University of Pretoria
Department of Economics Working Paper Series

Predicting International Equity Returns: Evidence from Time-Varying Parameter Vector Autoregressive Models

Rangan Gupta
University of Pretoria
Florian Huber
WU Vienna University
Philipp Piribauer
Austrian Institute of Economic Research (WIFO)
Working Paper: 2018-26
April 2018

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413

Predicting International Equity Returns: Evidence from Time-Varying Parameter Vector Autoregressive Models

Rangan Gupta*[1], Florian Huber[2], and Philipp Piribauer[3]

[1]Department of Economics, University of Pretoria
[2]Department of Economics, WU Vienna University of Economics and Business
[3]Austrian Institute of Economic Research (WIFO)

April 18, 2018

Abstract

In this paper, we forecast monthly stock returns of eight advanced economies using a time-varying parameter vector autoregressive model (TVP-VAR). Compared to standard TVP-VARs, our proposed model automatically detects whether time-variation in the parameters is needed through the introduction of a latent threshold process that is driven by the absolute size of parameter changes. The advantage of this framework is that it can dynamically detect whether a given regression coefficient is constant or time-varying during distinct time periods. We moreover compare the performance of this model with a wide range of nested alternative time-varying and constant parameter VAR models. Our results indicate that the threshold TVP-VAR outperforms its competitors in terms of point and density forecasts. A portfolio allocation exercise confirms the superiority of our proposed framework. In addition, a copula-based analysis also indicates that it pays off to adopt a multivariate modeling framework, especially during periods of stress, like the recent financial crisis.

Keywords: International equity markets, Time-varying vector autoregression, Point and density forecasts, Portfolio allocation.

JEL Codes: C32, G10, G17.

1 Introduction

The existing literature on forecasting international stock returns (for developed and developing economies alike), based on a wide array of models and predictors is vast (see, for example, Rapach

*Corresponding author: Rangan Gupta, Department of Economics, University of Pretoria, Pretoria, 0002, South Africa; E-mail address: rangan.gupta@up.ac.za.

et al. 2005, Welch and Goyal 2008, Rapach et al. 2013, , Rapach and Zhou 2013, Sousa et al. 2016, Aye et al. 2017, Jordan et al. 2017a, Jordan et al. 2017b, among others). While practitioners in finance require real-time forecasts of stock returns for asset allocation, academics are particularly 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). However, stock return forecasting is highly challenging, since it inherently contains a sizeable unpredictable component. The resulting predictive performances therefore usually strongly depend on the chosen indices, sample periods, models and potential predictors.

Recent literature has identified at least two common features of successful models as a means to outperforming standard benchmark specifications in terms of predictive accuracy. First, a large information set in terms of a vast number of predictors (i.e., macroeconomic, financial, technical, institutional, behavioral; see Rapach and Wohar 2006, Rapach et al. 2010, Gupta et al. 2014, Gupta et al. 2017a, or Gupta et al. 2017b) appears to be required in order to successfully challenge standard random walk forecasts. Second, stock return predictions using a-theoretical techniques (which tend to exploit information on the recent behaviour of stock prices based on statistical approaches, and machine learning and computational intelligence techniques) typically tend to perform better than theoretically motivated empirical models (see, for example, Chen et al. 2003; Enke and Thawornwong 2005).

Against this backdrop, and particularly building on the latter point mentioned above, the objective of our paper is to forecast stock returns of eight developed markets (Canada, France, Germany, Italy, Japan, Switzerland, the United Kingdom, and the United States) based on time-varying parameter vector autoregressive (TVP-VAR) models. The choice of these eight equity markets are quite natural given their importance in the global economy, with these countries representing nearly two-third of global net wealth, and nearly half of world output. Note that the decision to look at only the past value of stock returns of the various economies in the model emanates from the evidence in favor of increased co-movement between asset prices, and stock markets in particular, due to financial integration across economies (Diebold and Yilmaz 2009; Diebold and Yilmaz 2009). While this regularity has received considerable attention in the academic literature on the dynamics of stock markets, for some reason, this has not been exploited to its fullest in the forecasting literature. An exception to this is the recent work by Huber et al. (2017), which advocates the use of large Bayesian vector autoregressive (BVAR) model specifications with common stochastic volatility as a means to forecasting monthly global equity indices. The time-varying specification of the covariance structure moreover accounts for sudden shifts in the level of volatility. In an out-of-sample exercise, the proposed model specification is moreover shown to markedly outperform the random walk for both point and density forecasts. In addition, it is well-established that stock market movements serve as a leading indicator for the wider economy (Stock and Watson 2003; Gupta and Hartley 2013; Plakandaras et al. 2017). Hence, we do not incorporate the information of any other predictors in our multivariate models, barring the lagged stock returns of the domestic and foreign economies. This helps us to ensure that our search for the best-suited forecasting model for stock returns can be provided independent of fundamentals.

Following Huber et al. (2017), but realizing the well-established fact that stock returns evolve in a nonlinear fashion (McMillan 2005), we extend the constant parameter approach of Huber et al. (2017), into a time-varying framework. Specifically, we use a flexible variant of a TVP-VAR model. This framework, as recently proposed by Huber et al. 2016, allows one to dynamically detect whether a given regression coefficient is constant or time-varying by combining ideas from the literature of latent threshold and mixture innovation models. In particular, this treshold TVP-VAR (TTVP-VAR) approach introduces a set of latent thresholds that controls the degree of time-variation separately for each parameter at each point in time. The proposed framework incorporates a wide variety of competing models, like the standard time-varying parameter model, a change-point model with an unknown number of regimes, mixtures between different models, and also the simple constant parameter model. Finally, to assess systematically, in a data-driven fashion, which predictors should be included in the model, Huber et al. 2016 impose a set of Normal-Gamma priors on the initial state of the system.

While this is our primary proposed framework for forecasting equity returns of the eight advanced economies, we compare the performance of this model with a wide range of nested alternative time-varying and constant parameter vector autoregressive models. To the best of our knowledge, this is the first paper to produce point and density forecasts of equity returns using time-varying approaches, and in particular, a threshold time-varying vector autoregressive model.

The remainder of the paper is organized as follows: Section 2 outlines the main econometric model used in our forecasting exercise, while Section 3 presents the data and results, with the latter including also a portfolio exercise. Finally, Section 4 concludes the paper.

2 Econometric framework

2.1 A time-varying parameter VAR for modeling international equity returns

Our goal is to construct a model that incorporates international linkages in financial markets. To this end, we postulate that the growth rate of a set of N equity price indices in $\{y_t\}_{t=1}^T$ follows a time-varying parameter VAR,

$$y_t = (I_M \otimes x_t')\beta_t + Q_t v_t, \quad v_t \sim \mathcal{N}(0, H_t), \tag{2.1}$$

with

- $x_t = (y_{t-1}, \dots, y_{t-p}, 1)'$ being a $M = pN + 1$ vector of lagged endogenous variables as well as an intercept term,[1]

[1] In the empirical application, we use a single lag of y_t based on computational reasons.

- β_t is a $K = NM$-dimensional vector of dynamic regression coefficients. Notice that β_t is in principle allowed to evolve over time according to some specific law of motion described in Section ADD.
- v_t is a set of N white noise shocks that follow a multivariate Gaussian distribution and Q_t is a lower uni-triangular (i.e. lower triangular with unit diagonal) matrix of dimension $N \times N$. We store the $v = N(N-1)/2$ free elements in Q_t in a v-dimensional vector q_t.
- Finally, $H_t = \text{diag}(e^{h_{1t}}, \ldots, e^{h_{Nt}})$ is a diagonal matrix that stores the variance parameters.

The model in Eq. (2.1) is the observation equation of a multivariate state space model that has been proposed in Primiceri (2005) and Cogley and Sargent (2005). Notice that the elements in β_t, v_t and $h_t = (h_{1t}, \ldots, h_{Nt})$ feature a specific law of motion. Typically, researchers assume that the states evolve according to a random walk with an unrestricted state innovation variance covariance matrix. This, however, potentially leads to overfitting issues that might be detrimental for forecasting accuracy. To circumvent such issues, we follow cite TTVP and use a parsimonious law of motion for the latent states.

2.2 A parsimonious law of motion for the coefficients

We complete the model description by outlining a law of motion for β_t, q_t and h_t. Following Huber et al. (2016) we assume that the elements of $\xi_t = (\beta_t', q_t')'$, ξ_{it} $(i = 1, \ldots, v+K)$ follow a random walk process,

$$\xi_{it} = \xi_{it-1} + \sqrt{\theta}_{jt}\eta_{jt}, \quad \eta_{jt} \sim \mathcal{N}(0,1), \tag{2.2}$$

where θ_{jt} is a time-varying process innovation variance that follows

$$\theta_{jt} = d_{jt}\sigma^2_{j0,t} + (1-d_{jt})\sigma^2_{j1,t}, \tag{2.3}$$

with $\sigma^2_{j0,t} \gg \sigma^2_{j1,t}$ and

$$d_{jt} = \begin{cases} 1 & \text{if} \quad |\Delta\xi_{it}| > c_j \\ 0 & \text{if} \quad |\Delta\xi_{it}| \leq c_j. \end{cases} \tag{2.4}$$

Equations (2.3) and (2.4) imply that if the absolute change in ξ_{it} is sufficiently large (i.e. exceeds a threshold c_j), the indicator d_j equals unity and a rather large process innovation variance $\sigma^2_{j0,t}$ is adopted. By contrast, if the change is too small (i.e. below c_j), the process innovation variance is close to zero ($\sigma^2_{j1,t} \approx 0$) and thus the change in the parameters ξ_{jt} is small, i.e. $|\Delta\xi_{jt}| \approx 0$. One key advantage of the proposed specification is that if parameter movements appear to be rather small, we effectively zero them out while we allow for large swings. This strikes a balance between using a model with a few regimes as opposed to a model with many regimes ($T-1$ in the case of an unrestricted TVP model). In this paper, we follow TTVP and set $\sigma^2_{j1,t} = 10^{-5} \times \hat{\sigma}^2_{jt}$, with $\hat{\sigma}^2_{jt}$ denoting the OLS variance from a time-invariant VAR model.

For h_t, we follow Kastner and Frühwirth-Schnatter (2014) and assume that the log-volatilities follow an AR(1) process,

$$h_{jt} = \mu_j + \rho_j(h_{jt-1} - \mu_j) + \epsilon_{jt}, \quad \epsilon_{jt} \sim \mathcal{N}(0, \varsigma_j^2), \tag{2.5}$$

for $j = 1, \dots, N$. Hereby we let μ_j denote the unconditional mean of h_{jt}, ρ_j the autoregressive parameter and ς_j^2 is the variance of the log-volatility process.

2.3 Prior specification

Our prior setup closely follows Hotz-Behofsits et al. (2018). More specifically, we use weakly informative Gamma priors on $\sigma^{-2}_{j1,t} \sim \mathcal{G}(0.001, 0.001)$ and uniform priors on the thresholds,

$$c_j|\sigma_{j0,t} \sim \mathcal{U}(\pi_0\sigma_{j0,t}, \pi_1\sigma_{j0,t}). \tag{2.6}$$

We specify $\pi_0 = 0.1$ and $\pi_1 = 1.5$. This prior choice bounds the thresholds away from zero, implying that high frequency movements in ξ_{jt} are effectively shrunk towards zero.

On the initial state ξ_{j0} we use a normal-gamma (NG) shrinkage prior (see Griffin and Brown, 2010),

$$\xi_{j0}|\tau_j^2 \sim \mathcal{N}(0, \tau_j^2), \quad \tau_j^2 \sim \mathcal{G}(\delta, \delta\lambda/2), \lambda \sim \mathcal{G}(n_0, n_1). \tag{2.7}$$

The scaling parameters τ_j^2 follow a Gamma distribution that depends on δ and λ. The hyperparameter δ controls the excess kurtosis of the marginal prior obtained by integrating out the local scaling parameters τ_j^2. Small values imply a heavy tailed prior that allows for non-zero values of ξ_{j0} in the presence of a large global shrinkage parameter λ. The parameter λ pulls all elements in ξ_0 to zero. Given its importance we use an additional Gamma prior and consequently infer λ from the data. In what follows we set $\delta = 0.1$ and $n_0 = n_1 = 0.01$, introducing significant amounts of shrinkage but at the same time allow for heavy tails and thus sufficient flexibility to capture signals.

Finally, we follow Kastner and Frühwirth-Schnatter (2014) and use a weakly informative Gaussian prior on μ_j, a Beta prior on $(\rho_j + 1)/2 \sim \mathcal{B}(25, 5)$ and a non-conjugate Gamma prior on $\varsigma_j^2 \sim \mathcal{G}(1/2, 1/2)$. This choice translates into a Gaussian prior on $\pm\varsigma_j \sim \mathcal{N}(0, 1)$.

Estimation of the model is carried out using Markov chain Monte Carlo (MCMC) techniques. Our MCMC algorithm simulates the latent states on an equation-by-equation basis using forward-filtering backward-sampling (FFBS) techniques (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994). The thresholds are simulated using a Griddy Gibbs step that is based on constructing an approximation to the cumulative distribution function of the conditional posterior of c_j and then perform inverse transform sampling. Here it suffices to say that this is computationally straightforward since the conditional posterior is proportional to the density of a univariate Gaussian distribution times the uniform prior. The state innovation variances $\sigma^2_{j0,t}$ are simulated from an inverted Gamma distribution that takes a standard form. Last, the log-volatilities and the parameters of the state equation are obtained by using the algorithm outlined in Kastner and Frühwirth-Schnatter (2014) and implemented in the R package stochvol (Kastner, 2016).

The algorithm is repeated 20,000 times with the first 15,000 draws being discarded as burn-in. Convergence appears to be no issue with average inefficiency factors well below 30 in almost all cases. Repeated estimation based on randomly initializing certain coefficients also indicates that our algorithm performs well empirically.

3 Forecasting international equity returns

3.1 Data overview and model specification

The data used in this paper are monthly stock price indices of eight industrialized economies namely, Canada (S&P 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 United Kingdom (FTSE All Share Index), and the United States (S&P500 Index). The data on the indices are derived from the Global Financial Database. The stock price indices are converted into log returns, i.e., the first-difference of the natural logarithm of the indices multiplied by 100 to convert into percentage. The period covered in our paper is 1997:M03 to 2017:M02. While the end date corresponds to data availability at the time of writing this paper, the starting date is chosen to ensure that we have a decade each of data around the global financial crisis of 2007. Also note that, 1997 is the onset of the Dot-com bubble, and hence is a good starting point with the equity markets being in turmoil.

3.2 Competing models and design of the forecasting exercise

Our empirical forecasting design is recursive. This implies that we specify an initial estimation period of 1997:M03 to 2002:M02, and compute the one-step-ahead predictive densities for 2002:M03. The initial estimation period is then subsequently expanded by a single month and this procedure is repeated until the final observation in the sample (2017:02) is reached. Forecasts are then evaluated using log predictive scores (LPS) motivated in, for instance, Geweke and Amisano (2010).

To assess the merits of our empirical model we include a wide range of nested alternatives. The first one is a variant of the TVP-VAR with SV proposed in Primiceri (2005). The main differences stem from the fact that we use shrinkage priors on the initial state of the system and Gamma priors on the inverse of the state innovation variances. Notice that this model is nested within our approach by setting $d_{jt} = 1 \; \forall j, t$. The next model considered is a TVP-VAR with SV but with shrinkage priors on all regions of the parameter space. This model introduces a NG shrinkage prior on both, the initial state and the state innovation variances.[2] Moreover, we include three constant parameter VARs with SV. The first one uses a non-conjugate Minnesota prior where the tuning parameters are integrated out in a Bayesian fashion and the second one is a VAR with a NG shrinkage prior (see Huber and Feldkircher, 2017). The third specification is the VAR equipped with a SSVS prior proposed in George et al. (2008). Finally, inclusion of an AR(1) model and a random walk enable assessing whether using a multivariate framework pays off in predictive terms. All models considered feature a stochastic volatility specification in the errors.

3.3 Forecasting results

Table 1 presents a summarization of the out-of-sample predictive performance for the model frameworks under scrutiny. The bottom panel of the table provides summary metrics for the model-specific forecasting performance in terms of point forecasts. Specifically, we focus on well-known root mean-squared forecasting errors (RMSE) as a means to compare the predictive accuracy among the models considered. For all the models under scrutiny, point forecasts have been produced using the respective posterior mean estimates of the parameters of interest. The summary metrics for the point predictions provided in the bottom panel of the table are moreover standardized with respect to random-walk forecasts. Values below unity thus indicate predictive outperformance relative to random-walk predictions in terms of point predictions, while values above one indicate underperformance.

The top panel of Table 1 provide summary metrics for the respective out-of-sample forecast performance in terms of marginal log predictive scores (LPS). While standard RMSEs only focus on point forecasts, log predictive scores provide a well known measure for comparing forecast performance by explicitly accounting for the entire predictive density. As such, log predictive scores aim to enrich the comparison of predictive performances by also accounting for predictive uncertainty. Similar to the point forecasts, Table 1 also presents the respective measures for the density predicitons relative to random-walk forecasts. Specifically, negative values indicate underperformance relative to the random-walk benchmark, and conversely, positive values for the respective density forecasts indicate outperformance.

As expected, overall results for point predictions show that random-walk forecasts are particularly poor as compared to the other specifications. RMSEs of almost all indices under scrutiny appear to be well below unity. Also density forecasts show a relatively poor predictive performance of the (stochastic volatility-augmented) random walk specifications. The autoregressive specification with stochastic volatility (AR-SV), however, appears to perform much better. The table shows that the proposed modeling framework sketched above (TTVP) appears to perform particularly well

[2] For a detailed description of the model and the prior setup, see Feldkircher et al. (2017) and Hotz-Behofsits et al. (2018).

in terms of producing both accurate point predictions as well as density predictions. Almost in all indices under scrutiny, the proposed specifications ranks among the best performing approaches.

Interestingly, the estimation framework without a threshold (TVP) appears to produce forecasts which are much more imprecise. For all equity indices considered, TVP markedly underperforms relative TTVP. This finding holds true for both RMSEs and log predictive scores. Accounting for potential structural breaks in the equity trajectories thus appears particularly beneficial. Due to potential problems of overfitting the data, for some series, the standard TVP setup even hardly manages to outperform random walk predictions. The over-parameterization problem in the standard TVP specification results in the poorest predictive performance among the competing specifications under scrutiny (other than the random walk).

While a stand-alone TVP parameterization seems to perform relatively poor due to overfitting problems, the additional use of Bayesian shrinkage in terms of a normal-gamma prior (TVP NG) appears to markedly improve forecast performance. For all country-specific indices in the sample, TVP NG produces more precise out-of-sample forecasts as compared to TVP for both RMSEs as well as log predictive scores. However, comparing results for TVP NG with those of TTVP shows that TTVP tends to slightly outperform the former. This outperformance is particularly pronounced when focusing on point predictions. For all indices considered TTVP produces slightly lower overall RMSEs as compared to TVP NG. This finding also translates to density forecast metrics. Only for United Kingdom, TVP NG appears to slightly outperform TTVP in terms of log predictive scores. The joint performance for density forecasts, however, also indicates TTVP as performing best.

Turning attention to the remaining constant-parameter specifications (Minn-VAR, NG-VAR and SSVS) shows that all of these appear to outperform the standard TVP specification both in terms of point as well as density predictions. Among the three specifications, the vector autoregressive model using a Minnesota prior specification (Minn-VAR) appears to produce the most precise forecasts. The constant-parameter VAR specification with a normal-gamma shrinkage prior specification (NG-VAR) only slightly underperforms relative to the Minnesota specification. A much more notable, however, not particularly remarkable drop in forecast performance relative to the Minnesota specification, can be seen in the predictive summary metrics for the VAR specification using a stochastic search variable selection (SSVS) prior.

Comparing the normal-gamma shrinkage setting with time-variation in the parameters (TVP NG), with its constant-parameter counterpart (NG-VAR) reveals that the former slightly outperforms the latter in terms of joint log predictive scores. However, inspection of the index-specific predictive performances in terms of log predictive scores reveals that the constant parameter setting slightly outperforms TVP NG in almost all cases. Turning attention to the metrics on point predictions corroborate these findings. NG-VAR produces lower RMSEs as compared to TVP NG for all indices under scrutiny.

While Table 1 presents overall metrics for out-of-sample forecast performance, Figure 1 depicts the evolution of the cumulative log predictive scores for the specifications under scrutiny over

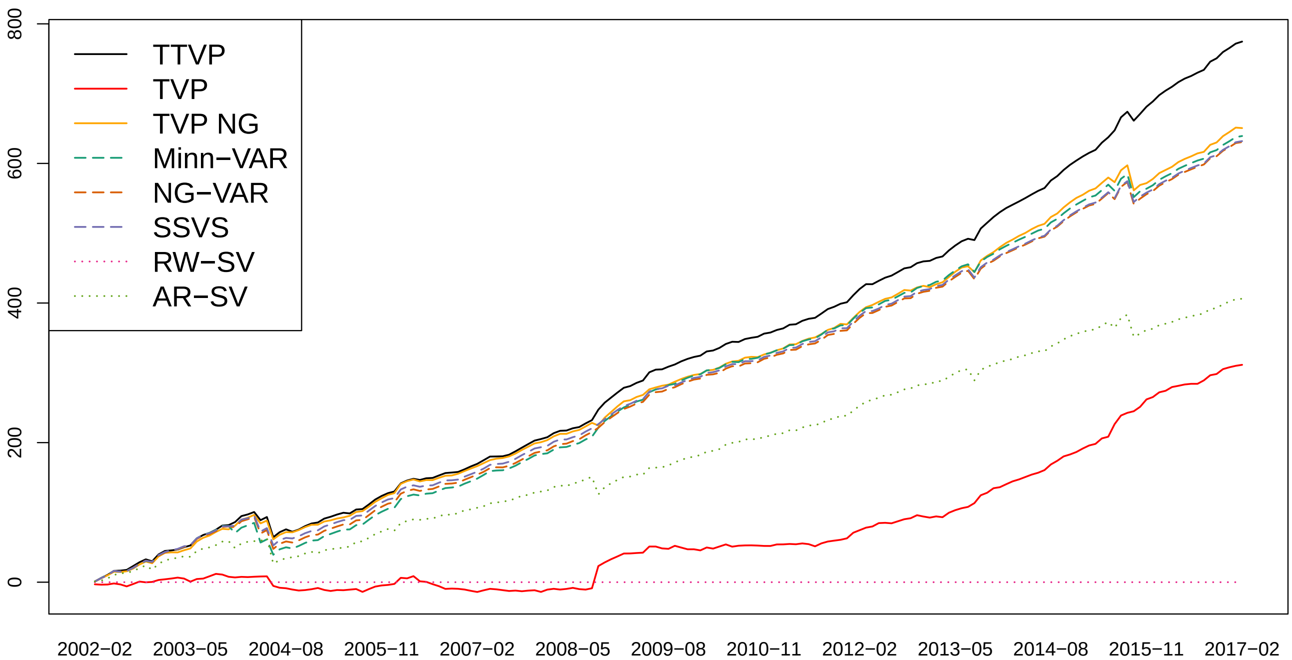


Fig. 1: Joint log predictive scores relative to the RW-SV model

Table 1: Evaluation of point and density forecasts

	TTVP	TVP	TVP NG	Minn-VAR	NG-VAR	SSVS	AR-SV
			Log predictive scores				
Joint	774.593	311.458	650.631	639.294	630.625	631.966	406.095
UK	27.991	-16.812	28.293	40.246	37.624	40.142	36.797
CA	47.283	-9.371	46.907	47.618	47.486	43.990	50.564
FR	62.280	12.192	58.153	60.542	58.912	59.206	65.620
JP	50.883	-1.012	47.285	48.972	50.877	47.789	54.153
DE	66.972	14.706	47.888	62.608	62.697	64.259	55.989
IT	57.147	10.255	38.962	30.521	41.295	38.934	46.380
US	37.727	-26.596	33.194	28.591	28.288	26.950	40.829
CH	46.966	-15.535	46.768	45.382	44.029	39.767	50.837
			Root mean square errors				
UK	0.720	0.941	0.726	0.719	0.720	0.721	0.721
CA	0.727	1.013	0.733	0.722	0.722	0.726	0.727
FR	0.711	0.905	0.721	0.701	0.705	0.703	0.700
JP	0.728	1.040	0.743	0.720	0.723	0.723	0.723
DE	0.710	0.905	0.711	0.696	0.696	0.697	0.702
IT	0.730	0.917	0.735	0.715	0.732	0.742	0.732
US	0.745	1.023	0.752	0.739	0.740	0.740	0.748
CH	0.697	0.936	0.700	0.687	0.692	0.692	0.694

time. The model-specific cumulative log predictive scores are measured relative to the random-walk benchmark, which is given by the zero line. Country-specific performance profiles are provided in the Appendix. Figure 1 corroborates the overall finding of Table 1, showing a clear outperformance of the proposed model framework (TTVP) as compared to the alternative specifications. In the beginning of the sampling period, a time-varying parameter specification with a normal-gamma shrinkage prior (TVP NG) appeared to slightly outperform TTVP. However, around the year 2009, TTVP appeared to supersede the alternative specifications in terms of forecast performance. During the economic and financial turmoils (2003/2004 and 2008/2009) the modeling framework without the threshold specification (TVP), appeared to perform particularly well in terms of density predictions. This is mainly due to the overall larger variance in the predictive densities of TVP, resulting in a less severe penalization of large forecast errors as compared to competing specifications.

3.4 Dissecting the log predictive likelihood

The previous section highlighted sustained predictive gains for most multivariate models considered. One additional question that typically arises centers on whether the gains stem from a positive feedback on the univariate marginal predictive densities or arise from joint modeling all elements in y_t simultaneously. To this end, we follow Dovern et al. (2016) and decompose the joint log predictive likelihood of model j as follows,

$$\log p(y_{t+1}|y_{1:t}, \mathcal{M}_j) = \sum_{j=1}^{N} \log p(y_{jt+1}|y_{1:t}, \mathcal{M}_j) + \log c\{P^{-1}(y_{1t+1}), \ldots, P^{-1}(y_{Nt+1})\}. \tag{3.1}$$

Herewith, we let $p(y_{t+1}|y_{1:t}, \mathcal{M}_j)$ denote the predictive likelihood of model j, $\mathcal{M}_j$, evaluated at the outcome y^o_{t+1}, $p(y_{jt+1}|y_{1:t}, \mathcal{M}_j)$ are the N univariate marginal predictive likelihoods, $P(y_{jt+1})$ is the corresponding cdf and $c(\bullet)$ denotes the probability density function of a Gaussian copula. Eq. (3.1) indicates that the joint LPS can be decomposed in terms of a sequence of univariate marginal log scores and a copula term that establishes the covariance structure of the predictive density.

Fig. 5 provides a graphical representation of Eq. (3.1) for the three top performing models over time. A few findings are worth mentioning. First, notice that across the (multivariate) models, we observe a pronounced degree of time variation in the contribution of the copula term to the overall predictive likelihood. Especially during periods that are characterized by economic stress such as the recent financial crisis, it seems to pay off to adopt a multivariate model and to exploit information from the cross section in an effective manner. This effect is especially pronounced during the recent financial crisis, where the marginal log predictive scores have been grossly negative and the copula term contributed positive to the joint predictive performance of the model. Second, we also observe some periods with a negative contribution, especially during the first part of the sample. This hints towards periods where most multivariate models fail to adequately recover the predictive covariance structure. Third, and finally, the strong overall predictive performance of the TTVP specification is complemented by a particularly accurate modeling of the predictive covariance structure.

3.5 A portfolio exercise

Next, we assess whether using our multivariate models also leads to better economic performance, as measured by (annualized) Sharpe ratios. More specifically, we assume that the models considered are used to guide the behavior of an investor who aims to invest in all N markets considered. To this end, two trading strategies are consequently adopted.

The first one is based on the well-known global minimum-variance portfolio (GMV) strategy that aims to minimize the portfolio variance. Let $p_{it|t-1}$ denote the mean of the one-step-ahead predictive density of model i, $P_{it|t-1}$ the corresponding predictive variance and w_{it} is a N-dimensional model-specific weight vector (all quantities condition on information up to time $t-1$. The optimization problem is

$$\begin{aligned} &\underset{w_{it}}{\text{minimize}} \quad w_{it} P_{it|t-1} w'_{it} \\ &\text{subject to} \quad \sum_{j=1}^{N} w_{it} = 1. \end{aligned} \tag{3.2}$$

The second strategy, labeled the target global minimum variance portfolio (TGMV) augments Eq. (3.2) by an additional constraint. This constraint states that the weights are chosen such that the portfolio variance is minimized subject to a pre-specified target return τ^*. In what follows we choose three target returns $\tau^* \in \{\frac{10\%}{12}, \frac{15\%}{12}, \frac{20\%}{12}\}$.

Table 2 shows annualized Sharpe ratios across models. In general, using economic evaluation criteria corroborate the findings based on LPS described above. Considering model performance for the different strategies suggests that the TVP NG specification yields the highest Sharpe ratio for the GMV strategy, being closely tracked by a simple AR(1) model with SV. This, however, does not carry over to the TGMV strategy. Here we find that a VAR equipped with the SSVS prior excels. The TTVP specification as well as the remaining linear VARs also display a rather strong performance. Notice that, consistent with the findings based on LPS, the standard TVP model seems to display an inferior performance across the space of different portfolio allocation schemes. Comparing the results for different target returns yields comparable insights except that for $\tau^* = \frac{20\%}{12}$ we find that the TVP NG and the AR(1) model yield Sharpe ratios that are qualitatively identical, outperforming all remaining models by quite large margins.

	GMV	TGMV		
	GMV	$\tau^* = \frac{10\%}{12}$	$\frac{15\%}{12}$	$\frac{20\%}{12}$
TTVP	0.782	0.856	0.808	0.538
TVP	0.613	0.724	0.726	0.715
TVP NG	0.787	0.787	0.787	0.787
Minn-VAR	0.801	0.805	0.820	0.476
NG-VAR	0.829	0.836	0.828	0.569
SSVS	0.828	0.863	0.872	0.628
RW-SV	0.771	0.695	0.687	0.649
AR-SV	0.787	0.787	0.787	0.787

Table 2: Annualized Sharpe ratios across different portfolio allocation strategies

4 Concluding remarks

The empirical regularity of increased co-movement between stock markets due to financial integration across economies has not been exploited to its fullest in the forecasting literature. Against this backdrop, this paper forecasts stock return of eight developed markets namely, Canada, France, Germany, Italy, Japan, Switzerland, the United Kingdom, and the United States based on time-varying parameter vector autoregressive models. In the process, this paper puts forth a novel approach to estimate time-varying parameter models in a Bayesian framework by assuming that the state innovations follow a threshold model, with the threshold variable being the absolute period-on-period change of the corresponding states. This implies that if the (proposed) change is sufficiently large, the corresponding variance is set to a positive value, otherwise it is set close to zero, suggesting that the states remain virtually constant. In the process, this threshold time-varying parameter vector autoregressive (TTVP-VAR) model with stochastic volatility is capable of discriminating between a wide array of competing models, especially that feature many to few structural breaks in the regression parameters. While this is our primary proposed framework for forecasting equity returns of the eight advanced economies, we compare the performance of this model with a wide range of nested alternative time-varying and constant parameter vector autoregressive models. Based on an out-of-sample period of 2002:M03 to 2017:M02, given an in-sample of 1997:M03 to 2002:M02, our results indicate that the TTVP-VAR outperforms its competitors for both point and density forecasts. A portfolio allocation exercise also confirms the superiority of our proposed model. We also observed sustained predictive gains for most multivariate models considered. Given this, and the fact that the joint log predictive scores can be decomposed in terms of a sequence of univariate marginal log scores and a copula term, we indicate that it pays off to adopt a multivariate model and to exploit information from the cross section in an effective manner, especially during periods of stress, like the recent financial crisis. This observation in turn validates our decision to take multivariate approach under the premise of increased co-movement between stock markets due to financial integration.

As part of future research, one could extend our analysis to other financial markets, like bonds and currencies, as well as to commodity markets. Given spillovers across asset classes, it would make sense to incorporate the various asset classes together in the vector autoregressive model proposed here, and then conduct a forecasting exercise.

References

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

Carter CK and Kohn R (1994) On Gibbs sampling for state space models. *Biometrika* 81(3), 541–553

Chen AS, Leung MT and Daouk H (2003) Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Computers & Operations Research* 30(6), 901–923

Cogley T and Sargent TJ (2005) Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic dynamics* 8(2), 262–302

Diebold FX and Yilmaz K (2009) Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal* 119(534), 158–171

Dovern J, Feldkircher M and Huber F (2016) Does joint modelling of the world economy pay off? Evaluating global forecasts from a Bayesian GVAR. *Journal of Economic Dynamics and Control* 70, 86–100

Enke D and Thawornwong S (2005) The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications* 29(4), 927–940

Feldkircher M, Huber F and Kastner G (2017) Sophisticated and small versus simple and sizeable: When does it pay off to introduce drifting coefficients in Bayesian VARs? *arXiv preprint arXiv:1711.00564*

Frühwirth-Schnatter S (1994) Data augmentation and dynamic linear models. *Journal of Time Series Analysis* 15(2), 183–202

George EI, Sun D and Ni S (2008) Bayesian stochastic search for VAR model restrictions. *Journal of Econometrics* 142(1), 553–580

Geweke J and Amisano G (2010) Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting* 26(2), 216–230

Griffin JE and Brown PJ (2010) Inference with normal-gamma prior distributions in regression problems. *Bayesian Analysis* 5(1), 171–188

Gupta R, Hammoudeh S, Modise MP and Nguyen D (2014) Can Economic Uncertainty, Financial Stress and Consumer Sentiments Predict U.S. Equity Premium? *Journal of International Financial Markets, Institutions and Money* 33(3), 367–378

Gupta R and Hartley F (2013) The role of asset prices in forecasting inflation and output in South Africa. *Journal of Emerging Market Finance* 12(3), 239–291

Gupta R, Majumdar A and Wohar ME (2017a) The Role of Current Account Balance in Forecasting the US Equity Premium: Evidence from a Quantile Predictive Regression Approach. *Open Economies Review* 28(2), 47–59

Gupta R, Mwamba Mwamba JW and Wohar ME (2017b) The role of partisan conflict in forecasting the US equity premium: A nonparametric approach. *Finance Research Letters* (https://doi.org/10.1016/j.frl.2017.10.023)

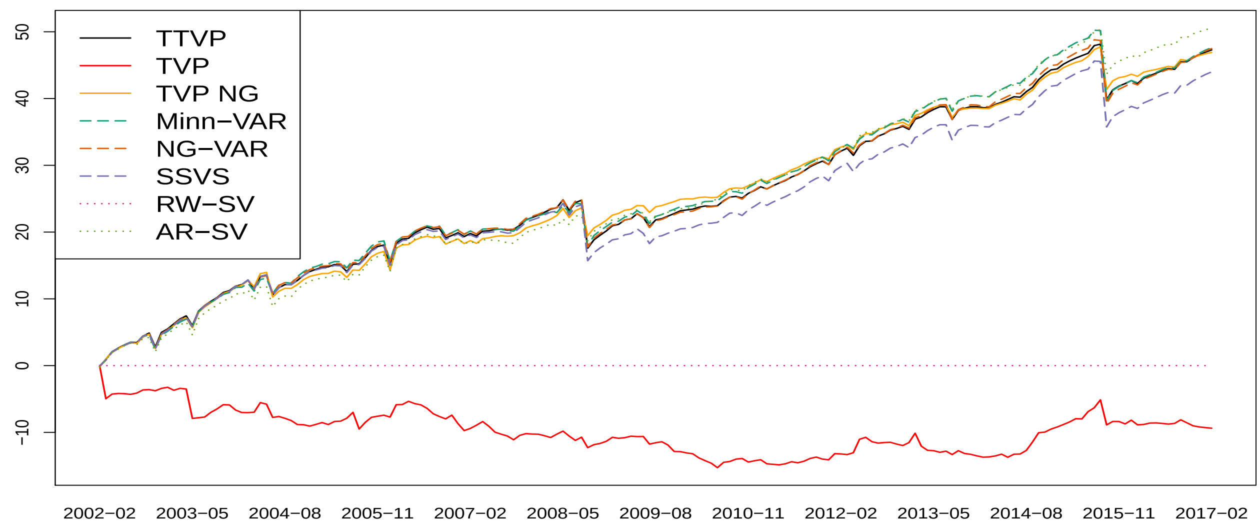
Hotz-Behofsits C, Huber F and Zörner TO (2018) Predicting crypto-currencies using sparse non-Gaussian state space models. *arXiv preprint arXiv:1801.06373*

Huber F and Feldkircher M (2017) Adaptive shrinkage in Bayesian vector autoregressive models. *Journal of Business & Economic Statistics* , 1–13

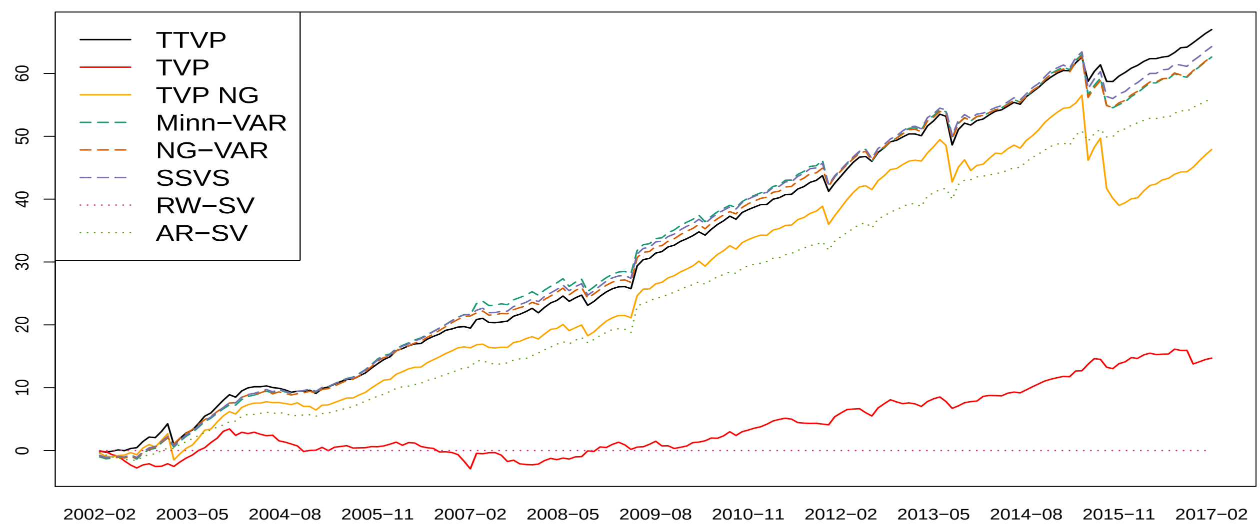
Huber F, Kastner G and Feldkircher M (2016) A New Approach To Time Varying Parameters in Vector Autoregressive Models. *Papers 1607.04532, arXiv.org*

Huber F, Krisztin T and Piribauer P (2017) Forecasting Global Equity Indices Using Large Bayesian

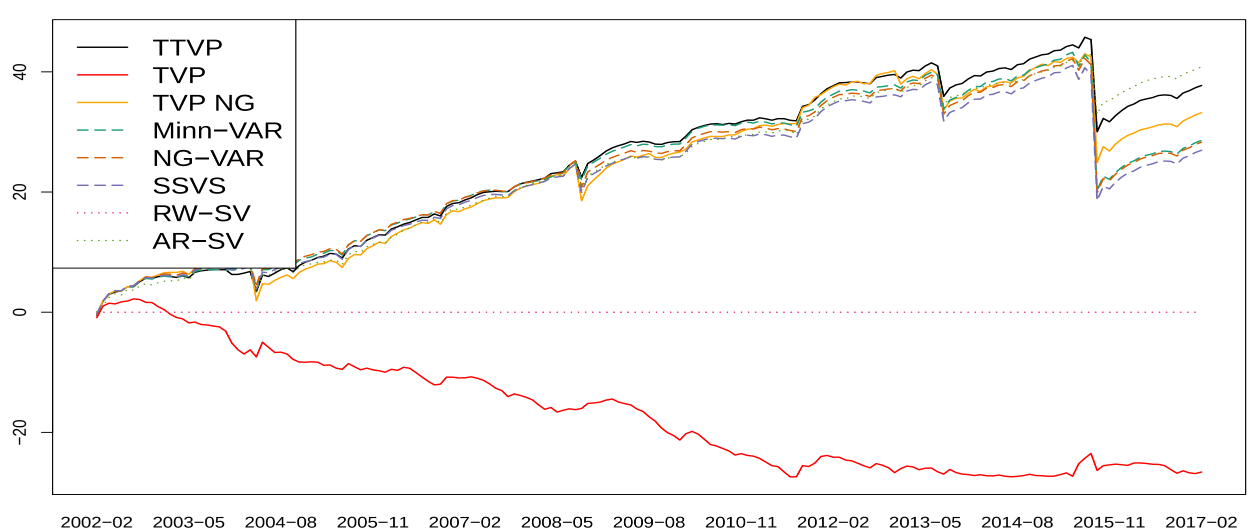
VARs. *Bulletin of Economic Research* 69(3), 288–308
Jordan SJ, Vivian A and Wohar ME (2017a) Forecasting market returns: bagging or combining? *International Journal of Forecasting* 33(1), 102–120
Jordan SJ, Vivian A and Wohar ME (2017b) Stock returns forecasting with metals: sentiment vs. fundamentals. *The European Journal of Finance* (https://doi.org/10.1080/1351847X.2017.1323770), 1–24
Kastner G (2016) Dealing with Stochastic Volatility in Time Series Using the R Package stochvol. *Journal of Statistical Software* 69(5), 1–30
Kastner G and Frühwirth-Schnatter S (2014) Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis* 76, 408–423
McMillan DG (2005) Non-linear dynamics in international stock market returns. *Review of Financial Economics* 14(1), 81–91
Plakandaras V, Cunado J, Gupta R and Wohar ME (2017) Do leading indicators forecast US recessions? A nonlinear re-evaluation using historical data. *International Finance* 20(3), 289–316
Primiceri GE (2005) Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies* 72(3), 821–852
Rapach D and Zhou G (2013) Forecasting stock returns. In *Handbook of Economic Forecasting*, volume 2. Elsevier, 328–383
Rapach DE, Strauss JK and Zhou G (2010) Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies* 23(2), 821–862
Rapach DE, Strauss JK 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 DE and Wohar ME (2006) In-Sample vs. Out-of-Sample Tests of Stock Return Predictability in the Context of Data Mining. *Journal of Empirical Finance* 13(2), 231–247
Rapach DE, Wohar ME and Rangvid J (2005) Macro variables and international stock return predictability. *International Journal of Forecasting* 21(1), 137–166
Sousa RM, Vivian A and Wohar ME (2016) Predicting asset returns in the BRICS: The role of macroeconomic and fundamental predictors. *International Review of Economics & Finance* 41, 122–143
Stock JH and Watson MW (2003) Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* 41(3), 788–829
Welch I and Goyal A (2008) A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *Review of Financial Studies* 21(4), 1455–1508



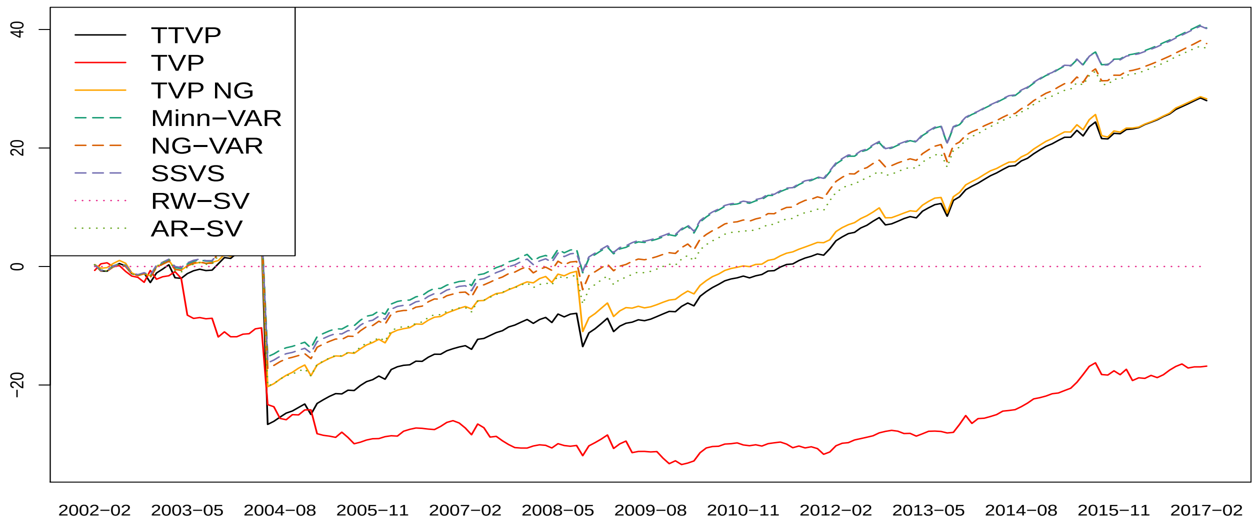
(a) Canada



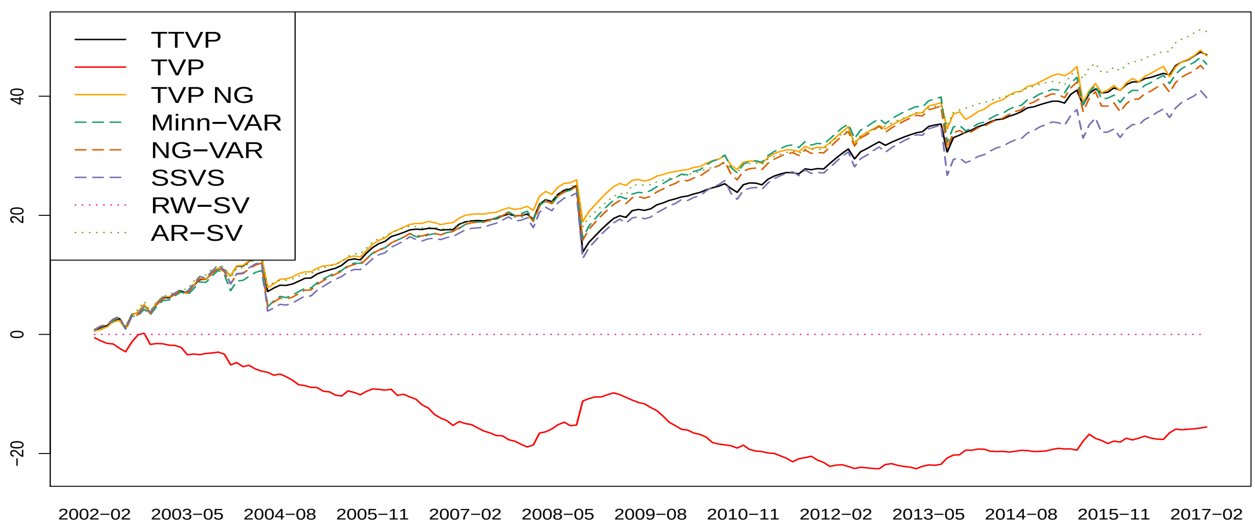
(b) Germany



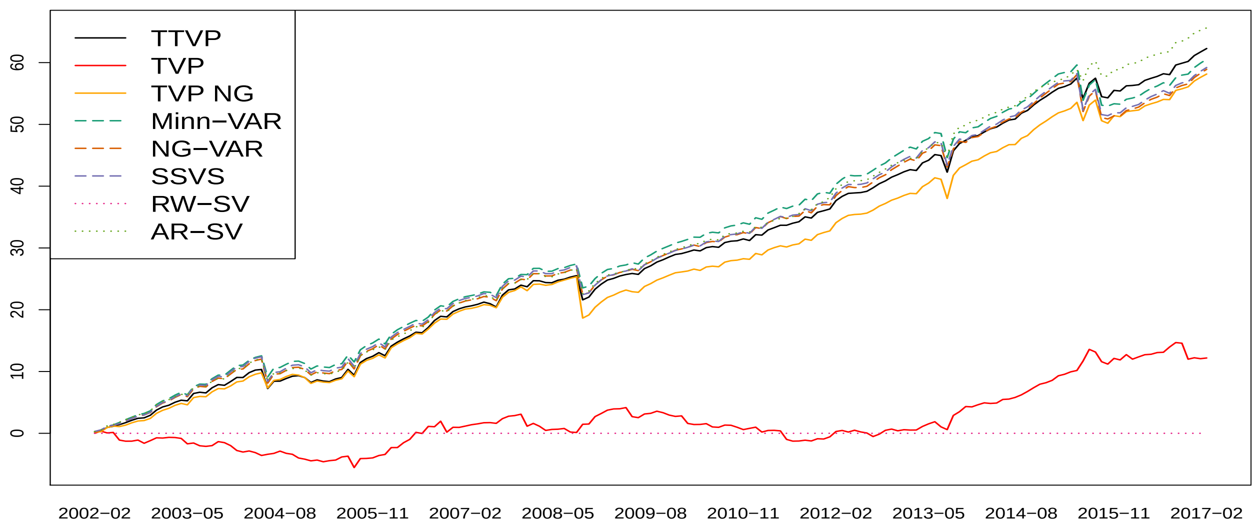
(c) United States



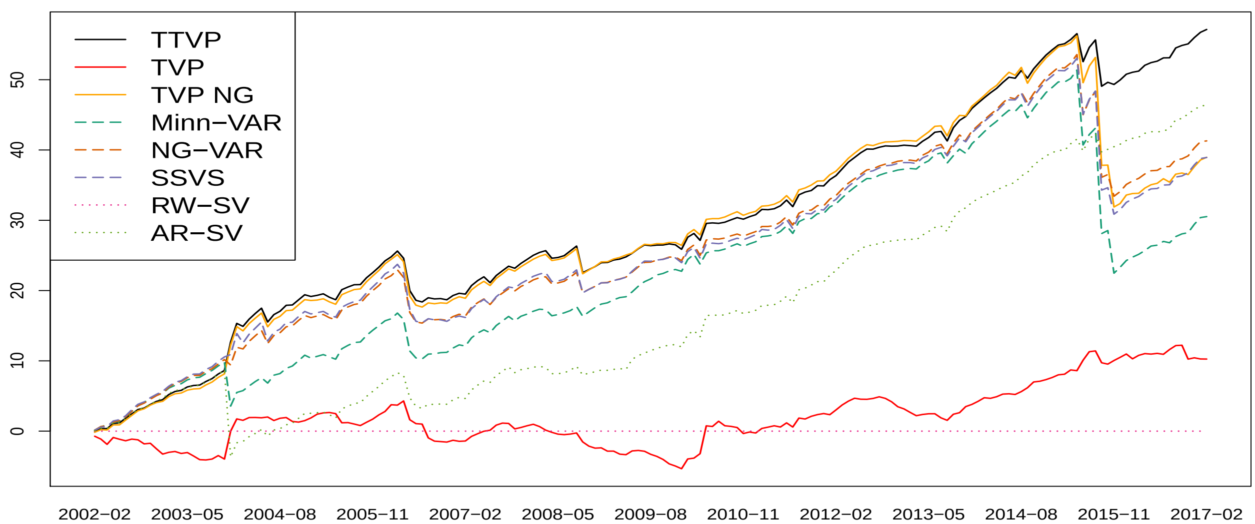
(a) United Kingdom



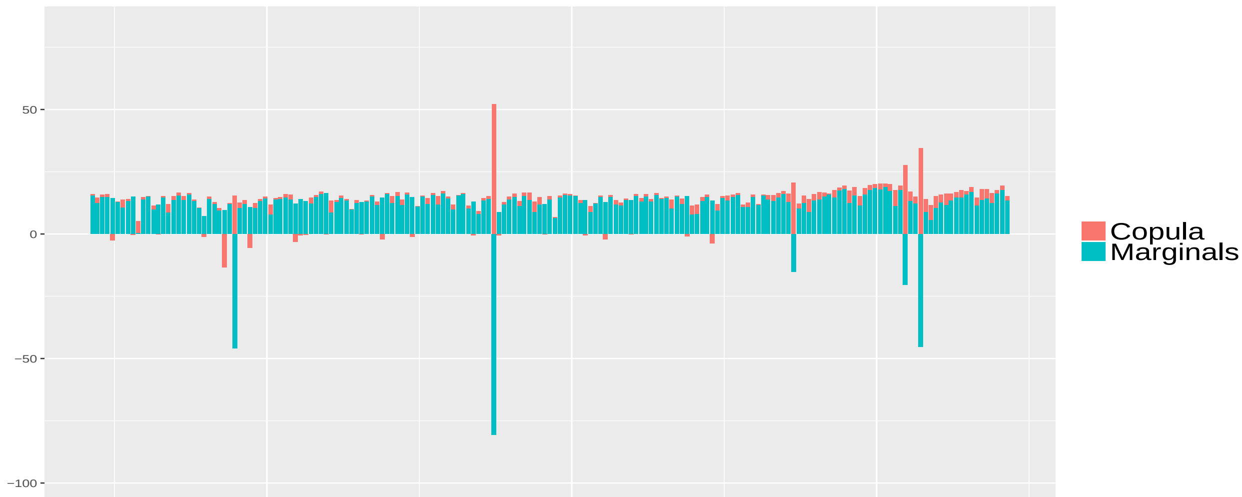
(b) Switzerland



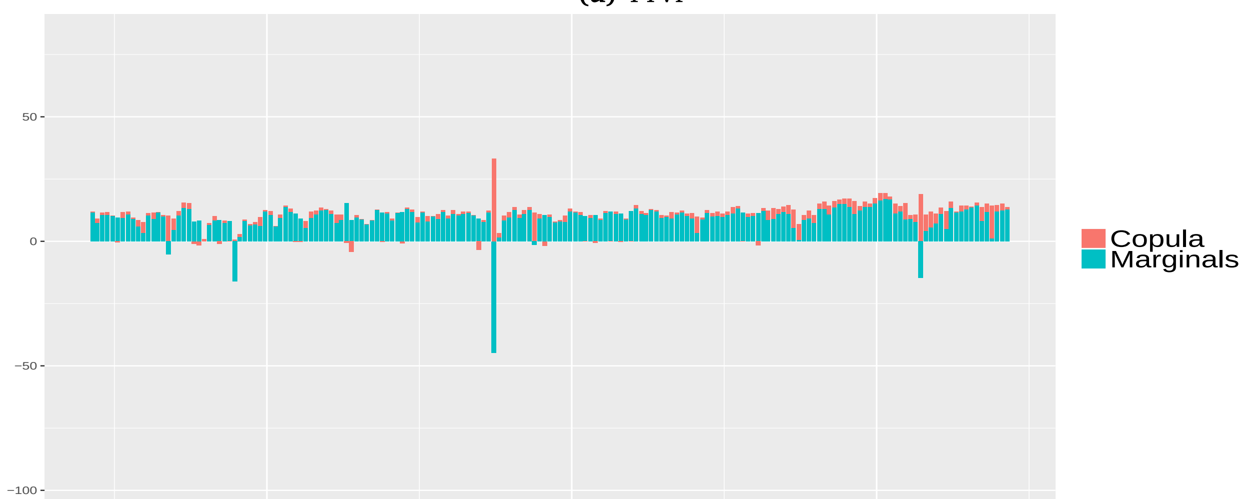
(a) France



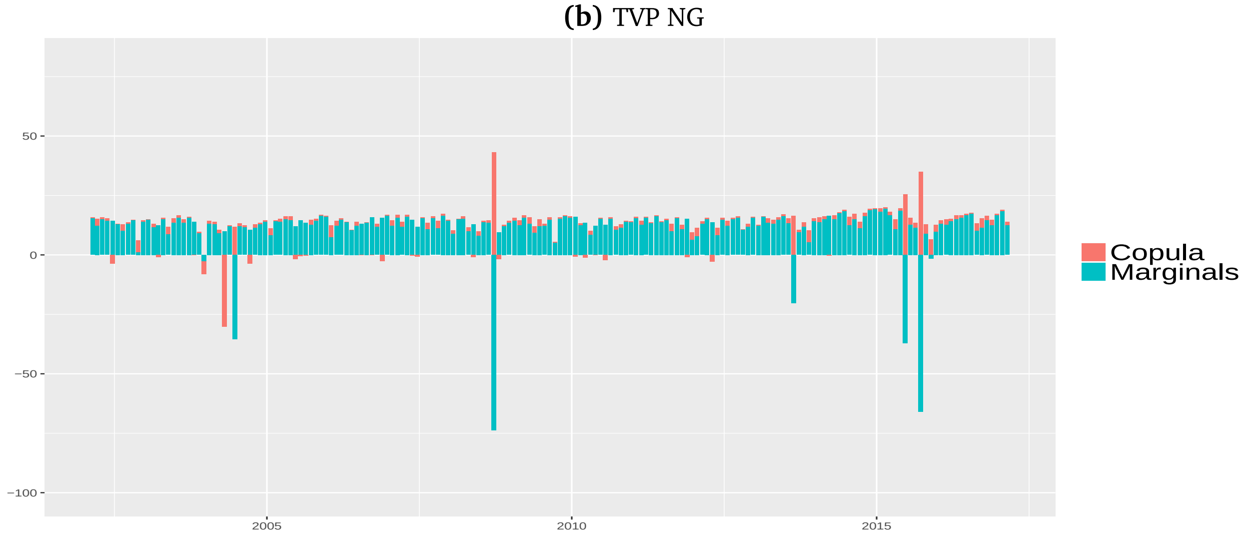
(b) Italy



(a) TTVP



(b) TVP NG



(c) Minn-VAR

Fig. 5: Contribution of the copula term to the joint predictive likelihood