Tail Risks and Forecastability of Stock Returns of Advanced Economies: Evidence from Centuries of Data

Afees A. Salisu*, Rangan Gupta** and Ahamuefula E. Ogbonna***

Abstract

This study examines the out-of-sample predictability of market risks measured as tail risks for stock returns of eight (Canada, France, Germany, Japan, Italy, Switzerland, the United Kingdom (UK), and the United States (US)) advanced countries using a long-range monthly data of over a century. We follow the Conditional Autoregressive Value at Risk (CAViaR) of Engle and Manganelli (2004) to measure the tail risks since it utilizes the tail distribution rather the whole distribution. Consequently, we produce results for both 1% and 5% VaRs across four variants (Adaptive, Symmetric absolute value, Asymmetric slope and Indirect GARCH) of the CAViaR. Thereafter, we use relevant model diagnostics such as the Dynamic Quantile test (DQ) test and %Hits to determine the model that best fits the data. The results obtained are then used in the return predictability following the Westerlund and Narayan (2012, 2015) method which allows us to account for some salient features such as persistence, endogeneity and conditional heteroscedasticity effects. We consequently partition our models into three variants (one-predictor, two-predictor and three-predictor models) and examine their forecast performance in contrast with a driftless random walk model. Three findings are discernible from the empirical analysis. First, we find that the choice of VaR matters when determining the "best" fit CAViaR model for each return series as the outcome seems to differ between 1% and 5% VaRs. Second, the predictive model that incorporates both stock tail risk and oil tail risk produces better forecast outcomes than the one with own tail risk indicating the significance of both domestic and global risks in the return predictability of advanced countries.

Keywords: Stock returns; Tail risks; Forecasting; Advanced equity markets

JEL Codes: C22; G15; G17; Q02

1. Introduction

On one hand, practitioners in finance require forecasts of stock returns for asset allocation. On the other hand, academics are interested in stock return forecasts since they have important implications for producing robust measures of market efficiency, which in turn, helps to produce more realistic asset pricing models (Rapach and Zhou, 2013). Naturally, the existing literature on

* Corresponding author. Centre for Econometric and Allied Research, University of Ibadan, Ibadan, Nigeria. Email: adebare1@yahoo.com.

** Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa; Email address: rangan.gupta@up.ac.za.

*** Centre for Econometric and Allied Research and Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria. Emails: ae.ogbonna@cear.org.ng; ogbonnaephraim@gmail.com.

forecasting international stock returns, based on a wide array of models and predictors, is vast (see for example, Rapach et al., (2005), Gupta et al., (2017, 2020), Jordan et al., (2017, 2018), Christou and Gupta (2020), Christou et al., (forthcoming), and references cited therein).

In this regard, over the last two decades, in the wake of multiple periods of financial distress like the burst of the dot-com bubble, the Lehman default, the "Great Recession" followed by the European debt crisis and the Chinese stock market crash, and more recently the outbreak of the COVID-19 pandemic, studies have delved into the issue of computing tail risk of the equity (asset) market (Engle and Manganelli, 2004; Bollerslev and Todorov, 2011; Bollerslev et al., 2015; Cremers et al., 2015; Baker et al., 2020), and then analysing its predictive role for stock returns (Kelly and Jiang, 2014; Chevapatrakul et al., 2019; Hollstein et al., 2019; Andersen et al., 2020). Note that, tail risk is the additional risk which, commonly observed, fat-tailed asset return distributions have relative to normal distributions (Li and Rose, 2009).

We aim to extend this burgeoning literature on the forecastability of stock returns based on information content of tail risks by undertaking a historical perspective for advanced economies namely, Canada, France, Germany, Italy, Japan, Switzerland, the United Kingdom (UK), and the United States (US). In particular, unlike existing studies, which concentrate on post-World War II data, we cover the longest possible data available on the evolution of the tail risks and the corresponding predictability of the equity markets of these eight countries, and in the process, avoid the issue of the sample selection bias in such analyses. Specifically speaking, while our monthly data ends in 2020, it starts at 1915, 1898, 1870, 1905, 1914, 1916, 1693, and 1791 for Canada, France, Germany, Italy, Japan, Switzerland, the UK and the US respectively, and hence covers the entire available historical information of the stock prices of these markets, as well as many extreme events (such as the South Sea Bubble; series of bank panics over 1785 to early 1900; the Spanish Flu; the two World Wars; the "Great Depression"; the oil shocks; Black Monday; the Asian Financial Crisis, besides the one already mentioned above) associated with global crises in history.[1] Note that, besides availability of historical stock market data, the choice of these mature equity markets is primarily motivated by their importance in the global economy, with these countries representing nearly two-third of global net wealth, and nearly half of world output (Das et al., 2019).

[1] The reader is referred to Table A2 of Boubaker et al., (2020) for a detailed summary of global crises starting from the 14th century.

At this stage, it must be pointed out that there are primarily two approaches for computing tail risks; one associated with option implied measures, while the other based on the underlying returns data. Understandably, due to unavailability of such long-spans of historical data on options, we take the second route, whereby we estimate tail risk with the use of Value at Risk (VaR) by employing conditional autoregressive VaR specifications as in Engle and Manganelli (2004). In this regard, the models considered are: (i) the adaptive model; (ii) the symmetric slope model, (iii) the asymmetric slope model, and; (iv) the indirect generalized autoregressive conditional heteroscedasticity (GARCH) model with an autoregressive mean. Then, the specific tail risks model that best-fits the returns data statistically, is used to forecast stock returns based on an out-of-sample forecasting exercise, given that the ultimate test of any predictive model (in terms of the econometric methodologies and the predictors used) is in its forecasting performance (Campbell, 2008). In addition, to control for a global shock in terms of tail risks, we investigate the role of the West Texas Intermediate (WTI) oil tail risks from 1859, besides own tail risks in forecasting equity returns in these eight countries. The inclusion of oil tail risks in the model along with own tail risks make sense, given the large global literature that exists regarding the oil-stock nexus (see Degiannakis et al., (2018), and Smyth and Narayan (2018) for detailed reviews in this regard), besides some recent evidence of tail risks interconnectedness between these two markets (Mensi et al., 2017). Finally, given the US being a dominant driving force for other international equity markets (Rapach et al., 2013; Aye et al., 2017), as well as historical evidence of tail risk spillovers from the US on to the remaining seven markets (Ji et al., 2020), we also analyse the forecasting ability of US tail risks over and above domestic tail and oil risks for Canada, France, Germany, Italy, Japan, Switzerland, the UK.

To the best of our knowledge, this is the first paper to obtain estimates of tail risks of equity markets of eight advanced economies using over centuries of monthly data, and then incorporate its role in forecasting the stock returns of this market, by controlling for oil tail risks as well as tail risks of the US (when it comes to the remaining seven stock returns). In terms of the forecasting analysis, we rely on a returns predictability framework following Westerlund and Narayan (2012, 2015), which allows us to account for persistence, endogeneity and conditional heterocsedasticity effects, which are typical features of most financial series. The remainder of the paper is organized as follows: Section 2 describes the methodology and the data, Section 3 discusses the econometric results associated with the estimation of the tail risks, and also the various forecasting exercises. Finally, Section 4 concludes.

2. Methodology and Data

2.1 Methodology

The methodology essentially hinges on the risk-return hypothesis expressed in standard theories of finance such as Capital Asset Pricing Model (CAPM), the Arbitrage Pricing Theory and Fisher hypothesis, which assume that returns respond to market (systematic) risk and unsystematic risks (see Fama and French, 2004, Kumar, 2016; among others). Our focus lies in the former since it has market-wide effects and by extension can affect a large number of assets unlike the latter which is limited to a few assets. Consequently, we formulate a predictive model for stock return series where the risks associated with the stock market (country-specific) as well as those associated with the global economy such as oil market risks[2], serve as predictors. We follow the approach of Engle and Manganelli (2004) to estimate the stock market and oil market risks technically described as tail risks as the modelling concentrates attention on the asymptotic form of the tail, rather than modeling the whole distribution. This approach involves a conditional autoregressive quantile specification of Value at risk (VaR)[3], which is also termed conditional autoregressive value at risk (CAViaR)[4] and provides an alternative measure of market (systematic) risk used by financial institutions. Rather than modelling the whole distribution, Engle and Manganelli (2004)[5] provide a different approach to quantile estimation of VaR. A generic CAViaR specification is given as:

---

[2] There is a huge body of literature linking stock returns to movements in oil price (see for a recent survey, Smyth and Narayan, 2018, Salisu, Swaray and Oloko, 2019).

[3] Several attractions to the use of Value at risk (VaR) as a standard measure of market risk are well documented in Engle and Manganelli (2004). Chief among these attractions is its conceptual simplicity as it reduces the market risk associated with any portfolio to a single (monetary) amount.

[4] The new approach is designed to overcome the statistical problem inherent in the standard VaR method. Since VaR is simply a particular quantile of future portfolio values, conditional on current information, and because the distribution of portfolio returns typically changes over time, the challenge is to find a suitable model for time-varying conditional quantiles, an issue that is ignored in the standard VaR but incorporated in the CAViaR.

[5] There are other approaches of modelling tail risks (see Boudoukh, Richardson and Whitelaw, 1998; Danielsson and de Vries, 2000), however we favour the one proposed by Engle and Manganelli (2004) given the inherent shortcomings in the previous approaches and the ability of the latter to overcome them. For instance, the approach proposed by Danielsson and de Vries (2000) is not "extreme enough" to capture the tail of the distribution and more importantly, the quantile models are nested in a framework of iid variables, which is not consistent with the characteristics of most financial series, and, consequently, the risk of a portfolio may not vary with the conditioning information set (Engle and Manganelli, 2004).

$$f_t(\beta) = \beta_0 + \sum_{i=1}^{q} \beta_i f_{t-i}(\beta) + \sum_{j=1}^{r} \beta_j l(x_{t-j}) \quad [1]$$

where $f_t(\beta) \equiv f_t(x_{t-1}, \beta_\theta)$ denote the time t θ-quantile of the distribution of portfolio returns formed at $t-1$. Note that θ subscript is supressed from β_θ as in equation [1] for notational convenience. Also, $p = q + r + 1$ is the dimension of β and l is a function of a finite number of lagged values of observables. The autoregressive terms $\beta_i f_{t-i}(\beta)$, $i = 1, \ldots, q$, ensure that the quantile changes "smoothly" over time. The role of $l(x_{t-j})$ is to link $f_t(\beta)$ to observable variables that belong to the information set. We estimate four variants of the tail risks namely Adaptive, Symmetric absolute value, Asymmetric slope and Indirect GARCH and are respectively specified as follows:

Adaptive:

$$f_t(\beta_1) = f_{t-1}(\beta_1) + \beta_1 \left\{ \left[1 + \exp\left(G\left[y_{t-1} - f_{t-1}(\beta_1) \right] \right) \right]^{-1} - \theta \right\} \quad [2]$$

Symmetric absolute value:

$$f_t(\beta) = \beta_1 + \beta_2 f_{t-1}(\beta) + \beta_3 |y_{t-1}| \quad [3]$$

Asymmetric slope:

$$f_t(\beta) = \beta_1 + \beta_2 f_{t-1}(\beta) + \beta_3 (y_{t-1})^+ + \beta_4 (y_{t-1})^- \quad [4]$$

Indirect GARCH (1,1):

$$f_t(\beta) = \left(\beta_1 + \beta_2 f_{t-1}^2(\beta) + \beta_3 y_{t-1}^2 \right)^{1/2} \quad [5]$$

where G is some positive finite number which makes the model a smoothed version of a step function and the last term in equation [2] converges almost surely to $\beta_1 \left[I\left(y_{t-1} \le f_{t-1}(\beta_1) \right) - \theta \right]$ if $G \to \infty$ with $I(\cdot)$ representing the indicator function. Note that equations [3] and [5] are symmetric in nature while equation [4] is asymmetric as the response to positive and negative returns is identical for the former category but differs for the latter. While the adaptive model has a unit coefficient on the lagged VaR, the other three are mean reverting implying that the coefficient on the lagged VaR is not constrained to be 1.

We subject all the return series to CAViaR test under the four alternative specifications as previously mentioned. We follow the procedure of Engle and Manganelli (2004) and thus, we produce results for both 1% and 5% VaRs across the four variants of the CAViaR. Thereafter, we use relevant model diagnostics such as the Dynamic Quantile test (DQ) test and %Hits[6] to determine the model that best fits the data. The results obtained are then used in the return predictability following the Westerlund and Narayan (2012, 2015) methods, which allows us to account for some salient features such as persistence, endogeneity and conditional heterocsedasticity effects typical of most financial series and relies on the following predictive model partitioned into three variants as follows: [7]

Case I: Here, we examine the relative forecast performance of own tail risk in contrast with a driftless random walk. In other words, this is a single predictor model as it only accommodates own tail risk in the return predictability.

$$r_t = \omega + \phi tr_{t-1}^r + \alpha\left(tr_t^r - \rho_o tr_{t-1}^r\right) + \varepsilon_t \qquad [6]$$

Case II: Here, we extend equation (1) to include oil tail risk given the strong connection between oil and stock (see Narayan and Gupta, 2015; Salisu and Isah, 2017; Smyth and Narayan, 2018; Salisu, Swaray and Oloko, 2019, among others). The idea is to see whether including the oil tail risk will improve the return predictability.

$$r_t = \omega + \phi_1 tr_{t-1}^r + \alpha_1\left(tr_t^r - \rho_o tr_{t-1}^r\right) + \phi_2 tr_{t-1}^o + \alpha_2\left(tr_t^o - \rho_o tr_{t-1}^o\right) + \varepsilon_t \qquad [7]$$

Case III: We further extend equation (2) to include US stock tail risk. This motivated by the possible spillover effects of financial contagion from the US to the rest of the world including other advanced countries (see Chen, Mancini-Griffoli and Sahay, 2014; Georgiadis, 2015; and di Giovanni, 2021).

$$\begin{aligned} r_t = {} & \omega + \phi_1 tr_{t-1}^r + \alpha_1\left(tr_t^r - \rho_o tr_{t-1}^r\right) + \phi_2 tr_{t-1}^o + \alpha_2\left(tr_t^o - \rho_o tr_{t-1}^o\right) \\ & + \phi_3 tr_{t-1}^{us} + \alpha_2\left(tr_t^{us} - \rho_o tr_{t-1}^{us}\right) + \varepsilon_t \end{aligned} \qquad [8]$$

.

[6] These are standard test statistics for evaluating the relative performance of the alternative specifications of CAViaR test.

[7] See Westerlund and Narayan (2015) for computational details while several applications are evident in the literature as regards the use of this methodology for stock return predictability (see for example, Bannigidadmath and Narayan, 2015; Narayan and Bannigidadmath, 2015; Narayan and Gupta, 2015; Phan, Sharma, and Narayan, 2015; Devpura, Narayan, and Sharma, 2018; Salisu, Swaray and Oloko, 2019; Salisu, Raheem and Ndako, 2019, among others).

where r_t is the log return of stock price indexes at period t, computed as $100 * \Delta \log(p_t)$, p_t being the stock price index; ω is the intercept; tr is the tail risk obtained as the one that best fits the data. Note that the superscript on the tail risk defines the return series used in calculating it, thus, superscripts " r " and " o " denote the tail risks for the stock return series (i.e. the dependent variable) and oil returns respectively. The oil return series is computed as the log return of the West Texas Intermediate (WTI)[8] crude oil price, and ε_t is the zero mean idiosyncratic error term. With reference to equation (1), we include an additional term - $\alpha\left(tr_t^r - \rho_o tr_{t-1}^r\right)$ in the predictive model in addition to the lagged predictor series - ϕtr_{t-1}^r in order to resolve any inherent endogeneity bias resulting from the correlation between the predictor series and the error term as well as any potential persistence effect.[9] This same procedure is followed for other equations while the technical details justifying this inclusion are well documented in Westerlund and Narayan (2012, 2015). Finally, given the use of monthly frequency over a long range data covering more than a century, the need to account for conditional heteroscedasticity effect becomes crucial and this is done by pre-estimating the predictive model with the conventional GARCH-type model and pre-weighting all the data with the inverse of standard deviation obtained from the latter. The resulting equation is estimated with the Ordinary Least Squares to obtain the feasible quasi generalized least squares estimates.

For the forecast evaluation, we essentially focus on the out-of-sample forecast performance since the literature is replete with studies on in-sample predictability whose outcome cannot be used to generalize for the out-of-sample predictability and more importantly forecast accuracy of return series is better determined with out-of- sample forecasts (see Narayan and Gupta, 2015; Salisu, Swaray and Oloko, 2019). As conventional for time series forecasting of financial series, we use the drifless random walk as the benchmark and its forecast performance is compared with the tail risk-based predictive models. We employ both single (Root Mean Square Forecast Error) and pairwise forecast measure using the Clark and West (2007) while the 75:25 data split, is

[8] WTI according to Narayan and Gupta (2015) is considered a good reflector of movements in global oil prices.
[9] Some preliminary tests are rendered in this regard to establish the presence of these effects and the results can be provided by the authors upon request.

respectively used to split the data into 75% of the full sample for the in-sample estimation and the remaining 25% for the out-of-sample forecast.[10]

2.2 Data sources

The data used in this paper are monthly stock price indices of eight industrialized economies namely, Canada (SandP TSX 300 Composite Index), France (CAC All-Tradable Index), Germany (CDAX Composite Index), Italy (Banca Commerciale Italiana Index), Japan (Nikkei 225 Index), Switzerland (All Share Stock Index), the UK (FTSE All Share Index), and the US (SandP500 Index). The oil price data corresponds to that of the WTI. The data on the indices and oil price are derived from Global Financial Data.[11] The stock price indices and oil price are converted into log returns in percentage, i.e., the first-difference of the natural logarithm of the indices multiplied by 100. Table 1 below summarizes the sample period of the data set, with UK having the longest coverage starting at 1693, and Switzerland the shortest span, with the stock index beginning at 1916. But, overall, all stock indices and the oil price covers more than a century of monthly data.

Table 1: Data scope for stock returns for the selected countries and oil price returns[12]

	Start Period	End Period	Nob
Canada	1915:M02	2020:M10	1269
France	1898:M01	2020:M10	1474
Germany	1870:M01	2020:M10	1810
Italy	1905:M02	2020:M10	1389
Japan	1914:M08	2020:M10	1275
Switzerland	1916:M02	2020:M10	1257
UK	1693:M02	2020:M10	3933
US	1791:M09	2020:M10	2750
Oil price	1859:M10	2020:M10	1933

Note: Nob is the number of observations.

We present some descriptive statistics for all the return series s in Table 2. While the German stock returns seem to be the most volatile among the selected advanced countries and has

[10] Note that there is no theoretical guidance in the literature for data splitting in forecast analysis, however, studies have adopted 25:75, 50:50 and 75:25 respectively between the in-sample and out-of-sample forecasts (see Narayan and Gupta, 2015) and the outcome is observed to be insensitive to the choice of data split (see Narayan and Gupta, 2015; Salisu et al., 2019b).

[11] https://globalfinancialdata.com/.

[12] Note that we utilize the entire data sample for own tail risk analysis while the period is adjusted appropriately for multi-predictor models.

the most peaked distribution, we find all the return series to be heavy tailed given the leptokurtic nature of the kurtosis statistics and are also negatively skewed with the exception of Italy and Japan. It is therefore not surprising why they are non-normal judging by the Jarque-Bera test and therefore limiting the measurement of the market to the distribution of the tail rather than the whole distribution is justified. Nonetheless, we offer additional empirical support in the next section.

Table 2: Summary statistics for the stock return series

	Mean	Std. Dev.	Skewness	Kurtosis	JB test	Nobs
Canada	0.3938	4.5149	-1.1477	9.3824	2432.453***	1269
France	0.5305	5.1299	-0.1983	5.3890	360.1903***	1474
Germany	0.2261	7.0812	-4.7270	115.9331	968595.4***	1810
Italy	0.4116	6.7489	0.9140	9.5998	2714.240***	1389
Japan	0.5351	6.0914	0.3278	10.1517	2739.980***	1275
Switzerland	0.2960	4.3251	-0.6021	8.3236	1560.383***	1257
UK	0.1181	3.9901	-0.5552	56.5928	470881.3***	3933
US	0.2504	3.8523	-0.6695	15.0064	16722.92***	2750

Note: Nobs = Number of observations; *** denotes significance at 1% level.

3. Empirical results

3.1 Determining the tail risk that "best" fits the data

To generate the tail risk data for relevant return series, we estimate 1% and 5% VaRs, using the four CAViaR specifications described in the preceding section and several diagnostics are computed to enable us determine the "best" CAViaR specification for each return series. The analyses are rendered for both in-sample and out-of-sample periods and the diagnostics for performance evaluation are obtained accordingly. Since our study essentially focuses on the out-of-sample predictability of the tail risks, we limit our discussion of results to this period. The results are presented in Tables 3, 4, 5 and 6 respectively for 1% VaR of Symmetric Absolute Value, Asymmetric Slope, Indirect GARCH and Adaptive specifications while Tables 7, 8, 9, and 10 are 5% VaR respectively presented in the same order as 1% VaR.[13] The criteria used are the DQ test and %Hits for the out-of-sample and for the "best" tail risk variant, we expect the %Hits to be relatively 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In

[13] The computational and theoretical procedures for the implementation of the four variants of the CAViaR test are well presented in the Engle and and Manganelli (2004). We are also grateful to these authors for providing useful Matlab odes for CAViaR estimation.

the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance, the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter. For easy reference, the "best" choice for each stock return series is put in bold and we find that the choice of VaR matters as the performance seems to differ between 1% VaR and 5% VaR. Thus, our out-of-sample forecast analysis is carried out for both VaRs in order to further test whether same conclusion would be obtained for the tail risks predictability. Nonetheless, all the return series exhibit volatility clustering as measured by the statistically significant coefficient (Beta2) on the autoregressive term. This outcome further confirms that the phenomenon of clustering of volatilities is relevant also in the tails (see also Engle and Manganelli, 2004) and thus, pre-weighting the data with the inverse of standard deviation obtained from the conventional GARCH-type model is justified. Some graphical illustrations are provided in Figures 1 and 2 for the 1% and 5% VaRs respectively and some relative co-movements can be teased out between the tail risks and the return series, in addition to the volatile nature of the series.

3.2 Out-of-Sample forecast evaluation of tail risks

This section presents and discusses the forecast performance results. After generating our tail risk data for the relevant return series and determining the "best" CAViaR specification for each series, we extend our analysis to include an out-of-sample forecast evaluation. To validate our analysis, we perform a 3-layer robustness test. First, we evaluate our sample using different variants of the predictive model namely; one-predictor model (own tail risk), two-predictor model (own tail risk and oil tail risk) and three-predictor model (own tail risk, oil tail risk and US tail risk). Second, we analyse the sample at 1% and 5% VaRs and lastly, we spread our out-of-sample evaluation across multiple horizons (h = 6, 12 and 24 months), employing the RMSE and Clark and West forecast methods. We find that in all the countries, barring a few (UK and Japan), our proposed models outperform the benchmark model (driftless randomwalk model). Importantly, of the three proposed models, the two-predictor model performs best on the average particularly when the 5% VaR tail risks are utilized as over 70% of all the cases considered show superior performance over the benchmark model judging by the Clark and West (2007) test (see Tables 11, 13, 15 for one-predictor, two-predictor and three-predictor models, respectively). There are two implications of this finding. One, the inclusion of oil tail risk in the own-tail risk-predictive model improves the forecast accuracy over and above the benchmark model. This further validates the inclusion of both domestic risk (stock tail risk) and global risk (oil tail risk) in the predictive model of stock returns for improved forecast accuracy of the model. Our findings are further corroborated by the root mean square forecast error (rmsfe) statistics for the two predictor which are consistently smaller than those of the one-predictor and three-predictor models (see Tables 12, 14, 16, for one-predictor, two-predictor and three-predictor models, respectively). Overall, this conclusion echoes the findings of earlier works like Wang et al., (2018), Salisu et al. (2019) and Alqahtani et al. (2020) which also confirm improved out-of-sample predictability of oil for stock returns. Two, we find that the inclusion of the US stock tail risk does not seem to improve the forecast performance of the two-predictor model of the other advanced countries. In addition, our result shows that the model at 5% VaR represents a significant improvement over 1% VaR for both forecast estimates (CW and RMSE). Also, the forecasting prowess is mixed between short and long forecast horizons, while it improves over long forecast horizons for some countries (such as Canada, France, Italy and US for a one-predictor 1% tail risk model), it does only at short forecast horizons for some others (such as Germany, Switzerland and UK). Regardless, the overarching evidence shows that the benchmark model is least preferred to the proposed tail-risk models.

4. Conclusion

This study examines the out-of-sample predictability of market risks measured as tail risks for stock returns of eight advanced countries using a long-range monthly data of over a century. The methodology essentially hinges on the risk-return hypothesis expressed in standard theories of finance such as Capital Asset Pricing Model (CAPM), the Arbitrage Pricing Theory and Fisher hypothesis, which assume that returns respond to market (systematic) risk and unsystematic risks. Our focus lies in the former since it has market-wide effects and by extension can affect a large number of assets unlike the latter which is limited to a few assets. We follow the Conditional Autoregressive Value at Risk (CAViaR) of Engle and Manganelli (2004) to measure the tail risks since it utilizes the tail distribution rather the whole distribution. Consequently, we produce results for both 1% and 5% VaRs across four variants (Adaptive, Symmetric absolute value, Asymmetric slope and Indirect GARCH) of the CAViaR. Thereafter, we use relevant model diagnostics such as the Dynamic Quantile test (DQ) test and %Hits to determine the model that best fits the data. The results obtained are then used in the return predictability following the Westerlund and Narayan (2012, 2015) method which allows us to account for some salient features such as persistence, endogeneity and conditional heteroscedasticity effects. For the forecast evaluation, we essentially focus on the out-of-sample forecast performance since the literature is replete with studies on in-sample predictability whose outcome cannot be used to generalize for the out-of-sample predictability and more importantly forecast accuracy of return series is better determined with out-of- sample forecasts. As conventional for time series forecasting of financial series, we use the driftless random walk as the benchmark and its forecast performance is compared with the tail risk-based predictive models. We consequently partition our models into three variants (one-predictor, two-predictor and three-predictor models) and examine their forecast performance in contrast with the driftless random walk model. Three findings are discernible from the empirical analysis. First, we find that the choice of VaR matters when determining the "best" fit CAViaR model for each return series as the outcome seems to differ between 1% and 5% VaRs. Second, the predictive model that incorporates both stock tail risk and oil tail risk produces better forecast outcomes than the one with own tail risk indicating the significance of both domestic and global risks in the return predictability of advanced countries. Understandably, our results highlight that investors need to account for tail risks in their portfolio decisions, which are based on accurate forecasts of stock returns. Also from the perspective of academicians, our results suggest that stock markets are at least weakly inefficient, and the role of local and global tail risks must be incorporated into asset pricing models. Finally, with stock market movements being a predictor of the real economy (Stock and Watson, 2003), policy authorities would need to closely monitor tail risks in the equity markets to get an understanding of the future movements in output and inflation, and accordingly design policy responses.

An extension of this study, based on data availability, it would be interesting to investigate the role of historical tail risks in predicting emerging equity markets, and possibly also other asset and commodity markets.

References

Alqahtani, A., Bouri, E., and Vo, X. V. (2020). Predictability of GCC stock returns: The role of geopolitical risk and crude oil returns. *Economic Analysis and Policy*, *68*, 239-249.

Andersen, T.G., Fusari, N., Todorov, V. (2020). The pricing of tail risk and the equity premium: Evidence from international option markets. Journal of Business & Economic Statistics, 38(3), 662-678.

Aye G.C., Balcilar, M., and Gupta, R. (2017). International stock return predictability: Is the role of US time-varying? Empirica, 44(1), 121–146.

Baker, S., Bloom, N., Davis, S. J., Kost, K., Sammon, M., and Viratyosin, T. (2020). The unprecedented stock market reaction to COVID-19. The Review of Asset Pricing Studies, 10(4), 742-758.

Bannigidadmath, D., and P. Narayan. 2015. Stock return predictability and determinants of predictability and profits. Emerging Markets Review 26:153–73.

Bollerslev, T., and Todorov, V. (2011). Tails, fears, and risk premia. Journal of Finance, 66(6), 2165–2211.

Bollerslev, T., Todorov, V., and Xu, L. (2015). Tail risk premia and return predictability. Journal of Financial Economics, 118(1), 113–134.

Boubaker, H., Cunado, J., Gil-Alana, L.A., and Gupta, R. (2020). Global crises and gold as a safe haven: Evidence from over seven and a half centuries of data. Physica A: Statistical Mechanics and its Applications, 540, 123093.

Boudoukh, J., Richardson, M., and Whitelaw, R. F. (1998). The Best of Both Worlds. Risk, 11, 64-67.

Chen, J., Mancini-Griffoli, T. and Sahay, R. (2014). Spillovers from United States Monetary Policy on Emerging Markets: Different This Time? IMF Working Papers 14/240.

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

Campbell, J.Y. (2008). Viewpoint: Estimating the equity premium. Canadian Journal of Economics, 41(1), 1-21.

Chevapatrakul, T., Xu, Z., and Yao, K. (2019). The impact of tail risk on stock market returns: The role of market sentiment. International Review of Economics & Finance, 59(C), 289-301.

Christou, C., and Gupta, R. (2020). Forecasting equity premium in a panel of OECD countries: The role of economic policy uncertainty. The Quarterly Review of Economics and Finance, 76(C), 243-248.

Christou, C., Gupta, R., and Jawadi, F. (Forthcoming). Does Inequality Help in Forecasting Equity Premium in a Panel of G7 Countries? The North American Journal of Economics and Finance.

Cremers, M., Halling, M., and Weinbaum, D. (2015). Aggregate jump and volatility risk in the cross-section of stock returns. Journal of Finance, 70 (2), 577–614.

Danielsson, J., and de Vries, C. G. (2000). Value-at-Riska nd Extreme Returns. Annales d'Economiee t de Statistique, 60, 239-270.

Das, S., Demirer, R., Gupta, R., and Mangisa, S. (2019). The effect of global crises on stock market correlations: Evidence from scalar regressions via functional data analysis. Structural Change and Economic Dynamics, 50(C), 132-147.

Degiannakis, S.A., Filis, G., and Arora, V. (2018). Oil Prices and Stock Markets: A Review of the Theory and Empirical Evidence. The Energy Journal, 39(5), https://doi.org/10.5547/01956574.39.5.sdeg.
Devpura, N., P. K. Narayan, and S. S. Sharma. 2018. Is stock return predictability time varying? Journal of International Financial Markets, Institutions and Money 52:152–72.
Engle, R.F. and Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. Journal of Business and Economic Statistics, 22(4), 367-381.
Fama, E.F. and French, K.R. (2004). The CAPM: Theory and Evidence', Journal of Economic Perspectives, Vol 18, No. 3, 2004.
Gupta, R., Huner, F., and Piribauer, P. (2020). Predicting international equity returns: Evidence from time-varying parameter vector autoregressive models. International Review of Financial Analysis, 68, 101456.
Gupta, R., Majumdar, A., and Wohar, M.E. (2017). The Role of Current Account Balance in Forecasting the US Equity Premium: Evidence From a Quantile Predictive Regression Approach. Open Economies Review, 28(1), 47-59.
Hollstein, F., Nguyen, D.B.B., Prokopczuk, M., and Wese Simen, C. (2019). International Tail Risk and World Fear. Journal of International Money and Finance, 93, 244-259.
Ji, Q., Liu, B-Y., Cunado, J., and Gupta, R. (2020). Risk spillover between the US and the remaining G7 stock markets using time-varying copulas with Markov switching : evidence from over a century of data. The North American Journal of Economics and Finance, 51, 100846.
Jordan S.J., Vivian, A., and Wohar, M.E. (2017). Forecasting market returns: bagging or combining? International Journal of Forecasting, 33(1), 102–120.
Jordan S.J., Vivian, A., and Wohar, M.E. (2018). Stock returns forecasting with metals: sentiment vs. fundamentals. The European Journal of Finance, 24(6), 458-477.
Kelly, B., and Jiang, H. (2014). Tail risk and asset prices. Review of Financial Studies, 27(10), 2841–2871.
Kumar, R. (2016). Risk and return. In Kumar, Valuation Theories and Concepts, Academic Press, London, 47-72. doi:10.1016/b978-0-12-802303-7.00002-4.
Li, X-M., and Rose, L.C. (2009). The tail risk of emerging stock markets. Emerging Markets Review, 10, 242–256.
Mensi, W., Hammoudeh, S., Shahzad, S.J.H., and Shahbaz, M. (2017). Modeling systemic risk and dependence structure between oil and stock markets using a variational mode decomposition-based copula method. Journal of Banking & Finance, 75, 258-279.
Narayan, P. K., and D. Bannigidadmath. 2015. Are Indian stock returns predictable? Journal of Banking and Finance 58:506–31.
Narayan, P. K., and R. Gupta. 2015. Has oil price predicted stock returns for over a century? Energy Economics 48:18–23.
Phan, D. H. B., S. S. Sharma, and P. K. Narayan. 2015. Stock return forecasting: Some new evidence. International Review of Financial Analysis 40:38–51.
Rapach, D.E., Strauss, J.K., and Zhou, G. (2013). International Stock Return Predictability: What Is the Role of the United States? The Journal of Finance, 68(4), 1633-1662.
Rapach D.E., Wohar, M.E., and Rangvid, J. (2005). Macro variables and international stock return predictability. International Journal of Forecasting 21(1), 137–166

Rapach, D.E., and Zhou, G. (2013). Forecasting stock returns. Handbook of Economic Forecasting, 2 (Part A), Graham Elliott and Allan Timmermann (Eds.), Amsterdam: Elsevier, 328-383.

Salisu, A. A., I. D. Raheem, and U. B. Ndako. 2019. A sectoral analysis of asymmetric nexus between oil price and stock returns. International Review of Economics and Finance, 61 (C):241–59.

Salisu, A. A., R. Swaray, and T. F. Oloko. 2019. Improving the predictability of the oil–US stock nexus: The role of macroeconomic variables. Economic Modelling 76:153–71.

Salisu, A. A., and K. Isah. 2017. Revisiting the oil price and stock market nexus: A nonlinear Panel ARDL approach. Economic Modelling, 66 (C), 258-71.

Smyth, R., and Narayan, P.K. (2018). What do we know about oil prices and stock returns? International Review of Financial Analysis, 57, 148-156.

Stock J.H. and Watson M.W. (2003). Forecasting output and inflation: The role of asset prices. Journal of Economic Literature, 41(3), 788–829

Wang, Y., Wei, Y., Wu, C., and Yin, L. (2018). Oil and the short-term predictability of stock return volatility. *Journal of Empirical Finance*, *47*, 90-104.

Westerlund, J., and P. K. Narayan. 2012. Does the choice of estimator matter when forecasting returns? Journal of Banking and Finance 36:2632–40.

Westerlund, J., and P. K. Narayan. 2015. Testing for predictability in conditionally heteroskedastic stock returns. Journal of Finance and Economics 13 (2):342–75.

Table 3: Estimates and Relevant statistics for the country-specific CAViaR specification [Symmetric Absolute Value]

1% VaR	Canada	France	Germany	Italy	Japan	Switzerland	**UK**	**US**	Oil
Beta1	0.1620	0.8556	0.9621	0.3645	3.6076	1.7658	**0.4993**	**0.2037**	0.6428
Standard errors	0.2637	0.6527	0.5243	0.1861	2.5056	1.5846	**0.2047**	**0.1275**	0.2149
P values	0.2695	0.0950	0.0332	0.0251	0.0750	0.1326	**0.0074**	**0.0551**	0.0014
Beta2	0.9274	0.8172	0.7306	0.9141	0.6196	0.7683	**0.7913**	**0.8495**	0.8364
Standard errors	0.0467	0.1060	0.0701	0.0489	0.2238	0.1873	**0.0811**	**0.0228**	0.0258
P values	0.0000	0.0000	0.0000	0.0000	0.0028	0.0000	**0.0000**	**0.0000**	0.0000
Beta3	0.2730	0.3299	0.8816	0.1801	0.5378	0.4571	**0.8648**	**0.6528**	0.5609
Standard errors	0.1146	0.1411	0.1124	0.1091	0.2452	0.3762	**0.2511**	**0.1362**	0.0939
P values	0.0086	0.0097	0.0000	0.0494	0.0141	0.1122	**0.0003**	**0.0000**	0.0000
RQ	118.8014	132.6606	417.8987	144.1425	141.5537	114.9148	**375.1110**	**317.4198**	392.2562
Hits in-sample (%)	0.9103	1.0267	0.9615	1.1249	1.0323	1.0568	**1.0195**	**1.0667**	1.0695
Hits out-of-sample (%)	1.0000	2.6000	2.0000	2.6000	1.4000	0.8000	**0.8000**	**1.2000**	2.0000
DQ in-sample (P-values)	0.0502	0.9848	0.2546	0.9795	0.0300	0.0288	**0.7421**	**0.0061**	0.0002
DQ out-of-sample (P-values)	0.0024	0.0012	0.5698	0.0000	0.0169	0.9995	**0.9992**	**0.9923**	0.0018

Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases where more than one tail risk is not significant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Table 4: Estimates and Relevant statistics for the country-specific CAViaR specification [Asymmetric Slope]

1% VaR	Canada	France	Germany	Italy	Japan	Switzerland	UK	US	Oil
Beta1	**0.0613**	0.6385	1.6711	0.2042	**2.6652**	**2.4000**	0.0939	0.3725	**1.3322**
Standard errors	**0.3395**	0.4974	0.7044	0.1053	**2.2472**	**1.6124**	0.1514	0.1765	**0.3324**
P values	**0.4113**	0.0996	0.0088	0.0262	**0.1178**	**0.0683**	0.2676	0.0174	**0.0000**
Beta2	**0.9282**	0.8381	0.7393	0.8990	**0.6933**	**0.7108**	0.8055	0.8100	**0.7977**
Standard errors	**0.0495**	0.0792	0.0748	0.0386	**0.2071**	**0.1377**	0.0979	0.0222	**0.0254**
P values	**0.0000**	0.0000	0.0000	0.0000	**0.0004**	**0.0000**	0.0000	0.0000	**0.0000**
Beta3	**0.1550**	0.2733	0.2786	0.1738	**0.3989**	**0.0724**	0.6969	0.3388	**0.2415**
Standard errors	**0.0736**	0.0940	0.1182	0.0575	**0.2906**	**0.2543**	0.7662	0.1087	**0.0675**
P values	**0.0176**	0.0018	0.0092	0.0012	**0.0849**	**0.3879**	0.1815	0.0009	**0.0002**
Beta4	**0.4200**	0.3482	0.8195	0.3254	**0.4819**	**0.7125**	0.8547	0.9072	**0.7180**
Standard errors	**0.3136**	0.1595	0.1332	0.2017	**0.1730**	**0.3832**	0.2354	0.0996	**0.1603**
P values	**0.0902**	0.0145	0.0000	0.0533	**0.0027**	**0.0315**	0.0001	0.0000	**0.0000**
RQ	**117.6338**	131.5496	406.7026	142.4098	**141.7739**	**101.6143**	375.1387	303.0804	**357.0882**
Hits in-sample (%)	**1.3004**	1.0267	0.9615	1.0124	**1.0323**	**1.0568**	1.0195	1.0222	**1.0101**
Hits out-of-sample (%)	**1.2000**	2.4000	1.6000	2.0000	**1.4000**	**1.0000**	0.8000	1.6000	**1.6000**
DQ in-sample (P-values)	**0.1767**	0.9819	0.2488	0.9876	**0.0296**	**0.0327**	0.7424	0.4206	**0.1135**
DQ out-of-sample (P-values)	**0.0127**	0.0071	0.9637	0.0185	**0.0230**	**0.9371**	0.9987	0.8295	**0.9625**

Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Table 5: Estimates and Relevant statistics for the country-specific CAViaR specification [Indirect GARCH]

1% VaR	Canada	France	Germany	Italy	Japan	Switzerland	UK	US	Oil
Beta1	-0.2365	17.4096	2.9380	4.2515	112.2847	8.3057	1.4881	1.2126	1.0007
Standard errors	2.0906	11.0229	3.1199	2.4909	33.3347	6.9602	0.8930	0.8283	0.7132
P values	0.4550	0.0571	0.1732	0.0439	0.0004	0.1164	0.0478	0.0716	0.0803
Beta2	0.9308	0.7416	0.7379	0.9168	0.2032	0.8319	0.7840	0.7483	0.7519
Standard errors	0.0287	0.0858	0.0564	0.0157	0.1493	0.0528	0.0350	0.0175	0.0102
P values	0.0000	0.0000	0.0000	0.0000	0.0867	0.0000	0.0000	0.0000	0.0000
Beta3	0.6124	0.5325	2.1269	0.2596	2.0425	0.9473	1.8237	2.3942	1.7883
Standard errors	0.4759	0.3778	0.2801	0.5043	1.2260	2.1074	0.3515	0.4155	0.4637
P values	0.0991	0.0793	0.0000	0.3034	0.0479	0.3265	0.0000	0.0000	0.0001
RQ	119.1051	135.2723	417.5263	145.7389	139.4030	110.8515	372.1061	303.6876	379.3325
Hits in-sample (%)	0.9103	0.9240	1.0256	0.8999	0.9032	1.1889	0.9321	1.0222	1.2478
Hits out-of-sample (%)	1.4000	2.4000	1.6000	2.6000	1.4000	1.2000	1.4000	1.6000	2.8000
DQ in-sample (P-values)	0.9953	0.9943	0.3029	0.9946	0.9964	0.4324	0.6448	0.9239	0.2793
DQ out-of-sample (P-values)	0.0180	0.0071	0.9350	0.0000	0.0068	0.9965	0.8474	0.8273	0.0475

Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Table 6: Estimates and Relevant statistics for the country-specific CAViaR specification [Adaptive]

1% VaR	Canada	France	Germany	Italy	Japan	Switzerland	UK	US	Oil
Beta1	0.1832	**1.0776**	**0.8728**	**0.9274**	-1.4530	0.6622	3.5030	0.9539	20.1375
Standard errors	0.6032	**0.3763**	**1.4828**	**0.3098**	0.9286	0.8750	0.0002	0.1687	0.0000
P values	0.0249	**0.0021**	**0.2781**	**0.0014**	0.0588	0.2246	0.0000	0.0000	0.0000
RQ	150.7064	**150.6373**	**605.8372**	**189.9467**	175.3511	126.0767	497.0929	365.2065	427.7278
Hits in-sample (%)	0.5202	**1.2320**	**1.5385**	**1.3498**	0.9032	0.5284	0.9613	1.2000	0.6536
Hits out-of-sample (%)	1.2000	**1.6000**	**1.2000**	**1.2000**	0.4000	0.8000	0.8000	1.0000	2.0000
DQ in-sample (P-values)	1.0000	**0.1423**	**0.0000**	**0.9352**	0.0096	1.0000	0.0186	0.0000	0.0034
DQ out-of-sample (P-values)	0.0001	**0.6125**	**0.1622**	**0.0001**	0.9114	0.9028	0.9597	0.0001	0.0024

Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Table 7: Estimates and Relevant statistics for the country-specific CAViaR specification [Symmetric Absolute Value]

5% VaR	Canada	France	Germany	**Italy**	Japan	Switzerland	**UK**	US	Oil
Beta1	0.0453	0.3787	0.6884	**0.3717**	1.2006	0.3173	**0.1822**	0.3888	0.0894
Standard errors	0.1175	0.3324	0.4306	**0.2214**	0.4279	0.3273	**0.0487**	0.0781	0.0331
P values	0.3498	0.1273	0.0549	**0.0466**	0.0025	0.1662	**0.0001**	0.0000	0.0035
Beta2	0.9114	0.8062	0.6761	**0.8829**	0.6997	0.8459	**0.8724**	0.8365	0.8668
Standard errors	0.0494	0.0866	0.2714	**0.0511**	0.0977	0.0861	**0.0346**	0.0590	0.0386
P values	0.0000	0.0000	0.0064	**0.0000**	0.0000	0.0000	**0.0000**	0.0000	0.0000
Beta3	0.2385	0.3097	0.5322	**0.1740**	0.3234	0.2904	**0.2457**	0.2587	0.2934
Standard errors	0.1028	0.1023	0.5248	**0.0655**	0.1153	0.1096	**0.0872**	0.1465	0.0882
P values	0.0102	0.0012	0.1553	**0.0039**	0.0025	0.0040	**0.0024**	0.0387	0.0004
	410.010	471.774	1037.223	**516.298**	491.675		**1200.546**	1000.243	1311.754
RQ	3	1	3	**5**	9	388.5899	**4**	1	6
Hits in-sample (%)	4.9415	4.9281	5.0000	**4.9494**	5.0323	4.8877	**5.0393**	5.0667	5.0505
Hits out-of-sample (%)	3.6000	6.8000	8.0000	**5.4000**	8.4000	4.8000	**5.2000**	4.2000	9.2000
DQ in-sample (P-values)	0.0095	0.5626	0.0055	**0.1644**	0.1546	0.0912	**0.0432**	0.0000	0.0000
DQ out-of-sample (P-values)	0.6765	0.0006	0.0214	**0.3827**	0.0009	0.0635	**0.2409**	0.8609	0.0000

Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Table 8: Estimates and Relevant statistics for the country-specific CAViaR specification [Asymmetric Slope]

5% VaR	Canada	France	Germany	Italy	Japan	Switzerland	UK	US	Oil
Beta1	0.1807	0.3895	0.3995	0.2356	1.9007	**0.4018**	0.0863	**0.4731**	0.2309
Standard errors	0.3017	0.3436	0.1493	0.1931	0.9697	**0.3465**	0.0661	**0.0936**	0.0944
P values	0.2746	0.1285	0.0037	0.1111	0.0250	**0.1231**	0.0957	**0.0000**	0.0072
Beta2	0.8721	0.7842	0.7035	0.8784	0.4325	**0.8454**	0.8571	**0.7802**	0.7416
Standard errors	0.0560	0.0832	0.0618	0.0593	0.1718	**0.0701**	0.0417	**0.0344**	0.0377
P values	0.0000	0.0000	0.0000	0.0000	0.0059	**0.0000**	0.0000	**0.0000**	0.0000
Beta3	0.1026	0.3051	0.2794	0.1750	0.4224	**0.0265**	0.1984	**0.1227**	0.2059
Standard errors	0.1335	0.1157	0.0694	0.0670	0.0839	**0.0790**	0.1128	**0.0489**	0.0687
P values	0.2211	0.0042	0.0000	0.0045	0.0000	**0.3687**	0.0393	**0.0061**	0.0014
Beta4	0.3935	0.3346	0.6566	0.2043	0.9437	**0.4261**	0.3370	**0.4843**	0.7376
Standard errors	0.1482	0.1124	0.0894	0.1051	0.2473	**0.1558**	0.0932	**0.1251**	0.0892
P values	0.0040	0.0015	0.0000	0.0260	0.0001	**0.0031**	0.0001	**0.0001**	0.0000
RQ	403.8464	468.9917	1004.6588	513.3517	486.8819	**370.0134**	1191.0169	**959.5576**	1235.0178
Hits in-sample (%)	5.0715	5.0308	4.9359	4.9494	5.0323	**5.0198**	5.0102	**5.0222**	5.0505
Hits out-of-sample (%)	4.6000	6.8000	8.8000	5.6000	7.8000	**5.6000**	6.2000	**5.2000**	10.4000
DQ in-sample (P-values)	0.1984	0.6230	0.8672	0.2932	0.9611	**0.0967**	0.6598	**0.9228**	0.0001
DQ out-of-sample (P-values)	0.5878	0.0008	0.0445	0.3932	0.0036	**0.2418**	0.3818	**0.7885**	0.0011

Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Table 9: Estimates and Relevant statistics for the country-specific CAViaR specification [Indirect GARCH]

5% VaR	Canada	France	Germany	Italy	Japan	Switzerland	UK	US	Oil
Beta1	**-0.3143**	1.6207	3.4322	2.4524	5.1853	0.4625	0.6392	1.6810	0.0000
Standard errors	**0.5064**	1.2599	4.5867	1.6667	3.1135	1.4904	0.2347	0.5997	0.0000
P values	**0.2674**	0.0992	0.2271	0.0706	0.0479	0.3782	0.0032	0.0025	0.4858
Beta2	**0.9266**	0.7727	0.5370	0.8996	0.7923	0.8266	0.7853	0.7884	0.7529
Standard errors	**0.0243**	0.0370	0.2170	0.0275	0.0638	0.0492	0.0069	0.0239	0.0021
P values	**0.0000**	0.0000	0.0067	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Beta3	**0.2393**	0.4523	1.0224	0.1500	0.2606	0.5360	0.4834	0.4934	0.7438
Standard errors	**0.1241**	0.2759	0.7112	0.3218	0.1304	0.4306	0.3895	0.1401	0.0061
P values	**0.0269**	0.0506	0.0753	0.3206	0.0228	0.1066	0.1073	0.0002	0.0000
	412.560	473.753	1041.187	529.068	492.329		1211.760	995.037	1288.550
RQ	**5**	5	4	5	6	383.6714	9	0	9
Hits in-sample (%)	**5.0715**	5.1335	5.1282	5.0619	5.1613	5.0198	5.0685	5.1111	4.6940
Hits out-of-sample (%)	**4.8000**	7.6000	8.8000	5.8000	8.8000	5.0000	7.2000	4.4000	8.0000
DQ in-sample (P-values)	**0.0204**	0.9789	0.1830	0.0879	0.1566	0.4196	0.0071	0.0201	0.0000
DQ out-of-sample (P-values)	**0.4226**	0.0006	0.0011	0.5298	0.0000	0.0639	0.2068	0.9112	0.0000

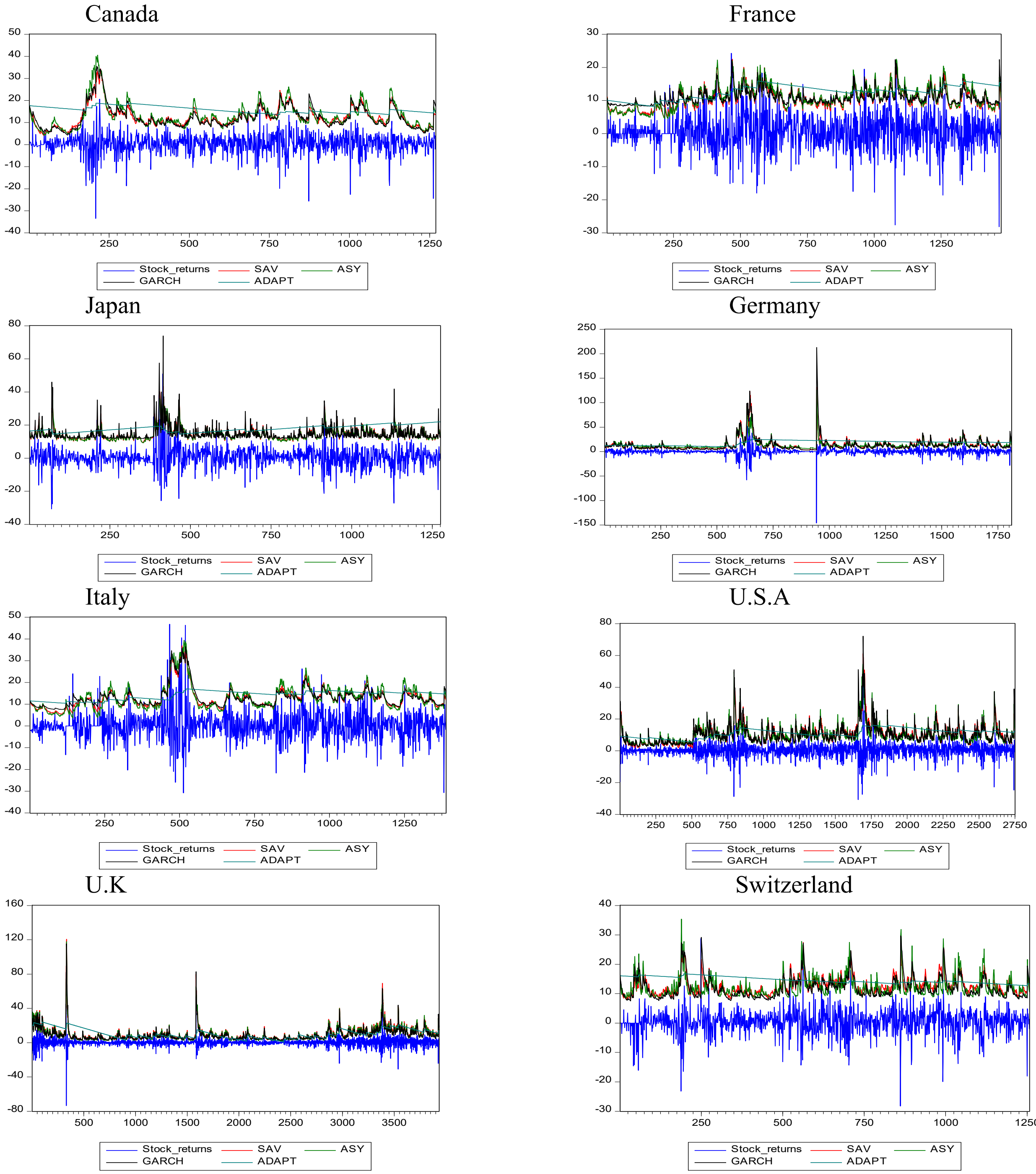
Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Table 10: Estimates and Relevant statistics for the country-specific CAViaR specification [Adaptive]

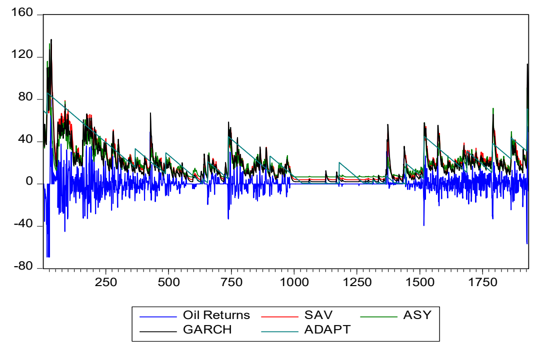
5% VaR	Canada	France	Germany	Italy	Japan	Switzerland	UK	US	Oil
Beta1	1.4507	**0.5713**	**1.3392**	1.8435	**0.9455**	0.5372	1.0434	1.2535	**4.0488**
Standard errors	0.0500	**0.1985**	**0.1230**	0.0796	**0.3706**	0.2009	0.0839	0.0178	**0.0000**
P values	0.0000	**0.0020**	**0.0000**	0.0000	**0.0054**	0.0038	0.0000	0.0000	**0.0000**
RQ	434.364	**495.032**	**1201.549**	573.767	**556.783**	416.570	1351.181	1030.172	**1363.500**
Hits in-sample (%)	4.9415	**5.5441**	**5.1282**	5.0619	**5.0323**	5.1519	4.9519	5.1556	**3.9810**
Hits out-of-sample (%)	4.6000	**5.2000**	**4.4000**	4.8000	**5.2000**	4.6000	4.0000	4.8000	**6.0000**
DQ in-sample (P-values)	0.0024	**0.0578**	**0.0000**	0.7744	**0.0000**	0.0000	0.0000	0.0008	**0.0000**
DQ out-of-sample (P-values)	0.0907	**0.0353**	**0.3211**	0.0016	**0.5897**	0.0045	0.0139	0.1712	**0.0000**

Note: significant at 5% formatted in bold. Yellow highlight indicates a country's tail risk that best "fits" the return series. The criteria used are the DQ test and %Hits for Out-of-Sample. For the "best" tail risk variant, we expect the %Hits to be 1% for 1% VaR and 5% for 5% VaR while DQ test statistic is not expected to be significant. In cases, where more than one tail risk is statistically insignificant in terms of the DQ test, then, we consider the tail risk with the closest value to the expected value for the % Hits. In the same vein, where all the tail risks are statistically significant, then, the %Hits becomes a major criterion except where some distinctions can still be made with the significant DQ test statistics. For instance the DQ test for Canada is significant for all the variants under the 1% VaR, however, both Asymmetric slope and Indirect GARCH survive the DQ test at the 1 percent level and comparing the %Hits, the former gives a better fit than the latter.

Fig. 1: Co-movement between stock returns (inclusive of oil returns) and tail risk of 1 percent VaR in selected countries

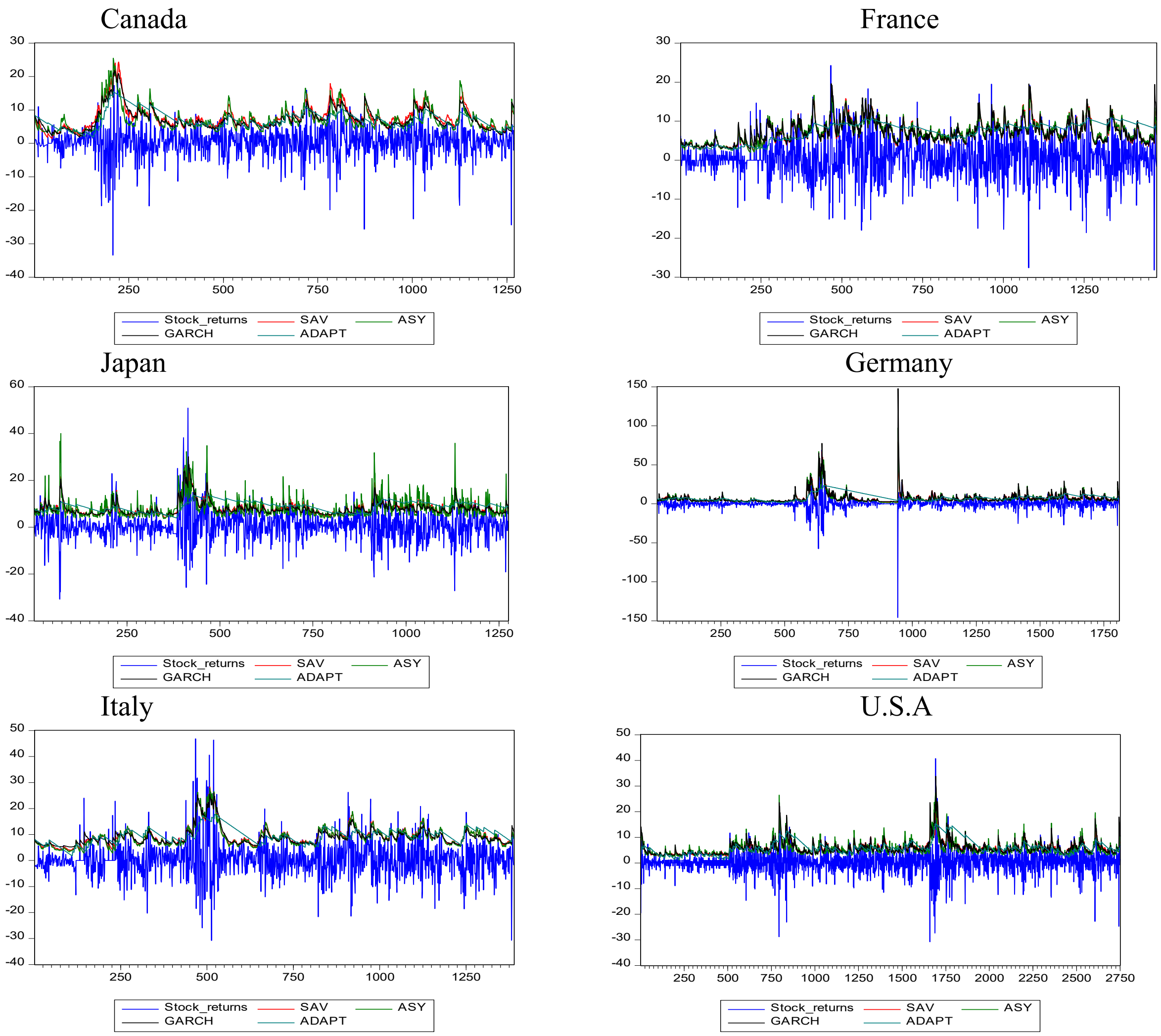


Oil

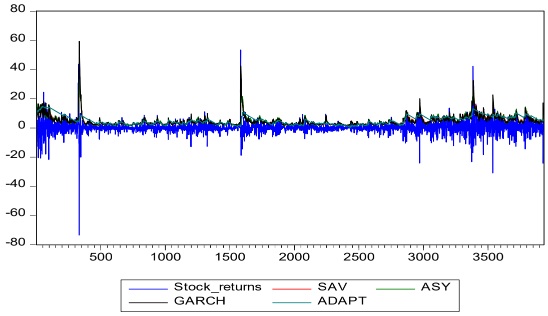


Note: SAV is Symmetric Absolute Value, ASY is Asymmetric Slope, GARCH is Indirect GARCH, ADAPT is Adaptive. While the returns series (stock and oil) are computed as 100*$\Delta \log(p_t)$ and p_t is the price level.

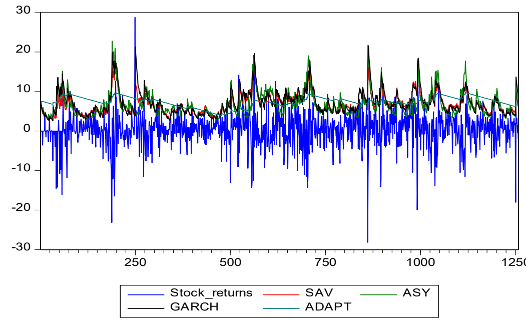
Fig. 2: Co-movement between stock returns (inclusive of oil returns) and tail risk of 5 percent VaR in selected countries



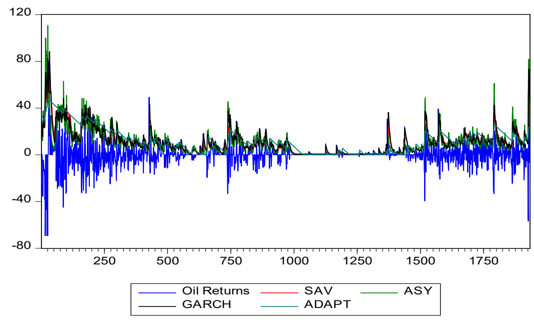
U.K



Switzerland



Oil



Note: SAV is Symmetric Absolute Value, ASY is Asymmetric Slope, GARCH is Indirect GARCH, ADAPT is Adaptive. While the returns series (stock and oil) are computed as 100*$\Delta\log(p_t)$ and p_t is the price level.

Table 11: Clark and west out-of-sample forecast (one predictor)

Out-of-sample 1%	Canada	France	Germany	Italy	Japan	Switzerland	UK	US
h = 6	5.1554^{b}	12.5751^{a}	5.5771	22.5835^{b}	22.0696^{a}	6.2924^{a}	-4.9909	3.0875
	(1.9323)	(3.8304)	(0.4360)	(2.3643)	(2.7312)	(2.2612)	(-1.1088)	(1.2763)
h = 12	5.1102^{a}	12.7250^{a}	5.9544	23.0755^{a}	22.0154^{a}	6.3852^{a}	-4.9078	3.1150^{c}
	(1.9251)	(3.8864)	(0.4674)	(2.4245)	(2.7362)	(2.3028)	(-1.0919)	(1.2913)
h = 24	5.2312^{a}	12.7280^{a}	4.4402	23.7868^{a}	22.3940^{a}	6.2528^{a}	-4.9145	3.1063^{c}
	(1.9937)	(3.9190)	(0.3495)	(2.5103)	(2.8107)	(2.2740)	(-1.0914)	(1.2948)
5%								
h = 6	4.7788^{b}	12.3085^{a}	19.4953^{a}	-1.1331	27.9832^{a}	6.1399^{a}	-4.6364	3.3465^{c}
	(1.8152)	(3.6969)	(1.9809)	(-0.0948)	(3.7951)	(2.2093)	(-1.0722)	(1.4826)
h = 12	4.7389	12.4770^{a}	19.9048^{a}	-0.8859	27.7157^{a}	6.2318^{a}	-4.5490	3.3764^{c}
	(1.8096)	(3.7562)	(2.0313)	(-0.0744)	(3.7613)	(2.2505)	(-1.0534)	(1.5000)
h = 24	4.8474^{a}	12.4920^{a}	20.4153^{a}	-1.8898	28.2968^{a}	6.1004^{a}	-4.6312	3.3900^{c}
	(1.8728)	(3.7905)	(2.0974)	(-0.1596)	(3.8682)	(2.2215)	(-1.0688)	(1.5141)

Note: For the Clark and West test, the null hypothesis of a zero coefficient is rejected if the t-statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test), and +2.00 for 0.01 test (for a one sided 0.01 test) (see Clark and West, 2007), and are denoted by $^{c,\,b}$ and ^{a}, respectively; and the values of the t-statistic are denoted in parentheses. One predictor here only accommodates own tail risk.

Table 12: RMSFE out-of-sample forecast (one predictor)

Out-of-sample								
1%	Canada	France	Germany	Italy	Japan	Switzerland	UK	US
h = 6	8.2549	11.0865	21.8755	22.9867	14.8858	10.4129	12.4780	8.0708
h = 12	8.2515	11.0820	21.8599	22.9752	14.8902	10.4073	12.4889	8.0620
h = 24	8.2245	11.0789	21.9491	22.9636	14.8803	10.4087	12.5607	8.0478
5%								
h = 6	8.2737	10.7148	17.8514	18.5724	19.0114	10.3844	13.0513	7.7297
h = 12	8.2676	10.7160	17.8469	18.5447	19.0270	10.3789	13.0632	7.7216
h = 24	8.2367	10.7235	17.8620	18.6193	19.0362	10.3802	13.1343	7.7104

Note: RMSFE denotes Root Mean Square Forecast Error. One predictor here only accommodates own tail risk.

Table 13: Clark and west out-of-sample forecast (two predictor)

Out-of-sample								
1%	Canada	France	Germany	Italy	Japan	Switzerland	UK	US
h = 6	5.2333^{b}	12.5421^{a}	3.0139	24.6791^{a}	23.0360^{a}	6.4827^{a}	4.9125^{b}	4.1733
	(1.9932)	(3.8378)	(0.2622)	(2.5748)	(2.9101)	(2.3201)	(1.6699)	(1.2493)
h = 12	5.1866^{b}	12.6673^{a}	3.4415	25.0077^{a}	23.0527^{a}	6.5618^{a}	4.9390^{b}	4.1732
	(1.9855)	(3.8880)	(0.7637)	(2.6212)	(2.9271)	(2.3581)	(1.6847)	(1.2542)
h = 24	5.3057^{a}	12.6580^{a}	2.3494	25.4768^{a}	23.1736	6.4533	4.9755^{b}	4.2767^{c}
	(2.0547)	(3.9173)	(0.2059)	(2.6909)	(2.9744)^{a}	(2.3417)^{a}	(1.7052)	(1.2953)
5%								
h = 6	5.0227^{b}	12.2706^{b}	19.3054^{b}	-0.8102	27.8035^{b}	6.4800^{b}	5.4542^{b}	5.1721^{b}
	(1.9526)	(3.6872)	(1.9298)	(-0.0701)	(3.6979)	(2.3023)	(1.8620)	(1.6570)
h = 12	4.9791^{b}	12.4602^{b}	19.7448^{b}	-0.4950	27.5861^{a}	6.5324^{a}	5.4842^{b}	5.1694^{b}
	(1.9459)	(3.7481)	(1.9821)	(-0.0430)	(3.6828)	(2.3294)	(1.8786)	(1.6627)
h = 24	5.0848^{a}	12.5078^{a}	20.2896^{a}	-1.4129	27.7404^{a}	6.4038^{a}	5.5497^{b}	5.2647^{b}
	(2.0105)	(3.7875)	(2.0498)	(-0.1234)	(3.7429)	(2.3018)	(1.9112)	(1.7065)

Note: For the Clark and West test, the null hypothesis of a zero coefficient is rejected if the t-statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test), and +2.00 for 0.01 test (for a one sided 0.01 test) (see Clark and West, 2007), and are denoted by $^{c,\,b}$ and ^{a}, respectively; and the values of the t-statistic are denoted in parentheses. Two predictors accommodate both own and oil tail risks.

Table 14: RMSFE out-of-sample forecast (two predictors)

Out-of-sample								
1%	Canada	France	Germany	Italy	Japan	Switzerland	UK	US
h = 6	8.1980	11.0535	20.8465	22.2557	14.5415	10.6307	8.2284	8.7361
h = 12	8.1943	11.0456	20.8501	22.2240	14.5290	10.6195	8.2277	8.7220
h = 24	8.1671	11.0401	20.9287	22.1558	14.4879	10.5984	8.2661	8.6929
5%								
h = 6	8.1776	10.5989	17.9195	17.6957	18.3450	10.4274	8.0932	8.2140
h = 12	8.1715	10.6101	17.9198	17.6803	18.3270	10.4223	8.0937	8.2004
h = 24	8.1407	10.6475	17.9447	17.7572	18.2569	10.4263	8.1142	8.1741

Note: RMSE denotes Root Mean Square Forecast Error. Two predictors accommodate both own and oil tail risks.

Table 15: Clark and west out-of-sample forecast (three predictors)

Out-of-sample							
1%	Canada	France	Germany	Italy	Japan	Switzerland	UK
h = 6	5.2057[b]	11.8642[a]	3.5521	24.7908[b]	24.0207[a]	5.7808[a]	5.2133[b]
	(1.9492)	(3.5888)	(0.2916)	(2.5849)	(2.9754)	(2.0595)	(1.6834)
h = 12	5.1608[b]	11.9571[a]	3.9357	25.1258[a]	24.1725[a]	5.8535[a]	5.2150[b]
	(1.9426)	(3.6309)	(0.3245)	(2.6320)	(3.0102)	(2.0944)	(1.6901)
h = 24	5.2773[a]	11.9241[a]	2.4688	25.6213[a]	24.4054[a]	5.7671[a]	5.2835[b]
	(2.0096)	(3.6539)	(0.2041)	(2.7042)	(3.0723)	(2.0848)	(1.7224)
5%							
h = 6	5.8172[a]	12.2598[a]	18.1457[b]	-2.7135	28.1094[a]	6.5613[a]	3.6983
	(2.3357)	(3.6620)	(1.8294)	(-0.2247)	(3.7376)	(2.3349)	(1.2102)
h = 12	5.7729[a]	12.4447[a]	18.6081[b]	-2.3831	28.0391[a]	6.6039[a]	3.7251
	(2.3301)	(3.7222)	(1.8837)	(-0.1982)	(3.7445)	(2.3590)	(1.2236)
h = 24	5.8679[a]	12.4804[a]	19.1876[b]	-3.8130	28.1180[a]	6.4728[a]	3.7524
	(2.3962)	(3.7595)	(1.9538)	(-0.3182)	(3.7957)	(2.3307)	(1.2405)

Note: For the Clark and West test, the null hypothesis of a zero coefficient is rejected if the t-statistic is greater than +1.282 (for a one sided 0.10 test), +1.645 (for a one sided 0.05 test), and +2.00 for 0.01 test (for a one sided 0.01 test) (see Clark and West, 2007), and are denoted by c, b and a, respectively; and the values of the t-statistic are denoted in parentheses. Three predictors accommodate own tail risk, oil tail risk and US stock tail risk. Therefore, the return predictability for US stock returns is suppressed in this case.

Table 16: RMSFE out-of-sample forecast (three predictors)

Out-of-sample							
1%	Canada	France	Germany	Italy	Japan	Switzerland	UK
h = 6	8.2522	11.0193	20.0022	22.2484	14.6626	10.5232	8.2735
h = 12	8.2464	11.0049	19.9943	22.2176	14.6432	10.5112	8.2688
h = 24	8.2181	10.9834	20.0841	22.1509	14.5942	10.4826	8.2910
5%							
h = 6	8.2001	10.6718	17.8780	18.9380	17.8287	10.3047	8.1160
h = 12	8.1935	10.6801	17.8825	18.9196	17.8034	10.3004	8.1107
h = 24	8.1629	10.7072	17.9159	19.0403	17.7291	10.3020	8.1154

Note: RMSE denotes Root Mean Square Forecast Error. Three predictors accommodate own tail risk, oil tail risk and US stock tail risk. Therefore, the return predictability for US stock returns is suppressed in this case.