Finance Research Letters 69 (2024) 106179

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Forecasting U.S. recessions using over 150 years of data: Stock-market moments versus oil-market moments[☆]

Elie Bouri[a,b,*], Rangan Gupta[c], Christian Pierdzioch[d], Onur Polat[e]

[a] School of Business, Lebanese American University, Lebanon
[b] Korea University Business School, Seoul, Korea
[c] Department of Economics, University of Pretoria, Pretoria, 0002, South Africa
[d] Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany
[e] Department of Public Finance, Bilecik Şeyh Edebali University, Bilecik, Turkiye

ARTICLE INFO

JEL classification:
C53
E32
E37
G17

Keywords:
Recessions
Stock-market and oil-market moments
Forecasting
Shrinkage estimators
AUC statistics

ABSTRACT

Using monthly data from 1871 to 2024 and logistic models with shrinkage estimators, we compare the contribution of stock and oil-market moments (returns, volatility, skewness, and kurtosis) to the accuracy of out-of-sample forecasts of U.S. recessions at various forecast horizons, while controlling for standard macroeconomic predictors and the total connectedness indexes of the moments. Adding stock-market moments to the potential predictors improves significantly the accuracy of out-of-sample forecasts at an intermediate forecast horizon, where the lagged recession dummy, term spread, and stock returns are top predictors. Oil-market moments and connectedness indexes do not contribute much to forecast accuracy.

1. Introduction

A large and significant literature has been developed in recent years studying in detail the role played by the first and second moments of stock and crude oil returns in predicting recessions and/or economic activity of the United States (U.S.) based on post World War II (WWII) data (see, for example, Liu and Moench, 2016; Gupta et al., 2017, Gupta and Wohar, 2017, Kilian and Vigfusson (2017), Huang and Startz (2020), Loycsa et al. (2021), and Gupta et al. (2022) for detailed reviews of this literature). This paper adds to this prominent literature in multiple ways.

Firstly, we go beyond stock and crude oil returns and volatilities by considering the role of skewness and (excess) kurtosis for forecasting recessions. The third and fourth moments of asset return distribution, over and above their first and second moments, known to proxy for rare disaster events (Mei et al., 2017; Gupta et al., 2023) could trigger economic slowdowns (Manela and Moreira, 2017; Pierdzioch and Gupta, 2020).[1]

☆ We would like to thank an anonymous referee for many helpful comments. Any remaining errors are solely ours.

* Corresponding author.
E-mail addresses: elie.elbouri@lau.edu.lb (E. Bouri), rangan.gupta@up.ac.za (R. Gupta), macroeconomics@hsu-hh.de (C. Pierdzioch), onur.polat@bilecik.edu.tr (O. Polat).

[1] Furthermore, many researchers have documented the complex relationship between stock and/or oil markets and the economy, and have provided evidence of strong deviation of the returns distribution of these two assets from a normal distribution for various shocks and crisis periods. Accordingly, non-zero skewness and excess kurtosis are stylized features of stock and oil returns, suggesting that returns and volatility do not suffice to describe fully the non-normal returns distributions. In this regard, existing studies highlight the importance of higher-order moments reflecting asymmetric fat-tail risks, and extreme values of returns (Kim et al., 2014; Langlois, 2020; Carnero et al., 2023), and so higher-order moments may also be useful to model and forecast the economic cycle.

https://doi.org/10.1016/j.frl.2024.106179
Received 9 August 2024; Received in revised form 9 September 2024; Accepted 22 September 2024
Available online 30 September 2024
1544-6123/

Secondly, given the well-established evidence on within and across interlinkages between not only returns and volatility (Balcilar et al., 2015, 2017), but also involving skewness and kurtosis of oil and stock prices (Lei et al., 2021; Zhang et al., 2022; Nekhili and Bouri, 2023), we analyze whether time-varying connectedness across the four moments can play a significant role in forecasting recessions. Thirdly, instead of limiting the analysis to just post-WWII recessions, we use the longest possible monthly data sample of 1871 to 2024, where the starting point is driven by data availability of important predictors, i.e., the term spread, the inflation rate, and the short-term interest rate (Plakandaras et al., 2019). Finally, to ensure our results are robust, we also conduct a forecasting experiment over 1900 to 2024 with other available metrics of rare disasters such as geopolitical risks, climate risks and shortages, i.e., supply constraints (Burriel et al., 2024), which in turn might carry additional forecasting content beyond stock and crude oil moments.

Given the use of relatively large number of predictors, involving standard macroeconomic variables, the moments of oil and stock markets (considered individually and in form of their connectedness indexes), as well as proxies of rare disaster risks, we employ shrinkage estimators to estimate Logit models, and then use the estimated models to produce out-of-sample forecasts of U.S. recessions. To the best of our knowledge, this is a first such attempt using over 150 years of monthly data, with the only other relevant paper being that of Plakandaras et al. (2017), who investigate only the role of the first moment (returns) of oil and stock prices for U.S. recessions over 1871 to 2016.

2. The data

Data are monthly from January 1871 to March 2024. We measure recessions as a binary 0/1-variable based on the classification of economic cycles by the National Bureau of Economic Research (NBER).[2] Because the NBER typically announces recessions with a delay, we essentially predict future months that the NBER will eventually classify as recessions. Regarding the standard macroeconomic predictors, we include the term spread (10 year government bond yield minus short-term interest rate), the short-term, and the inflation rate (month-on-month growth in the consumer price index (CPI)).[3] CPI-deflated real values of the West Texas Intermediate (WTI) oil prices and the S&P 500 stock index are used to compute log-real returns,[4] which, in turn, are then fitted into the Autoregressive Conditional Density (ACD) model (Hansen, 1994) to obtain conditional volatility, conditional skewness (reflecting returns asymmetry), and conditional (excess) kurtosis (capturing extreme values of returns).[5]

We also consider various total connectedness indices (TCI). These include TCI for the returns (TCI oil-stock returns), volatility (TCI oil-stock volatility), skewness (TCI oil-stock skewness), and kurtosis (TCI oil-stock kurtosis) of oil and stock markets; TCI for the four moments of stock market (TCI stock moments) and crude oil market (TCI oil moments); and the TCI of the eight moments together across oil and stock markets (TCI oil-stock all moments). All these seven TCIs are estimated using the approach of Diebold and Yilmaz (2009, 2012).[6]

Fig. 1 shows the macroeconomic and financial data that we study in our empirical research.

3. Empirical methods

The starting point for studying potential predictors of recessions is a binary outcome forecasting model, such as the widely-studied Logit regression model (see, for example, Estrella and Mishkin, 1998), can be represented as follows:

$$\log \frac{P(R_{t+h}=1|X_t=x_t)}{P(R_{t+h}=0|X_t=x_t)\hat{\text{E}}} = \beta_0 + \beta^T x_{m,t} + \gamma^T x_{f,t}, \tag{1}$$

where $P(R_{t+h}=1|X_t=x_t)$ denotes the probability of a recession, R, in period $t+h$ (where $h \geq 1$ denotes the forecast horizon), given that the predictors take on the value x_t in period t, β_0 denotes an intercept coefficient, β^T is the (transposed) vector of coefficients associated with macroeconomic predictors, $x_{m,t}$, commonly considered in earlier literature on recession forecasting, and; γ^T is the coefficient vector associated with the financial market predictors, $x_{f,t}$. We consider four forecast horizons: 3, 6, 9, and 12 months

[2] See http://www.nber.org/cycles.html.

[3] Data for the short-term rate until December 2023 come from the website of Professor Amit Goyal: https://sites.google.com/view/agoyal145, and then updated data until March 2024 are constructed using the 3-month Treasury bill rate from the FRED database of the Federal Reserve Bank of St. Louis. As far as CPI and the yield on 10 year government bond are concerned, they are sourced from the website of Professor Robert J. Shiller: http://www.econ.yale.edu/~shiller/data.htm.

[4] Nominal values of both these series are obtained from Global Financial Data (GFD): https://globalfinancialdata.com/.

[5] The reader is referred to Ahmed et al. (2024) for a compact summary of this model. For the estimation of the ACD model, we rely on the R add-on package "racd" (Ghalanos, 2014). In this regard, we must stress that, besides the power of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) processes in modeling the conditional volatility of stock and crude oil returns, the presence of extreme shocks and crisis periods, which characterize stock and oil markets, implies that returns strongly deviate from a normal distribution. This suggests the inappropriateness of using volatility as a full measure of return variability (Bouri et al., 2021). Therefore, the dynamics of skewness and kurtosis become relevant, which necessitates alternative GARCH models capable of capturing the excess volatility shocks. The ACD model comprises the Autoregressive Moving Average (ARMA)-GARCH process augmented with additional parameters capturing time-varying conditional skewness and conditional (excess) kurtosis. We estimate the ACD model with skew-student or normal inverse Gaussian innovations, which helps building a reasonable model that serves the purpose of our analysis.

[6] We utilize the R codes developed by Dr. David Gabauer, which are available at: https://github.com/GabauerDavid/ConnectednessApproach. In order to obtain time-varying indexes of connectedness starting in January, 1871, we apply rolling-window estimation to the spillover model of Diebold and Yilmaz (2009, 2012) on the oil-market and stock-market moments derived from the ACD model going back to November 1859 — the earliest available common data point for the ACD-based estimates of the moments of real log-returns of oil and equity. Note that, since monthly data on CPI is not available before 1871, we use annual version of the same, sourced from GFD, to deflate the nominal values of monthly oil and stock prices from 1859 to 1870.

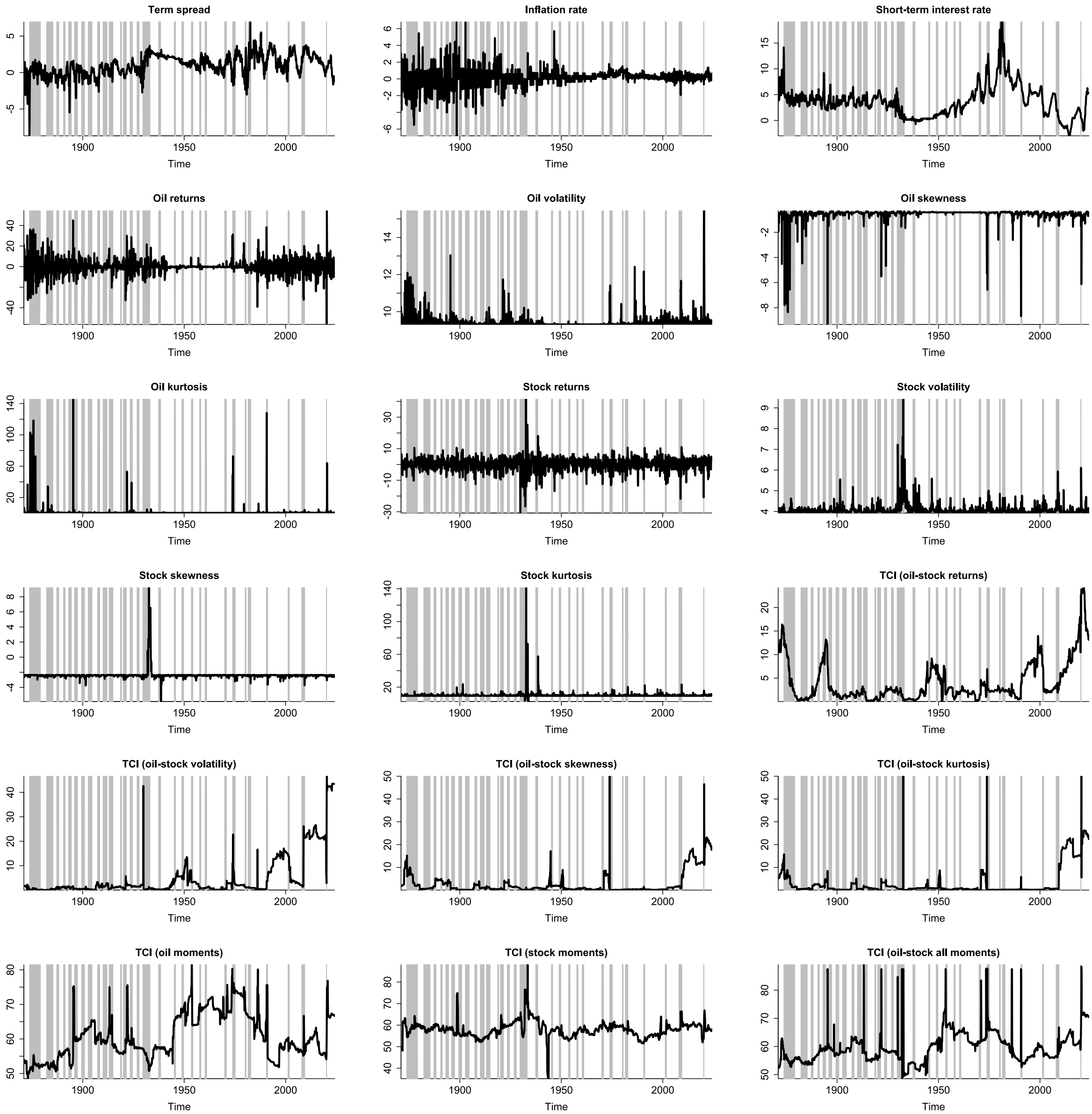

Fig. 1. Data. Note: Shaded areas denote recessions.

(corresponding to one to four quarters). In our case, the macroeconomic predictors, $x_{m,t}$, comprise a lagged recession dummy (not the contemporaneous one so as to account in a simplified way for a potential information lag), the term-spread, the inflation rate, and the short-term interest rate. Regarding the financial market predictors, $x_{f,t}$, we consider three different constellations:

1. Oil returns, oil volatility, oil skewness, and oil kurtosis.
2. Stock market returns, stock volatility, stock skewness, and stock kurtosis.
3. TCI oil-stock returns, TCI oil-stock volatility, TCI oil-stock skewness, TCI oil-stock kurtosis, TCI oil moments, TCI stock moments, and TCI oil-stock all moments.

We combine these three constellations along with the macroeconomic predictors to set up five different forecasting models:

(a) Benchmark: A parsimonious Benchmark model that features only the macroeconomic predictors.
(b) Benchmark-oil-moments: This model extends the Benchmark model to include the oil-market moments.
(c) Benchmark-stock-moments: This model extends the Benchmark model to include the stock-market moments.
(d) Benchmark-both-moments: This model extends the Benchmark model to include both oil- the stock-market moments.

(e) Benchmark-both-moments-TCI: This model extends the Benchmark-both-moments model to include the various TCIs.

Given the use of a relatively large number of predictors, we employ shrinkage estimators to estimate our forecasting Logit regression model. The penalized Logit regression model minimizes the following objective function with respect to the parameters:

$$\min \left[-(1/N) \sum_{t=1}^{N} y_t(\beta_0 + \beta^T x_{m,t} + \gamma^T x_{f,t}) - \log(1 + \exp(\beta_0 + \beta^T x_{m,t} + \gamma^T x_{f,t}\hat{\text{E}})) \right] + \lambda \left[(1-\alpha)\|B\|_2^2/2 + \alpha\|B\|_1 \right], \tag{2}$$

where $B = (\beta, \gamma)$ and N denotes the number of observations used for estimation; y_{t+h} denotes a recession dummy in period $t+h$; $\alpha = 1$ in case of the Lasso estimator; $\alpha = 0$ in case of a Ridge regression estimator, and; $\alpha \in (0, 1)$ in case of an elastic net estimator. When studying an elastic net estimator, we set $\alpha = 0.5$, right in the middle between the Lasso and the Ridge regression estimator. We select the shrinkage parameter, λ, using 10-fold cross validation, where we select the shrinkage parameter that maximizes the area (AUC) under the receiver operating characteristic (ROC) curve.[7]

Maximizing the AUC by cross validation is a natural choice because the AUC statistic is also used to assess the predictive performance of the forecasting models. For a recent application of the AUC statistic in the context of recession forecasting, see Pierdzioch and Gupta (2020), among others, and for an introductory exposition, refer to Greiner et al. (2000). For a “perfect” forecasting model, we have AUC = 1. For a “not better than pure noise” forecasting model, we have AUC = 0.5. A larger AUC statistic, thus, signals a better performance of a forecasting model and, hence, we also can compare the predictive accuracy of two forecasting models in terms of the AUC statistic.[8]

4. Empirical results

Table 1 summarizes the results for the Lasso estimator, an elastic net, and the Ridge regression estimator. For all three estimators, the AUC statistics are reported for the respective benchmark model (AUC1) and the alternative model (AUC2) along with the results of the DeLong et al. (1988) test and the corresponding p-value. The AUC statistics show a clear tendency to decrease when the forecast horizon increases from the short ($h = 3$) to the long ($h = 12$) forecast horizon. For the short ($h = 3$) forecast horizons, the majority of test results is insignificant. Hence, adding stock-market moments or oil-market moments (or both) to the array of potential predictors does not systematically improve directional accuracy of out-of-sample forecasts of recessions. For the intermediate forecast horizon ($h = 6$), in contrast, the results change. The main message conveyed by the results for the intermediate forecast horizon is that stock-market moments matter for the directional accuracy of out-of-sample forecasts, where the differences between the AUC1 and AUC2 statistics are quite noticeable when we consider the model without moments or the model that features oil-market moments as the benchmark. For the two long forecast horizons, in turn, the test results again are insignificant, or they have a positive sign, with few exceptions only (for example, adding stock-market moments to the array of predictors yields a significant test result at $h = 9$ when we use the Ridge regression estimator). Finally, adding the TCIs to a model that already contains stock-market moments and oil-market moments deteriorates the forecasting performance.

We report in Table A.1 at the end of the paper (Appendix) some additional out-of-sample results for sensitivity, specificity, and the Matthews correlation coefficient. The results show that stock-market moments increase the correlation coefficient at $h = 6$ under the Lasso and elastic net estimator, and at $h = 9$ for all three shrinkage estimators. For $h = 12$ the oil-market moments and stock-market moments have a comparable effect on the correlation coefficient, while for $h = 3$ stock-market moments raise the correlation coefficient under the elastic net estimator. Hence, as in Table 1, we observe the contribution of stock-market moments to forecasting performance mainly for an intermediate forecast horizon.

In Fig. 2, we plot, for the forecast horizons $h = 3, 6$, the dynamic paths of the AUC statistics, which, in turn, are constructed by using the first three decades of the out-of-sample period to generate an initial estimate of the AUC statistics of the forecasting models under inspection, and then recursively adding results for the next months until we reach the end of the forecasting period. We observe that the forecasting models yield a comparable and stable performance at the short forecast horizon ($h = 3$), with the differences between the AUC statistics in general being small. At the intermediate forecast horizon ($h = 6$) the differences between the AUC statistics become noticeable and the AUC statistics exhibit a discernible tendency to increase over time, implying that directional forecast accuracy has improved in the later parts of the sample period. For the Lasso and the elastic net estimators, the model that features stock-market moments clearly dominates the other models in terms of the AUC statistic, where it is also clear that it is the model featuring stock-market moments in the array of potential predictors that explains the comparatively good performance of the model that includes both stock- and oil-market moments. In particular, the difference between the paths of the AUC statistic obtained for the model without any moments and the model with oil-market moments are hardly visible in the figure. The model with stock-market moments, in contrast, clearly outperforms the benchmark model and the model with oil-market moments under the Lasso and elastic net estimators. Incorporating the TCIs into the pool of potential predictors, in turn, deteriorates directional forecasting accuracy.

[7] Results for 3-fold cross validation are qualitatively similar to the results we report in Table 1 and are available from the authors upon reasonable request. It should also been noted that the results for the simplified Logit models that we present in Fig. A.2 and Table A.2 at the end of the paper (Appendix) do not depend on the cross-validation scheme being used.

[8] We use the R computing environment (R Core Team, 2023), the R add-on package “glmnet” (Friedman et al., 2010; Tay et al., 2023) for estimation of the logistic regression models by means of the shrinkage estimators, and the R add-on package “pROC” (Robin et al., 2011) for computation of the AUC statistics.

Table 1
Out-of-sample results.

Estimator	lasso				Elastic net				Ridge regression			
Statistic	AUC1	AUC2	Test	pval	AUC1	AUC2	Test	pval	AUC1	AUC2	Test	pval
	$h=3$				$h=3$				$h=3$			
Benchmark vs. Benchmark-oil-moments	0.8837	0.8878	−0.9760	0.3290	0.8866	0.8822	0.9243	0.3554	0.8978	0.8966	0.7164	0.4737
Benchmark vs. Benchmark-stock-moments	0.8837	0.8880	−0.5650	0.5721	0.8866	0.8816	0.5178	0.6046	0.8978	0.8957	0.3478	0.7280
Benchmark-oil-moments vs. Benchmark-stock-moments	0.8878	0.8880	−0.0192	0.9846	0.8822	0.8816	0.0665	0.9470	0.8966	0.8957	0.1493	0.8813
Benchmark-oil-moments vs. Benchmark-both-moments	0.8878	0.8932	−0.6512	0.5149	0.8822	0.8913	−0.9902	0.3221	0.8966	0.8983	−0.2962	0.7671
Benchmark-stock-moments vs. Benchmark-both-moments	0.8880	0.8932	−1.8722	0.0612	0.8816	0.8913	−2.8227	0.0048	0.8957	0.8983	−0.6742	0.5002
Benchmark-both-moments vs. Benchmark-both-moments-TCI	0.8932	0.8770	2.4711	0.0135	0.8913	0.8740	2.5079	0.0121	0.8983	0.8868	2.1141	0.0345
	$h=6$				$h=6$				$h=6$			
Benchmark vs. Benchmark-oil-moments	0.7900	0.7886	0.2072	0.8359	0.7869	0.7887	−0.3024	0.7623	0.8182	0.8146	2.0621	0.0392
Benchmark vs. Benchmark-stock-moments	0.7900	0.8186	−3.3922	0.0007	0.7869	0.8213	−3.3761	0.0007	0.8182	0.8274	−1.6481	0.0993
Benchmark-oil-moments vs. Benchmark-stock-moments	0.7886	0.8186	−3.5581	0.0004	0.7887	0.8213	−3.4519	0.0006	0.8146	0.8274	−2.1810	0.0292
Benchmark-oil-moments vs. Benchmark-both-moments	0.7886	0.8210	−3.7787	0.0002	0.7887	0.8241	−3.7119	0.0002	0.8146	0.8269	−1.9179	0.0551
Benchmark-stock-moments vs. Benchmark-both-moments	0.8186	0.8210	−0.5502	0.5822	0.8213	0.8241	−0.6233	0.5331	0.8274	0.8269	0.1640	0.8697
Benchmark-both-moments vs. Benchmark-both-moments-TCI	0.8210	0.7888	3.4574	0.0005	0.8241	0.7672	5.3735	0.0000	0.8269	0.8024	3.4867	0.0005
	$h=9$				$h=9$				$h=9$			
Benchmark vs. Benchmark-oil-moments	0.7630	0.7508	1.9616	0.0498	0.7557	0.7500	0.7777	0.4367	0.7457	0.7344	1.2752	0.2022
Benchmark vs. Benchmark-stock-moments	0.7630	0.7647	−0.1952	0.8452	0.7557	0.7640	−0.8405	0.4007	0.7457	0.7644	−1.7365	0.0825
Benchmark-oil-moments vs. Benchmark-stock-moments	0.7508	0.7647	−1.5562	0.1197	0.7500	0.7640	−1.3693	0.1709	0.7344	0.7644	−2.6087	0.0091
Benchmark-oil-moments vs. Benchmark-both-moments	0.7508	0.7593	−0.9704	0.3318	0.7500	0.7620	−1.2265	0.2200	0.7344	0.7609	−2.3567	0.0184
Benchmark-oil-moments vs. Benchmark-both-moments	0.7647	0.7593	2.0901	0.0366	0.7640	0.7620	0.6982	0.4851	0.7644	0.7609	0.9854	0.3244
Benchmark-both-moments vs. Benchmark-both-moments-TCI	0.7593	0.7133	4.0221	0.0001	0.7620	0.7151	4.1440	0.0000	0.7609	0.7298	2.9975	0.0027
	$h=12$				$h=12$				$h=12$			
Benchmark vs. Benchmark-oil-moment	0.7365	0.7274	0.9258	0.3546	0.7324	0.7314	0.0922	0.9265	0.7366	0.7341	0.3070	0.7589
Benchmark vs. Benchmark-stock-moments	0.7365	0.7356	0.0849	0.9323	0.7324	0.7358	−0.2921	0.7702	0.7366	0.7467	−1.2980	0.1943
Benchmark-oil-moments vs. Benchmark-stock-moments	0.7274	0.7356	−0.9005	0.3678	0.7314	0.7358	−0.4253	0.6706	0.7341	0.7467	−1.3087	0.1906
Benchmark-oil-moments vs. Benchmark-both-moments	0.7274	0.7218	0.6285	0.5297	0.7314	0.7164	1.4942	0.1351	0.7341	0.7406	−0.7080	0.4789
Benchmark-oil-moments vs. Benchmark-both-moments	0.7356	0.7218	2.4349	0.0149	0.7358	0.7164	3.3690	0.0008	0.7467	0.7406	1.0875	0.2768
Benchmark-both-moments vs. Benchmark-both-moments-TCI	0.7218	0.6842	2.7147	0.0066	0.7164	0.6836	2.3820	0.0172	0.7406	0.6893	3.7087	0.0002

Note: AUC1 denotes the AUC statistic of the Benchmark model. AUC2 denotes the AUC statistic of the alternative model. Initial estimates of the forecasting model use data up to and including 1899/12. The estimation window is then recursively expanded by adding step-by-step one month of data until the end of the sample period is reached.

To sum up, the dynamic paths of the AUC statistics are consistent with the results reported in Table 1 and, in addition, they show that: (i) the ranking of the various models is quite stable over time, and; (ii) the test results reported in the tables do not merely reflect a clustering of forecasting performance during specific historical episodes.

In Table 2, we consider how often the predictors are included in the estimated forecasting models, by focusing on the elastic net estimator. The Ridge regression estimator includes all predictors all the time (but shrinks the coefficients toward zero), while the Lasso estimator tends to select a forecasting model that is somewhat more parsimonious than the elastic-net model, but the results, available upon request from the authors, are not too different. The numbers we report in the table are in percent and indicate how

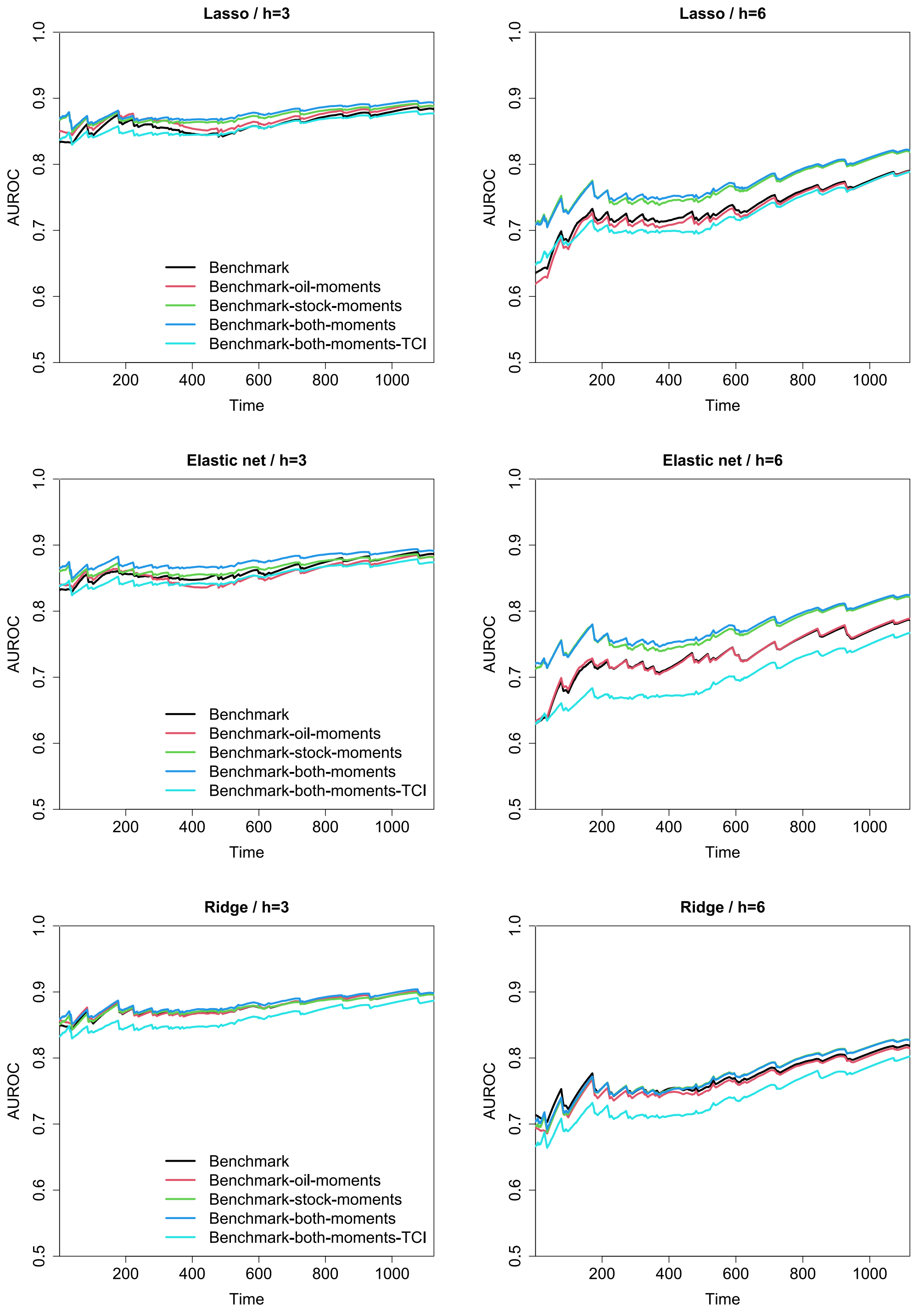

Fig. 2. Dynamics of AUC.
Note: Initial estimates of the forecasting model use data up to and including 1899/12. The estimation window is then recursively expanded by adding step-by-step one month of data until the end of the sample period is reached. The first estimate of the AUC statistic uses the first three decades of out-of-sample forecasts of the probability of a recession thus generated. The time path of the AUC statistic obtains by adding step-by-step one out-of-sample forecast until the end of the sample period is reached.

often a predictor is being included by the elastic net estimator in the forecasting model that can select among all predictors (that is, we report results for the benchmark-both-moments-TCI model). Two results stand out. First, the forecasting models become more

Table 2
Inclusion of the predictors in the forecasting models.

Predictor	$h = 3$	$h = 6$	$h = 9$	$h = 12$
Recession dummy (lagged)	100.0000	100.0000	100.0000	99.4547
Term spread	99.8653	91.4131	85.2003	92.0927
Inflation rate	7.9461	43.6105	88.9341	99.3865
Short-term interest rate	77.1044	93.5767	98.8459	97.6142
Oil returns	1.5488	37.2549	93.2111	99.4547
Oil volatility	47.2054	39.3509	84.4535	91.0020
Oil skewness	2.8283	46.5855	82.6205	100.0000
Oil kurtosis	0.7407	35.1589	79.0224	71.0293
Stock returns	100.0000	94.3205	100.0000	100.0000
Stock volatility	33.4007	45.0980	91.8534	95.9782
Stock skewness	34.8822	48.6139	98.2349	99.5910
Stock kurtosis	19.7980	30.1555	83.9104	85.8214
TCI (oil-stock returns)	3.0303	35.3617	86.5580	98.5685
TCI (oil-stock volatility)	23.7710	46.1122	83.3673	95.8419
TCI (oil-stock skewness)	4.1077	49.1548	99.5927	100.0000
TCI (oil-stock kurtosis)	5.9933	37.3225	94.6368	98.9775
TCI (oil moments)	9.4276	50.2366	92.7359	98.5685
TCI (stock moments)	23.7710	37.3225	88.0516	97.6823
TCI (oil-stock all moments)	12.2559	39.0128	93.5506	98.9775

Note: Numbers are in percent and indicate how often a predictor is being included by the elastic-net estimator in the forecasting model that can select among all predictors (Benchmark-both-moments-TCI model). Initial estimates of the forecasting model use data up to and including 1899/12. The estimation window is then recursively expanded by adding step-by-step one month of data until the end of the sample period is reached.

complex as the forecast horizon increases. Second, and more important, the top predictors at the interesting intermediate ($h = 6$) forecast horizon are the lagged recession dummy, stock-market returns, the short-term interest rate, and the term spread.

As an extension, we consider a relatively short sample period for which, data for several potentially influential additional control variables are available. Specifically, for our analysis of the shorter sample period ranging from 1900 January to March 2024, the additional control variables are intended to capture the class of disaster events that have attracted considerable attention in recent research. We, thus, control for newspapers-based measures of geopolitical risks due to acts and threats (Caldara and Iacoviello, 2022),[9] two measures of climate risks,[10] and four (energy, food, industry and labor) shortages indexes (Caldara et al., 2024).[11]

We summarize the results for the shorter sample period in Fig. A.1 in Appendix. We plot the recursively estimated dynamic paths of the AUC statistic for a benchmark model and a rival model. The benchmark model features as potential predictors the two measures of geopolitical risks, the two metrics of climate risks, the four shortage indexes, a lagged recession dummy, the term spread, the inflation rate, a short-term interest rate, and the four oil moments. We call this model the extended benchmark model. The rival model extends the array of potential predictors to include the stock-market moments. The dynamic paths of the AUC statistic clearly show that the rival model produces better out-of-sample forecasts than the extended benchmark model except at $h = 12$, a finding that lends further support to the observations made for the long sample period. The results for the shorter sample period, however, go beyond the findings for the long sample period in the sense that, the rival model outperforms the extended benchmark model not only at $h = 6$, but also at $h = 3$ and, to a lesser extent, at $h = 9$. In fact, the DeLong et al. (1988) test, as applied to the entire out-of-sample forecasting period, gives highly significant test results for $h = 3$ (test = −4.5548, p-value = < 0.0001), for $h = 6$ (test = −3.3023, p-value = 0.0010), and for $h = 9$ (test = −2.976, p-value = 0.0029).

Another interesting issue regards a potential look-ahead bias that the full-sample estimates of the ACD moments may inject into our results. In this regard, we note that using conditional moments should mitigate a potential look-ahead bias. In addition, we note that the key predictors (at the intermediate forecast horizon) are the lagged recession dummy, the short-term interest rate, the term spread, and stock-market returns, where the latter, by definition, are not subject to a look-ahead bias. Notwithstanding, in order to address this issue in some more detail, we focus on the key predictors plus oil returns and re-estimate our forecasting models as Logit models. We summarize the resulting AUC statistics in Fig. A.2 and the corresponding Matthews correlation coefficients in Table A.2. The Logit results show that the contribution of stock returns to forecasting accuracy dominates that of oil returns, and that this dominance now even strengthens well beyond the $h = 6$ forecast horizon.

5. Concluding remarks

Our results, using over 150 years (1871–2024) of monthly data, corroborate much significant previous research that the term spread is an important predictor of recessions in the U.S., and further indicate that stock-market moments in general, and stock

[9] The data is available for download from https://www.matteoiacoviello.com/gpr.htm.

[10] Like Choi et al. (2020), we use the average of standardized abnormal deviations in temperature, precipitation, number of heating and cooling days, with all raw data sourced from the National Climatic Data Center (NCDC): https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/national/time-series. For the first metric of climate risks, the standardized abnormal deviations enter with the same sign. For the second metric, a negative value of abnormal deviations for heating days is used when computing the average.

[11] The data can be accessed from: https://www.matteoiacoviello.com/shortages.html.

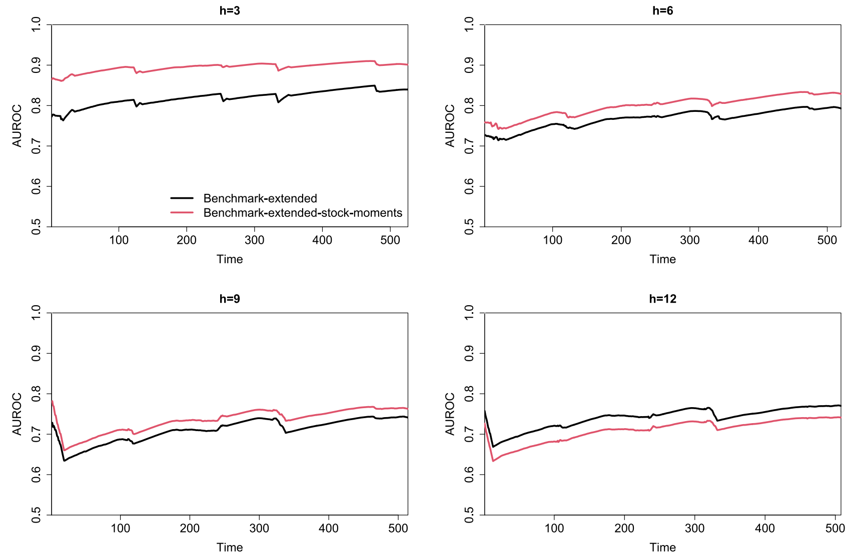

Fig. A.1. Results for a shorter sample period.
Note: Initial estimates of the forecasting model use data up to and including 1949/12. The estimation window is then recursively expanded by adding step-by-step one month of data until the end of the sample period is reached. The first estimate of the AUC statistic uses the first three decades of out-of-sample forecasts of the probability of a recession thus generated. The time path of the AUC statistic obtains by adding step-by-step one out-of-sample forecast until the end of the sample period is reached. Estimates are obtained using an elastic net estimator.

returns in particular, are among the top predictors of U.S. recessions, mainly when we consider the 6-month-ahead forecasting horizon. While similar results have been observed by earlier researchers (Plakandaras et al., 2017; Pierdzioch and Gupta, 2020), we have derived our results for a long sample period that spans over 150 years of data, as well as a shorter sample period for robustness purposes starting in 1900, which also includes data on rare disaster events (geopolitical risks, climate risks, and supply constraints) that have attracted increasing attention of researchers in recent years. Given our utilization of a large number of macroeconomic and financial predictors, we have traced out the incremental contribution of oil- and stock-market moments, as well as their degree of (inter)connectedness, using a logistic model estimated by means of shrinkage estimators (Lasso, elastic net, Ridge regression).

Our econometric exercise suggests that movements in stock markets are exogenous impulses, i.e., they serve as leading indicators for the business cycles of the U.S., rather than being endogenous response to the macroeconomic conditions, as Ludvigson et al. (2021) have stressed recently. In light of this line of reasoning, the recently much discussed issue of whether the Federal Reserve should account for stock market stability in designing monetary policy clearly is pertinent, especially in the wake of forthcoming recessions associated with bearish stock markets (Kurov et al., 2022). Building on our econometric exercise, as part of future research, contingent on the availability of long-span data, it is interesting to perform a comparative analysis for other developed and developing economies to obtain a general picture regarding the relative roles of oil- and stock-market moments in predicting historical recessions.

CRediT authorship contribution statement

Elie Bouri: Writing – review & editing, Writing – original draft, Formal analysis. **Rangan Gupta:** Writing – original draft, Supervision, Project administration, Data curation, Conceptualization. **Christian Pierdzioch:** Writing – original draft, Software, Methodology, Formal analysis. **Onur Polat:** Writing – original draft, Methodology, Investigation.

Data availability

Data will be made available on request.

Appendix

See Figs. A.1 and A.2 and Tables A.1 and A.2.

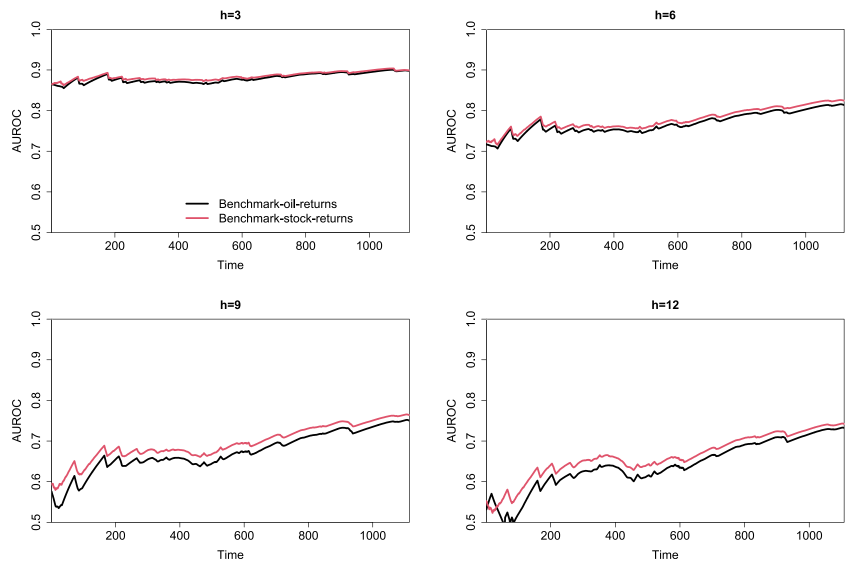

Fig. A.2. Results for the Logit models.
Note: Initial estimates of the forecasting model use data up to and including 1949/12. The estimation window is then recursively expanded by adding step-by-step one month of data until the end of the sample period is reached. The first estimate of the AUC statistic uses the first three decades of out-of-sample forecasts of the probability of a recession thus generated. The time path of the AUC statistic obtains by adding step-by-step one out-of-sample forecast until the end of the sample period is reached. Estimates are obtained using a Logit model.

Table A.1
Additional out-of-sample results.

Estimator	Lasso			Elastic net			Ridge regression		
Statistic	SE	SP	MCC	SE	SP	MCC	SE	SP	MCC
	$h = 3$			$h = 3$			$h = 3$		
Benchmark	0.7649	0.9046	0.6542	0.8612	0.7129	0.4946	0.8215	0.8587	0.6300
Benchmark-oil-moments	0.7592	0.9117	0.6614	0.8130	0.7420	0.4845	0.8017	0.8693	0.6293
Benchmark-stock-moments	0.7734	0.8931	0.6426	0.7932	0.8737	0.6290	0.8612	0.7429	0.5248
Benchmark-both-moments	0.7734	0.8949	0.6453	0.7904	0.8763	0.6306	0.8470	0.8083	0.5858
Benchmark-both-moments-TCI	0.7677	0.8993	0.6479	0.7734	0.8772	0.6185	0.8045	0.8383	0.5894
	$h = 6$			$h = 6$			$h = 6$		
Benchmark	0.9856	0.2111	0.2246	1.0000	0.0468	0.1067	0.8674	0.6440	0.4339
Benchmark-oil-moments	0.9885	0.2217	0.2358	1.0000	0.0406	0.0992	0.8588	0.6422	0.4251
Benchmark-stock-moments	0.8674	0.5742	0.3749	0.9078	0.5097	0.3595	0.8732	0.5963	0.3981
Benchmark-both-moments	0.8790	0.5892	0.3971	0.9078	0.5053	0.3561	0.8617	0.6051	0.3956
Benchmark-both-moments-TCI	0.8012	0.6325	0.3682	0.7983	0.5565	0.3011	0.7954	0.6608	0.3889
	$h = 9$			$h = 9$			$h = 9$		
Benchmark	0.9653	0.3310	0.2859	0.9624	0.2928	0.2560	0.9711	0.3185	0.2834
Benchmark-oil-moments	0.9595	0.3292	0.2786	0.9711	0.3017	0.2717	0.9827	0.2893	0.2759
Benchmark-stock-moments	0.9017	0.4854	0.3358	0.8960	0.4809	0.3272	0.9075	0.4428	0.3093
Benchmark-both-moments	0.8988	0.4641	0.3172	0.8988	0.4552	0.3105	0.9133	0.4117	0.2918
Benchmark-both-moments-TCI	0.7890	0.5280	0.2699	0.7977	0.5244	0.2745	0.8092	0.5004	0.2654
	$h = 12$			$h = 12$			$h = 12$		
Benchmark	0.9827	0.1392	0.1650	0.9855	0.1151	0.1483	0.9769	0.2578	0.2479
Benchmark-oil-moments	0.9595	0.2926	0.2532	0.9682	0.2881	0.2595	0.9682	0.2721	0.2481
Benchmark-stock-moments	0.9335	0.3559	0.2715	0.9191	0.3568	0.2577	0.9538	0.3095	0.2590
Benchmark-both-moments	0.9393	0.3417	0.2670	0.9249	0.3488	0.2576	0.9595	0.3122	0.2671
Benchmark-both-moments-TCI	0.8035	0.4612	0.2295	0.7948	0.4657	0.2253	0.8266	0.4246	0.2215

Note: The threshold is set equal to the unconditional probability of a recession (approximately 23%) during the out-of-sample forecasting period. SE denotes sensitivity. SP denotes specificity. MCC denotes the Matthews correlation coefficient. Initial estimates of the forecasting model use data up to and including 1899/12. The estimation window is then recursively expanded by adding step-by-step one month of data until the end of the sample period is reached.

Table A.2
Matthews correlation coefficients for the Logit models.

Model	$h = 3$	$h = 6$	$h = 9$	$h = 12$
Benchmark	0.6262	0.3472	0.2128	0.0901
Benchmark-oil-returns	0.6262	0.3500	0.2055	0.1426
Benchmark-stock-returns	0.6549	0.3800	0.2595	0.2265
Benchmark-both-returns	0.6527	0.3809	0.2582	0.2361

References

Ahmed, R., Bouri, E., Hosseini, S.M., Shahzad, S.J.H., 2024. Spillover in higherorder moments across carbon and energy markets: A portfolio view. Eur. Financial Manag. http://dx.doi.org/10.1111/eufm.12482.

Balcilar, M., Gupta, R., Miller, S.M., 2015. Regime switching model of US crude oil and stock market prices: 1859 to 2013. Energy Econ. 49, 317–327.

Balcilar, M., Gupta, R., Wohar, M.E., 2017. Common cycles and common trends in the stock and oil markets: Evidence from more than 150 years of data. Energy Econ. 61, 72–86.

Bouri, E., Lei, X., Jalkh, N., Xu, Y., Zhang, H., 2021. Spillovers in higher moments and jumps across US stock and strategic commodity markets. Resour. Policy 72, 102060.

Burriel, P., Kataryniuk, I., Pérez, M., Viani, F., 2024. A new supply bottlenecks index based on newspaper data. Int. J. Central Bank. 20 (2), 17–69.

Caldara, D., Iacoviello, M., 2022. Measuring geopolitical risk. Amer. Econ. Rev. 112 (4), 1194–1225.

Caldara, D., Iacoviello, M., Yu, D., 2024. Measuring Shortages Since 1900. Mimeo, Available for download from: https://www.matteoiacoviello.com/research_files/SHORTAGE_PAPER.pdf.

Carnero, M.A., León, A., Ñíguez, T.M., 2023. Skewness in energy returns: Estimation, testing and implications for tail risk. Q. Rev. Econ. Finance 90, 178–189.

Choi, D., Gao, Z., Jiang, W., 2020. Attention to global warming. Rev. Financ. Stud. 33 (3), 1112–1145.

DeLong, E.L., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 44 (3), 837–845.

Diebold, F.X., Yilmaz, K., 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. Econom. J. 119 (534), 158–171.

Diebold, F.X., Yilmaz, K., 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. Int. J. Forecast. 28 (1), 57–66.

Estrella, A., Mishkin, F.S., 1998. Predicting U.S. recessions: Financial variables as leading indicators. Rev. Econ. Stat. 80 (1), 45–61.

Friedman, J., Tibshirani, R., Hastie, T., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1–22.

Ghalanos, A., 2014. RACD: Autoregressive conditional density models. r package version 1.0-5. Available for download from: https://bitbucket.org/alexiosg/.

Greiner, M., Pfeiffer, D., Smith, R.D., 2000. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prevent. Vet. Med. 45 (1–2), 23–41.

Gupta, R., Ji, Q., Pierdzioch, C., Plakandaras, V., 2023. Forecasting the conditional distribution of realized volatility of oil price returns: The role of skewness over 1859 to 2023. Finance Res. Lett. 58 (Part C), 104501.

Gupta, R., Olson, E., Wohar, M.E., 2017. Forecasting key US macroeconomic variables with a factor-augmented qual VAR. J. Forecast. 36 (6), 640–650.

Gupta, R., Pierdzioch, C., Salisu, A.A., 2022. Oil-price uncertainty and the U.K. unemployment rate: A forecasting experiment with random forests using 150 years of data. Resour. Policy 77, 102662.

Gupta, R., Wohar, M.E., 2017. Forecasting oil and stock returns with a qual VAR using over 150 years of data. Energy Econ. 62, 181–186.

Hansen, B.E., 1994. Autoregressive conditional density estimation. Internat. Econom. Rev. 35 (3), 705–730.

Huang, Y.-F., Startz, R., 2020. Improved recession dating using stock market volatility. Int. J. Forecast. 36 (2), 507–514.

Kilian, L., Vigfusson, R.J., 2017. The role of oil price shocks in causing U.S. recessions. J. Money Credit Bank. 49 (8), 1747–1776.

Kim, W.C., Fabozzi, F.J., Cheridito, P., Fox, C., 2014. Controlling portfolio skewness and kurtosis without directly optimizing third and fourth moments. Econom. Lett. 122 (2), 154–158.

Kurov, A., Olson, E., Zaynutdinova, G.R., 2022. When does the fed care about stock prices? J. Bank. Financ. 142, 106556.

Langlois, H., 2020. Measuring skewness premia. J. Financ. Econ. 135 (2), 399–424.

Lei, X., Jalkh, N., Yahua, X., Zhang, H., 2021. Spillovers in higher moments and jumps across US stock and strategic commodity markets. Resour. Policy 72, 102060.

Liu, W., Moench, E., 2016. What predicts US recessions? Int. J. Forecast. 32 (4), 1138–1150.

Loycsa, Š., Výrost, T., Plíhal, T., 2021. A tale of tails: New evidence on the growth-return nexus. Finance Res. Lett. 38, 101526.

Ludvigson, S.C., Ma, S., Ng, S., 2021. Uncertainty and business cycles: Exogenous impulse or endogenous response? Am. Econ. J.: Macroecon. 13 (4), 369–410.

Manela, A., Moreira, A., 2017. News implied volatility and disaster concerns. J. Financ. Econ. 123 (1), 137–162.

Mei, D., Liu, J., Ma, F., Chen, W., 2017. Forecasting stock market volatility: Do realized skewness and kurtosis help? Phys. A 481, 153–159.

Nekhili, R., Bouri, E., 2023. Higher-order moments and co-moments' contribution to spillover analysis and portfolio risk management. Energy Econ. 119, 106596.

Pierdzioch, C., Gupta, R., 2020. Uncertainty and forecasts of U.S. recessions. Stud. Nonlinear Dyn. Econom. 24 (4), 20180083.

Plakandaras, V., Cunado, J., Gupta, R., Wohar, M.E., 2017. Do leading indicators forecast U.S. recessions? A nonlinear re-evaluation using historical data. Int. Finance 20 (3), 289–316.

Plakandaras, V., Gogas, P., Papadimitriou, T., Gupta, R., 2019. A re-evaluation of the term spread as a leading indicator. Int. Rev. Econ. Finance 64, 476–492.

R Core Team, 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, Available for download from: https://www.R-project.org/.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., 2011. pROC: An open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12, 77, http://www.biomedcentral.com/1471-2105/12/77/.

Tay, J.K., Narasimhan, B., Hastie, T., 2023. Elastic net regularization paths for all generalized linear models. J. Stat. Softw. 106 (1), 1–31.

Zhang, H., Jin, C., Bouri, E., Gao, W., Xu, Y., 2022. Realized higher-order moments spillovers between commodity and stock markets: Evidence from China. J. Commod. Mark. 100275.