Equity Premium Predictability: Combination Forecasts versus Multivariate Regression Predictions ☆

Claudia Champagne [a], Stéphane Chrétien [b,*], Frank Coggins [c]

[a] *Associate Professor of Finance, Department of Finance*
[c] *Full Professor of Finance, Department of Finance*
Faculté d'administration, Université de Sherbrooke, Chaire Desjardins en finance responsable and GReFA, 2500 Boul.Université, Sherbrooke, QC, Canada, J1K 2R1

[b] *Full Professor of Finance and Investors Group Chair in Financial Planning*
Finance, Insurance and Real Estate Department
Faculty of Business Administration, Laval University, CRREP, GReFA and LABIFUL
Pavillon Palasis-Prince, 2325, rue de la Terrasse, Quebec City, QC, Canada, G1V 0A6

January 2018

Abstract

This paper examines the combination forecast and multivariate regression approaches for equity premium predictability. We evaluate 27 specifications with a unique Canadian database to avoid the data mining inherent in using common U.S. data. We find significant predictive evidence for most models. In sample, multivariate regression predictions perform better than combination forecasts, although regression results display evidence of instability and overfitting. Out of sample, combination forecasts are superior when relying on many individual models, but imposing economic restrictions on multivariate regression predictions yields similar performance. Both approaches show that incorporating information from numerous variables improves forecasting precision and economic value.

JEL Classifications: G11, G12, G14, C53

Keywords: Equity Premium Predictability; Combination Forecast; Multivariate Regression Prediction

---

☆ We thank Marie-Hélène Gagnon, Amit Goyal, Philippe Grégoire, Alexandre Kopoin, Colin Smith and seminar participants at the 57th Annual Meeting of the *Société canadienne de science économique* in Ottawa and the 14th Annual International Conference on Finance in Athens for helpful comments, and Leticia Eugenia Corona Callejas, Frédéric Gagnon and Stéphanie Hamel for excellent research assistance. This work was supported by the Social Sciences and Humanities Research Council (Chrétien), the Investors Group Chair in Financial Planning and Faculty of Business Administration at Laval University (Chrétien), and the *Chaire Desjardins en finance responsable* and *Faculté d'administration*, *Université de Sherbrooke* (Champagne and Coggins).

* Corresponding author. Voice: 1 418 656-2131, ext. 3380. E-mail: stephane.chretien@fsa.ulaval.ca.

Equity Premium Predictability: Combination Forecasts versus Multivariate Regression Predictions

Abstract

This paper examines the combination forecast and multivariate regression approaches for equity premium predictability. We evaluate 27 specifications with a unique Canadian database to avoid the data mining inherent in using common U.S. data. We find significant predictive evidence for most models. In sample, multivariate regression predictions perform better than combination forecasts, although regression results display evidence of instability and overfitting. Out of sample, combination forecasts are superior when relying on many individual models, but imposing economic restrictions on multivariate regression predictions yields similar performance. Both approaches show that incorporating information from numerous variables improves forecasting precision and economic value.

JEL Classification: G11, G12, G14, C53

Keywords: Premium Predictability; Combination Forecast; Multivariate Regression Prediction

1. Introduction

Forecasting the equity premium is a topic of great importance for academics and practitioners in economics and finance. Numerous empirical studies conclude that the equity premium is predictable with information variables (also called predictive/state variables or market indicators) like the Treasury bill yield, term premium, credit (or default) premium, dividend yield or dividend-price ratio, etc.[1] The documented predictability has led to the use of information variables in numerous conditional asset pricing applications.[2] In practice, the topic is highly relevant for market timing strategists and business cycle analysts.

However, equity premium predictability remains controversial. Arguments against predictability include data mining, small sample bias, spurious regression, model instability, poor predictive ability and low economic value (see, among others, Nelson and Kim, 1993; Pesaran and Timmermann, 1995; Bossaerts and Hillion, 1999; Stambaugh, 1999; Ferson, Sarkissian, and Simin, 2003; Goyal and Welch, 2003; Paye and Timmermann, 2006; Ang and Bekaert, 2007; Timmermann, 2008; Welch and Goyal, 2008; Turner, 2015; McLean and Pontiff, 2016), but rebuttals to most of these arguments exist (see, among others, Lewellen, 2004; Marquering and Verbeek, 2004; Campbell and Thompson, 2008; Cochrane, 2008; Rapach, Strauss, and Zhou, 2010; Cenesizoglu and Timmermann, 2012; Pettenuzzo, Timmermann, and Valkanov, 2014; Maio, 2016; Li and Tsiakas, 2017; Luo and Zhang, 2016).

Most existing studies use a single information variable (or a small number of variables) and focus on in-sample (IS) tests. However, recently, an influential strand of literature considers a large number of predictive variables and out-of-sample (OS) results. Investigating more than 15

[1] The vast literature on return predictability is too voluminous to cover fully. The classic articles that first study the predictive ability of the most commonly-used information variables include Keim and Stambaugh (1986), Campbell (1987), Campbell and Shiller (1988), and Fama and French (1988, 1989).

[2] A non-exhaustive list of examples include asset pricing modelling and testing (Ferson and Harvey, 1991; etc.), performance evaluation (Ferson and Schadt, 1996; etc.) and asset allocation (Kandel and Stambaugh, 1996; etc.).

variables, Welch and Goyal (2008) emphasize the poor OS predictive power and instability of univariate models, concluding that these models would not have helped investors to profitably time the market. In contrast, using a similar set of variables but focusing on a multivariate approach based on the combination of predictions from univariate models, Rapach, Strauss, and Zhou (2010) show that it is possible to achieve significant and stable OS results with large economic values (measured by utility gains) for a mean-variance investor.

The promising results of Rapach, Strauss, and Zhou (2010) suggest that more analysis of the combination forecast approach is needed and, in general, call for a greater emphasis on multivariate approaches for equity premium predictions. This paper explores the statistical and economic performance of two of the most common multivariate forecasting approaches in predicting the equity premium: the combination forecast (CF) approach and the multivariate regression (MV) approach. The CF approach combines individual predictions of the equity premium from the information variables to obtain a forecast. The MV approach instead forecasts with the predicted values from a multivariate regression of the equity premium on a subset of information variables selected with economic or statistical criteria.[3] Our analysis is based on the study of 27 different specifications of these approaches and has two distinctive features.

First, we make a thoughtful selection of the various specifications to examine unresolved issues. For example, we look at the impact of the number of predictive variables considered in the models. The CF approach relies on diversifying the noise in individual predictions. The number of predictions should affect this diversification benefit, similar to the role of the number of assets

[3] For the CF approach, recent reviews of the literature include Timmermann (2006) and Aiolfi, Capistrán, and Timmermann (2011). Mamaysky, Spiegel, and Zhang (2007) and Detzel and Strauss (2016) are examples in finance of forecasting with model combination. For the MV approach, it is generally covered in most standard econometrics textbooks. Hocking (1976) reviews the statistical variable selection methods relevant for this paper. Elliott and Timmermann (2016) cover both approaches and many more. Other multivariate approaches have been successful in predicting the equity premium, but are beyond the scope of this paper. For example, Ludvigson and Ng (2007) and Chrétien and Kopoin (2014) use a dynamic factor analysis approach that extracts common factors from hundreds of economic and financial variables (a so-called data rich environment) to successfully predict the equity premium.

in portfolio diversification. On the other hand, adding insignificant individual predictions to the combination could be hurtful, as such addition reduces the weight placed on the best performing predictions. There is currently little evidence on the sensitivity of the CF approach for equity premium forecasting to the number of individual predictions. Since the combining weights used in the CF approach can also influence this issue, we implement five weighting schemes, including schemes that overweight the individual forecasts with the best historical performance.

For the MV approach, increasing the number of variables should increase the IS statistical performance, but may produce model instability and data overfitting, which typically lead to poor OS performance. Welch and Goyal (2008), Rapach, Strauss, and Zhou (2010) and Li and Tsiakas (2017) illustrate these issues by showing that a "kitchen sink" model that includes all their information variables obtain a very poor OS performance. Hence low-dimension MV models could perform better out of sample. We thus consider models with a low number of variables either motivated by the finance literature or selected by statistical (general-to-specific or specific-to-general) methods. However, another way to reduce the impact of these issues is to impose economically-motivated restrictions on the MV forecasts. Campbell and Thompson (2008) and Pettenuzzo, Timmermann, and Valkanov (2014) argue that imposing relevant constraints has a positive performance impact for univariate models. We provide a new investigation of the potentially stronger impact that such constraints could play on the performance of MV models.

Second, we use an empirical setup that allows a direct comparison between the approaches and provide new out-of-U.S.-sample evidence. The setup is based on common performance criteria and data, and on a comprehensive Canadian sample that alleviates data mining issues related to the almost exclusive use of U.S. data in existing studies. For both IS and OS results, we evaluate the statistical performance of the predictions with a R^2 statistic that compares the mean squared prediction errors of a predictive model and the historical mean benchmark model. We also evaluate the economic value of the predictions by finding the utility gain for a mean-variance investor from using a predictive model instead of the historical mean model. We finally examine model instability by looking at results across sub-periods. Campbell and Thompson (2008) show that using statistical and economic criteria is important for a complete assessment of predictive models, as models with weak statistical results can still provide large utility gains. Similarly, Cenesizoglu and Timmermann (2012) find that it is common for models to have negative statistical performance and yet simultaneously add economic value. Our results provide further evidence on the relation between statistical and economic performance criteria.

Our Canadian dataset includes monthly data on 36 information variables covering the period from February 1950 to June 2015. Such a long time series allows us to use the 1950-1969 period as a pre-evaluation period to select 25 relevant variables, and to keep the 1970-2015 period as an OS evaluation period. With the 25 variables, we also define seven different-size subsets of variables to form the predictive models, motivated by the finance literature or statistical selection methods. While there are plenty of return predictability studies with U.S. data, there is almost no evidence with Canadian data.[4] The uniqueness of our dataset and the lack of existing Canadian evidence ensure that our analysis provides new findings which are relatively free from the data mining issue and represent a useful robustness check on important U.S. evidence.

The main findings from our empirical results are as follows. First, we find strong evidence of IS, OS and economically-valuable predictability for most CF and MV models. In sample, all models achieve statistically significant R^2 statistics and positive economic values. Over the full

[4] For equity premium predictions with Canadian data, no article uses the CF approach and only four articles use the MV approach: Solnik (1993), Ferson and Harvey (1994), Carmichael and Samson (1996) and Korkie and Turtle (1998). However, these studies present no OS results and no economic value results. We are able to identify eight other articles with some Canadian return predictability evidence: Rapach, Wohar, and Rangvid (2005), Guo (2006b), Paye and Timmermann (2006), Deaves, Miu, and White (2008), Hjalmarsson (2010), Chrétien and Coggins (2017), Rapach, Strauss, and Zhou (2013), and Chrétien and Kopoin (2014). The first six papers examine univariate models and the last two papers study multivariate models based on U.S. information variables or the dynamic factor analysis.

evaluation period, the average annualized utility gain across models is 2.8%. Out of sample, some models lose their statistical significance, but all models still provide positive economic value, with an average of 1.9%. The best performing CF and MV models obtain a relatively comparable OS performance over time, with R^2 statistics between 1% and 2% and utility gains above 3% from 1970 to 2015. These large and positive utility gains indicate that a mean-variance investor (with a risk aversion coefficient of three) would be willing to pay considerable management fees to have access to the forecasting models relative to an historical average model.

Second, our results relying on a large number of individual forecasts confirm the findings of Rapach, Strauss, and Zhou (2010) on the forecasting ability of the CF approach. However, we document that the approach loses its OS predictive ability when it uses a small number of variables, as commonly done in the literature. Also, relative to CF models, we find that MV models obtain better IS results and comparable OS economic values. Furthermore, we show that the negative forecasting impact of the instability and overfitting problems that plague the MV approach, exemplified by the "kitchen sink" results mentioned earlier, can be largely reduced by imposing economically-motivated restrictions on the MV forecasts. In particular, we find that the biggest positive effect on statistical performance is provided by a restriction on the highest value of the expected equity premium, which can be motivated by a reasonable maximum market Sharpe ratio, as proposed by Pettenuzzo, Timmermann, and Valkanov (2014). Ultimately, in its version with constrained forecasts, the MV approach can perform as well as the CF approach.

Finally, our results document the importance of emphasizing the best individual predictors as input variables in multivariate models. Specifically, among the subsets of variables considered, we find that a subset made of nine variables with the best statistical performance in the pre-evaluation period provides the most accurate and economically valuable predictions for both approaches. Similarly, the CF weighting schemes that overweight individual forecasts with the best historical performance also obtain the best predictive results. Although we present new support for a five-variable Canadian MV model proposed by Korkie and Turtle (1998), we do not find OS evidence in favor of a MV model based on the most frequent choice of variables in U.S. studies, namely the dividend-price ratio, the relative Treasury bill yield, the term premium and the credit premium (e.g., see Ferson and Schadt, 1996; Chordia and Shivakumar, 2002; Santa-Clara and Valkanov, 2003). Hence, our findings raise questions on the robustness and value of this commonly-used model for capturing time-varying expected equity premia outside the U.S.

The rest of the paper is divided as follow. The next section describes the data and selection of information variables. Section 3 provides the methodology for the estimation and evaluation of the predictive model forecasts. Section 4 presents the empirical results. Section 5 concludes.

2. Data and Information Variable Selection

It is more convenient to describe our variables before presenting the methodology. This section introduces our Canadian equity premium variable and the set of information variables considered in the multivariate models, and present their descriptive statistics. All variables are sampled at monthly frequency and the dataset covers the period from February 1950 to June 2015.

The main evaluation period of the predictive models is from January 1970 to June 2015. We keep the first 20 years (from February 1950 to December 1969) as a pre-evaluation period that contains enough data to preselect information variables and models, and to obtain a reliable initial OS forecast. To analyse the stability of the predictive relations, we also examine two sub-periods, from 1970 to 1991 and from 1992 to 2015. We split in 1992 because it marks an important monetary policy change by the Bank of Canada. On February 26, 1991, the Bank announced that it would move to a target inflation policy, with a first target of 3% set for 1992.[5]

[5] Chrétien and Coggins (2017) find that the change in inflationary context is a source of instability for univariate predictive models of the Canadian equity premium. Welch and Goyal (2008) find a similar result in the U.S.

2.1. Equity Premium

In the predictive models, the dependent variable is always the Canadian equity premium (EQP), which is the excess return of the Canadian equity market index over the risk-free rate. The equity market returns from February 1950 to January 1956 are the total returns (including dividends) on the value-weighted equity market index from the Canadian Financial Markets Research Centre (CFMRC). From February 1956, we use the total returns on the S&P/TSX Composite Index (previously known as the TSE Composite Index) from the CFMRC or Datastream databases. We switch to the S&P/TSX Composite Index as soon as data are available as we have access to its dividend yield and price-earnings ratio. The risk-free rate is the one-month return on the three-month Government of Canada Treasury bills, taken from the CFMRC database.

Figure 1 shows the monthly EQP from 1950 to 2015. Vertical dashed lines split the series into the pre-evaluation period, first sub-period and second sub-period. We can easily locate the important market downturns since 1950, including the recession-linked corrections of August-October 1957 and April-May 1970, the oil shock of 1973-1974, the sharp decline associated with high interest rates and inflation concerns of the early 1980s, the October 1987 crash, the Russian debt default of August 1998, the burst of the tech bubble at the end of 2000 and the start of 2001 (led in Canada by the decline of Nortel Networks Inc.), the September 2001 terrorism attack and the subprime crisis of September-October 2008.

{INSERT FIGURE 1 ABOUT HERE}

2.2. Information Variables

We choose the independent variables among a set of 25 information variables that are lagged by one period compare to EQP. They are further identified by a name beginning with "Z" and can be classified into three categories. Table 1 presents an overview of the variables by giving their category, name, short description, sample start date and data sources. Appendix A provides the construction details and precise sources. Seven variables are *equity market characteristics*:

- the dividend yield (ZDY);
- the dividend-price ratio (ZDP);
- the price-earnings ratio (ZPE);
- the previous equity premium (ZEQP);
- the share volume growth (ZVOLG);
- the issuing activity (ZISSUE);
- the January dummy (ZJAN).

Twelve variables are based on *interest rates, yield spreads or exchange rates*:

- the Treasury bill yield monthly variation (ZTBILLv);
- the Treasury bill yield relative to its twelve-month moving average (ZTBILLr);
- the long-term Government bond yield (ZLTGOV);
- the long-term Government bond yield monthly variation (ZLTGOVv);
- the long-term Gov. bond yield relative to its twelve-month moving average (ZLTGOVr);
- the term premium (ZTERM);
- the credit premium (ZCREDIT);
- the short-term credit premium (ZCREDITs);
- the return-based credit premium (ZCREDITr);
- the Canada/U.S. exchange rate (ZFX);
- the Canada/U.S. exchange rate monthly variation (ZFXv);
- the Canada/US exchange rate relative to its twelve-month moving average (ZFXr).

Six variables are aggregate economic indicators:

- the inflation rate (ZINF);
- the industrial production growth (ZPRODG);
- the unemployment rate (ZUNEMP);
- the money supply growth (ZMONEYG);
- the gross domestic product growth (ZGDPG);
- the Composite leading indicator growth (ZLEAD).

{INSERT TABLE 1 ABOUT HERE}

These 25 variables were selected as follows. We start with a set of 36 potential information variables, based on variables that have already been used or could make sense in a Canadian context, or that are common in U.S. studies (see also Chrétien and Coggins (2017)). Given our focus on multivariate models, we then use the following variable selection decisions. First, we eliminate the four variables with the fewest available observations, including no observation in the pre-evaluation period.[6] Second, we exclude seven variables based on their correlations in the pre-evaluation period to alleviate multicollinearity problems. Specifically, using data from 1950 to 1969, we identify all pairs of variables that are highly correlated (i.e., have an absolute value of their correlation above 0.75) and keep only the most commonly-used variable in each pair.[7]

2.3. Descriptive Statistics

Table 2 shows the full-sample mean, standard deviation, minimum, maximum, excess kurtosis and skewness, as well as the sub-period means of the variables. The Canadian equity premium has a monthly average of 0.46% (an annualized value of 5.47%) and a standard deviation of 4.33% from February 1950 to June 2015. The annualized mean EQP is 8.32% from 1950 to 1969, 2.07% from 1970 to 1991 and 6.24 % from 1992 to 2015. The minimum is -23.53% in October 1987 and the maximum is 15.84% in January 1975 as the market recovers from the oil shock recession. The excess kurtosis of 2.67 and skewness of -0.70 are similar to the ones observed for the U.S. equity premium. Figure 2 illustrates the distribution of the Canadian equity premium by showing histograms of the series for the full period (panel A) and the sub-periods (panel B).

{INSERT TABLE 2 ABOUT HERE}

{INSERT FIGURE 2 ABOUT HERE}

The information variables also have typical descriptive statistics. The most interesting element is perhaps the various economic environments provided by the sub-periods. For example,

---

[6] The excluded variables are Canadian versions of the realized stock variance variable of Guo (2006a, 2006b) and the cross-sectional beta price of risk variable of Polk, Thompson, and Vuolteenaho (2006), and the Fisher Commodity Energy Price Index growth and the Fisher Commodity Metals and Minerals Price Index growth.

[7] The excluded variables are the forward-looking dividend yield, the earnings-price ratio, the dollar equity volume growth, the Government of Canada Treasury bill yield, the Bank of Canada prime rate, its variation and its value relative to its twelve-month moving average. Among the 25 remaining variables, the largest pre-evaluation period correlations are 0.691 (between ZTBILLr and ZLTGOVr), 0.614 (between ZLTGOVv and ZCREDITr), -0.617 (between ZFX and ZUNEMP), -0.626 (between ZTBILLr and ZTERM) and -0.635 (between ZDP and ZPE). We recognize that choosing predictors based on an initial sample implicitly makes assumptions for the rest of the sample, but this selection strategy avoids the use of data on which we rely to evaluate the predictive models.

the 1970-1991 sub-period is characterized by the oil shock of 1973-1974 and the subsequent inflationary concerns of the 1970s and 1980s. It resulted in a mean inflation rate approximately three times higher than in the other two sub-periods, as well as higher means for the money supply growth and the long-term Government bond yield. The 1992-2015 sub-period is instead characterized by a controlled inflation environment, with an inflation target at 2% since 1995.

With U.S. data, Welch and Goyal (2008) find that the OS predictive ability of many information variables deteriorates after the oil shock. Chrétien and Coggins (2017) find a similar result in Canada, but also report that variables that are significant predictors in the inflationary sub-period (in the other sub-periods) tend to lose their significance in the other sub-periods (in the inflationary sub-period). This last evidence is promising in terms of the ability of multivariate models to forecast the Canadian equity premium. As argued by Rapach, Strauss, and Zhou (2010), one potential benefit of these models is that they can adequately combine variables with complementary predictive ability in different economic contexts.

3. Econometric Methodology

We examine the monthly predictive ability of multivariate predictive models for the Canadian equity premium with models based on the MV and CF approaches. We consider IS and OS analyses to evaluate the performance of the models, with a full evaluation period that goes from 1970 to 2015. We also assess the economic value of their predictions for a mean-variance investor. We first describe the predictive models and then present the forecast evaluation criteria.

3.1. Predictive Models

This section describes the 27 models under consideration: seven models based on the MV approach and 20 models based on the CF approach. Let r_{t+1} be the equity premium at time $t + 1$ and let $z_{i,t}$ be an information variable i at time t. Let N be the total number of information variables considered (forming the set $\mathbb{N}$). As discussed in the previous section, in this paper, N = 25 and $\mathbb{N}$ = {ZDY, ZDP, ZPE, ZEQP, ZVOLG, ZISSUE, ZJAN, ZTBILLv, ZTBILLr, ZLTGOV, ZLTGOVv, ZLTGOVr, ZTERM, ZCREDIT, ZCREDITs, ZCREDITr, ZFX, ZFXv, ZFXr, ZINF, ZPRODG, ZUNEMP, ZMONEYG, ZGDPG, ZLEAD}. Each model specifies an approach j and a selected subset $\mathbb{P}_k \subseteq \mathbb{N}$ of K variables to obtain an econometric prediction $\hat{r}_{j\text{-}k,t+1}$. We next present the IS versus OS estimation strategies, the MV approach, the CF approach, the selected subsets of information variables and a summary of the predictive models under consideration.

3.1.1. In-Sample versus Out-of-Sample Predictions

We consider IS and OS versions of the predictions. Most of the predictability literature focuses on IS results. As argued by Pesaran and Timmermann (1995), Welch and Goyal (2008) and others, examining OS predictability is a useful model diagnostic and is relevant for decision makers working in real time. Let T be the number of observations available in the full evaluation period. Let S be the number of observations available in the pre-evaluation period. Let T_1 and T_2 be the numbers of observations available in the first and second sub-periods of the evaluation period. In this paper, T = 546, S = 239, T_1 = 264 and T_2 = 282. The total number of observations is thus $T + S$, with observations denoted as $t = -S, -S + 1, \ldots, -1, 0, 1, \ldots, T_1 - 1, \ldots, T - 1$.

For IS predictability, the model coefficients are estimated once in the evaluation period, thus using the data $\{r_{t+1}, z_{1,t}, \ldots, z_{N,t}\}_{t=0}^{T-1}$ in regressions to obtain the predicted values $\{\hat{r}_{j\text{-}k,t+1}\}_{t=0}^{T-1}$ for the full period. Similarly, the data $\{r_{t+1}, z_{1,t}, \ldots, z_{N,t}\}_{t=0}^{T_1-1}$ and $\{r_{t+1}, z_{1,t}, \ldots, z_{N,t}\}_{t=T_1-1}^{T-1}$ lead to the predicted values for each sub-period. For OS predictability, the coefficients are re-estimated for each prediction to account for new data available at the time of the forecast. Following Rapach, Strauss, and Zhou (2010), we focus on the recursive estimation window method. This method exploits all available data with a window size growing over time. Hence, we obtain the forecast

$\hat{r}_{j\text{-}k,t+1}$ by using the latest available variables $\{z_{1,t}, \dots, z_{N,t}\}$ and estimates from regressions with the data $\{r_{s+1}, z_{1,s}, \dots, z_{N,s}\}_{s=-S}^{t-1}$.[8] For robustness, we also use a rolling method with a fixed-size window of 240 months (or less if fewer observations are available).

3.1.2. Multivariate Regression Approach

The MV approach, *j* = MV, is a straightforward extension of the univariate approach emphasized by Welch and Goyal (2008). It uses a selected subset $\mathbb{P}_k$ of the information variables in the following multivariate regression framework:

$$r_{t+1} = \alpha_{\text{MV-}k} + \sum_{i \in \mathbb{P}_k} \beta_{i,\text{MV-}k} \times z_{i,t} + \varepsilon_{\text{MV-}k,t+1}. \tag{1}$$

Using ordinary least square (OLS) estimates of the coefficients, the regression forecast $\hat{r}_{\text{MV-}k,t+1}$ made at time *t* from model MV-*k* corresponds to the predicted value of the regression:

$$\hat{r}_{\text{MV-}k,t+1} = \hat{\alpha}_{\text{MV-}k} + \sum_{i \in \mathbb{P}_k} \hat{\beta}_{i,\text{MV-}k} \times z_{i,t}. \tag{2}$$

With an appropriate subset of variables, the MV approach can lead to a full-fledge predictive model of the equity premium that accounts for the interaction between information variables and can be easily interpreted economically. However, it can also result in data overfitting and suffer from multicollinearity problems. Hence, although it typically obtains the best IS performance, its OS performance is the subject of more debates and seems to deteriorate when too many variables are in the model. For example, Welch and Goyal (2008) and Rapach, Strauss, and Zhou (2010) show that a “kitchen sink” model that uses 15 variables simultaneously performs poorly in OS

[8] In this paper, the pre-evaluation period from February 1950 to December 1969 contains up to S = 239 observations. The estimation window size available for the initial forecast is thus consistent with McCracken (2007), who advocates a minimum size of 240 observations. Welch and Goyal (2008) also use a window size of at least 240 observations. However, some information variables in our sample do not go back as much. For example, the variable with the lowest available historical sample is ZGDPG, which starts in March 1961, leaving 108 observations to estimate the initial OS forecast for MV models using this variable.

predictions of the U.S. equity premium. Our empirical results will shed light on the effect of the number of variables on the performance of the MV approach.

3.1.3. Combination Forecast Approach

The CF approach, j = CF$_w$, uses a weighting scheme w to combine individual predictions from the information variables to obtain a new forecast. Bates and Granger (1969) is the first to point out that such combinations can outperform the potentially noisy individual forecasts themselves through a portfolio diversification-type effect. Timmermann (2006) and Aiolfi, Capistrán, and Timmermann (2011) provide relevant surveys. Rapach, Strauss, and Zhou (2010) argue that it works well in predicting the U.S. equity premium with 15 variables. However, as the noise of individual forecasts is likely to be better diversified when many forecasts are combined, it remains an open question as to whether the approach is useful with a small number of variables.

The approach begins with the following individual predictive regression for each of the K information variables in a selected subset $\mathbb{P}_k$:

$$r_{t+1} = a_i + b_i \times z_{i,t} + e_{i,t+1}. \tag{3}$$

Let $\hat{r}_{i,t+1} = \hat{a}_i + \hat{b}_i \times z_{i,t}$ be the corresponding predicted equity premium using the OLS estimates of the coefficients. The combination forecast $\hat{r}_{\mathrm{CF}_w\text{-}k,t+1}$ made at time t by model CF$_w$-k is a weighted average of the K individual predictions:

$$\hat{r}_{\mathrm{CF}_w\text{-}k,t+1} = \sum_{i \in \mathbb{P}_k} \omega_{i,\mathrm{CF}_w\text{-}k,t} \times \hat{r}_{i,t+1}. \tag{4}$$

The weight $\omega_{i,\mathrm{CF}_w\text{-}k,t}$ represents the combining weight assigned at time t by model CF$_w$-k to the individual prediction from variable i. Following Rapach, Strauss, and Zhou (2010), we consider five weighting schemes w based on different choices of weights to obtain $\hat{r}_{\mathrm{CF}_w\text{-}k,t+1}$. The first three schemes use simple methods based on the mean and median:

- w = MEAN: This scheme uses an equally-weighted average of the individual predictions, so that the weights are equal to 1 / K.
- w = MED: This scheme takes the median of the individual predictions.
- w = TRIM: This trimmed-mean scheme sets the weights to zero to the highest and lowest individual predictions, and to 1 / (K – 2) for the remaining predictions.

The next two schemes use weights that are functions of the historical performance of the individual predictions, as proposed by Stock and Watson (2004). For OS results, to obtain *ex ante* weights, we rely on the OS performance of each variable over a holdout period that starts at observation m and ends in the month before the CF prediction. Specifically, with discount factor θ, the discount mean square prediction error (DMSPE) method considers the following weights:

$$\omega_{i,\mathrm{CF}_{\mathrm{DMSPE}}\text{-}k,t} = \frac{\varphi_{i,t}^{-1}}{\sum_{i\in\mathbb{P}_k}\varphi_{i,t}^{-1}}, \text{ where } \varphi_{i,t} = \sum_{s=m}^{t-1}\theta^{t-1-s}\left(\hat{r}_{i,s+1} - r_{s+1}\right)^2. \tag{5}$$

In this paper, we use 60 months as initial holdout period (thus relying on the individual OS predictions from January 1965 to December 1969 to form the weights for the January 1970 combination forecast) and consider two different values for the discount factor:

- w = DMSPE1: This scheme takes θ = 1, so that there is no discounting.
- w = DMSPE2: This scheme takes θ = 0.9, so that greater importance is attached to the recent accuracy of the individual predictions.

For IS results, we implement a variation of the DMSPE1 scheme that accounts for the IS performance of each individual prediction over the period investigated. Specifically, for the full evaluation period, the IS DMSPE1 method considers the following fixed weights:

$$\omega_{i,\mathrm{CF}_{\mathrm{DMSPE1}}\text{-}k,t} = \frac{\varphi_{i}^{-1}}{\sum_{i\in\mathbb{P}_k}\varphi_{i}^{-1}}, \text{ where } \varphi_{i} = \sum_{t=0}^{T-1}\left(\hat{r}_{i,t+1} - r_{t+1}\right)^2. \tag{6}$$

For sub-period results, the summation in φ_i goes from 0 to $T_1 - 1$ for the first sub-period and from $T_1 - 1$ to $T - 1$ for the second sub-period.

3.1.4. Subsets of Information Variables

We consider the following seven subsets of information variables to form $\mathbb{P}_k$, using either the relevant literature or different model selection methods. The first four subsets are considered in both predictive approaches, while the last three subsets are relevant only for the MV approach.

- *k* = CAN: This set uses five variables based on the results of Korkie and Turtle (1998) and Chrétien and Coggins (2009). They find that the dividend yield, a January dummy variable, the Treasury bill yield variation, the long-term Government bond yield and the Canada/US exchange rate (i.e., $\mathbb{P}_{\text{CAN}}$ = {ZDY, ZJAN, ZTBILLv, ZLTGOV, ZFX}) are relevant information variables in a Canadian context. The detailed predictive performance of a multivariate model with these variables has never been carefully examined.
- *k* = US: This set uses four variables that have become the most common choices of researchers in U.S. applications of conditional models, namely the dividend-price ratio, the relative Treasury bill yield, the term premium (or spread) and the credit/default premium (i.e., $\mathbb{P}_{\text{US}}$ = {ZDP, ZTBILLr, ZTERM, ZCREDIT}). Ferson and Schadt (1996), Chordia and Shivakumar (2002), Santa-Clara and Valkanov (2003) are a few examples of articles using these variables. The ability of their Canadian counterparts in predicting the Canadian equity premium has not been examined in a multivariate setup.[9]
- *k* = BV: This set uses nine variables that have significant IS univariate predictive power for the Canadian equity premium in the period from 1950 to 1969. These “best variables” in our pre-evaluation period are the dividend-price ratio, the previous equity premium, a

[9] Chrétien and Coggins (2017) look at their *individual* predictive ability in a Canadian context and find that only the relative Treasury bill yield and term spread are significant IS, with the former significant OS as well.

January dummy variable, the relative Treasury bill yield, the long-term Government bond yield, the relative long-term Government bond yield, the term premium, the money supply growth and the leading indicator growth (i.e., $\mathbb{P}_{\mathrm{BV}}$ = {ZDP, ZEQP, ZJAN, ZTBILLr, ZLTGOV, ZLTGOVr, ZTERM, ZMONEYG, ZLEAD}). Table 3 presents the results used to identify these nine variables by showing the adjusted R^2 statistic and F-statistic (with its significance) of the 25 IS univariate predictive regressions.[10]

{INSERT TABLE 3 ABOUT HERE}

- k = ALL: This set uses all information variables in our sample, so that $\mathbb{P}_{\mathrm{ALL}} = \mathbb{N}$. Welch and Goyal (2008) call a MV model with such a choice a "kitchen sink" model.
- k = FOR: This set includes variables identified by a forward selection technique. This specific-to-general technique begins with no variable in the set used for a MV model. It then calculates F-statistics for all variables, one by one, that reflect each variable's contribution to the model if it is included and adds the variable that has the largest statistic to the model. Next it calculates the statistics again for the variables still remaining outside the model to identify a second variable for inclusion. This selection process is repeated until all remaining variables obtain F-statistics with p-values greater than 10%. Thus, variables are added one by one to the model until no remaining variable produces a significant statistic. Once a variable is in the model, it stays.
- k = STEP: This set includes variables identified by a stepwise selection technique, a modification of the forward selection technique in which variables already in the model do not necessarily stay there. As in the forward selection method, variables are added one by one to the model, and the F-statistic for a variable to be added must be significant at

[10] This selection strategy is admittedly not an optimal way to choose predictors to maximize multivariate predictive performance. But it still provides a relevant set of variables and avoids the look-ahead bias that would exist if we use data from the evaluation period to select predictors.

the 10% level. After a variable is added, however, the stepwise technique looks at all the variables already included in the model and deletes any variable that does not produce a *F*-statistic significant at the 10% level. Only after this check is made and the necessary deletions are accomplished can another variable be added to the model. The process ends when no variable outside the model has a significant *F*-statistic and every variable in the model is significant, or when the variable to be added to the model is the one just deleted.

- k = BACK: This set includes variables identified by a backward elimination technique. This general-to-specific technique begins by calculating the *F*-statistic for a model with all information variables. Then variables are deleted one by one until all variables remaining in the model produce *F*-statistics significant at the 10% level. At each step, the variable showing the smallest contribution to the model is deleted.

The subsets of variables of the MV-FOR, MV-STEP and MV-BACK models can change through time, as we re-implement the model selection techniques in each sub-period for IS results and when new data become available for OS results. They are fixed for the other models.

3.1.5. Summary of the Predictive Models

Given our two approaches, j = {MV, CF_w}, and seven subsets of information variables, k = {CAN, US, BV, ALL, FOR, STEP, BACK}, we consider a total of 27 predictive models. Seven models are based on the MV approach: MV-CAN, MV-US, MV-BV, MV-ALL, MV-FOR, MV-STEP and MV-BACK. Twenty models are based on the CF approach: CF_{MEAN}-CAN, CF_{MED}-CAN, CF_{TRIM}-CAN, CF_{DMSPE1}-CAN, CF_{DMSPE2}-CAN, CF_{MEAN}-US, CF_{MED}-US, CF_{TRIM}-US, CF_{DMSPE1}-US, CF_{DMSPE2}-US, CF_{MEAN}-BV, CF_{MED}-BV, CF_{TRIM}-BV, CF_{DMSPE1}-BV, CF_{DMSPE2}-BV, CF_{MEAN}-ALL, CF_{MED}-ALL, CF_{TRIM}-ALL, CF_{DMSPE1}-ALL, and CF_{DMSPE2}-ALL.

These models represent a good mix of the large number of potential multivariate models that can be formed from a set of information variables. They consider two of the most prominent forecasting approaches. The MV approach takes into account the interactions between information variables in estimating the regression coefficients. It thus captures the information in one variable while controlling for the others. But it can suffer from data overfitting, instability and multicollinearity problems, which are more common when many variables are included. By comparison, the CF approach relies on simple univariate regression coefficients and attempts to "diversify" the noise in individual forecasts by forming a weighted-average prediction. It does not however clearly specify the optimal weights for such an average and the number of variables needed for sufficient noise diversification. Our empirical results will examine these issues.

The subsets of variables are motivated from the finance literature (for CAN and US), are statistically selected (for BV, FOR, STEP and BACK), or avoid the model selection problem altogether (for ALL). The five weighting schemes for the CF approach allows an examination of its sensitivity to the choice of combining weights. They offer a representative choice of the large number of possible weighting schemes. The MEAN, MED and TRIM schemes are simple averaging methods. The DMSPE1 and DMSPE2 schemes account for the historical performance of the individual forecasts. While more elaborate schemes exist, Timmermann (2006) concludes that simple combining methods typically outperform more complicated methods.

3.2. Forecast Evaluation

This section describes the main criteria to evaluate and compare the predictive performance of the models in sample, out of sample and in economic value. Intuitively, these criteria examine if the forecast from a predictive model, $\hat{r}_{j\text{-}k,t+1}$, lead to smaller squared errors and better expected utility than a benchmark forecast based on the historical mean, $\bar{r}_{t+1}$, where the forecasts are either

IS or OS predictions.[11] The criteria follow those in Welch and Goyal (2008), Campbell and Thompson (2008) and Rapach, Strauss, and Zhou (2010) to make our analysis easily comparable.

3.2.1. Statistical Performance of the Predictions

To evaluate the statistical performance of the predictions, we use the following R^2 statistic:

$$R^2 = 1 - \frac{\sum_{t=0}^{T-1}(r_{t+1} - \hat{r}_{j\text{-}k,t+1})^2}{\sum_{t=0}^{T-1}(r_{t+1} - \bar{r}_{t+1})^2}. \tag{7}$$

This statistic is comparable to the familiar R^2 statistic from an OLS regression. In fact, for IS results of the MV approach, it is precisely equal to the R^2 statistic of the regression. When $R^2 > 0$, the predictive model forecast outperforms the historical average forecast as its mean squared prediction error (MSPE) is smaller. Although always non negative for IS results, the R^2 statistic can be negative for OS results.

To test for the null hypothesis that $R^2 \leq 0$ versus the alternative hypothesis that $R^2 > 0$, we follow Rapach, Strauss, and Zhou (2010) by using the MSPE-adjusted statistic of Clark and West (2007), which generates asymptotically valid inferences when comparing forecasts from nested linear models.[12] The MSPE-adjusted statistic can be found by first calculating

$$f_{t+1} = (r_{t+1} - \bar{r}_{t+1})^2 - \left[(r_{t+1} - \hat{r}_{j\text{-}k,t+1})^2 - (\bar{r}_{t+1} - \hat{r}_{j\text{-}k,t+1})^2\right] \tag{8}$$

and then regressing f_{t+1} on a constant. The significance of the statistic is the *p*-value from a standard normal distribution of a one-sided (upper-tail) test on the *t*-statistic of the constant.

[11] The historical mean forecast is the full-sample historical mean for IS performance and the historical mean in the estimation window for OS performance.

[12] As discussed by Rapach, Strauss, and Zhou (2010), the MSPE-adjusted statistic is an adjusted version of the Diebold and Mariano (1995) and West (1996) statistic. This last statistic is well behaved for nonnested models, but Clark and McCracken (2001) and West (1996) find that it has a nonstandard distribution and can be severely undersized when comparing nested linear models, as is the case in this paper. Using simulations with a variety of sample sizes, Clark and West (2007) find that the MSPE-adjusted statistic performs relatively well in terms of size and power when comparing forecasts from nested linear models.

3.2.2. Economic Value of the Predictions

As emphasized by Campbell and Thompson (2008), predictions with low explanatory power can still yield economically meaningful results for investors. Hence, we provide an assessment of the utility value of the forecasts for a mean-variance investor by following Marquering and Verbeek (2004), Campbell and Thompson (2008) and Rapach, Strauss, and Zhou (2010).

Specifically, to evaluate the economic value of the predictions, we first compute the optimal equity allocations at time t, ω_t, for a mean-variance investor with risk aversion parameter γ, based on a rolling-window estimate of the variance of stock returns, denoted as s_{t+1}^2, and either the predictive model forecast $\hat{r}_{j\text{-}k,t+1}$ or the historical mean forecast $\bar{r}_{t+1}$:

$$\widehat{\omega}_t = \frac{1}{\gamma}\left(\frac{\hat{r}_{j\text{-}k,t+1}}{s_{t+1}^2}\right), \qquad \overline{\omega}_t = \frac{1}{\gamma}\left(\frac{\bar{r}_{t+1}}{s_{t+1}^2}\right). \tag{9}$$

We obtain our results by using a ten-year rolling window for s_{t+1}^2, by restricting allocations to be between 0% and 150% to prevent extreme investments and rule out negative predictions, and by selecting γ = 3 (although other reasonable values lead to qualitatively similar results).

Let $\hat{\mu}$ and $\hat{\sigma}$ be the sample mean and standard deviation, respectively, of the returns on the portfolio with equity allocation $\widehat{\omega}_t$ and risk-free allocation (1 – $\widehat{\omega}_t$). Let $\bar{\mu}$ and $\bar{\sigma}$ be the corresponding statistics when using equity allocation $\overline{\omega}_t$. Then the realized average utility levels for a mean-variance investor in these strategies are given by:

$$\widehat{U} = \hat{\mu} - \frac{1}{2}\gamma\hat{\sigma}^2, \qquad \overline{U} = \bar{\mu} - \frac{1}{2}\gamma\bar{\sigma}^2. \tag{10}$$

The economic value of the predictions is equal to the utility gain from investing based on the predictive model, multiplied by 1200 to express it in average annualized percentage return:

$$Gain = \left(\widehat{U} - \overline{U}\right) \times 1200. \tag{11}$$

An interpretation of this utility gain (or certain equivalent return) is that it represents the annual fee that a mean-variance investor would be willing to pay for access to the additional information in the predictive model forecasts relative to information in the historical mean forecasts alone.

4. Empirical Results

This section presents the IS and OS empirical results. Then, it examines OS results when economically-motivated restrictions are imposed on the forecasts.

4.1. In-Sample Predictability Results

Table 4 shows the results for the R^2 statistic, its significance using the MSPE-adjusted statistic of Clark and West (2007) and the average annualized utility gain for each predictive model in the full evaluation period (1970-2015) and in the two sub-periods (1970-1991 and 1992-2015).[13]

{INSERT TABLE 4 ABOUT HERE}

The table reveals that all models are able to significantly predict the Canadian equity premium and generate economic value for a mean-variance investor. In the full period, the R^2 statistics are between 0.28% (see the CF_{MED}-CAN model) and 1.40% (see the CF_{MEAN}-BV model) for the CF models, and between 1.17% (see the MV-US model) and 8.42% (see the MV-ALL model) for the MV models, with all values statistically different from zero. The utility gains vary from 0.78% (see the CF_{MED}-CAN model) to 3.08% (see the CF_{MEAN}-BV model) for the CF models, and from 2.41% (see the MV-US model) to 8.32% (see the MV-ALL model) for the MV models. The information in 20 out of 23 models would be worth a fee of at least 1% per year for a mean-variance investor. Although the results vary across models, based on the IS results, the MV models outperform the CF models. These general findings are similar in the sub-periods, although the predictions are better in the 1970-1991 sub-period than in the 1992-2015 sub-period.

[13] Recall that the CF_{DMSPE2} models are not implemented for IS predictability.

For most CF models, the R^2 statistics are below 1%, suggesting that their explanatory power is not high. In comparison, Chrétien and Coggins (2017) find that 12 of the 25 univariate predictive models based on the variables individually have full sample IS R^2 statistics below the smallest value (0.28%) found for the CF models, and only 5 models have R^2 statistics above 1%. Comparing the CF models across weighting schemes, the MEAN and DMSPE1 schemes consistently obtain superior performance than the MED and TRIM schemes, except in sub-period results for the CF-US models.

It is expected that IS results for MV models improve with the number of variables included. The findings for CF models do not indicate such a clear relation, as the R^2 statistics and utility gains are larger for the CF-BV models (9 variables) than for the CF-ALL models (25 variables). Nevertheless, the CF-CAN and CF-US models (5 and 4 variables) generally perform worst. Such finding suggests that combining a low number of forecasts leads to less efficient diversification of the noise in individual forecasts. For MV models, it is useful to compare models with variables motivated by the finance literature (i.e., the MV-CAN and MV-US models) to models with variables selected through a statistical procedure (i.e., the MV-FOR, MV-STEP and MV-BACK models). Although the numbers of variables included are comparable, table 4 shows superior performance for the statistically motivated models, suggesting that the selection procedures are able to identify significant variables, at least in sample.

Overall, the IS results document that, based on the R^2 statistic and utility gain criteria, the MV models outperform the CF models. However, a high R^2 statistic could be the result of data overfitting or multicollinearity problems. Furthermore, one has to be careful about comparing IS R^2 statistics across models with a different number of variables. Finally, similar performance results across sub-periods do not necessarily indicate that the MV models are stable through time.

To further investigate the MV models and provide information on the most useful predictors in a multivariate setting, table 5 shows the value and significance of the coefficient estimates associated with the constant and each variable in the models, as well as the R^2 statistic, the adjusted R^2 statistic, denoted Adj R^2, the F-statistic (to test the hypothesis that all coefficients are equal to zero) and its significance. Heteroskedasticity and autocorrelation consistent Newey and West (1987) standard errors are used to establish the significance of the estimates.[14]

{INSERT TABLE 5 ABOUT HERE}

Table 5 shows that, according to the F-statistic, the models with variables selected through a statistical procedure (i.e., the MV-FOR, MV-STEP and MV-BACK models) are the only significant models for the full period and across sub-periods. Their adjusted R^2 statistics also suggest that they have the highest explanatory power. However, the variables selected by their procedure vary greatly across periods. For example, in the full period, the results for the MV-FOR model indicate that the equity premium is negatively related to the lagged variation in the long-term Government bond yield (ZLTGOVv) and the lagged relative long-term Government bond yield (ZLTGOVr), and positively related to the lagged return-based credit premium (ZCREDITr) and the lagged GDP growth (ZGDPG), with all estimates being significant. But the variables selected include only ZLTGOVr (significant) in the first sub-period, and the previous equity premium (ZEQP, significant) and ZGDPG (insignificant) in the second sub-period. This instability could be problematic for the OS performance of the models. Also, while the selection procedures behind the MV-FOR and MV-STEP models lead to almost identical subsets of variables, the general-to-specific technique of the MV-BACK model retains a much larger number of variables, raising the potential issues of multicollinearity and data overfitting.

[14] The number of lags is set according to the formula Int{$4(T/100)^{1/4}$} following Granger, Hyung and Jeon (2001).

These issues as well as model instability can also be raised for the MV-BV and MV-ALL models, which are characterized by a large number of variables. Their explanatory power is greatly reduced when looking at the adjusted R^2 statistic instead of the R^2 statistic, and their F-statistics show significance at the 5% level only for the full period. For the MV-CAN model, which was developed by Korkie and Turtle (1998) with data up to 1993, the significance of its coefficients and F-statistics disappears in the second sub-period. Finally, for the MV-US model, which is based on commonly used variables in U.S. studies, the results show no significant coefficient or F-statistic, casting doubts on the usefulness of the model outside the U.S.

4.2. Out-of-Sample Predictability Results

The IS results find that MV models outperform CF models, but there is evidence of model instability and data overfitting. Rapach, Strauss, and Zhou (2010) argue that such instability lead a “kitchen-sink” model like the MV-ALL model to fare worse than a more stable CF model based on many variables, like the CF-ALL models, in real time. This section investigates this argument further by examining OS results. Similar in format to table 4, table 6 shows the OS results for the R^2 statistic, its significance and the average annualized utility gain for each predictive model in the full evaluation period (1970-2015) and in the two sub-periods (1970-1991 and 1992-2015). The results use the recursive method for specifying the regression estimation window.

{INSERT TABLE 6 ABOUT HERE}

Table 6 shows that the CF models now outperform the MV models in the OS analysis. In the full period, the CF models obtain R^2 statistics between 0.11% (see the CF_{MED}-CAN model) and 1.86% (see the CF_{DMSPE2}-BV model), with 18 out of 20 values being statistically significant. Their utility gains vary between 0.21% (see the CF_{MED}-CAN model) and 3.78% (see the CF_{DMSPE2}-BV model), with economic values above 1% for all but two models. In comparison, Chrétien and Coggins (2017) find that only six of the 25 univariate predictive models based on the variables individually have OS R^2 statistics above 0.11%.

The CF models based on the highest number of variables (i.e., the CF-BV and CF-ALL models) are able to significantly predict the Canadian equity premium in the full period and across both sub-periods, but the other CF models are not significant in the second sub-period. Similar to the IS results, the R^2 statistics and utility gains for the CF-BV models are larger than those for the CF-ALL models, indicating that a pre-selection of variables, instead of an all-inclusive selection, could be beneficial for the CF approach. However, the results for the CF-CAN and CF-US models still indicate that using a low number of variables decreases the performance of the CF approach. Comparing the CF models across weighting schemes, the best-to-worst schemes for OS forecasting are DMSPE2, DMSPE1, MEAN, TRIM and MED for all subsets of variables, except for the US subset. Hence, a scheme that accounts for the historical performance of the individual predictions, with greater importance given to their recent accuracy, is the best choice in predicting the monthly Canadian equity premium with the CF approach.

To allow a diagnosis of the performance of the predictions through time, figure 3 provides a graphical analysis of the CF results by showing the evolution of the cumulative squared forecast errors differences, $\sum_{t=0}^{T-1}(r_{t+1}-\bar{r}_{t+1})^2-(r_{t+1}-\hat{r}_{t+1})^2$. When the slope of the line is positive (negative), the predictive model performs better (worse) than the historical mean model, as its squared errors are smaller (larger). The figure further allows seeing the performance of the predictive models for any desired period, by comparing the values of the line at the start and end of the period. For example, if the value of the line is lower (higher) in 2000 than in 2015, then the OS R^2 statistic for that period is positive (negative). Welch and Goyal (2008) and Rapach, Strauss, and Zhou (2010) also emphasize such graphical analysis for their results.

{INSERT FIGURE 3 ABOUT HERE}

Figure 3 shows that most lines are generally increasing though time, indicating that the predictive models are relatively constant in beating the historical mean model. This pattern is particularly strong for models that combine a large number of individual predictions (i.e., the CF-BV and CF-ALL models) and for models that use the DMSPE2 weighting scheme. The only apparent period of difficulty for the predictive models is from the early 1970s to the early 1980s, a period characterized by the oil shock, inflation concerns and increasing interest rates. The CF-BV and CF-ALL models show almost constant outperformance since the mid-1980s.

In contrast to the results for CF models, table 6 shows that the MV models have negative OS R^2 statistics, indicating that they cannot beat the historical mean model. For the full period, the values go from -14.77% (see the MV-ALL model) to -0.77% (see the MV-CAN model). Similar to Rapach, Strauss, and Zhou (2010), we find that the "kitchen-sink" model performs very poorly in real time. As discussed previously, this model presents signs of data overfitting and instability. Models with variables selected through a statistical procedure (i.e., the MV-FOR, MV-STEP and MV-BACK models) are other models with a similar diagnostic. Their results in table 6 show that their OS performance is much worse than their IS performance. Hence, while statistically-motivated MV models have superior IS results, economically-motivated MV models with a small number of variables fare better in real time. All MV models nevertheless provide positive utility gains that are comparable to those of CF models, suggesting that their predictions add value in the allocation decisions of a mean-variance investor. We will examine this conflicting result in the next subsection. Panel A of figure 4 confirms graphically the negative performance of the MV models, which is particularly negative from the early 1970s to the early 1980s.

{INSERT FIGURE 4 ABOUT HERE}

Table 7 is similar to table 6, but provides the OS results using the rolling method with a fixed-size estimation window of 240 months. As in Welch and Goyal (2008) and Chrétien and Coggins (2017), the findings are qualitatively the same for the recursive and rolling methods. However, the performance in terms of R^2 statistic and utility gain is generally lower with the rolling method than with the recursive method.

{INSERT TABLE 7 ABOUT HERE}

4.3. Predictability Results under Economically-Motivated Forecast Restrictions

One puzzling aspect of the results of tables 6 and 7 is that MV models provide positive utility gains although they have negative R^2 statistics. Using a large set of U.S. stock return prediction models with time-varying mean and volatility, Cenesizoglu and Timmermann (2012) also find that it is common for models to have negative statistical performance and yet simultaneously add economic value. Specifically, they document that, although both performance criteria are correlated, disagreement in sign between a model's statistical and economic performance arises in more than half of their 60-model comparisons.

The main reason for this apparent discrepancy is that the utility gain calculation restricts the equity allocations to values between 0% and 150%. When no restriction is imposed on the equity allocations, as shown by Campbell and Thompson (2008), the difference in mean returns $\hat{\mu} - \bar{\mu}$, and hence the related utility gain, should have the same sign as the R^2 statistic. In results not tabulated, we effectively find that OS utility gains become negative for MV models when the restriction is not imposed.[15] In contrast, for the OS results of CF models and for the IS results of all models, the utility gains remain similar and are always the same sign as the R^2 statistics, whether we impose the restriction or not. Hence, imposing restrictions on forecasts from MV

[15] Specifically, with unconstrained allocations, we find gains of -5.1%, -17.8%, -12.0%, -43.7%, -14.7%, -13.9% and -19.3% for the MV-CAN, MV-US, MV-BV, MV-ALL, MV-FOR, MV-STEP and MV-BACK models, respectively.

models is a promising avenue to improve their OS statistical performance. Such avenue is explored by Campbell and Thompson (2008), Pettenuzzo, Timmermann, and Valkanov (2014) and Li and Tsiakas (2017), who advocate imposing economically-motivated restrictions on equity premium forecasts and show that predictions constrained to be nonnegative have better explanatory power.[16]

Given the standard deviation of 4.6% for the equity premium over the evaluation period and the assumption that γ = 3, equation (9) implies that the forecasts leading to equity allocations between 0% and 150% are assumed to be at least 0% and at most (approximately) 1%. Economically, these boundaries are sensible. The motivation for the 0% boundary is that the expected equity premium should be positive for risk-averse investors. The 1% boundary implies an expected monthly equity market Sharpe ratio of at most 0.22. This value is similar to the maximum market Sharpe ratio advocated by MacKinlay (1995), who proposes a squared annual market Sharpe ratio of at most 0.6 (equivalent to a monthly Sharpe ratio of 0.226).[17]

Table 8 shows the OS R^2 statistic results (using the recursive method) when the forecasts are restricted to be: 1- nonnegative (denoted $\hat{r} \geq 0\%$); 2- less than or equal to 1% (denoted $\hat{r} \leq 1\%$); 3- nonnegative and less than or equal to 1% (denoted $0\% \leq \hat{r} \leq 1\%$).[18] Table 8 finds that economically-motivated forecast restrictions have different impact for MV models versus CF models. For CF models, the restrictions do not lead to material changes in their R^2 statistics, a

[16] In addition to a nonnegative restriction, Campbell and Thompson (2008) and Li and Tsiakas (2017) consider economic constraints on the coefficients, Pettenuzzo, Timmermann, and Valkanov (2014) impose a maximum value on the market Sharpe ratio and Li and Tsiakas (2017) implement statistical constraints via shrinkage estimation.

[17] In different contexts, Ross (1976), Cochrane and Saá-Requejo (2000), Pettenuzzo, Timmermann, and Valkanov (2014) and Chrétien and Kammoun (2017) also use a maximum Sharpe ratio restriction that implies monthly values between 0.24 and 0.29.

[18] Specifically, for the nonnegative restriction, the forecasts are set to 0 when the model predictions are negative. For the less than or equal to 1% restriction, the forecasts are set to 1% when the model predictions are greater than 1%. This truncation approach is similar to the one proposed by Campbell and Thompson (2008) and Li and Tsiakas (2017). Pettenuzzo, Timmermann, and Valkanov (2014) instead impose restrictions in the estimation to obtain constrained OLS estimates, and then use predicted values from the constrained OLS regression as predictions.

conclusion also reached by Rapach, Strauss, and Zhou (2010). By “diversifying away” the noise in individual predictions, the CF approach produces forecasts that rarely take extreme values.

{INSERT TABLE 8 ABOUT HERE}

Instead, for MV models, economic restrictions greatly improve their statistical performance, suggesting that their instability and overfitting issues can be partly alleviated with relevant forecast constraints. Table 8 indicates that there is a role in this improvement for both the non-negativity constraint and the less than or equal to 1% constraint. However, the former restriction leads only to a small performance increase not sufficient to render the R^2 statistics positive. Cenesizoglu and Timmermann (2012) also find that a non-negativity constraint has little effect on the disagreement between statistical and economic performance criteria. But their analysis neglects to examine the impact of restricting predictions that are too high. Table 8 shows that such a constraint alone greatly improves the R^2 statistic of MV models.

When both economic restrictions are considered jointly, the full evaluation period results in table 8 show that MV models obtain statistically significant OS R^2 statistics, except for the MV-US model. Furthermore, the R^2 statistics position the MV models only behind the CF-BV models for the best OS statistical performance.[19] The results for MV models remain similarly strong and significant in the second sub-period, but only the MV-CAN, MV-BV and MV-BACK models have significant R^2 statistics in the first sub-period. Panel B of figure 4 confirms graphically the largely improved performance of MV models when imposing both restrictions, compared to their unconstrained versions in panel A. However, it also shows that their performance is more volatile than that of CF models. Furthermore, a strong performance after 2003 is mainly responsible for the positive R^2 statistics of the MV-ALL, MV-FOR, MV-STEP and MV-BACK models. It is

[19] The utility gains reported in table 6 also position the MV models as second behind the CF-BV models for OS economic performance.

clear from the figure that the MV-CAN and MV-BV models with economically-motivated forecast restrictions imposed are the most reliable predictive models for the MV approach.

5. Conclusion

Using the CF approach to predict the equity premium, Rapach, Strauss, and Zhou (2010) show that it is possible to achieve significant and stable OS results with large economic values (measured by utility gains) for a mean-variance investor. We extend their results by exploring the statistical and economic performance of their approach in comparison with the MV approach. Our analysis is based on the study of 27 strategically-selected specifications of these approaches, as well as an empirical setup that allows their direct comparison and exploits a unique Canadian database to provide new out-of-U.S.-sample evidence.

Empirically, we find significant evidence of IS, OS and economically-valuable predictability for most models. The best performing CF and MV models obtain a relatively stable performance over time, with OS R^2 statistics between 1% and 2% and annualized utility gains above 3% over the period from 1970 to 2015. Hence a mean-variance investor (with a risk aversion coefficient of three) would be willing to a pay a considerable management fee to have access to these forecasting models relative to an historical average benchmark model. The OS results for these best models highlight the importance of choosing a reasonably large and relevant subset of information variables, instead of using all available variables or just a small number of them, and the significance of imposing economically-motivated restrictions on the MV forecasts.

Specifically, among the different subsets of variables considered, we find that a subset made of nine variables that are best performers in the pre-evaluation period, termed the BV subset, provides the most accurate and economically valuable predictions for the CF and MV approaches. CF models that use all 25 available variables also obtain statistically and economically significant results. But their relatively lower performance indicates that, against a diversification argument, including many individual predictions irrespective of their forecasting ability, instead of pre-selecting significant ones, is hurtful. The superior results for combining weight schemes that overweight individual forecasts with the best historical performance also confirm the importance of emphasizing the best individual predictors. When only four or five variables are considered, the OS results of the CF models greatly deteriorate, suggesting that noise in individual predictions is not diversified adequately in low-dimension models.

In contrast, MV models with a low number of variables obtain a better OS performance than the "kitchen sink" MV model using all information variables. The "kitchen sink" model and the MV models selected through a forward, stepwise or backward procedure display strong evidence of overfitting and instability, which contribute to their poor OS statistical performance. However, imposing relevant economic restrictions on MV predictions reduces considerably the impact of these issues and improves dramatically the performance of all MV models. In particular, a restriction on the highest value of the expected equity premium, which can be motivated by a reasonable maximum market Sharpe ratio, provides the biggest effect on performance.

Although the best performing MV model uses the BV subset, we also find supporting evidence for a Canadian MV model proposed by Korkie and Turtle (1998), based on the dividend yield, a January dummy, the Treasury bill yield variation, the long-term Government bond yield and the Canada/U.S. exchange rate. Specifically, the restricted forecasts from this model obtain statistically significant OS performance, although the results are weaker in the last sub-period. We do not find evidence in favor of a MV model based on the most common choice of variables in U.S. studies, namely the dividend-price ratio, the relative Treasury bill yield, the term premium and the credit premium. The insignificant OS results of this model raise questions on its robustness and value for capturing the conditional equity premium outside the U.S.

Overall, our findings of significant equity premium predictability outline numerous possible ways of using multivariate models for improved forecasting with information variables. The basic results of Rapach, Strauss, and Zhou (2010) as well as our extended results call for a greater emphasis on multivariate approaches (either with combination forecast or multivariate regression) for equity premium predictions, as well as the development of conditional asset pricing applications that better exploit the ability of multivariate models to capture time-varying economic conditions.

Appendix A: Construction Details and Sources of the Information Variables

This appendix gives a detailed description of the construction and sources of the information variables, regrouped by their category.

A.1. Equity Market Characteristic Variables

Seven information variables are related to equity valuation ratios and market-related variables (ZDY, ZDP, ZPE, ZEQP, ZVOLG, ZISSUE and ZJAN):

Dividend Yield (ZDY): The realized dividend yield (ZDY) for the S&P/TSX Composite Index is computed from the difference between the one-year total return of the index and its one-year price return. The data come from the CFMRC database and start in February 1957.

Dividend-Price Ratio (ZDP): We calculate the dividend-price ratio as the realized dividend yield multiplied by the value of the index one year prior and divided by the current value of the index. The data go back to February 1957.

Price-Earnings Ratio (ZPE): The price-earnings ratio of the S&P/TSX Composite Index is obtained from the CANSIM database as series V122629. It is available starting in February 1956. It corresponds to the current market price divided by the earnings in the latest fiscal year.[20]

Previous Equity Premium (ZEQP): ZEQP is the lagged EQP and captures the predictive information in the equity premium of the previous month.

Volume Growth (ZVOLG): The volume of shares growth variable (ZVOLG) is the growth in the monthly number of shares transacted on the Toronto Stock Exchange. The data come from the CANSIM database as series V37413 and start in March 1953.

Issuing Activity (ZISSUE): To compute the corporate issuing activity, we first compute the dollar amount of net equity issuing activity, following Welch and Goyal (2008) as: Net Issue at

[20] From August 2001 to July 2002, the ratio was not listed due to negative earnings. We replace the missing value with the maximum value of the price-earnings ratio prior to August 2001.

month t = Market Capitalisation at month t – Market Capitalisation at month t-1 × (1 + Market Capital Gain Return at month t). The issuing activity variable (ZISSUE) is the net equity expansion defined as the ratio of the twelve-month moving sums of net issues divided by the current market capitalisation. From February 1951 to September 2000, the price and number of shares for each stock in the CFMRC database is used to obtain the total market capitalisation. From October 2000, we take the market capitalisation of the S&P/TSX Composite Index from Datastream. ZISSUE is related to a variable proposed by Boudoukh, Michely, Richardson, and Roberts (2007) in the U.S. They show that since the adoption of SEC rule 10b-18 in 1982, there has been an explosion of share repurchase transactions. They argue that this distribution channel has caused the dividend yield to lose its predictive power, but a payout yield that includes share repurchases still provides significant predictions. Kooli and L'Her (2010) similarly find a decline in dividend paying firms and a significant increase in share repurchase programs in Canada.

January Dummy (ZJAN): This dummy variable (not lagged) is set to 1 in January and 0 for the other months and captures the so-called January effect predicting higher returns in January.

A.2. Interest Rate, Yield Spread or Exchange Rate Variables

Twelve information variables are related to interest rates (ZTBILLv, ZTBILLr, ZLTGOV, ZLTGOVv, ZLTGOVr, ZTERM, ZCREDIT, ZCREDITs, ZCREDITr, ZFX, ZFXv and ZFXr).

Treasury Bill Yields (ZTBILLv, ZTBILLr): The Treasury bill yield is the annualized yield-to-maturity of the three-month Government of Canada Treasury bill. It is taken from the CFMRC database and is also available as series V122541 in the CANSIM database. As the Treasury bill yield is highly persistent, we consider its monthly variation (ZTBILLv) and its value relative to its twelve-month moving average (ZTBILLr).

Long-Term Government Bond Yields (ZLTGOV, ZLTGOVv, ZLTGOVr): The long-term Government yield is the average yield-to-maturity of the Government of Canada Treasury bonds with a maturity of ten years or more. It is available in the CFMRC database and in the CANSIM database as series V122487. We use its value (ZLTGOV), its monthly variation (ZLTGOVv) and its value relative to its twelve-month moving average (ZLTGOVr).

Term Premium (ZTERM): ZTERM is ZLTGOV minus the lagged Treasury bill yield.

Credit Premium (ZCREDIT, ZCREDITs, ZCREDITr): We consider three credit premium (or default spread) variables. The first credit premium is the difference between the lagged yield on long-term corporate bonds and ZLTGOV. To construct a long history of the corporate yields, we combine three series. From February 1950 to October 1977, we use the series V35752 from the CANSIM database, the Scotia-McLeod Canada Long-Term All-Corporate Yield Index. From November 1977 to June 2007, we take the Scotia Capital Canada All-Corporations Long-Term bond yield series from CFMRC, also available as series V122518 in the CANSIM database. From July 2007, we take the yield from the Bank of America Merrill Lynch Canada Corporate Bond 10Y+ Index (series MLCCTPL) from Datastream. In an effort to avoid mixing three series, a second yield spread variable is computed as the difference between the yield on the three-month prime corporate paper (series V122491) and the Treasury bill yield. This short-term credit premium variable (ZCREDITs) goes back to February 1956. Finally, we form a return-based credit premium variable (ZCREDITr) as the difference between long-term corporate bond and long-term Government bond returns. For the corporate bond returns, we use series V35754 (the Scotia-McLeod Canada Long-Term All-Corporate Total Return Index) from December 1950 to October 2002. From November 2002, we take the total returns from the Bank of America Merrill Lynch Canada Corporate Bond 10Y+ Index (series MLCCTPL). We obtain the Government bond returns from the CFMRC database.

CAD/USD Exchange Rates (ZFX, ZFXv, ZFXr): Canada’s largest trading partner is by far the U.S. The spot exchange rate in Canadian dollars per U.S. dollar is collected from the CFMRC database and is also available as series V37426 from the CANSIM database. As it is highly persistent, we use its lagged value (ZFX), its lagged variation (ZFXv) and its lagged value relative to its twelve-month moving average (ZFXr).[21] ZFX, ZFXv and ZFXr start in November 1950, December 1950 and October 1951, respectively.

A.3. Macroeconomic Variables

Six information variables are based on macroeconomic indicators (ZINF, ZPRODG, ZUNEMP, ZMONEY, ZGDPG and ZLEAD).

Inflation Rate (ZINF): The inflation rate is the monthly growth in the Consumer Price Index (CPI) obtained from the CANSIM database as series V41690973.

Industrial Production Growth (ZPRODG): The monthly industrial production growth is the monthly growth in the Industrial Production Index (IPI) extracted from the CANSIM database as series V53384745 or Datastream. It is available from March 1956.

Unemployment Rate (ZUNEMP): The unemployment rate data are collected from Datastream (code CNOUN014R) and go back to February 1960. It is also available as series V2064894 from the CANSIM database from January 1975.

Money Supply Growth (ZMONEYG): The money supply growth is the monthly growth of the Money Supply Index obtained as series V37173 from the CANSIM database. The money supply variable represents the unadjusted currency outside banks.

Gross Domestic Product Growth (ZGDPG): The gross domestic product (GDP) growth is the monthly growth in the seasonally adjusted GDP for all industries. We construct the ZGDPG variable with two series (V329529 and V65201483) from the CANSIM database. The first one,

[21] We also considered the three-month forward CAD/USD exchange rate (series V37437) used by Korkie and Turtle (1998). However, given its correlation of 0.9997 with the spot exchange rate, we discard it from further analysis.

the GDP at factor cost in 1992 constant prices, allows going back to March 1961, but is now discontinued. The second one, the GDP at basic prices in 2007 constant prices, is used as soon as possible so that it is behind the ZGDPG variable from March 1997. Although they differ slightly in their methodology, the series produce growths correlated at 0.90 in their common time span.

Composite Leading Indicator Growth (ZLEAD): The composite leading indicator (CLI) growth is mainly the monthly growth of the unsmoothed CLI available as series V7687 from the CANSIM database. According to Statistics Canada, the CLI is "comprised of ten components which lead cyclical activity and together represent all major categories of GDP. It thus reflects a variety of mechanisms that can cause business cycles."[22] The components are an housing index, the business and personal services employment, the TSE 300 Index, the money supply M1, the U.S. Composite Leading Indicator, the average work week hours, the new orders in durable goods, the shipments/inventories of finished goods, the furniture and appliance sales and other durable goods sales. The series starts in May 1952, but is discontinued since April 2012. Thereafter, we use a CLI from the OECD Main Economic Indicators, available in Datastream as series CNCYLEADT. The CLIs have a correlation of 0.97 in their common time span.

[22] For this explanation and more details on the CLI series, see the Statistics Canada website at the following address: www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=1601&lang=fr&db=imdb&adm=8&dis=2.

References

Aiolfi, M., Capistrán, C., and Timmermann, A., 2011, "Forecast combinations," in: Clements, M.P., and Hendry, D.F. (eds.), *The Oxford Handbook of Economic Forecasting*, Oxford, United Kingdom: Oxford University Press, Ch. 12, 355–390.

Ang, A., Bekaert, G., 2007, "Return predictability: Is it there?" *Review of Financial Studies*, 20, 651–707.

Bates, J.M., and Granger, C.W.J., 1969, "The combination of forecasts," *Operational Research Quarterly*, 20, 451–468.

Bossaerts, P., and Hillion P., 1999, "Implementing statistical criteria to select return forecasting models: what do we learn?," *Review of Financial Studies*, 12, 405–428.

Boudoukh, J., Michely, R., Richardson, M. and Roberts, M., 2007, "On the importance of measuring payout yield: Implications for empirical asset pricing," *Journal of Finance*, 62, 877–915.

Campbell, J.Y., 1987, "Stock returns and the term structure," *Journal of Financial Economics*, 18, 373–399.

Campbell, J.Y., and Shiller, R., 1988, "Stock prices, earnings, and expected dividends," *Journal of Finance*, 43, 661–676.

Campbell, J.Y., and Thompson, S.B., 2008, "Predicting the equity premium out of sample: Can anything beat the historical average?" *Review of Financial Studies*, 21, p.1509–1531.

Carmichael, B., and Samson, L., 1996, "The determination of risk premiums and the integration of the Canadian and American stock markets (in French)," *Canadian Journal of Economics*, 24, 595–614.

Cenesizoglu, T., and Timmermann, A., 2012, "Do return prediction models add economic value?" *Journal of Banking and Finance*, 36, 2974–2987.

Chordia, T., and Shivakumar, L., 2002, "Momentum, business cycle, and time-varying expected returns," *Journal of Finance*, 57, 985–1019.

Chrétien, S., and Coggins, F., 2009, "Election outcomes and financial market returns in Canada," *North American Journal of Economics and Finance*, 20, 1–23.

Chrétien, S., and Coggins, F., 2017, "Additional evidence on information variables and equity premium predictability," Working Paper, Laval University.

Chrétien, S., and Kammoun, M., 2017, "Mutual fund performance evaluation and best clienteles," *Journal of Financial and Quantitative Analysis*, 52, 1577–1604.

Chrétien, S., and Kopoin, A., 2014, "Dynamic factor analysis and equity premium predictability: Canadian evidence," Working Paper, Laval University.

Clark, T.E., and McCracken, M.W., 2001, "Tests of equal forecast accuracy and encompassing for nested models," *Journal of Econometrics*, 105, 85–110.

Clark, T.E., and West, K.D., 2007, "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics*, 138, 291–311.

Cochrane, J.H., 2008, "The dog that did not bark: A defense of return predictability," *Review of Financial Studies*, 21, 1533–1575.

Cochrane, J.H., and Saá-Requejo, J., 2000, "Beyond arbitrage: good-deal asset price bounds in incomplete markets," *Journal of Political Economy*, 108, 79–119.

Deaves, R., Miu, P., and White, C.B., 2008, "Canadian stock market multiples and their predictive content," *International Review of Economics and Finance*, 17, 457–466.

Detzel, A., and Strauss, J., 2016, "The dog has barked for a long time: dividend growth is predictable," Working Paper, University of Denver.

Diebold, F.X., and Mariano, R.S., 1995, "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Elliott, G., and Timmermann, A., 2016, *Economic Forecasting*, Princeton University Press.

Fama, E.F., and French, K.R., 1988, "Dividend yields and expected stock returns," *Journal of Financial Economics*, 22, 3–25.

Fama, E.F., and French, K.R., 1989, "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, 25, 23–49.

Ferson, W.E., and Harvey, C.R., 1991, "The variation of economic risk premiums," *Journal of Political Economy*, 99, 385–415.

Ferson, W.E., and Harvey, C.R., 1994, "An exploratory investigation of the fundamental determinants of national equity market returns," in: Frankel, J.A. (ed.), *Internationalization of Equity Markets*, Chicago, United States: University of Chicago Press, 59–148.

Ferson, W.E., Sarkissian, S., and Simin, T.S., 2003, "Spurious regression in financial economics," *Journal of Finance*, 58, 1393–1413.

Ferson, W.E., and Schadt, R.W., 1996, "Measuring fund strategy and performance in changing economic conditions," *Journal of Finance*, 51, 425–461.

Goyal, A., and Welch, I., 2003, "Predicting the equity premium with dividend ratios," *Management Science*, 49, 639–654.

Granger, C.W.J., Hyung, N., and Jeon, Y., 2001, "Spurious regressions with stationary series," *Applied Economics*, 33, 899–904.

Guo, H., 2006a, "On the out-of-sample predictability of stock market returns," *Journal of Business*, 79, 645–670.

Guo, H., 2006b, “The risk-return relation in international stock markets,” *The Financial Review*, 41, 565–587.

Hjalmarsson, E., 2010, “Predicting global stock returns,” *Journal of Financial and Quantitative Analysis*, 45, 49–80.

Hocking, R.R., 1976, “The analysis and selection of variables in a linear regression,” *Biometrics*, 32, 1–50.

Kandel, S., and Stambaugh, R.F., 1996, “On the predictability of stock returns: an asset allocation perspective,” *Journal of Finance*, 51, 385–424.

Keim, D., and Stambaugh, R.F., 1986, “Predicting returns in the stock and bond markets,” *Journal of Financial Economics*, 17, 357–390.

Kooli, M., and L’Her, J.-F., 2010, “Dividends versus share repurchases evidence from Canada: 1985-2003,” *The Financial Review*, 45, 57–81.

Korkie, B., and Turtle, H., 1998, “The Canadian investment opportunity set, 1967-1993,” *Canadian Journal of Administrative Science*, 15, 213–229.

Lewellen, J., 2004, “Predicting returns with financial ratios,” *Journal of Financial Economics*, 74, 209–237.

Li, J., and Tsiakas, I., 2017, “Equity premium prediction: The role of economic and statistical constraints,” *Journal of Financial Markets*, 36, 56–75.

Ludvigson, S., and Ng, S., 2007, “Empirical risk return relation: A factor analysis approach,” *Journal of Financial Economics*, 83, 171–222.

Luo, X., and Zhang, J.E., 2016, “Expected stock returns and forward variance,” *Journal of Financial Markets*, forthcoming.

MacKinlay, A.C., 1995, “Multifactor models do not explain deviations from the CAPM,” *Journal of Financial Economics*, 38, 3–28.

Maio, P., 2016, “Cross-sectional return dispersion and the equity premium,” *Journal of Financial Markets*, 29, 87–109.

Mamaysky, H., Spiegel, M., and Zhang, H., 2007, “Improved forecasting of mutual fund alphas and betas,” *Review of Finance*, 11, 359–400.

Marquering, W., and Verbeek, M., 2004, “The economic value of predicting stock index returns and volatility,” *Journal of Financial and Quantitative Analysis*, 39, 407–429.

McClean, R.D., and Pontiff, J., 2016, “Does academic research destroy stock return predictability?” *Journal of Finance*, 71, 5–32.

Nelson, C., and Kim, M., 1993, “Predictable stock returns: the role of small sample bias,” *Journal of Finance*, 48, 641–661.

Newey, W.K., and West, K.D., 1987, "A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

Paye, B.S., and Timmermann, A., 2006, "Instability of return prediction models," *Journal of Empirical Finance*, 13, 274–315.

Pettenuzzo, D., Timmermann, A., and Valkanov, R., 2014, "Forecasting stock returns under economic constraints," *Journal of Financial Economics*, 114, 517–553.

Pesaran, M.H., and Timmermann, A., 1995, "Predictability of stock returns: robustness and economic significance," *Journal of Finance*, 50, 1201–1228.

Polk, C., Thompson, S., and Vuolteehaho, T., 2006, "Cross-sectional forecasts of the equity premium," *Journal of Financial Economics*, 81, 101–141.

Rapach, D., Strauss, J., 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.

Rapach, D., Strauss, J., and Zhou, G., 2013, "International stock return predictability: what is the role of the United States?" *Journal of Finance*, 68, 1633–1662.

Rapach, D., Wohar, M., and Rangvid, J., 2005, "Macro variables and international return predictability," *International Journal of Forecasting*, 21, 137–166.

Ross, S.A., 1976, "The arbitrage theory of capital asset pricing," *Journal of Economics Theory*, 13, 341–360.

Santa-Clara, P., and Valkanov, R., 2003, "The presidential puzzle: political cycles and the stock market," *Journal of Finance*, 58, 1841–1872.

Solnik, B., 1993, "The performance of international asset allocation strategies using conditioning information," *Journal of Empirical Finance*, 1, 33–55.

Stambaugh, R.F., 1999, "Predictive regressions," *Journal of Financial Economics*, 54, 375–421.

Timmermann, A., 2006, "Forecast Combinations," in: Elliott, G., Granger, C.W.J., and Timmermann, A. (eds.), *Handbook of Economic Forecasting*, Amsterdam, The Netherlands: Elsevier, Ch. 4, 135–196.

Timmermann, A., 2008, "Elusive return predictability," *International Journal of Forecasting*, 20, 15–27.

Turner, J.A., 2015, "Casting doubt on the predictability of stock returns in real time: Bayesian model averaging using realistic priors," *Review of Finance*, 19, 785–821.

Welch, I., and Goyal, A., 2008, "A comprehensive look at the empirical performance of equity premium prediction," *Review of Financial Studies*, 21, 1455–1508.

West, K.D., 1996, "Asymptotic inference about predictive ability," *Econometrica*, 64, 1067–1084.

Table 1: Overview of the Information Variables

Category	Variable	Description	Start	Data Sources
Equity Market Characteristic Variables	ZDY	Dividend Yield	1957:02	CFMRC
	ZDP	Dividend-Price Ratio	1957:02	CFMRC
	ZPE	Price-Earnings Ratio	1956:02	CANSIM V122629
	ZEQP	Previous Equity Premium	1950:03	Lagged value of EQP
	ZVOLG	Volume Growth	1953:03	CANSIM V37413
	ZISSUE	Issuing Activity	1951:02	CFMRC or Datastream
	ZJAN	January Dummy	1950:02	
Interest Rate, Yield Spread or Exchange Rate Variables	ZTBILLv	Treasury Bill Yield Variation	1950:02	Computed from TBILL, CFMRC or CANSIM V122541
	ZTBILLr	Treasury Bill Yield Relative	1950:02	Computed from TBILL
	ZLTGOV	Long Gov Bond Yield	1950:02	CFMRC or CANSIM V122487
	ZLTGOVv	Long Gov Bond Yield Variation	1950:02	Computed from ZLTGOV
	ZLTGOVr	Long Gov Bond Yield Relative	1950:02	Computed from ZLTGOV
	ZTERM	Term Premium	1950:02	Computed from ZLTGOV and ZTBILL
	ZCREDIT	Credit Premium	1950:02	CANSIM V35752, V122518, Datastream MLCCTPL and ZLTGOV
	ZCREDITs	Credit Premium Short	1956:02	CANSIM V122491 and ZTBILL
	ZCREDITr	Credit Premium Return	1950:12	CANSIM V35754, Datastream MLCCTPL, CFMRC
	ZFX	CAD/USD Rate	1950:11	CFMRC or CANSIM V37426
	ZFXv	CAD/USD Rate Variation	1950:12	Computed from ZFX
	ZFXr	CAD/USD Rate Relative	1951:10	Computed from ZFX
Macro-economic Variables	ZINF	Inflation Rate	1950:02	CANSIM V41690973
	ZPRODG	Industrial Production Growth	1956:03	CANSIM V53384745 or Datastream
	ZUNEMP	Unemployment Rate	1960:02	CANSIM V2064894 or Datastream CNOUN014R
	ZMONEYG	Money Supply Growth	1950:02	CANSIM V37173
	ZGDPG	GDP Growth	1961:03	CANSIM V329529 and V65201483
	ZLEAD	Leading Indicator Growth	1952:05	CANSIM V7687, Datastream CNCYLEADT

NOTES: This table presents an overview of the 25 information variables used in the multivariate predictive models. The column labelled Start gives the sample start date of the monthly observations of the variables. The end date is June 2015 for all variables. The column labelled Data Sources gives information on the sources of the series used to construct the information variables. CFMRC represents the Canadian Financial Markets Research Centre database. CANSIM represents the Canadian Socioeconomic database from Statistic Canada. Datastream represents the Thomson Reuters Datastream database.

Table 2: Descriptive Statistics of the Equity Premium and Information Variables

	Full Sample						1950-1969	1970-1991	1992-2015
Variable	Mean	Std Dev	Min	Max	Excess Kurt	Skew-ness	Mean	Mean	Mean
EQP	0.0046	0.043	-0.235	0.158	2.670	-0.700	0.0069	0.0017	0.0052
ZDY	0.0345	0.012	0.009	0.085	1.033	0.754	0.0369	0.0429	0.0254
ZDP	0.0323	0.010	0.009	0.086	0.605	0.296	0.0350	0.0399	0.0237
ZPE	24.786	35.008	6.580	254.98	32.163	5.482	17.024	13.732	39.732
ZEQP	0.0046	0.043	-0.235	0.158	2.673	-0.703	0.0070	0.0016	0.0054
ZVOLG	0.0293	0.235	-0.577	1.885	7.132	1.515	0.0370	0.0324	0.0208
ZISSUE	0.0273	0.082	-0.223	0.543	14.322	2.524	0.0259	0.0352	0.0209
ZJAN	0.0828	0.276	0.000	1.000	7.221	3.034	0.0795	0.0833	0.0851
ZTBILLv	0.0000	0.005	-0.036	0.033	13.637	0.102	0.0003	0.0000	-0.0002
ZTBILLr	0.0000	0.011	-0.039	0.045	2.230	0.105	0.0016	0.0002	-0.0015
ZLTGOV	0.0671	0.031	0.018	0.177	0.094	0.749	0.0466	0.1012	0.0525
ZLTGOVv	0.0000	0.003	-0.023	0.020	12.364	-0.171	0.0002	0.0000	-0.0003
ZLTGOVr	-0.0001	0.005	-0.026	0.032	5.978	0.616	0.0011	0.0004	-0.0015
ZTERM	0.0142	0.014	-0.043	0.044	1.697	-0.909	0.0141	0.0082	0.0199
ZCREDIT	0.0105	0.005	0.002	0.037	3.880	1.481	0.0070	0.0104	0.0134
ZCREDITs	0.0049	0.006	-0.002	0.039	5.930	2.196	0.0077	0.0060	0.0023
ZCREDITr	0.0005	0.013	-0.067	0.092	5.973	0.033	0.0006	0.0005	0.0005
ZFX	1.1537	0.168	0.948	1.600	-0.350	0.813	1.0235	1.1515	1.2618
ZFXv	0.0002	0.014	-0.074	0.126	9.551	0.518	0.0001	0.0002	0.0003
ZFXr	0.0010	0.035	-0.158	0.170	4.213	0.023	0.0005	0.0013	0.0012
ZINF	0.0030	0.005	-0.013	0.026	1.755	0.626	0.0021	0.0054	0.0015
ZPRODG	0.0027	0.007	-0.029	0.038	3.863	0.351	0.0012	0.0049	0.0016
ZUNEMP	0.0752	0.022	0.024	0.141	-0.113	0.219	0.0506	0.0813	0.0798
ZMONEYG	0.0053	0.016	-0.051	0.053	1.817	-0.670	0.0043	0.0073	0.0045
ZGDPG	0.0026	0.005	-0.015	0.022	1.100	0.004	0.0045	0.0025	0.0021
ZLEAD	0.0032	0.009	-0.028	0.035	0.691	-0.188	0.0027	0.0030	0.0038

NOTES: This table presents the full-sample mean, standard deviation, minimum, maximum, excess kurtosis and skewness, as well as the mean in three sub-periods for the variables in the study. The data are at monthly frequency and cover the period from February 1950 to June 2015. The variable EQP is the Canadian equity premium. The variables beginning by Z are the information variables and have been lagged by one month. The information variables are described in table 1 and in appendix A.

Table 3: Univariate In-Sample Forecasting Results for 1950-1969

	1950-1969	
Variable	Adj R^2	F-stat
ZDY	-0.10%	0.86
ZDP	6.02%	11.01 ***
ZPE	0.19%	1.33
ZEQP	2.16%	6.28 **
ZVOLG	-0.30%	0.41
ZISSUE	0.15%	1.36
ZJAN	1.43%	4.49 **
ZTBILLv	-0.42%	0.00
ZTBILLr	0.56%	2.36 *
ZLTGOV	1.06%	3.59 *
ZLTGOVv	-0.35%	0.17
ZLTGOVr	1.98%	5.86 **
ZTERM	1.76%	5.30 **
ZCREDIT	-0.02%	0.96
ZCREDITs	-0.24%	0.61
ZCREDITr	-0.01%	1.00
ZFX	-0.41%	0.06
ZFXv	-0.28%	0.36
ZFXr	-0.31%	0.33
ZINF	-0.40%	0.06
ZPRODG	-0.58%	0.04
ZUNEMP	0.00	0.60
ZMONEYG	0.84%	3.05 *
ZGDPG	0.43%	1.48
ZLEAD	2.74%	7.00 ***

NOTES: This table presents in-sample predictive regression results for the 25 information variables in the pre-evaluation period. The data are at monthly frequency and cover the period from February 1950 to December 1969. The sample start date for each regression depends on data availability for each information variable. The information variables (including their sample start date) are described in table 1 and in appendix A. The columns labelled Adj R^2 and F-stat give the adjusted R^2 statistic and F-statistic of the predictive regressions. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 4: In-Sample Forecasting Results

	1970-2015		1970-1991		1992-2015	
Model	R^2	Gain	R^2	Gain	R^2	Gain
MV-CAN	2.87% ***	3.73%	5.33% ***	4.97%	1.24% **	4.14%
MV-US	1.17% ***	2.41%	1.92% **	2.94%	1.31% *	3.25%
MV-BV	4.16% ***	5.16%	6.12% ***	3.90%	4.02% ***	4.65%
MV-ALL	8.42% ***	8.32%	12.08% ***	10.46%	11.45% ***	10.10%
MV-FOR	5.31% ***	6.60%	4.19% ***	3.50%	3.50% ***	5.16%
MV-STEP	4.88% ***	5.32%	4.19% ***	3.50%	3.50% ***	5.16%
MV-BACK	6.69% ***	6.23%	6.85% ***	7.60%	5.04% ***	6.45%
CF_{MEAN}-CAN	0.69% ***	1.43%	1.42% ***	2.35%	0.26% *	0.87%
CF_{MED}-CAN	0.28% **	0.78%	0.94% ***	1.99%	0.06%	0.21%
CF_{TRIM}-CAN	0.36% ***	0.79%	0.97% ***	1.93%	0.05%	0.26%
CF_{DMSPE1}-CAN	0.70% ***	1.43%	1.42% ***	2.35%	0.26% *	0.87%
CF_{DMSPE2}-CAN						
CF_{MEAN}-US	0.58% **	1.30%	0.95% **	1.68%	0.46% *	1.54%
CF_{MED}-US	0.47% **	1.05%	0.92% **	1.70%	0.48% *	1.54%
CF_{TRIM}-US	0.47% **	1.05%	0.92% **	1.70%	0.48% *	1.54%
CF_{DMSPE1}-US	0.53% **	1.30%	0.72% *	1.63%	0.23%	0.95%
CF_{DMSPE2}-US						
CF_{MEAN}-BV	1.40% ***	3.08%	2.01% ***	3.13%	1.13% ***	3.10%
CF_{MED}-BV	0.96% ***	1.85%	1.57% ***	3.03%	0.25%	1.11%
CF_{TRIM}-BV	1.08% ***	2.49%	1.57% ***	2.85%	0.80% ***	2.34%
CF_{DMSPE1}-BV	1.38% ***	3.03%	1.92% ***	3.08%	1.06% **	2.85%
CF_{DMSPE2}-BV						
CF_{MEAN}-ALL	0.97% ***	2.23%	1.25% ***	2.36%	1.04% ***	2.54%
CF_{MED}-ALL	0.32% ***	0.82%	0.52% ***	1.28%	0.29% ***	0.69%
CF_{TRIM}-ALL	0.80% ***	1.89%	1.03% ***	2.10%	0.89% ***	2.19%
CF_{DMSPE1}-ALL	0.97% ***	2.24%	1.27% ***	2.37%	1.05% ***	2.56%
CF_{DMSPE2}-ALL						

NOTES: This table presents the IS equity premium forecasting results for 27 predictive models based on either the MV approach (denoted as the MV-*k* models) or the CF approach (denoted as the CF_w-*k* models). The MV models are described in section 3.1.2. The CF_w models and their associated weighting schemes, w = {MEAN, MED, TRIM, DMSPE1, DMSPE2}, are described in section 3.1.3. The subsets of information variables used in the models, k = {CAN, US, BV, ALL, FOR, STEP, BACK}, are described in section 3.1.4. The results are provided for the full evaluation period (1970-2015), first sub-period (1970-1991) and second sub-period (1992-2015). The columns labelled R^2 give the R^2 statistic of equation (7). The columns labelled Gain give the utility gain for a mean-variance investor with risk aversion coefficient of three, or the management fee (in annualized percentage return) that such an investor would be willing to pay to have access to the forecasting models relative to a historical average benchmark model. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 5: In-Sample Multivariate Regression Results

	MV-CAN			MV-US			MV-BV			MV-ALL			MV-FOR			MV-STEP			MV-BACK		
	1970-2015	1970-1991	1992-2015	1970-2015	1970-1991	1992-2015	1970-2015	1970-1991	1992-2015	1970-2015	1970-1991	1992-2015	1970-2015	1970-1991	1992-2015	1970-2015	1970-1991	1992-2015	1970-2015	1970-1991	1992-2015
Constant	-0.03	-0.03	-0.05	0.00	-0.01	0.00	-0.01	-0.01	-0.01	-0.06 *	-0.07	-0.06	0.00	0.00	0.00	0.00	0.00	0.00	-0.05 **	0.03 *	0.00
ZDY	0.46 **	0.36	0.81							-0.45	0.17	-1.31									
ZDP				0.17	0.43	0.57	0.30	0.41	0.27	1.27 **	0.10	3.81 ***							0.68 **		
ZPE										0.00 **	0.00 **	0.00 **							0.00 **	0.00 ***	
ZEQP							0.04	-0.03	0.12	-0.02	-0.10	0.07			0.15 **			0.15 **			
ZVOLG										0.00	0.00	-0.01									
ZISSUE										-0.02	-0.05 **	0.02									
ZJAN	0.01	0.02	0.00				0.01	0.01	0.01	0.01	0.01	0.01									
ZTBILLv	-0.95 ***	-0.98 **	-0.77							-0.64	-0.95	-0.73									
ZTBILLr				-0.16	-0.26	0.12	0.28	0.37	0.23	0.32	0.49	0.33									
ZLTGOV	-0.20 **	-0.43 **	-0.15				-0.06	-0.09	-0.02	-0.32 **	-0.68 **	-0.14							-0.24 ***	-0.67 ***	
ZLTGOVv										-1.49 *	-1.42	-1.83	-1.82 ***			-1.74 ***			-2.22 ***		-3.36 **
ZLTGOVr							-1.05 **	-1.49 ***	-0.26	-0.43	-0.35	-0.35	-0.65 *	-1.26 ***		-0.79 **	-1.26 ***				
ZTERM				0.24	0.17	0.23	0.18	0.11	0.24	0.13	0.10	0.35									
ZCREDIT				-0.32	-0.78	-0.60				-0.38	2.05	-1.45									
ZCREDITs										0.09	0.28	-0.89									
ZCREDITr										0.31	0.39	0.19	0.40 *			0.43 *			0.37 *		0.54 **
ZFX	0.03 *	0.05 **	0.03							0.04 *	0.04	0.06 **							0.04 **		
ZFXv										-0.07	0.20	-0.09									
ZFXr										-0.04	-0.13	-0.01									
ZINF										0.19	0.32	-0.49									
ZPRODG										0.15	0.21	0.08									
ZUNEMP										0.19	0.84 **	-0.58								0.78 ***	
ZMONEYG							-0.04	0.02	-0.25	0.03	0.25	-0.37									
ZGDPG										0.79	0.52	1.49 *	0.79 *		1.19	0.94 **		1.19	0.77		1.44 *
ZLEAD							0.36	0.27	0.39	0.41	0.37	0.39	0.37						0.45 *		
R^2	2.87%	5.33%	1.24%	1.17%	1.92%	1.31%	4.16%	6.12%	4.02%	8.42%	12.08%	11.45%	5.31%	4.19%	3.50%	4.88%	4.19%	3.50%	6.69%	6.85%	5.04%
Adj R^2	1.97%	3.50%	-0.55%	0.44%	0.41%	-0.12%	2.55%	2.80%	0.85%	4.02%	2.84%	2.80%	4.43%	3.83%	2.80%	4.18%	3.83%	2.80%	5.30%	5.77%	4.02%
F-stat	3.19 ***	2.91 **	0.69	1.60	1.27	0.92	2.59 ***	1.84 *	1.27	1.91 ***	1.31	1.32	6.05 ***	11.46 ***	5.05 ***	6.94 ***	11.46 ***	5.05 ***	4.81 ***	6.37 ***	4.92 ***

NOTES: This table presents the IS predictive regression coefficient estimates and results for 7 predictive models based on the MV approach (denoted as the MV-*k* models). The MV models are described in section 3.1.2. The subsets of information variables used in the models, k = {CAN, US, BV, ALL, FOR, STEP, BACK}, are described in section 3.1.4. The results are provided for the full evaluation period (1970-2015), first sub-period (1970-1991) and second sub-period (1992-2015). The top portion of the table gives the value and significance of the coefficient estimate associated with the constant and each information variable in the models. The bottom portion of the table gives the R^2 statistic, the adjusted R^2 statistic and the F-statistic (to test the null hypothesis that all variable coefficients are equal to zero). ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 6: Out-of-Sample Forecasting Results

	1970-2015		1970-1991		1992-2015	
Model	R^2	Gain	R^2	Gain	R^2	Gain
MV-CAN	-0.77%	2.33%	-0.28%	2.72%	-1.41%	1.96%
MV-US	-3.38%	0.80%	-4.43%	0.34%	-2.00%	1.23%
MV-BV	-2.84%	3.47%	-4.85%	3.75%	-0.19%	3.21%
MV-ALL	-14.77%	2.27%	-23.45%	1.43%	-3.36%	3.05%
MV-FOR	-6.10%	2.01%	-9.72%	0.24%	-1.35%	3.65%
MV-STEP	-5.84%	1.98%	-9.22%	0.24%	-1.40%	3.59%
MV-BACK	-7.63%	2.16%	-9.56%	2.28%	-5.09%	2.04%
CF_{MEAN}-CAN	0.41% *	1.09%	0.71% *	2.18%	0.01%	0.07%
CF_{MED}-CAN	0.11%	0.21%	0.22%	0.57%	-0.04%	-0.11%
CF_{TRIM}-CAN	0.15%	0.25%	0.30%	0.80%	-0.05%	-0.27%
CF_{DMSPE1}-CAN	0.43% *	1.13%	0.75% *	2.26%	0.02%	0.08%
CF_{DMSPE2}-CAN	0.72% **	1.63%	1.15% **	2.98%	0.16%	0.36%
CF_{MEAN}-US	0.49% *	1.37%	0.76% *	1.86%	0.14%	0.93%
CF_{MED}-US	0.87% ***	2.20%	1.22% **	3.05%	0.41%	1.40%
CF_{TRIM}-US	0.87% ***	2.20%	1.22% **	3.05%	0.41%	1.40%
CF_{DMSPE1}-US	0.58% *	1.95%	0.86% *	2.90%	0.21%	1.06%
CF_{DMSPE2}-US	0.90% **	1.84%	1.27% **	2.40%	0.40%	1.33%
CF_{MEAN}-BV	1.36% ***	3.08%	1.79% **	4.19%	0.80% **	2.05%
CF_{MED}-BV	1.12% ***	2.42%	1.40% **	3.02%	0.74% **	1.86%
CF_{TRIM}-BV	1.13% ***	2.70%	1.40% **	3.47%	0.78% ***	1.97%
CF_{DMSPE1}-BV	1.41% ***	3.26%	1.82% **	4.44%	0.87% ***	2.15%
CF_{DMSPE2}-BV	1.86% ***	3.78%	2.48% ***	5.12%	1.05% ***	2.52%
CF_{MEAN}-ALL	0.62% ***	1.70%	0.71% **	2.35%	0.50% ***	1.10%
CF_{MED}-ALL	0.40% ***	0.99%	0.49% ***	1.31%	0.28% ***	0.70%
CF_{TRIM}-ALL	0.48% ***	1.39%	0.50% *	1.83%	0.46% ***	0.97%
CF_{DMSPE1}-ALL	0.64% ***	1.74%	0.75% **	2.42%	0.51% **	1.11%
CF_{DMSPE2}-ALL	0.97% ***	2.36%	1.20% ***	3.31%	0.67% ***	1.47%

NOTES: This table presents the OS equity premium forecasting results for 27 predictive models estimated with a recursive window scheme and based on either the MV approach (denoted as the MV-*k* models) or the CF approach (denoted as the CF_w-*k* models). The MV models are described in section 3.1.2. The CF_w models and their associated weighting schemes, *w* = {MEAN, MED, TRIM, DMSPE1, DMSPE2}, are described in section 3.1.3. The subsets of information variables used in the models, *k* = {CAN, US, BV, ALL, FOR, STEP, BACK}, are described in section 3.1.4. The results are provided for the full evaluation period (1970-2015), first sub-period (1970-1991) and second sub-period (1992-2015). The columns labelled R^2 give the R^2 statistic of equation (7). The columns labelled Gain give the utility gain for a mean-variance investor with risk aversion coefficient of three, or the management fee (in annualized percentage return) that such an investor would be willing to pay to have access to the forecasting models relative to a historical average benchmark model. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 7: Out-of-Sample Forecasting Results with a Rolling Estimation Window

	1970-2015		1970-1991		1992-2015	
Model	R^2	Gain	R^2	Gain	R^2	Gain
MV-CAN	-2.27%	1.98%	-2.15%	2.33%	-2.42%	1.64%
MV-US	-5.19%	0.50%	-5.54%	0.25%	-4.72%	0.74%
MV-BV	-4.70%	2.28%	-6.27%	3.55%	-2.64%	1.10%
MV-ALL	-23.00%	0.77%	-25.74%	1.14%	-19.42%	0.43%
MV-FOR	-8.68%	0.05%	-10.23%	0.41%	-6.64%	-0.29%
MV-STEP	-8.09%	0.12%	-9.66%	0.41%	-6.03%	-0.16%
MV-BACK	-12.09%	2.11%	-11.95%	2.02%	-12.27%	2.19%
CF_{MEAN}-CAN	0.12%	0.21%	0.37%	1.22%	-0.21%	-0.75%
CF_{MED}-CAN	0.25%	0.29%	0.15%	-0.04%	0.38% *	0.60%
CF_{TRIM}-CAN	0.14%	-0.11%	0.21%	0.37%	0.05%	-0.57%
CF_{DMSPE1}-CAN	0.15%	0.25%	0.41%	1.29%	-0.20%	-0.74%
CF_{DMSPE2}-CAN	0.47% *	0.76%	0.78% *	1.96%	0.06%	-0.36%
CF_{MEAN}-US	0.22%	0.86%	0.34%	1.05%	0.07%	0.68%
CF_{MED}-US	0.51% **	1.53%	0.75% **	2.27%	0.19%	0.84%
CF_{TRIM}-US	0.51% **	1.53%	0.75% **	2.27%	0.19%	0.84%
CF_{DMSPE1}-US	0.39%	1.44%	0.43%	2.18%	0.34%	0.75%
CF_{DMSPE2}-US	0.64% *	1.27%	0.78% *	1.51%	0.46%	1.05%
CF_{MEAN}-BV	1.17% ***	2.34%	1.42% **	3.04%	0.84% **	1.68%
CF_{MED}-BV	0.90% ***	1.64%	1.25% **	2.55%	0.44% *	0.79%
CF_{TRIM}-BV	1.08% ***	2.44%	1.15% **	2.89%	0.99% **	2.03%
CF_{DMSPE1}-BV	1.25% ***	2.60%	1.44% **	3.42%	0.99% **	1.83%
CF_{DMSPE2}-BV	1.65% ***	3.09%	2.02% ***	3.86%	1.17% **	2.38%
CF_{MEAN}-ALL	0.41% *	0.84%	0.51% *	1.54%	0.27%	0.19%
CF_{MED}-ALL	0.33% **	0.44%	0.43% **	0.75%	0.19%	0.15%
CF_{TRIM}-ALL	0.35% *	0.75%	0.39% *	1.14%	0.29%	0.39%
CF_{DMSPE1}-ALL	0.43% **	0.89%	0.54% *	1.62%	0.28%	0.21%
CF_{DMSPE2}-ALL	0.79% ***	1.56%	0.95% ***	2.44%	0.57%	0.74%

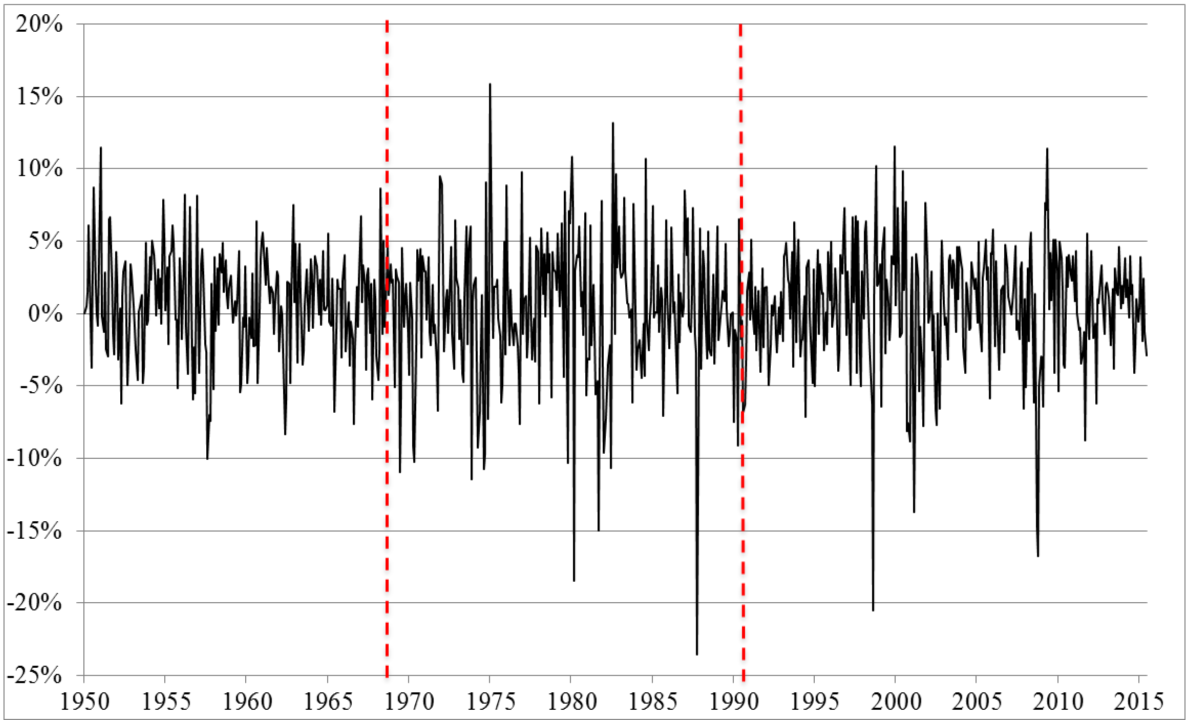
NOTES: This table presents the OS equity premium forecasting results for 27 predictive models estimated with a rolling (240-month) estimation window scheme and based on either the MV approach (denoted as the MV-*k* models) or the CF approach (denoted as the CF_w-*k* models). The MV models are described in section 3.1.2. The CF_w models and their associated weighting schemes, *w* = {MEAN, MED, TRIM, DMSPE1, DMSPE2}, are described in section 3.1.3. The subsets of information variables used in the models, *k* = {CAN, US, BV, ALL, FOR, STEP, BACK}, are described in section 3.1.4. The results are provided for the full evaluation period (1970-2015), first sub-period (1970-1991) and second sub-period (1992-2015). The columns labelled R^2 give the R^2 statistic of equation (7). The columns labelled Gain give the utility gain for a mean-variance investor with risk aversion coefficient of three, or the management fee (in annualized percentage return) that such an investor would be willing to pay to have access to the forecasting models relative to a historical average benchmark model. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 8: Out-of-Sample Forecasting Results with Restricted Forecasts

	R^2 for 1970-2015			R^2 for 1970-1991			R^2 for 1992-2015		
Model	$\hat{r} \geq 0\%$	$\hat{r} \leq 1\%$	$0\% \leq \hat{r} \leq 1\%$	$\hat{r} \geq 0\%$	$\hat{r} \leq 1\%$	$0\% \leq \hat{r} \leq 1\%$	$\hat{r} \geq 0\%$	$\hat{r} \leq 1\%$	$0\% \leq \hat{r} \leq 1\%$
MV-CAN	-0.65%	1.04% ***	1.16% ***	-0.29%	1.42% **	1.41% **	-1.11%	0.54% *	0.84% *
MV-US	-1.82%	-1.65%	-0.09%	-3.37%	-1.38%	-0.32%	0.22%	-2.00%	0.22%
MV-BV	-1.91%	0.15% **	1.07% ***	-3.65%	-0.03%	1.18% **	0.37% **	0.38% *	0.94% **
MV-ALL	-10.73%	-3.29%	0.75% **	-18.11%	-4.83%	0.50%	-1.02%	-1.27%	1.07% **
MV-FOR	-2.87%	-2.41%	0.83% **	-6.29%	-3.21%	0.21%	1.64%	-1.35%	1.64% ***
MV-STEP	-2.88%	-2.15%	0.81% **	-6.29%	-2.71%	0.21%	1.61% ***	-1.42%	1.59% ***
MV-BACK	-3.36%	-3.58%	0.69% **	-5.64%	-3.06%	0.86% *	-0.36%	-4.26%	0.47% *
CF_{MEAN}-CAN	0.45% *	0.40% *	0.44% *	0.77% *	0.70% *	0.76% *	0.02%	0.01%	0.02%
CF_{MED}-CAN	0.11%	0.11%	0.11%	0.23%	0.22%	0.23%	-0.04%	-0.04%	-0.04%
CF_{TRIM}-CAN	0.16%	0.15%	0.16%	0.31%	0.30%	0.31%	-0.05%	-0.05%	-0.05%
CF_{DMSPE1}-CAN	0.47% **	0.42% *	0.46% *	0.81% *	0.73% *	0.79% *	0.03%	0.02%	0.03%
CF_{DMSPE2}-CAN	0.73% **	0.71% **	0.72% **	1.17% **	1.12% **	1.14% **	0.16%	0.16%	0.16%
CF_{MEAN}-US	0.38% *	0.52% *	0.41% *	0.57%	0.82% *	0.63% *	0.12%	0.14%	0.12%
CF_{MED}-US	0.85% ***	0.91% ***	0.89% ***	1.19% **	1.28% **	1.25% **	0.41%	0.41%	0.41%
CF_{TRIM}-US	0.85% ***	0.91% ***	0.89% ***	1.19% **	1.28% **	1.25% **	0.41%	0.41%	0.41%
CF_{DMSPE1}-US	0.54% *	0.58% *	0.54% *	0.81% *	0.86% *	0.81% *	0.19%	0.21%	0.19%
CF_{DMSPE2}-US	0.72% **	0.82% **	0.64% **	1.02% **	1.14% **	0.89% **	0.32%	0.40%	0.32%
CF_{MEAN}-BV	1.28% ***	1.32% ***	1.24% ***	1.64% ***	1.72% **	1.58% ***	0.80% **	0.80% **	0.80% **
CF_{MED}-BV	1.07% ***	1.14% ***	1.09% ***	1.31% ***	1.43% **	1.34% ***	0.75% **	0.75% **	0.76% **
CF_{TRIM}-BV	1.13% ***	1.14% ***	1.13% ***	1.39% ***	1.40% **	1.40% ***	0.78% ***	0.78% ***	0.78% ***
CF_{DMSPE1}-BV	1.30% ***	1.39% ***	1.28% ***	1.64% ***	1.79% **	1.60% ***	0.87% ***	0.87% ***	0.87% ***
CF_{DMSPE2}-BV	1.70% ***	1.77% ***	1.61% ***	2.20% ***	2.31% ***	2.03% ***	1.05% ***	1.05% ***	1.05% ***
CF_{MEAN}-ALL	0.63% ***	0.63% ***	0.63% ***	0.73% **	0.72% **	0.74% **	0.50% ***	0.50% ***	0.50% ***
CF_{MED}-ALL	0.41% ***	0.40% ***	0.41% ***	0.50% ***	0.49% ***	0.50% ***	0.28% ***	0.28% ***	0.28% ***
CF_{TRIM}-ALL	0.50% ***	0.48% ***	0.50% ***	0.53% **	0.50% *	0.53% **	0.46% ***	0.46% ***	0.46% ***
CF_{DMSPE1}-ALL	0.65% ***	0.65% ***	0.66% ***	0.76% **	0.75% **	0.77% **	0.51% ***	0.51% ***	0.51% ***
CF_{DMSPE2}-ALL	0.97% ***	0.97% ***	0.97% ***	1.19% ***	1.20% ***	1.19% ***	0.67% ***	0.67% ***	0.67% ***

NOTES: This table presents the OS R^2 statistic results for 27 predictive models with forecasts restricted to be nonnegative (denoted $\hat{r} \geq 0\%$), less than or equal to 1% (denoted $\hat{r} \leq 1\%$), or nonnegative and less than or equal to 1% (denoted $0\% \leq \hat{r} \leq 1\%$). The models are estimated with a recursive window scheme and based on either the MV approach (denoted as the MV-*k* models) or the CF approach (denoted as the CF_w-*k* models). The MV models are described in section 3.1.2. The CF_w models and their associated weighting schemes, w = {MEAN, MED, TRIM, DMSPE1, DMSPE2}, are described in section 3.1.3. The subsets of information variables used in the models, k = {CAN, US, BV, ALL, FOR, STEP, BACK}, are described in section 3.1.4. The results are provided for the full evaluation period (1970-2015), first sub-period (1970-1991) and second sub-period (1992-2015). ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

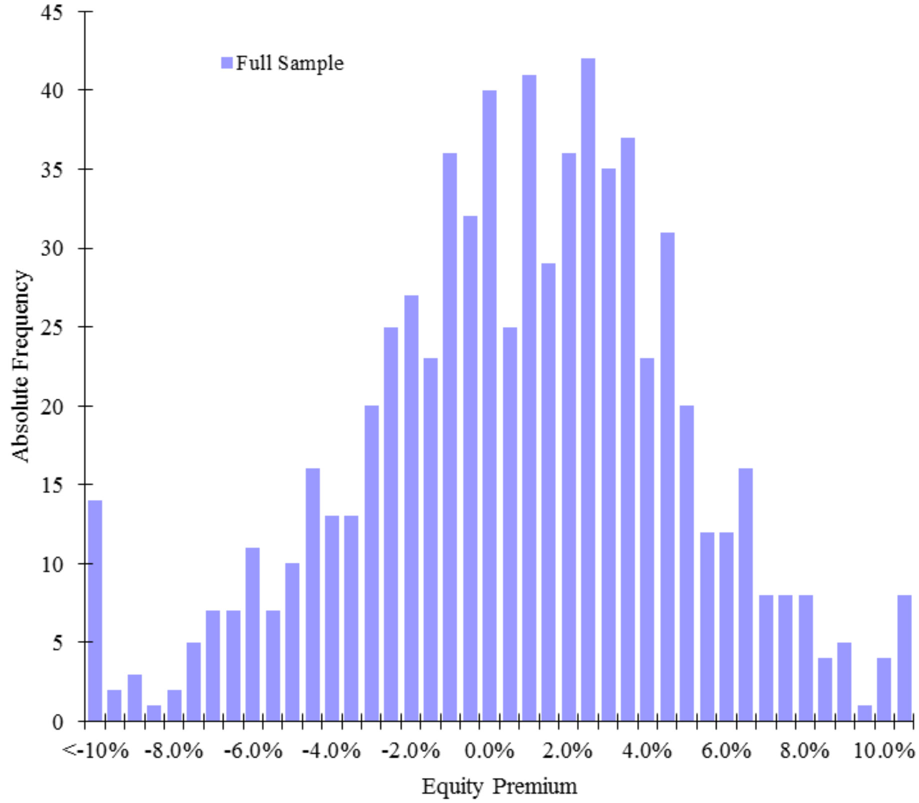
Figure 1: Canadian Equity Premium from February 1950 to June 2015



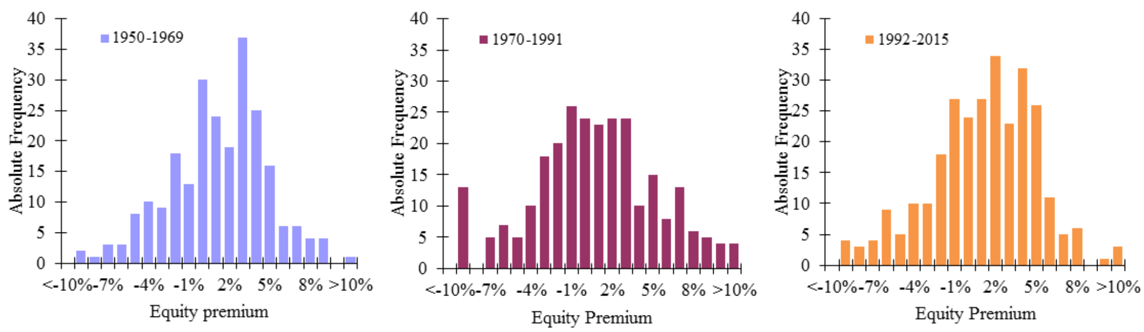
NOTES: This figure shows the evolution of the monthly realized equity premium in Canada from February 1950 to June 2015. Vertical dashed lines split the series into the pre-evaluation period, the first sub-period and the second sub-period.

Figure 2: Histograms of the Canadian Equity Premium from February 1950 to June 2015

Panel A: Full Sample

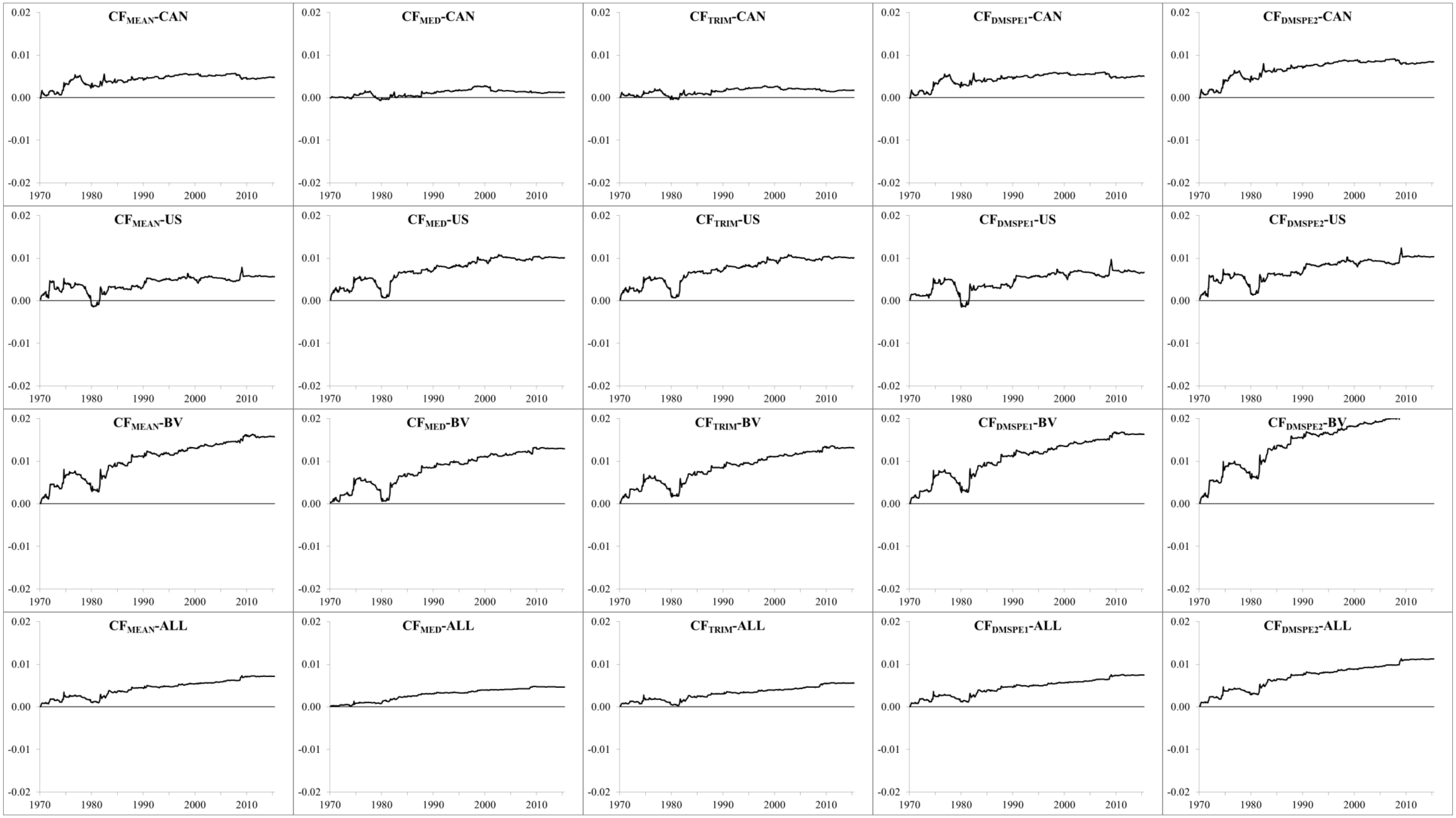


Panel B: Sub-period samples



NOTES: This figure illustrates the distribution of the monthly realized equity premium in Canada by showing histograms for the full sample (panel A) and for sub-period samples (panel B). The full sample covers the period from February 1950 to June 2015. The sub-period samples include the pre-evaluation period (1950-1969), first sub-period (1970-1991) and second sub-period (1992-2015).

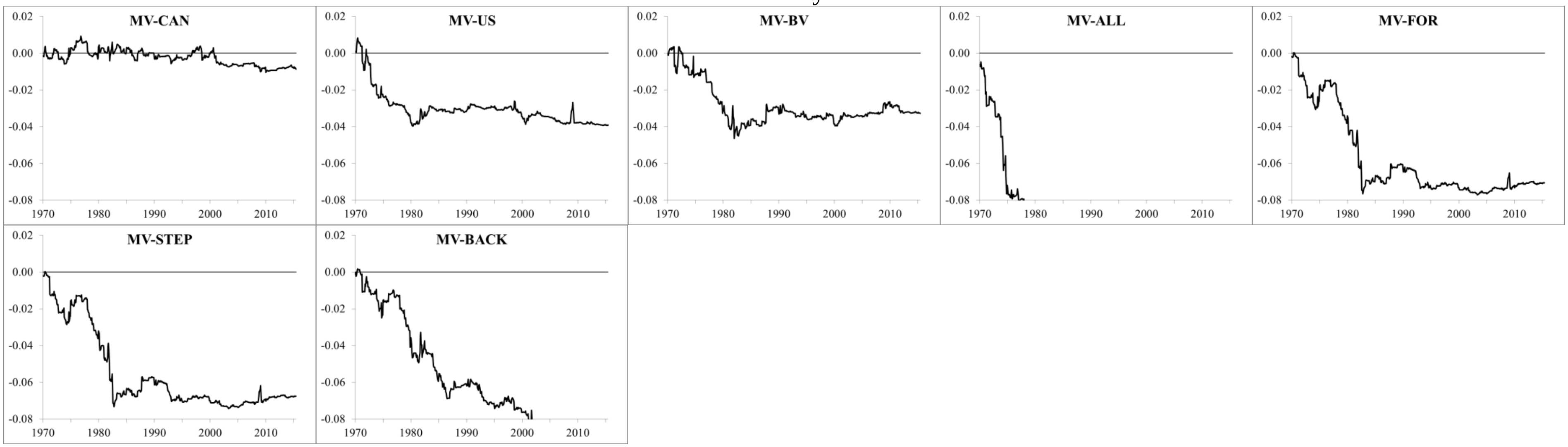
Figure 3: Differences in Cumulative Squared Forecast Errors for the Combination Forecast Models



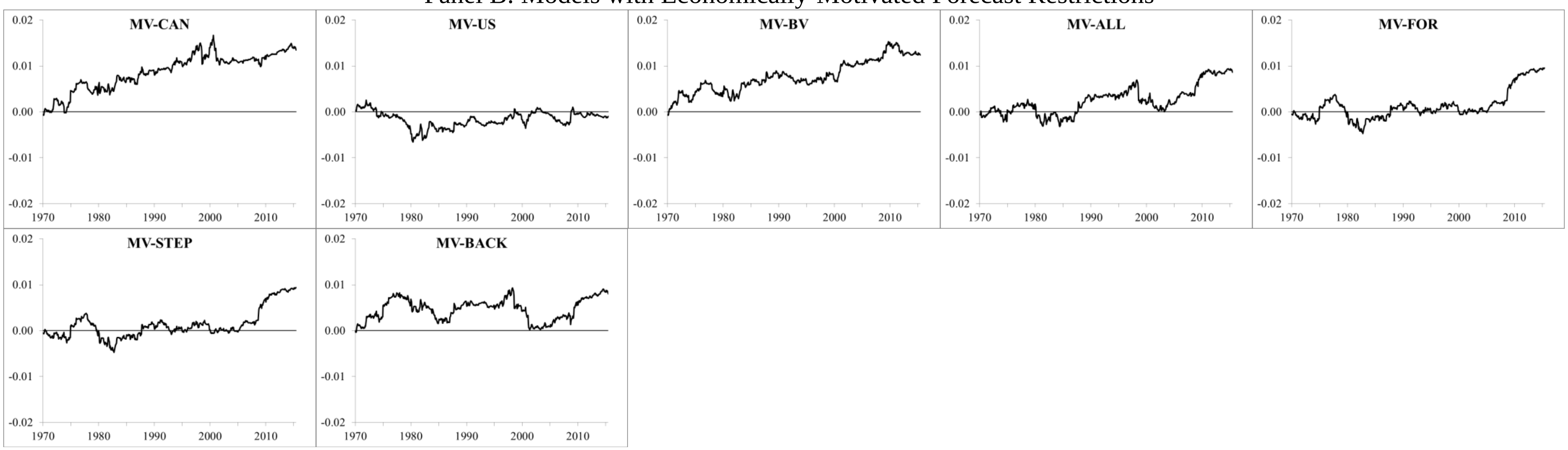
NOTES: This figure shows the OS statistical performance through time of the monthly forecasts from the CF models. The performance from January 1970 to June 2015 is illustrated by the cumulative squared forecast error differences between the historical average benchmark model and the predictive model noted in each graph. An increase (a decrease) in a line indicates better performance by the predictive (historical average) model.

Figure 4: Differences in Cumulative Squared Forecast Errors for the Multivariate Regression Models

Panel A: Models without Economically-Motivated Forecast Restrictions



Panel B: Models with Economically-Motivated Forecast Restrictions



NOTES: This figure shows the OS performance through time of the monthly forecasts from the MV models without (panel A) or with (panel B) economically-motivated forecast restrictions. The performance from January 1970 to June 2015 is illustrated by the cumulative squared forecast error differences between the historical average benchmark model and the predictive model noted in each graph. In panel B, the predictive model forecasts are set to 0% when negative and to 1% when greater than such value. An increase (a decrease) in a line indicates better performance by the predictive (historical average) model.