapach et al. (2016). Overall, the mixed models predict stock returns more steadily than the existing models.

6.5 Alternative asset weight limitation

In the above experiments of portfolio performance, we limit the stocks weight to lie in an interval of [-0.5, 1.5] as in Rapach et al. (2016) and Zhang et al. (2019). Now, we use alternative asset weight limitation within a range of 0 and 1.5, which is commonly used in the existing literatures (e.g., Campbell and Thompson (2008); Rapach et al. (2010); Neely et al. (2014); Huang et al. (2015) and Jiang et al. (2019)).

Insert Table 11 about here

Table 11 reports the results of the existing models and mixed models with asset weight limited to the range of 0 and 1.5. Consistent with the performance in Table 3, the mixed models result in a portfolio with higher gains in CER than those of the existing models. In addition, the results of the CT constraint model, Generalized Non-outlier model (GN), the Non-outlier Positive model (NP) obtain the same gains in CER as in in Table 3. The reason is that these three models have a non-negative constraint. Of course, the gains in CER of the Non-outlier Momentum model (NM) and MOP in Wang et al. (2018) are slightly less than the results with asset weight limited to the range of [-0.5, 1.5].

7. Conclusion

are more stable than the existing models in predicting stock return. To saver space, we do not report those results, but they are available upon request.

In this paper, we develop three mixed models, including the Generalized Non-outlier model (GN), the Non-outlier Positive model (NP) and the Non-outlier Momentum model (NM), to exhibit the improvement of predictability to stock return using monthly data from 1927:01 to 2017:12.

Compared with the existing models, the stock return forecasts by mixed models are more significant statistically and economically than the related counterparts. A mean-variance investor is willing to employ mixed models to predict stock return and allocate portfolio. Our findings survive other extension analysis, namely, multivariate information of forecasts analysis. Further analysis demonstrates that the mixed models contribute to outperform the history average benchmark and the existing models.

Our findings also have meaningful implications for robustness analysis including business cycle analysis, portfolio exercise considered transaction cost, rolling windows predictions, and alternative out-of-sample period. Demonstrating that under different robustness conditions, the mixed models still do a better job in forecasting stock return than the existing models. Notably, when we link to business cycle, the mixed models not only outperform the existing models and the benchmark model over recessions, but also beats counterparts during expansion periods. This finding fills in a gap that the predictability of stock returns in most literature is concentrated over recessions. (see, e.g. Rapach et al. (2010); Neely et al. (2014); Huang et al. (2015); Jiang et al. (2019); Wang et al. (2019) and Zhang et al. (2019)). Additionally, the prediction performance of the mixed models is not affected by the prediction approach and the out-of-sample period.

Overall, all mixed models improve significantly the prediction ability for stock return. Among the proposed mixed models, the Non-outlier Momentum model (NM) is most significant both in statistical and economic terms. The main reason is that the momentum model in Wang et al. (2018) has better prediction ability.

Acknowledgements

This work was supported by the National Natural Science Foundation of China granted [71771030, 71671018, 11301041] and Scientific Research Fund of Hunan Provincial Education Department [grant number 19A007].

References

1. Ang, A., Bekaert, G. (2007). Stock return predictability: is it there? Review of Financial Studies, 20 (3), 651-707.
2. Bahrami, A., Shamsuddin, A., Uylangco, K. (2019). Are advanced emerging market stock returns predictable? A regime-switching forecast combination approach. Pacific-Basin Finance Journal, 55, 142-160.
3. Bollerslev, T., Marrone, J., Xu, L., and Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. Journal of Financial and Quantitative Analysis, 49, 633-661.
4. Bollerslev, T., Tauchen, G., Zhou, H. (2009). Expected stock returns and variance risk premia. Review of Financial Studies, 22 (11), 4463-4492.
5. Campbell, J.Y. (1987). Stock returns and the term structure. Journal of Financial Economics, 18 (2), 373-399.
6. Campbell, J., and Thompson, S. P. (2008). Predicting excess stock returns out of sample: can anything beat the historical average? Review of Financial Studies, 21(4), 1509-1531.
7. Campbell, J.Y., Vuolteenaho, T. (2004). Bad beta, good beta. American Economic Review, 94 (5), 1249-1275.
8. Chen, J., Jiang, F., and Tong, G. (2017). Economic policy uncertainty in China and stock market expected returns. Accounting and Finance, 57, 1265-1286.
9. Chiang, I.-H. E., and Hughen, W. K. (2017). Do oil futures prices predict stock returns? Journal of Banking and Finance, 79, 129-141.
10. Christiansen, C., Schmeling, M., Schrimpf, A. (2012). A comprehensive look at financial volatility prediction by economic variables. Journal of Applied Econometrics, 27, 956-977.
11. Clark, T. E., and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics, 138, 291-311.
12. Cochrane, J.H. (1999). New facts in finance. Technical Report. National Bureau of Economic Research.
13. Cochrane, J.H. (2007). The dog that did not bark: a defense of return predictability. Review of Financial Studies, 21 (4), 1533-1575.
14. Dangl, T., Halling, M. (2012). Predictive regressions with time-varying coefficients. Journal of Financial Economics, 106 (1), 157-181.
15. Dai, Z., and Zhu, H. (2019). Forecasting stock market returns by combining Sum-of-the-parts and ensemble empirical mode decomposition, Applied Economics. doi:10.1080/00036846. 2019.1688 244.
16. Dai, Z., Zhu, H., and Wen, F. (2019). Two nonparametric approaches to mean absolute deviation portfolio selection model, Journal of Industrial and Management Optimization,

doi:10.3934/jimo.2019054.

17. Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253-263.
18. Fama, E.F., French, K.R. (1988). Dividend yields and expected stock returns. Journal of Financial Economics, 22 (1), 3-25.
19. Fama, E.F., French, K.R. (1989). Business conditions and expected returns on stocks and bonds. Journal of Financial Economics, 25 (1), 23-49.
20. Fama, E.F., Schwert, G.W. (1977). Asset returns and inflation. Journal of Financial Econometrics, 5 (2), 115-146.
21. Feunou, B., Jahan-Parvar, M. R., and Okou, C. (2015). Downside variance risk premium. Journal of Financial Econometrics, 1-43.
22. Gao, L., Han, Y., Li, S. Z., and Zhou, G. (2018). Market intraday momentum. Journal of Financial Economics, 129, 394-414.
23. Goyal, A., I. Welch. (2003). Predicting the equity premium with dividend ratios. Management Science 49,639-654.
24. Goyal, A., I. Welch. (2008). A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies, 21, 1455-1508.
25. Guo, H. (2006). Time-varying risk premia and the cross section of stock returns. Journal of Banking and Finance, 30 (7), 2087-2107.
26. Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. Review of Financial Studies, 28, 791-837.
27. Jiang, F., Lee, J. A., Martin, X., and Zhou, G. (2019). Manager sentiment and stock returns. Journal of Financial Economics. 132(1), 126-149.
28. Kelly, B., Pruitt, S. (2013). Market expectations in the cross-section of present values. Journal of Finance, 68 (5), 1721-1756.
29. Kilic, M., and Shaliastovich, I. (2019). Good and bad variance premia and expected returns. Management Science. 65(6), 2445-2945.
30. Lettau, M., Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. Journal of Finance, 56 (3), 815-849.
31. Lewellen, J. (2004). Predicting returns with financial ratios. Journal of Financial Economics, 74 (2), 209-235.
32. Lin, H., Wu, C., and Zhou, G. (2018). Forecasting corporate bond returns with a large set of predictors: An iterated combination approach. Management Science, 64, 3971-4470.
33. Ludvigson, S.C., Ng, S. (2007). The empirical risk–return relation: a factor analysis approach. Journal of Financial Economics, 83 (1), 171-222.
34. Ma, F., Liu, J., Wahab, M. I. M., and Zhang, Y. (2018). Forecasting the aggregate oil price

volatility in a data-rich environment. Economic Modelling, 72, 320-332.

35. Manela, A., and Moreira, A. (2017). News implied volatility and disaster concerns. Journal of Financial Economics, 123, 137-162.
36. Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. Management Science, 60, 1772-1791.
37. Nonejad, N. (2018). Déjà vol oil? Predicting S&P 500 equity premium using crude oil price volatility: Evidence from old and recent time-series data. International Review of Financial Analysis, 58, 260-270.
38. Pettenuzzo, D., Timmermann, A., and Valkanov, R. (2014). Forecasting stock returns under economic constraints. Journal of Financial Economics, 114, 517-553.
39. Phan, D. H. B., Sharma, S. S., and Narayan, P. K. (2015). Stock return forecasting: Some new evidence. International Review of Financial Analysis, 40, 38-51.
40. Rapach, D., Zhou, G. (2013). Forecasting Stock Returns. In: Handbook of Economic Forecasting, 2. Elsevier, pp. 328-383.
41. Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. Review of Financial Studies, 23, 821-862.
42. Rapach, D.E., Ringgenberg, M.C., Zhou, G. (2016). Short interest and aggregate stock returns. Journal of Financial Economics, 121 (1), 46-65.
43. Wang, Y., Liu, L., Ma, F., and Diao, X. (2018). Momentum of return predictability. Journal of Empirical Finance, 45, 141-156.
44. Wang, Y., Pan, Z., Liu, L., Wu, C. (2019). Oil price increases and the predictability of equity premium. Journal of Banking and Finance 102, 43-58.
45. Welch, I., and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies, 21, 1455-1508.
46. Zhang, Y., Ma, F., Shi, B., and Huang, D. (2018). Forecasting the prices of crude oil: An iterated combination approach. Energy Economics, 70, 472-483.
47. Zhang, Y., Wei, Y., Ma, F., Yi, Y. (2019). Economic constraints and stock return predictability: A new approach. International Review of Financial Analysis 63, 1-9.
48. Zhu, X., Zhu, J. (2013). Predicting stock returns: a regime-switching combination approach and economic links. Journal of Banking and Finance 37 (11), 4120-4133.

Table 1: Summary statistics

	Mean	Std.Dev.	Min.	Median	Max.	Skewness	Kurtosis	ρ(1)
Ret	0.005	0.054	-0.339	0.010	0.346	-0.433	10.961	0.086
DP	-3.373	0.462	-4.524	-3.348	-1.873	-0.218	2.657	0.992
DY	-3.368	0.459	-4.531	-3.341	-1.913	-0.246	2.637	0.992
EP	-2.738	0.417	-4.836	-2.790	-1.775	-0.602	5.620	0.986
DE	-0.635	0.329	-1.244	-0.627	1.380	1.516	9.031	0.991
SVAR	0.003	0.006	0.000	0.001	0.071	5.794	46.652	0.633
BM	0.568	0.266	0.121	0.542	2.028	0.779	4.464	0.985
NTIS	0.017	0.026	-0.058	0.017	0.177	1.651	11.247	0.979
TBL	0.034	0.031	0.000	0.030	0.163	1.078	4.282	0.993
LTY	0.051	0.028	0.018	0.042	0.148	1.085	3.603	0.996
LTR	0.005	0.024	-0.112	0.003	0.152	0.589	7.689	0.043
TMS	0.017	0.013	-0.037	0.018	0.046	-0.286	3.165	0.961
DFY	0.011	0.007	0.003	0.009	0.056	2.481	11.870	0.975
DFR	0.000	0.014	-0.098	0.001	0.074	-0.387	10.790	-0.120
INFL	0.002	0.005	-0.021	0.002	0.059	1.078	16.818	0.481

Notes. This table provides the summary statistics for the excess stock return of the S&P 500 index (Ret) and 14 macroeconomic variables. All data is available at homepage of Amit Goyal at http://www.hec.unil.ch/agoyal/. The sample period is from January 1927 to December 2017. ρ(1) reflects the first-order autocorrelation.

Table 2: Out-of- sample forecast performance

Panel A: Existing models								
	Original	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
DP	-0.132*	0.040*	0.607**	0.199	0.224	0.603**	0.809***	0.857***
DY	-0.476*	0.011**	0.175**	0.086*	0.189*	0.953***	1.033***	1.084***
EP	-1.488*	-0.585**	-0.808**	0.048*	0.064*	0.443**	0.202*	0.180*
DE	-1.396	-1.142	-0.618	-0.441	-0.495	-0.433	-0.247	-0.037
SVAR	0.156	0.076	0.849	0.107	0.124	0.104	0.161	0.155
BM	-1.579	-1.100	-0.900	-0.177	-0.128	-0.186	-0.121	0.220*
NTIS	-0.529	-0.529	0.435*	-0.334	-0.165	0.166*	-0.063	0.055
TBL	0.084*	0.271*	0.876**	0.858**	0.343*	[illegible]	0.384*	0.431*
LTY	-0.639*	0.327**	0.121**	0.745**	0.368**	[illegible]	0.076*	0.419*
LTR	-0.797	-0.677	0.062	-0.448	-0.379	-0.815	-0.952	-0.799
TMS	0.088	0.084	0.996**	0.234*	[illegible]	0.193	0.204	-0.184
DFY	-0.168	-0.168	0.648	-0.045	0.015	-0.042	0.022	0.032
DFR	-0.248	-0.429	0.501	0.194	[illegible]	-0.281	-0.108	-0.277
INFL	-0.060	-0.043	0.741	0.030	-0.011	-0.019	-0.004	-0.033
Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM_AVG
DP	**0.779****	**0.325***	**0.938***	**0.963***	**1.342****	**1.548****	**1.596****	**1.385****
DY	**0.715****	**0.104***	**0.738****	**0.841****	**1.605****	**1.685*****	**1.736*****	**1.516****
EP	**0.132****	**0.689***	[illegible]	**0.745****	**1.124****	**0.882****	**0.861****	**1.217****
DE	**-0.364**	**0.752**	**0.329**	**0.275**	**0.337**	**0.523***	**0.763***	**0.565**
SVAR	**0.871***	**0.843**	0.800	0.817	0.753	0.810	0.804	0.800
BM	**-0.408**	**-0.006**	**0.502***	**0.551***	**0.493***	**0.558***	**0.899***	**0.828***
NTIS	0.435*	[illegible]	**0.466**	**0.634***	**0.966****	**0.737***	**0.895***	**0.840***
TBL	**1.062****	[illegible]	**1.624****	**1.108****	**1.186****	**1.149****	**1.197****	**1.388****
LTY	**1.086****	**0.739****	**1.501****	**1.123****	**1.276****	**0.832****	**1.175****	**1.431****
LTR	**0.182**	**0.507**	**0.452**	**0.522**	**0.086**	**-0.052**	**0.102**	**0.300**
TMS	0.991**	**1.187***	**1.034***	0.926*	0.993*	**1.004***	0.615	0.971*
DFY	0.648	**0.656**	**0.752**	**0.812**	**0.755**	**0.820**	**0.829**	**0.796**
DFR	**0.320**	**0.600**	**0.942***	**0.653**	**0.467**	**0.641**	**0.471**	**0.668**
INFL	**0.758**	**0.853***	**0.817**	**0.776**	**0.768**	**0.782**	**0.754**	**0.789**

Notes. This table reports the out-of-sample forecasting performance of existing economic models and mixed models based on 14 economic variables. The forecast performance is evaluated by out-of-sample R^2 (R^2_{OOS}) which is multiplied by 100 to denote the percent value. Original refers to original forecast with univariate predictive regression model employing a recursive estimation window, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019) economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We evaluate the statistical significance by Clark and West (2007) statistic with *, ** and *** denoting significance at 10%, 5% and 1% levels, respectively. The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Table 3: Portfolio performance

Panel A: Existing models								
	Original	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
DP	-0.368	-0.366	-0.259	1.066	0.841	1.385	1.924	1.932
DY	-0.734	-0.466	-0.701	0.909	0.978	2.234	2.313	2.223
EP	-0.064	0.244	-0.028	1.356	1.443	2.039	1.861	1.362
DE	-0.078	-0.072	0.068	0.337	0.443	0.318	0.269	0.266
SVAR	0.635	0.570	0.786	0.420	0.549	0.295	0.361	0.351
BM	-1.559	-1.209	-1.573	0.345	0.400	0.591	1.205	1.591
NTIS	0.285	0.287	0.629	0.377	0.145	0.396	0.096	0.223
TBL	0.814	1.040	1.047	1.761	1.436	1.528	1.192	1.331
LTY	0.735	0.999	0.903	2.048	1.428	1.389	1.449	1.780
LTR	-0.680	-0.586	-0.426	-0.006	0.034	-1.258	-1.281	-1.309
TMS	1.272	1.220	1.656	1.217	1.000	0.975	1.103	0.696
DFY	-0.315	-0.315	-0.040	-0.086	0.053	-0.047	0.053	0.049
DFR	-0.065	0.011	0.051	0.204	0.159	0.112	0.255	0.034
INFL	0.039	0.070	0.283	0.444	0.041	0.152	0.166	0.050
Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
DP	**-0.257**	-0.842	**1.176**	**0.950**	**1.495**	**2.034**	**2.042**	**1.676**
DY	**-0.433**	-1.055	0.942	**1.011**	**2.267**	**2.346**	**2.257**	**2.146**
EP	**0.281**	**0.236**	**1.203**	**1.479**	**2.076**	**1.898**	**1.399**	**2.399**
DE	**0.074**	**0.181**	**0.471**	**0.577**	**0.453**	**0.404**	**0.486**	**0.630**
SVAR	0.721	0.627	0.570	0.700	0.380	0.445	0.436	0.398
BM	**-1.206**	**-1.263**	0.331	**0.386**	**0.577**	**1.191**	**1.577**	**1.128**
NTIS	**0.631**	0.502	0.597	0.365	0.616	0.316	0.504	0.486
TBL	**1.273**	0.793	**1.931**	**1.607**	**1.699**	**1.363**	**1.501**	**1.757**
LTY	**1.167**	0.525	**2.205**	**1.586**	**2.046**	**1.606**	**1.937**	**2.023**
LTR	**-0.332**	**-0.445**	**0.366**	**0.405**	**-0.888**	**-0.910**	**-0.939**	**-0.309**
TMS	1.603	1.101	1.438	1.221	1.195	1.324	0.917	1.304
DFY	-0.040	-0.047	**0.130**	**0.269**	**0.168**	**0.269**	**0.265**	**0.224**
DFR	**0.127**	**0.117**	**0.320**	**0.275**	**0.228**	**0.371**	**0.150**	**0.211**
INFL	**0.314**	**0.379**	**0.645**	**0.242**	**0.353**	**0.366**	**0.250**	**0.348**

Notes. This table reports certainty equivalent return (CER) of existing economic models and mixed models based on 14 economic variables. Original refers to original forecast with univariate predictive regression model employing a recursive estimation window, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019)

economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We report CER differences between assets of interest and natural benchmarks multiplied by 1200 to show annual percentages. The stock weight is restricted within a range of -0.5 and 1.5 following Rapach et al. (2016) and Zhang et al. (2019). The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Table 4: Multivariate results

Panel A: Existing models								
	Original	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
Average	-0.478*	-0.086*	0.232**	0.224*	0.313*	0.509**	0.734**	1.047***
Diffusion index	0.241**	0.283**	1.132**	0.556 **	0.487**	0.523**	0.630**	1.001***
Mean	0.507***	0.416**	1.280**	0.478**	0.416**	0.514**	0.503**	0.555***
Median	0.400***	0.400***	1.159*	0.224**	0.227***	0.190***	0.146**	0.194***
Trimmed mean	0.452**	0.421**	1.215**	0.412**	0.365**	0.433**	0.460***	0.489***
DMSPE(1)	0.537***	0.458**	1.306**	0.500**	0.450**	0.558***	0.562***	0.634***
DMSPE(0.9)	0.517**	0.442**	1.285**	0.490**	0.440**	0.546**	0.551**	0.626***
Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
Average	**0.626****	0.086*	**0.934****	**1.023****	**1.22[illegible]****	**1.444****	**1.757*****	**1.445****
Diffusion index	**1.173****	1.006**	**1.446****	**1.377****	**1.414****	**1.521****	**1.892*****	**1.644****
Mean	1.205**	0.890*	1.231**	1.169*	1.264**	1.252**	**1.309****	**1.281****
Median	1.159*	0.974*	0.994*	0.997*	0.959*	0.916*	0.972*	0.996*
Trimmed mean	1.188**	0.944*	1.164	1.117*	1.186*	1.213**	1.244**	1.215 **
DMSPE(1)	1.242**	0.869*	1.249**	1.200**	**1.304****	**1.309****	**1.386****	**1.329****
DMSPE(0.9)	1.225**	0.865*	1.239**	1.189**	**0**	**1.298****	**1.377****	**1.319****

Notes. This table reports the multivariate results of existing multivariate model based on 14 macroeconomic predictors. The forecast performance is evaluated by out-of-sample R^2 (R^2_{OoS}) which is multiplied by 100 to denote the percent value. In addition to average approach, employing the average of 14 predictor to forecast stock return, there are another five popular combination methods including the diffusion index, mean, median, trimmed mean, DMSPE(1) and DMSPE(0.9) following Pettenuzzo et al. (2014). Original refers to original forecast with univariate predictive regression model employing a recursive estimation window, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019) economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We evaluate the statistical significance by Clark and West (2007) statistic with *, ** and *** denoting significance at 10%, 5% and 1% levels, respectively. The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Table 5: Out-of-sample performance over business cycle with existing models

	Original	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
Recessions								
DP	1.783**	1.768**	2.494**	1.546**	1.231*	1.610**	1.779**	1.605**
DY	2.700**	2.612***	3.336**	2.278**	2.172**	2.501***	2.761***	2.363**
EP	-2.195	-0.956	-1.746	1.652	1.979	1.459	0.046	-0.158
DE	-2.637	-1.831	-1.978	0.039	0.463	-0.046	-0.471	-0.443
SVAR	0.666	0.412	0.873	0.811	0.751	0.403	0.411	0.408
BM	-0.370	-0.390	0.223	1.729*	2.386*	0.951	0.716	0.521
NTIS	-3.712	-3.712	-2.477	-0.410	0.242	0.342	0.052	-0.202
TBL	1.293	0.668	2.162	2.890	2.442	1.[illegible]89	1.745	1.952
LTY	0.591	0.894	1.376	2.807	2.823	1.7[illegible]2	1.012	1.016
LTR	0.882	0.877	1.497	0.244	0.330	0.122	-0.114	0.476
TMS	0.892	0.945	1.823*	1.232*	1.31[illegible]**	1.041	1.169*	0.897
DFY	-0.168	-0.168	0.675	0.066	0.250*	0.198	0.182	0.134
DFR	-0.583	-1.044	0.111	1.449	0.52[illegible]	-0.492	-0.706	-0.652
INFL	-0.415	-0.358	0.409	0.007	-0.154	0.160	-0.003	-0.142
Expansions								
DP	-0.828	-0.588	-0.079	-0.290	-0.142	0.237	0.456	0.585*
DY	-1.631	-0.934	-0.97[illegible]	-0.710	-0.532	0.391	0.405	0.619*
EP	-1.232	-0.451	-0.4[illegible]7	-0.535	-0.632	0.074	0.258	0.303
DE	-0.945	-0.892	[illegible]0.12[illegible]	-0.615	-0.843	-0.574	-0.166	0.110
SVAR	-0.029	-0.046	0.841	-0.148	-0.104	-0.005	0.070	0.063
BM	-2.018	-1.358	1.308	-0.870	-1.041	-0.599	-0.426	0.111
NTIS	0.627**	0.628**	1.492**	-0.307	-0.314	0.102	-0.104	0.148
TBL	-0.355	0.127	0.408	0.120	-0.420	-0.150	-0.111	-0.121
LTY	-1.085	0.[illegible]21	-0.335	-0.004	-0.525	0.083	-0.264	0.202
LTR	-1.407	-1.242	-0.459	-0.700	-0.636	-1.156	-1.257	-1.262
TMS	-0.204	-0.229	0.695	-0.129	-0.307	-0.115	-0.147	-0.577
DFY	-0.168	-0.168	0.639	-0.085	-0.071	-0.129	-0.036	-0.005
DFR	-0.126	-0.205	0.642	-0.262	-0.320	-0.204	0.110	-0.141
INFL	0.069	0.071	0.862	0.038	0.041	-0.084	-0.005	0.007

Notes. This table reports the out-of-sample forecasting performance over business cycle based on existing models. The forecast performance is evaluated by out-of-sample R^2 (R^2_{OoS}) which is multiplied by 100 to denote the percent value.. Original refers to original forecast with univariate predictive regression model employing a recursive estimation window, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019)

economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. We evaluate the statistical significance by Clark and West (2007) statistic with *, ** and *** denoting significance at 10%, 5% and 1% levels, respectively. The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Table 6: Out-of-sample performance over business cycle with mixed models

	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
Recessions								
DP	2.480**	2.055**	2.257**	1.943*	2.321**	2.491**	2.317**	2.299**
DY	3.248***	2.662**	2.914**	2.808**	3.137***	**3.397*** **	2.998**	3.118***
EP	**-0.368**	**1.765**	**2.101**	**2.429**	**1.908**	**0.495**	**0.291**	**1.809**
DE	**-1.172**	**0.830**	**0.698**	**1.121**	**0.613**	**0.188**	**0.328**	**0.682**
SVAR	**1.000**	**1.030**	**1.018**	**0.958**	**0.610**	**0.618**	**0.615**	**0.769**
BM	0.203	**1.808***	**2.322****	**2.979****	**1.543**	**1.308**	**1.114**	**2.013**
NTIS	-2.477	**0.945**	**0.361**	**1.013**	**1.113**	**0.823**	**0.569**	**0.831**
TBL	1.536	1.910*	**3.661***	**3.213***	**2.760**	**2.516**	**2.723**	**3.082**
LTY	**1.678**	**1.643**	**3.578**	**3.594***	**2.493**	[illegible]	**1.787**	**2.902**
LTR	1.492	1.236	1.015	1.101	0.893	0.657	1.247	1.066
TMS	**1.876***	1.792*	**2.003****	**2.090****	1.812*	**1.940***	1.668*	**1.918***
DFY	0.675	**0.756**	**0.837**	**1.021**	**0.969**	**0.953**	**0.905**	**0.940**
DFR	-0.350	**0.789**	**2.143**	**1.219**	[illegible]	-0.012	0.042	**0.760**
INFL	**0.467**	**1.257**	**0.778**	**0.617**	**0.931**	**0.768**	**0.629**	**0.757**
Expansion								
DP	**0.161**	**-0.304**	**0.459**	**0.607**	**0.987**	**1.205**	**1.334**	**1.053**
DY	**-0.206**	**-0.826**	**-0.053**	**0.126**	**1.048**	**1.063**	**1.277***	**0.933**
EP	**0.314***	**0.298**	**0.229**	**0.133**	**0.839***	**1.023***	**1.068***	**1.001**
DE	**-0.071**	**0.724**	[illegible]	**-0.033**	**0.237**	**0.645**	**0.921**	**0.523**
SVAR	0.824	0.775	[illegible]	0.766	0.805	**0.880**	0.873	0.811
BM	**-0.630**	**-0.665**	**-0.159**	**-0.331**	**0.111**	**0.285**	**0.821**	**0.397**
NTIS	**1.494****	0.949	0.503	0.497	0.912	0.706	1.013	0.843
TBL	**0.890**	**0.658**	**0.884**	**0.344**	**0.613**	**0.652**	**0.642**	**0.773**
LTY	**0.871**	[illegible]	**0.747**	**0.226**	**0.833**	**0.487**	**0.952**	**0.897**
LTR	**-0.295**	**0.242**	**0.248**	**0.311**	**-0.208**	**-0.309**	**-0.314**	**0.022**
TMS	0.669	**0.967**	0.681	0.503	**0.696**	0.663	0.233	0.626
DFY	0.639	0.620	**0.722**	**0.736**	0.678	**0.771**	**0.802**	**0.744**
DFR	0.563	0.532	0.506	0.448	0.564	**0.878**	0.627	0.635
INFL	**0.864**	0.705	0.831	0.833	0.708	0.788	0.800	0.801

Notes. This table reports the out-of-sample forecasting performance over business cycle based on existing models. The forecast performance is evaluated by out-of-sample R^2 (R^2_{OoS}) which is multiplied by 100 to denote the percent value. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We evaluate the statistical significance by Clark and

West (2007) statistic with *, ** and *** denoting significance at 10%, 5% and 1% levels, respectively. The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Table 7: Portfolio performance considered transaction cost

Panel A: Existing models								
	Original	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
DP	-0.510	-0.470	-0.413	-1.097	-0.318	0.644	1.369	1.411
DY	-1.119	-0.616	-1.090	-2.066	-0.596	1.164	1.437	1.428
EP	-0.396	0.069	-0.359	-0.962	0.187	1.284	1.200	0.653
DE	-0.119	-0.109	0.015	-0.326	0.111	0.068	0.024	0.066
SVAR	0.548	0.491	0.690	0.310	0.475	0.224	0.282	0.276
BM	-1.839	-1.379	-1.851	-2.284	-1.014	-0.280	0.439	0.908
NTIS	0.043	0.048	0.346	-0.548	-0.439	-0.057	-0.287	-0.127
TBL	0.619	0.939	0.822	0.520	0.767	1.074	0.752	0.980
LTY	0.539	0.951	0.685	0.196	0.465	1.307	0.916	1.397
LTR	-2.894	-2.712	-2.629	-1.336	-1.145	-2.402	-2.453	-2.503
TMS	0.998	0.959	1.335	0.357	0.450	0.537	0.708	0.271
DFY	-0.359	-0.359	-0.117	-0.269	-0.059	-0.132	-0.017	-0.030
DFR	-1.450	-1.351	-1.330	-0.619	-0.659	-0.643	-0.472	-0.656
INFL	-0.520	-0.485	-0.307	0.012	-0.367	-0.241	-0.201	-0.335

Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
DP	**-0.373**	-0.925	-1.000	**-0.223**	**0.739**	**1.464**	**1.507**	**0.941**
DY	**-0.587**	**-1.149**	-2.037	**-0.567**	**1.193**	**1.466**	**1.457**	**0.887**
EP	**0.106**	**0.096**	-0.925	**0.219**	**1.316**	**1.233**	**0.685**	**1.537**
DE	**0.025**	**0.155**	-0.202	**0.235**	**0.192**	**0.148**	**0.276**	**0.352**
SVAR	0.633	0.567	0.450	0.616	0.299	0.357	0.351	0.319
BM	**-1.376**	**-1.375**	-2.296	**-1.026**	**-0.392**	**0.427**	**0.896**	**0.079**
NTIS	**0.351**	**0.357**	-0.354	-0.246	**0.137**	-0.093	0.127	0.007
TBL	**1.143**	**0.874**	0.669	**0.916**	**1.224**	**0.901**	**1.130**	**1.201**
LTY	**1.097**	0.483	0.334	0.603	**1.446**	**1.054**	**1.535**	**1.261**
LTR	**-2.447**	**-1.795**	**-0.969**	**-0.779**	**-2.051**	**-2.102**	**-2.152**	**-1.494**
TMS	1.296	0.948	0.550	0.644	0.730	0.902	0.465	0.835
DFY	-0.117	**-0.093**	**-0.079**	**0.131**	**0.058**	**0.173**	**0.160**	**0.114**
DFR	**-1.231**	**-0.751**	**-0.498**	**-0.539**	**-0.522**	**-0.352**	**-0.535**	**-0.509**
INFL	**-0.272**	**-0.006**	**0.164**	**-0.191**	**-0.065**	**-0.025**	**-0.159**	**-0.037**

Notes. This table reports the certainty equivalent return (CER) of transaction costs at 50 basis points per transaction. Original refers to original forecast with univariate predictive regression model employing a recursive estimation window, while the CT, ZY, MoP(k) correspond to the

Campbell and Thompson (2008) economic constraint, Zhang et al. (2019) economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We report CER differences between assets of interest and natural benchmarks multiplied by 1200 to show annual percentages. The stock weight is restricted within a range of -0.5 and 1.5 following Rapach et al. (2016) and Zhang et al. (2019). The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Table 8: Out-of- sample forecast performance based on rolling estimation windows

Panel A: Existing models								
	Orignal	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
DP	-0.619	-0.441	0.158*	-0.032	0.191	-0.314	-0.013	0.224*
DY	-0.217	-0.083	0.456*	0.054	-0.055	0.033	0.190*	0.306*
EP	-1.849	-0.623	-1.177	-0.259	-0.512	-0.516	-0.362	-0.317
DE	-1.535	-0.341	-0.810	-0.518	-0.091	-0.787	-0.555	-0.322
SVAR	-1.751	-1.668	-0.440	-1.003	-1.066	-0.708	-0.112	-0.111
BM	-2.133	-1.466	-1.405	-0.723	-0.402	-0.582	-0.429	-0.127
NTIS	-1.418	-0.831	-0.801	-0.482	-0.447	-0.230	-0.071	-0.053
TBL	-2.407	-1.004	-1.705*	-0.124*	-0.731	-0.2[illegible]9*	0.683**	0.313*
LTY	-1.209	-0.094*	-0.506*	0.441*	0.589**	0.351*	0.032*	0.039*
LTR	-0.791	-0.548	0.102**	-0.034*	-0.723	-1.175	-1.247	-1.223
TMS	-1.043*	-0.812	-0.199**	-0.257*	-0.4[illegible]2	-0.523	-0.178*	0.082*
DFY	-1.943	-0.890	-1.412	-0.965	-0.67[illegible]	-0.725	-0.444	-0.839
DFR	-2.242	-1.755	-1.556	-1.133	-0.94[illegible]	-1.287	-0.728	-1.036
INFL	-0.753	-0.538	-0.020	-0.181	-0.016	0.245	0.228	0.352
Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
DP	**0.351***	**0.535**	**0.693***	**0.917***	**0.411**	**0.712***	**0.949***	**0.870***
DY	**0.678****	**0.652***	**0.70[illegible]***	**0.596***	**0.683***	**0.841***	**0.979****	**0.917***
EP	**0.075***	**0.403**	**0.[illegible]14***	**0.161**	**0.156**	**0.310**	**0.408***	**0.480**
DE	**0.415***	**0.960***	**0.201**	**0.629***	**-0.067**	**0.164**	**0.398***	**0.497***
SVAR	**-0.201**	**0.129**	**0.308**	**0.245**	**-0.114**	**0.483**	**0.484**	**0.348**
BM	**-0.712**	**-0.154**	**0.005**	**0.326**	**0.146**	**0.299**	**0.687***	**0.415**
NTIS	**-0.087**	**0.[illegible]97**	**0.142**	**0.176**	**0.394***	**0.553***	**0.563***	**0.533***
TBL	**-0.266**	**0.4[illegible]5**	**0.578****	**-0.029***	**0.483****	**1.385****	**1.028****	**1.065****
LTY	**0.620****	**0.657***	**1.144****	**1.292****	**1.055****	**0.736****	**0.780****	**1.270****
LTR	**0.345****	**0.838***	**0.859****	**0.196***	-0.282	-0.354	-0.330	0.258*
TMS	**0.032****	**1.097****	**0.522****	**0.338***	**0.257***	**0.602****	**0.862****	**0.776****
DFY	**-0.199**	**0.257**	**-0.434**	**-0.148**	**-0.193**	**0.087**	**-0.308**	**0.006**
DFR	**-1.050**	**0.324**	**-0.412**	**-0.228**	**-0.565**	**-0.006**	**-0.315**	**-0.187**
INFL	**0.195**	**0.511**	**0.580**	**0.746***	**0.978***	**0.961***	**1.085***	**1.020***

Notes. This table reports the out-of-sample forecasting performance of alternative prediction approach. The forecast performance is evaluated by out-of-sample R^2 (R^2_{OoS}) which is multiplied by 100 to denote the percent value. Original refers to original forecast with univariate predictive regression model employing a rolling estimation windows, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019) economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We evaluate the statistical significance by Clark and West (2007) statistic with *, ** and *** denoting significance at 10%, 5% and 1% levels, respectively. The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Table 9: Out-of- sample forecast performance (1979:01-2017:12)

Panel A: Existing models								
	Orignal	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
DP	-1.008	-0.720	0.230	-0.408	-0.214	-0.038	0.232	0.657*
DY	-1.527	-0.840	-0.435	-0.710	-0.405	0.219	0.446	0.768*
EP	-1.653	-0.417	-0.513	0.046	-0.544	0.390	-0.149	0.068
DE	-0.617	-0.192	0.686	-0.008	-0.339	-0.293	-0.499	-0.343
SVAR	0.328	0.194	1.488*	0.214	0.231	0.205	0.288	0.275
BM	-2.423	-1.597	-1.286	-0.748	-0.840	-0.370	-0.630	0.082
NTIS	-1.003	-1.002	0.612*	-0.655	-0.679	-0.062	-0.406	-0.130
TBL	-0.813	-0.385	0.513	0.321	-0.831	-0.164	-0.241	-0.160
LTY	-1.028	-0.273	0.244	0.189	-0.938	-0.241	-0.776	-0.401
LTR	-0.012	0.020	1.426*	-0.476	-0.148	-0.759	-1.053	-0.997
TMS	-0.130	-0.138	1.390*	0.191	[illegible]	0.178	-0.072	-0.428
DFY	-0.122	-0.122	1.245	-0.021	0.081	-0.001	0.051	0.054
DFR	0.249	-0.063	1.502*	0.440	0.072	-0.200	0.127	-0.115
INFL	-0.329	-0.322	1.012	-0.157	-0.330	-0.340	-0.155	-0.307

Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
DP	**0.518**	**0.409**	**0.836**	**1.024**	**1.200**	**1.470***	**1.895***	**1.392***
DY	**0.338**	**0.213**	**0.381**	**0.686**	**1.311***	**1.537***	**1.859***	**1.324***
EP	**0.785***	**1.183***	**1.186***	**0.596**	**1.530***	**0.991**	**1.208***	**1.338***
DE	**1.110**	**1.260**	**1.282**	**0.951**	**0.997**	**0.791**	**0.996**	**1.044**
SVAR	**1.524***	1.457	1.375	1.391	1.293	1.375	1.363	1.363
BM	**-0.438**	**0.063**	**0.390**	**0.297**	**0.767**	**0.508**	**1.219***	**0.866**
NTIS	0.612*	**[illegible]**	**0.684**	**0.660**	**1.277***	**0.933**	**1.277***	**1.113***
TBL	**0.940**	**[illegible]**	**1.603***	0.451	**1.118***	**1.040***	**1.121***	**1.213***
LTY	**0.999**	**0.812**	**1.455***	**0.327**	**1.025**	**0.490**	**0.864**	**0.950**
LTR	**1.459***	1.265	1.032	1.360*	0.749	0.455	0.512	0.911
TMS	1.382*	**1.627***	**1.530***	1.204*	**1.518***	1.267*	0.912	1.371*
DFY	1.245	**1.260**	**1.314**	**1.416**	**1.334**	**1.386**	**1.389**	**1.370**
DFR	1.191	1.410	**1.693***	1.325	1.053	1.381	1.139	1.365
INFL	**1.019**	**1.289**	**1.160**	0.988	0.978	**1.163**	1.011	1.070

Notes. This table reports the out-of-sample forecasting performance of alternative out-of-sample from 1979:01 to 2017:12. The forecast performance is evaluated by out-of-sample R^2 (R^2_{OoS}) which is multiplied by 100 to denote the percent value. Original refers to original forecast with univariate predictive regression model employing a rolling estimation windows, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019) economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We evaluate the statistical significance by Clark and West (2007) statistic with *, ** and *** denoting significance at 10%, 5% and 1% levels, respectively.

Table 10: Out-of- sample forecast performance (1989:01-2017:12)

Panel A: Existing models								
	Orignal	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
DP	-1.817	-1.321	-1.347	-1.010	-0.331	-0.143	0.300	0.654
DY	-2.718	-1.401	-2.298	-1.596	-0.522	0.177	0.450	0.658
EP	-1.932	-0.059	-1.635	-0.197	-0.204	0.592	0.880	1.208
DE	-0.632	-0.003	-0.196	-0.117	-0.097	-0.204	-0.211	-0.146
SVAR	0.632	0.434	0.769	0.426	0.458	0.313	0.431	0.415
BM	-2.858	-1.471	-2.466	-1.431	-0.818	-0.288	-0.042	0.606
NTIS	-2.264	-2.263	-1.447	-1.115	-0.446	0.297*	0.533**	-0.023
TBL	-0.223	-0.214	0.352	-0.180	0.159	-0.1[illegible]5	-0.261	-0.123
LTY	-0.006	-0.006	0.512	-0.197	-0.045	-0.169	-0.242	-0.062
LTR	0.022	-0.074	0.429	-0.729	-0.389	-0.745	-0.613	-0.625
TMS	-0.760	-0.760	-0.145	-0.567	0.01[illegible]	-0.225	-0.225	-0.541
DFY	-0.290	-0.290	0.267	-0.126	0.051	-0.107	-0.066	-0.044
DFR	-0.328	-0.788	0.131	0.155	-0.40[illegible]	-0.840	-0.508	-0.851
INFL	-0.523	-0.523	0.022	-0.459	-0.222	-0.361	-0.163	-0.262
Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
DP	**-0.851**	**-0.859**	**-0.540**	**0.139**	**0.327**	**0.770**	**1.124**	**0.513**
DY	**-0.981**	**-0.986**	**-1.176**	**-0.102**	**0.597**	**0.871**	**1.078***	**0.484**
EP	**0.329**	**0.216**	**0.100**	**0.093**	**0.889**	**1.177**	**1.504***	**1.001**
DE	**0.432**	**0.395**	**0.218**	**0.338**	**0.231**	**0.225**	**0.363***	**0.340**
SVAR	**0.823***	0.726	0.563	0.595	0.450	0.568	0.552*	0.551
BM	**-1.080**	**-1.0[illegible]**	**-1.039**	**-0.426**	**0.104**	**0.349**	**0.998***	**0.220**
NTIS	-1.447	**0.418**	**-0.605**	**0.063**	**0.807****	**1.043****	**0.486***	**0.478**
TBL	**0.360**	0.204	0.330	**0.669***	**0.404**	0.248	**0.386**	**0.441**
LTY	0.512	0.236	0.313	0.465	0.340	0.267	0.448	0.379
LTR	0.332	0.211	-0.220	0.120	-0.236	-0.103	-0.115	-0.058
TMS	-0.145	**0.556**	**-0.058**	**0.523**	**0.285**	**0.284**	**-0.032**	**0.270**
DFY	0.267	**0.383**	**0.383**	**0.458**	**0.403**	**0.443**	**0.465**	**0.432**
DFR	-0.330	**0.336**	**0.613**	**0.058**	-0.382	-0.050	-0.392	0.036
INFL	0.022	**0.481**	0.050	**0.287**	0.149	**0.347**	**0.248**	0.222

Notes. This table reports the out-of-sample forecasting performance of alternative out-of-sample from 1989:01 to 2017:12. The forecast performance is evaluated by out-of-sample R^2 (R^2_{OOS}) which is multiplied by 100 to denote the percent value. Original refers to original forecast with univariate predictive regression model employing a rolling estimation windows, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019) economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We evaluate the statistical significance by Clark and West (2007) statistic with *, ** and *** denoting significance at 10%, 5% and 1% levels, respectively.

Table 11: Portfolio performance with alternative asset weight limitation

Panel A: Existing models								
	Original	CT	ZY	MoP(1)	MoP(3)	MoP(6)	MoP(9)	MoP(12)
DP	-0.859	-0.366	-0.257	0.949	0.836	1.380	1.856	1.851
DY	-1.206	-0.466	-0.433	1.057	1.066	2.106	2.178	2.098
EP	-0.073	0.244	0.281	1.421	1.554	2.067	1.998	1.531
DE	-0.178	-0.072	0.074	0.340	0.446	0.322	0.273	0.270
SVAR	0.609	0.570	0.721	0.420	0.500	0.246	0.296	0.287
BM	-1.904	-1.209	-1.206	0.507	0.625	0.781	1.244	1.469
NTIS	0.285	0.287	0.631	0.376	0.146	0.394	0.097	0.225
TBL	1.005	1.040	1.273	1.810	1.443	1.485	1.283	1.447
LTY	0.505	0.999	1.167	2.172	1.540	1.900	1.469	1.848
LTR	-0.807	-0.586	-0.332	0.008	0.056	-1.186	-1.212	-1.247
TMS	1.220	1.220	1.603	1.169	0.962	0.924	1.053	0.653
DFY	-0.315	-0.315	-0.040	-0.086	0.053	-0.047	0.053	0.049
DFR	0.056	0.011	0.127	0.238	0.202	0.180	0.323	0.105
INFL	0.039	0.070	0.314	0.469	0.072	0.183	0.197	0.080

Panel B: Mixed models								
	GN	NP	NM(1)	NM(3)	NM(6)	NM(9)	NM(12)	NM-AVG
DP	-0.257	-0.842	1.059	0.945	1.490	1.966	1.961	1.633
DY	-0.433	-1.055	1.090	1.099	2.139	2.211	2.131	2.042
EP	0.281	0.236	1.457	1.591	2.104	2.035	1.568	2.310
DE	0.074	0.181	0.475	0.581	0.456	0.407	0.490	0.633
SVAR	0.721	0.627	0.570	0.651	0.331	0.381	0.371	0.419
BM	-1.206	-1.265	0.510	0.628	0.784	1.247	1.473	1.061
NTIS	0.631	0.502	0.596	0.366	0.615	0.317	0.506	0.488
TBL	1.273	0.993	1.981	1.614	1.660	1.454	1.618	1.799
LTY	1.167	0.525	2.330	1.697	2.057	1.627	2.006	1.959
LTR	-0.332	-0.445	0.379	0.427	-0.816	-0.842	-0.876	-0.264
TMS	1.603	1.101	1.389	1.183	1.145	1.273	0.874	1.252
DFY	-0.040	-0.047	0.130	0.269	0.168	0.269	0.265	0.224
DFR	0.127	0.117	0.354	0.318	0.296	0.439	0.221	0.248
INFL	0.314	0.379	0.670	0.272	0.383	0.397	0.281	0.373

Notes. This table reports certainty equivalent return (CER) of existing economic models and mixed models based on 14 economic variables. Original refers to original forecast with univariate predictive regression model employing a recursive estimation window, while the CT, ZY, MoP(k) correspond to the Campbell and Thompson (2008) economic constraint, Zhang et al. (2019)

economic constraint approach and momentum of return predictability strategy, where k is the look-back period with k=1,3,6,9,12, respectively. The GN, NP and NM(k) correspond to the mixed models, namely, he Generalized Non-outlier model, the Non-outlier Positive model and the Non-outlier Momentum model, respectively. We report CER differences between assets of interest and natural benchmarks multiplied by 1200 to show annual percentages. The stock weight is restricted within a range of 0 and 1.5 following Rapach et al. (2010), Neely et al. (2014). The in-sample period is 1927:01-1947:12, while the out-of-sample period is 1948:01-2017:12.

Author statement

Title: Stock return predictability from mixed model perspective

Dr. Zhifeng Dai has made substantial contributions to the conception or design of the work, and drafted the work or revised it.

Miss Huan Zhu has made substantial contributions to the acquisition, analysis, or interpretation of data for the work.

All persons who have made substantial contributions to the work are reported in the manuscript.

Dr. Zhifeng Dai and Miss Huan Zhu have approved the final version to be published.

Highlights

- We propose an efficient strategy to predict stock market returns by mixing existing forecasting models.
- Our strategies obtain more accurate return forecasts than existing forecasting models.
- An investor can realize sizeable economic gains using our new approach.
- The prediction performance of mixed forecasting models is robust to a series of extension test.