Understanding Term Structure of Variance Swap Rates, Market Return Predictability and Variance Swap Investments When Volatility can Jump

Yi Hong*

International Business School Suzhou, Xi'an Jiaotong-Liverpool University

Xing Jin†

Warwick Business School, University of Warwick

*International Business School Suzhou (IBSS), Xi'an Jiaotong Liverpool University, China; Telephone: +86 512-88161729; Email: yi.hong@xjtlu.edu.cn.

†Warwick Business School, Coventry, CV4 7AL, United Kingdom; Telephone: +44 2476575698; Email: Xing.Jin@wbs.ac.uk.

Understanding Term Structure of Variance Swap Rates, Market Return Predictability and Variance Swap Investments When Volatility can Jump

Abstract: This paper proposes a tractable self-exciting double-jump model for stock return and its variance processes, extending existing two-factor term structure models of variance swap rates in the literature to a new three-factor model. We investigate the capability of this new model in capturing the dynamics of stock return and the information possessed by the variance swap rates. Importantly, our three factors can significantly predict aggregate equity returns as opposed to the ones proposed in the traditional two-factor models in the literature. In stark contrast to the existing literature, our empirical results indicate that variance swap rates have more powerful predictive ability than variance risk premium for the market returns. Unlike the popular double-jump model in the literature, our new model allows us to derive closed-form solutions to the optimal variance swap investment up to solving a set of ordinary differential equations which greatly facilitates new understanding of volatility trading. Specifically, we find that the investor always takes a short-long-short strategy.

JEL Classification: G11

Keywords: Optimal portfolio selection, jump-diffusion models, self-exciting process, variance swap investments, market return predictability

1 Introduction

In order to exploit variance risk premium and hedge variance risk, the variance swap contracts have become the most actively traded variance-related derivative securities due to their direct exposure to volatility[1] and provided good investment opportunities for market

[1]As demonstrated by Liu and Pan (2003), an option can provide indirect exposure to volatility since the option price is also affected by the underlying stock. Our paper also differs from theirs claims in another important aspect in that they make a seemingly unrealistic assumption of constant jump size to solve the optimal portfolio choice in a complete market. In contrast, we do not make this assumption and solve the optimal portfolio choice in semi-closed form in an incomplete market.

participants[2]. Importantly, it has been documented that variance risk premium can significantly predict subsequent stock market returns. And as such, it is vitally important to model variance swap rate term structures, accurately estimate variance risk premium and analytically solve the optimal portfolio choice problem involving variance swaps. Primarily due to the analytical tractability, there is large literature on the variance swap modelling in continuous-time setting. In particular, by using PCA, Egloff, Leippold and Wu (2010) find that two stochastic variance risk factors govern the short and long end of the variance swap term structure variation, respectively. And they show that a two-factor pure-diffusion model can fit variance swap rates well. Furthermore, they explicitly solve the optimal dynamic portfolio choice problem and find that it is generally optimal for an investor to take a short position in a short-term variance swap and a long position in a long-term variance swap. In Filipović, Gourier and Mancini (2015), they show that a bivariate quadratic pure-diffusion model provides a good fit to variance swap rates. In contrast to Egloff, Leippold and Wu (2010), they find that an investor optimally takes a short position in a long-term variance swap to earn the significant negative variance risk premium and a long position in a short-term variance swap to hedge volatility risk. In fact, Filipović, Gourier and Mancini (2015) study the optimal portfolio choice problem involving variance swaps in quadratic pure-diffusion variance swap models and hence the closed-form solution as in Egloff, Leippold and Wu (2010) is unavailable. In particular, in the pure-diffusion models of both Egloff, Leippold and Wu (2010) and Filipović, Gourier and Mancini (2015), only two variance swaps are incorporated in the portfolio choice problem because any two variance swaps can span the two sources of risk in two-factor variance swap rate dynamics and thus a third variance swap is redundant.

[2]In Egloff, Leippold and Wu (2010), the market price of variance risk (γ^v) is around -16 as opposed to the much smaller (in magnitude) market price of return risk (γ^S) of 2.13. To demonstrate the attractiveness of the significantly negative variance risk premium, first, they show that the optimal portfolio weights on the stock index is mainly determined by the market price of return risk (γ^S) and positive in the absence of variance swaps. In contrast, in a model where an investor is allowed to trade S&P500 and variance swaps, they show that an investor takes a short-long strategy in variance swap to gain the negative variance risk premium while, despite the positive stock risk premium, she takes short positions in S&P500 to hedge the short position in a variance swap given the instantaneous negative correlation between the stock index and the variance. Second, as indicated by Table 7 in their paper, the investment strategies involving variance swaps have a Sharpe ratio of at least 1.20 while the Sharpe ratio of S&P500 is at most 0.5.

As widely documented strong empirical evidence, for example, Eraker, Johannes and Polson (2003), Chernov, Gallant, Ghysels and Tauchen (2003), Eraker (2004), Broadie, Chernov and Johannes (2007), Todorov (2009), and Bandi and Reno (2015) among others find strong evidence for co-jumps in volatility and stock returns, i.e., that a big jump in stock prices is likely to be associated with a big jump in volatility. These studies also suggest that jumps do play a key role in explaining the observed risk premium. Aït-Sahalia, Karaman and Mancini (2015) extend the two-factor model in Egloff, Leippold and Wu (2010) to the widely used double-jump model (called "SV2F-PJ-VJ" model in their paper) by incorporating jumps in both stock returns and volatility. Interestingly, like the aforementioned pure-diffusion models, the double-jump model is still a two-factor model. The reason for this is that in the double-jump model in Aït-Sahalia, Karaman and Mancini (2015), the intensity of jumps in variance process is an affine function of variance (e.g., the perfect correlation between the jump intensity and variance). Recently, Santa-Clara and Yan (2010) provides empirical evidence that jump intensity and the variance are largely uncorrelated and do not support the jump intensity model in Aït-Sahalia, Karaman and Mancini (2015).

The aforementioned models are all two-factor models for variance swap rates. Our factor analysis, however, shows that there are exactly three factors underlying variance swap rates. Thus, one of the objectives in the present paper is to propose a three-factor model for the term structure of the variance swap rates. More specifically, we propose a self-exciting double-jump model which disentangles the linear relation between the jump intensity and the variance processes. As a result, the jump intensity is a new factor in addition to the two factors in Egloff, Leippold and Wu (2010) and Aït-Sahalia, Karaman and Mancini (2015). More importantly, we show the following results. First, the excess expected log market return is a linear function of three factors in our model while it is a linear function of two factors in Egloff, Leippold and Wu (2010) and Aït-Sahalia, Karaman and Mancini (2015). Second, in contrast to their results, a third variance swap is not redundant given any two variance swaps available for trading.

The above observations bring us to ask the following questions: How much our three-factor model outperform their two-factor model for predicting market returns? If the link

between the jump intensity and variance in Aït-Sahalia, Karaman and Mancini (2015) is invalid, are two variance swaps still sufficient to span the variance swap market? If not, what does an optimal strategy involving three variance swaps look like as opposed to the short-long rule in Egloff, Leippold and Wu (2010)? In the meantime, what is the benefit from trading a third variance swap? What is the economic cost of ignoring jumps especially in volatility if an investor mistakenly uses a pure-diffusion model in variance swap investments?

More interestingly, we find that our 3-factor predictive model is equivalent to a predictive regression model with three variance swaps as predictors while the two-factor predictive regression model in Egloff, Leippold and Wu (2010) and Aït-Sahalia, Karaman and Mancini (2015) is the same as the one with two variance swaps. This finding prompts us to ask if the market return predictability can be further improved by using more variance swaps. We find that the use of three swap contracts can indeed better predict the stock returns over a range of investment horizons. To the best of our knowledge, we are the first to use the variance swap rates to predict market returns.

To summarize, in contrast to the double-jump stochastic volatility model in Aït-Sahalia, Karaman and Mancini (2015), our model has several attractive features. First, our model outperforms the two-factor jump-diffusion models (e.g., Egloff, Leippold and Wu (2010), Pan (2002) and Aït-Sahalia, Karaman and Mancini (2015)) over the out-of-sample period in terms of pricing variance swaps across all time-to-maturities and predicting stock returns over the time horizon up to two years. Second, unlike the double-jump stochastic volatility model in Aït-Sahalia, Karaman and Mancini (2015), given two variance swaps, a third variance swap is not redundant in the sense that it cannot be replicated by trading these two variance swaps. And thus, this makes the third variance swap valuable for investment. Furthermore, any three variance swaps can span the linear space generated by three sources of risks: short-run variance, long-run variance and jump. Third and more importantly, we obtain a closed-form solution to the optimal investment in variance swaps, which greatly facilitates the new understanding of volatility trading in the double-jump stochastic volatility model especially in comparison with the short-long strategy suggested by Egloff, Leippold and Wu (2010) and the long-short strategy proposed by Filipović, Gourier and Mancini (2015) in

pure-diffusion models. To the best of our knowledge, this paper is the first one that provides explicit solutions to the dynamic optimal investment in variance swaps in jump-diffusion models.

To evaluate the economic performance of our three-factor model in comparison with the proposed two-factor models in Egloff, Leippold and Wu (2010), Pan (2002) and Aït-Sahalia, Karaman and Mancini (2015), we run the predictability regressions for S&P500 index. In particular, the three factors in our model are used as predictors while the two factors in the reference models are used as predictors. We document that our three factors more significantly predict the excess aggregate equity returns than the two factors in the aforementioned jump-diffusion models, indicated by the adjusted-R^2 ranging from 7.0% to 37.7% with the time horizon from 1 month to 2 years in the out-of-sample period. More interestingly, we find that variance swap rates themselves, rather than the variance risk premiums advocated widely in the literature, have better predictability on the underlying stock returns. This finding reminds us to deeply understand the economic role of variance swap rates in asset pricing and portfolio selection.

Our predictive model is closely related to other recent studies where volatility related factors are used as predictors for aggregate equity returns. More specifically, Bollerslev, Tauchen and Zhou (2009) show the strong short-horizon predictive power of the variance premium for aggregate equity returns. Our three-factor model nests their model as a special case in the sense that in our parametric framework, the variance risk premium is a linear function of our three factors. Our model also relates to the one in Adrian and Rosenberg (2008) who propose a decomposition of volatility of market return into a short-run, quickly mean reverting component and a long-run, slowly evolving component and demonstrates the predictability of the two volatility factors for market returns. Our model differs from theirs in that we have three components especially with the jump component representing the possibility of rare events or tail risks. In recent literature, it has been documented that tail risks have powerful predictive ability for market returns. For example, in Bollerslev, Todorov and Xu (2015), they find that most of the predictability for the aggregate market portfolio previously ascribed to the variance risk premium stems from the jump tail risk

component.

To better understand the investment in variance swaps, we solve the optimal variance swap investment problem in comparison with the existing results in the literature. To be more specific, our empirical results show that for a trader who wants to chase up the dynamics of variance swap risk premia, it is always optimal to take a short-long-short strategy, namely, long positions in the medium-term variance swap contracts and short positions in both the short-term and the long-term contracts. This well explains the underlying reason why Egloff, Leippold and Wu (2010) and Filipović, Gourier and Mancini (2015) have achieved the apparently opposite optimal trading strategies in variance swaps. First, both strategies are suboptimal within our model, because they are insufficient to replicate the term structure of variance swap risk premia throughout the contract combinations, although the term structure of variance swap rates can be well calibrated in their two-factor variance risk models. As a result, the empirical analysis in Egloff, Leippold and Wu (2010) shows that an investor's optimal trading positions in two variance swap contracts depend not only on the relative magnitudes of two variance risk premia, but also on the maturity gap between two contracts. Second, in the optimal trading positions in Egloff, Leippold and Wu (2010), each of the two variance swap contracts indeed plays a dual role of both exploiting risk premium and hedging. This suggests that the trade does not make full use of information in the dynamics of variance swap risk premia. Their optimal trading strategies in variance swaps are rooted in the absence of a third variance swap contract. In the presence of this extra contract, each contract then can fully play a single role: either exploring risk premium or hedging. Therefore, our empirical analysis enhances the understanding of the optimal trading strategies involving two variance swaps in Egloff, Leippold and Wu (2010) and Filipović, Gourier and Mancini (2015).

Furthermore, as the suboptimal trading strategies identified in Egloff, Leippold and Wu (2010) cannot take full advantage of the information in the term structure of variance swap risk premia, for example, due to the ignorance of jumps in variance, these strategies may cause substantial economic costs to the trader who heavily participates in variance swap

trading.[3] This brings us to ask how much the value of economics cost could be by ignoring jumps in volatility. Surprisingly, we find in all examples tested that if our double-jump model is the true model, then the strategies obtained from the pure-diffusion model in Egloff, Leippold and Wu (2010) always violate the jump-induced constraint on jump exposure and thus lead to a 100 percent wealth equivalent loss by following the suboptimal strategies. Interestingly, we also find that suboptimal strategies due to inaccurate parameter estimation can also easily violate the jump-induced constraint on jump exposure and thus result in a 100 percent wealth equivalent loss. By contrast, in their pure-diffusion model, Zhou and Zhu (2012) show that ignoring parameter uncertainty can only lead to negligible economic costs. In short, our results along with those in Zhou and Zhu (2012) empirically illustrate these serious consequences caused by both model and parameter misspecification in variance swap investments, which further reveals new insights into the complexity of variance swap trading. Therefore, our study contributes to the literature by enhancing our understanding of variance swap trading when both stock price and volatility can jump.

The rest of the paper is organized as follows. Section 2, we apply the factor analysis and the principal component analysis (PCA) to identify the number of common factors driving the evolution of the variance swap market and select the informative contracts for the empirical analysis. In Section 3, we propose a new self-exciting process for jump intensity in a double-jump model, present the estimation of the pricing models is discussed in detail, and discuss the model performance. In Section 4, we conduct the analysis on the predictability of stock returns by incorporating the the information extracted from the variance swap market. In Section 5, we present the framework for investments in the variance swap market to explicitly solve the dynamic optimal investment problem in variance swaps and conduct a comprehensive analysis on investments in variance swaps by emphasizing the role of jumps in both return and volatility. We conclude in Section 6. All proofs are collected in Appendices.

[3]For example, the economic cost of using a one-factor Heston stochastic volatility model, as demonstrated in Zhou and Zhu (2012), can be as high as 70%.

2 Risk Factors and Informational Content in Variance Swap Market

The data used for the empirical analysis involve the S&P500 index returns and the set of variance swap contracts spanning from 1 month to 24 months during the period from November 4, 2008 to September 29, 2017. In order to obtain a comprehensive understanding of the informative role of variance swap contracts, we collect all the variance swap rates on the S&P500 index available at Bloomberg. All these contracts expire in 1,2, 3, 4, 5, 6, 9, 12, 18, and 24 months (so 10 time-to-maturities in total).

Given such a complete set of the variance swap contracts on the S&P500 index, two questions are of interest for both asset pricing and allocation:1) How many risk factors drive the term structure of variance swap rates in the sample period? 2) Which variance swap contracts possess most rich information that can be used to predict the dynamics of the S&P500 index, as implied by Bollerslev, Tauchen and Zhou (2009). The former question is important for model specification to effectively capture the dynamics of variance swap rates, while the latter facilitates investors to trade volatility by participating in the variance swap market.

2.1 Risk Factors in Variance Swap Market

To exploit the number of risk factors that drive the evolution of the variance swap market in the sample period, we conduct a factor analysis on the set of variance swap rates with the time-to-maturity from 1 month up to 24 months. Suppose that there are m common factors (or risk factors) driving the term structure of variance swap rates, we may have the following relation:

$$\mathbf{VS}_{\boldsymbol{\tau}} = \boldsymbol{\mu} + \boldsymbol{\Lambda} \times \mathbf{F} + \mathbf{e}, \tag{1}$$

where $\mathbf{VS}_{\boldsymbol{\tau}}$ represents a 10×1 vector each element of which presents an observed variance swap rate with the time-to-maturity $\tau \in \{1, 2, 3, 4, 5, 6, 9, 12, 18, 24\}$, $\boldsymbol{\mu}$ is a constant of means, $\boldsymbol{\Lambda}$ is a constant $10 \times m$ of factor loadings, while $\mathbf{F}$ presents a $m \times 1$ vector of common

factors, associated with an error term **e**. The numerical decomposition in Equation (1) on the sample set of the variance swap rates suggests that the maximum number of the common factors is less than 3, e.g., $m \leq 3$, with the factor loadings on all the VS contracts, reported in Table 1.[4] The log-likelihood ratios in Table 1 show that the setup of three common

Panel A: Factor Loadings for $m = 2$

	Time-to-Maturity $\boldsymbol{\tau}$ (months)										(Cum.) of Total
	1	2	3	4	5	6	9	12	18	24	Covariance (%)
$\mathbf{F}_1$	0.858	0.828	0.801	0.774	0.748	0.723	0.669	0.634	0.554	0.511	51.606
$\mathbf{F}_2$	0.498	0.557	0.598	0.632	0.663	0.690	0.742	0.772	0.831	0.857	99.610
Panel B: Factor Loadings for $m = 3$											
$\mathbf{F}_1$	**0.864**	**0.826**	0.796	0.769	0.741	0.716	0.663	0.629	0.551	0.509	51.13
$\mathbf{F}_2$	0.498	0.560	0.601	0.635	0.667	0.693	0.745	0.775	**0.833**	**0.859**	99.53
$\mathbf{F}_3$	-0.038	0.026	0.052	0.061	**0.068**	**0.069**	0.060	0.043	0.009	-0.007	99.76

Table 1: **Factor Loadings on VS contracts**. Equipped with the common factors, e.g., $m = 2, 3$, the loadings of these factors on all the 10 VS contracts on the S&P500 index over the period from November 4, 2008 to September 29, 2017 are reported. All the time-to-maturities are in months. The log-likelihood ratios are -7.053 and -6.486 for $m = 2$ and 3, respectively.

factors outperforms the one with two factors. Panel B suggests that the three factors are sufficient to capture the term structure of variance swap rates over the sample period. In total, they successfully explain 99.76% of the covariance of these swap rates, while the third factor makes the marginal contribution of 0.23%.

More specifically, the first factor has more loadings on those short-term VS contracts, e.g, the 1- and 2-m ones. In contrast, the second one has more loadings on those long-term VS contracts with the time-to-maturities of 18 and 24 months. Apart from these dominant factors, the third one need be explained in caution. First, its loadings on each VS contract are not much smaller, compared with the first two factors. Second, the signs of these loadings can be positive or negative, unlike those positive ones of the first two factors. Third, the third factor has relatively high loadings on the medium-term VS contracts (e.g., the 5- and 6-m ones), compared with its loadings on other contracts. In sum, it seems that the role of

[4]This factor analysis can be done in Matlab 2017a.

the third one is likely to compensate the loadings of the first two factors in case that they were over- or under-loaded on VS contracts, which cause the relatively high loadings of the third factor on those medium-term VS contracts.

Furthermore, the behavior entailed by these three common factors leads to the implications for the model specification and portfolio choice of variance swaps. On the one hand, These patterns suggests that in a favored pricing model for variance swaps, three risk factors are indeed required. Namely, the first two factor are used to capture the dynamics of variance swap rates in the short- and long-term separately, while the third one is to make compensations for unexpected changes across all maturities. As the dual of pricing, on the other hand, the results in Table 1 imply that three VS contracts with the short-, medium- and long-term maturities are enough to span the whole variance swap market, which may result in the optimal investment strategies in this market.

2.2 Selection of Informative Variance Swap Contracts

We further investigate the information content of variance swap contracts, which is actually ignored in the present literature. For example, the variance swap contracts with fixed time to maturities of 2, 3, 6, 12, and 24 months are used in Egloff, Leippold and Wu (2010), Aït-Sahalia, Karaman and Mancini (2015), and Filipović, Gourier and Mancini (2015), while the contracts with the time to maturities of 1, 3, 6, 12 and 24 months are employed in Li and Zinna (2017). Since a variance swap contract represents the most direct way of achieving exposures or hedging against variance risk, and can be replicated by dynamic trading in the portfolio of the underlying asset and options, it is therefore interesting to identify those variance swap contracts that contain most valuable information to predict future variance and according the S&P500 index returns.

For this purpose, we first define the longer k-period index returns as follows:

$$R_{t+k,t} = \sum_{j=1}^{k} r_{t+j}, \tag{2}$$

where r_{t+j} represents the daily excess return, namely, the difference between the daily log-

return of the S&P500 index and the T-bill rate from time $t+j-1$ to $t+j$, and so a "day" is referred to as the unit time interval here. To form up the monthly return, we employ the scale of 21 by assuming 21 working days in a calendar month. In this way, we synthesize the cumulative k-period returns with $k = 1, 3, 6, 12$, and 24 months. Inspired by Bollerslev, Tauchen and Zhou (2009), we then employ a novel method by simply regressing the k-period returns $R_{t+k,t}$ on the combinations of the variance swap contracts as follows:

$$R_{t+k,t} = \beta_0(k) + \boldsymbol{\beta}_1(k)\mathbf{VS}_{t,t+\tau} + u_{t+k,t} = \tilde{\beta}_0(k) + \boldsymbol{\Gamma}(k)\mathbf{X} + \tilde{u}_{t+k,t}, \tag{3}$$

where the term $\mathbf{VS}_{t,t+\tau}$ represents the combination of the variance swap contracts with the time-to-maturities of $\tau = 1, 2, 3, 4, 5, 6, 9, 12, 18$ and 24 months. To mitigate the impact of the multi-linearity among the VS contracts (also see Table 1), we further apply the principal component analysis (PCA) to these contracts so that a set of orthogonal factors, denoted as a matrix of $\mathbf{X}$, are obtained, and they capture the full variation of the VS rates. Accordingly, $\boldsymbol{\Gamma}$ represents the corresponding vector of coefficients for the orthogonal factors X. As a result, we use the explanatory power or the adjusted $R^2_{k,\tau}$ obtained from the regression (3) as the measure to gauge the information quality of the VS contracts. In particular, we apply the method proposed by Britten-Jones, Neuberger and Nolte (2011) (similar to Hodrick (1992)) to mitigate the problem of overlapping data embedded in the construction of the returns $R_{t+k,t}$.

To identify the optimal set of VS contracts that possess the richest information in terms of predictability over various forecasting horizons (k), we narrow down the scope of the investigation firstly by focusing on those VS contracts with $\tau = 1, 2, 3, 6, 12$ and 24 months widely studied in the literature and leave the rest for the further examination. Table 2 reports the adjusted-R^2s with various combinations of VS contracts. Although it is well-known that the R^2s would increase with the forecasting horizon using overlapping returns, Panel A clearly shows that the adjusted-R^2s steadily increase with more VS contracts in the set across all forecasting horizons, indicating richer information possessed by these VS contracts, while there is less information in the remaining four VS contracts (e.g,, $\tau = 4, 5, 9$

and 18 months), as shown in Panel C. Also, the maximum adjusted-R^2s generated by the combinations of 5 VS contracts out of 6 are very close to those R^2s by the full set of 6 VS contracts. Furthermore, Panel B reports the R^2s obtained from the regressions over all the combinations of 5 VS contracts. Amongst all these combinations, the set of the 1-, 2-, 3-, 6- and 24-month VS contracts may achieve the best predication power overall across all the horizons, indicating that these five VS contracts possess much richer information compared with other contracts.[5] In order to balance both the loss of information in VS contracts and the complexity of model calibration, we therefore choose these five VS contracts for the following empirical analysis.[6] [7]

The bottom panel of Figure 2 plots the dynamics of the 1-, 2-, 3-, 6- and 24-month VS contracts during the full sample period, while the top and middle panel illustrate the dynamics of the S&P500 index and its daily log-returns over time, separately. It clearly shows that the variance swap market is much more volatile than the equity market, indicating more jumps in the former. Moreover, the summary statistics of these VS contracts are reported in Table 3. Note that the variance swap rate mean in each maturity category is quoted in volatility percentage units, presenting an upward term structure of variance swap rates. Meanwhile, the realized variances across all the time-to-maturities, reported in Panel C, are estimated as follows:

$$RV_{t,t+\tau} = \frac{252}{n}\sum_{i=1}^{n}(\log\frac{S_{t_i}}{S_{t_{i-1}}})^2,$$

where n denotes the total daily observations within the time-to-maturity τ. Due to much less volatility in the S&P500 index during the post-crisis period, the averaged realized variance is lower than the mean of variance swap rates in each maturity category, and hence presents a hump term structure across the time-to-maturitiy.

[5]The results in Panel B also show that the traditional choices of VS contracts, e.g., $\tau = 2, 3, 6, 12, 24$ months or $1, 3, 6, 12, 24$ months, simply ignore the information embedded in VS contracts for forecasting in short term (e.g, $k = 1$ month).

[6]Since the gap between the R^2s in Panel A and ones in Panel D is mainly caused by the 18-month contract, we hence investigate its role in terms of forecasting by replacing the 12-month contract with it in the original 6 VS contracts, e.g., $\tau = 1, 2, 3, 6, 18, 24$ months. If we are restricted to choose five contracts out of six, the set of the 1-, 2-, 3-, 6- and 24-month VS contracts is still the best choice for our purpose.

[7]We also check the replication of the 1-, 2-, 3-, 6- and 24-m VS contracts on the remaining VS contracts through a number of regressions, and find that the extremely high R^2s are gained for the 4-, 5-, 9-, 12- and 18-m VS contracts with 99.90%, 99.94%, 99.83%, 99.75% and 99.79%, respectively.

Panel A: Maximum Adjusted-R^2 (%) for Combinations of 6 VS Contracts					
Number (♯) of Selected Contracts	Forecasting Horizon (k months)				
	1	3	6	12	24
1	1.10	2.55	19.24	27.47	36.26
2	3.92	9.00	20.60	28.34	37.16
3	10.66	12.16	23.21	29.45	38.55
4	12.47	12.41	23.81	29.77	38.64
5	12.52	12.62	23.82	29.92	38.64
6	12.53	12.72	23.79	30.06	38.75
Panel B: Adjusted-R^2 (%) for Combinations of 5 VS Contracts out of 6 in Panel A					
Set of 5 VS Contracts	Forecasting Horizon (k months)				
	1	3	6	12	24
(1, 2, 3, 6, 12)	11.58	12.24	23.22	29.67	38.06
(1, 2, 3, 6, 24)	**12.52**	**12.56**	**22.28**	**28.72**	**38.63**
(1, 2, 3, 12, 24)	10.03	9.90	23.78	29.86	38.64
(1, 2, 6, 12, 24)	12.47	9.42	22.48	29.66	38.29
(1, 3, 6, 12, 24)	7.19	12.62	23.82	29.92	38.62
(2, 3, 6, 12, 24)	2.35	10.24	22.22	29.77	38.49
Panel C: Adjusted-R^2 (%) for Remaining 4 VS Contracts					
Set of Remaining VS Contracts	Forecasting Horizon (k months)				
	1	3	6	12	24
(4, 5, 9, 18)	2.73	9.26	20.98	28.20	36.01
Panel D: Adjusted-R^2 (%) for All 10 VS Contracts					
Set of All VS Contracts	Forecasting Horizon (k months)				
	1	3	6	12	24
(1, 2, 3, 4, 5, 6, 9, 12, 18, 24)	15.37	14.45	24.70	33.97	46.14

Table 2: **Maximum Adjusted-R^2 for Combinations of VS Contracts**. All the 10 VS contracts on the S&P500 index at Bloomberg are collected over the period from November 4, 2008 to September 29, 2017. All the time-to-maturities are in months. Panel A reports the maximum adjusted-R^2s generated by the combinations of 6 VS contracts over various forecasting horizons, e.g.,$\tau = 1, 2, 3, 6, 12$ and 24 months and $k = 1, 3, 6, 12$, and 24 months, while Panel B presents the adjusted-R^2s of each of the combinations of 5 VS contracts out of 6. Panel C reports the adjusted-R^2 of the remaining VS contracts with $\tau = 4, 5, 9$ and 18 months. Panel D shows the maximum adjusted-R^2s generated by the set of all the 10 VS contracts (e.g., with $\tau = 1, 2, 3, 4, 5, 6, 9, 12, 18$ and 24 months).All the adjusted-R^2s are obtained by applying the approach proposed by Britten-Jones, Neuberger and Nolte (2012).

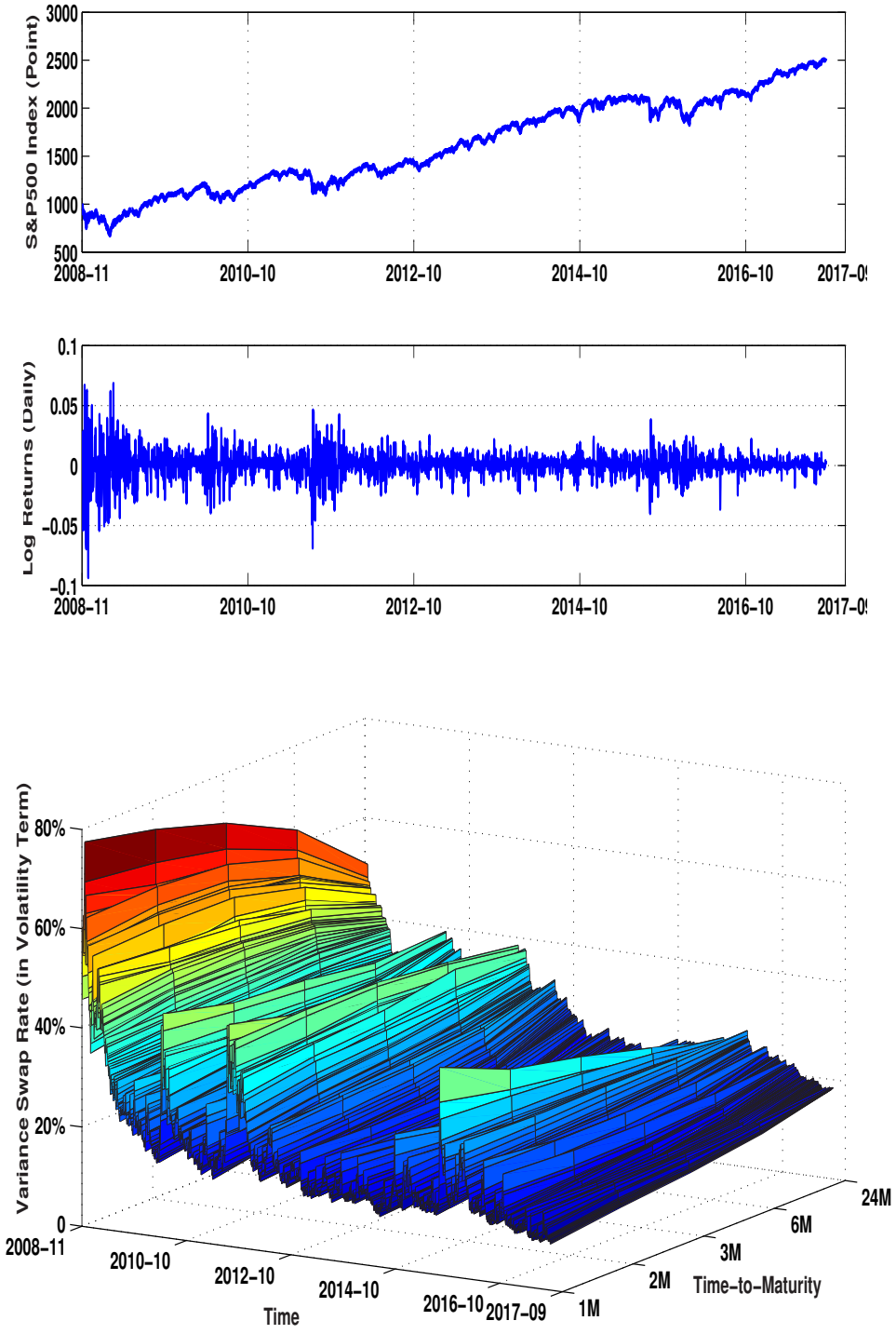

Figure 1: **Term Structure of Variance Swap Rates.** The dynamics of variance swap rates with 1-, 2-, 3-, 6- and 24-month time-to-maturity are plotted in volatility percentage units, e.g., $\sqrt{VS_{t,t+\tau}} \times 100$ from from November 4, 2008 to September 29, 2017. There are 2,242 daily observations for each time-to-maturity τ.

Panel A: Variance Swap Rates					
Time to Maturity	Mean	St. Dev.	Skew.	Kur.	AC(1)
1	19.262	8.931	2.218	9.121	0.978
2	20.427	8.765	2.141	8.538	0.986
3	21.293	8.512	2.025	7.946	0.989
6	22.772	7.701	1.704	6.364	0.992
24	25.771	5.945	1.075	3.998	0.995
Panel B: S&P 500 Index Returns					
Log-returns	0.039	0.185	-0.507	10.657	-0.063
Panel C: Realized Variances					
1	15.248	9.453	2.388	10.981	0.985
2	15.367	8.337	1.993	7.396	0.991
3	15.538	7.843	1.905	6.824	0.993
6	15.395	6.500	1.651	6.199	0.994
24	15.604	3.591	0.583	2.563	0.996

Table 3: **Summary Statistics of Variance Swap Rates**. All the variance swap rates are collected from Bloomberg during the full sample period is from November 4, 2008 to September 29, 2017. The descriptive statistics, including mean, stand deviation (St. Dev.), skewness (Skew.), kurtosis (Kurt.) and the first-order autocorrelation (AC1), are reported, while the mean of the variance swap rate (also the realized variance mean) in each maturity category is quoted in percentage, and time to maturities are quoted in months. Note that the mean and standard deviation of the log-returns in Panel B are reported in the annualized term. The realized variance is calculated as $RV_{t,t+\tau} = \frac{252}{n}\sum_{i=1}^{n}(\log\frac{S_{t_i}}{S_{t_{i-1}}})^2$ where n denotes the total observations within the time-to-maturity τ and the scale of 21 (e.g., presenting 21 work days in a calendar month) is used to assemble the realized variance over the time horizon $[t, t+\tau]$ with $n = 21 \times \tau$ for $\tau = 1, 2, 3, 6$ and 24 months.

3 A Self-Exciting Stochastic Volatility Model for Variance Swaps

The results in Section 2 motivate us to propose a self-exciting stochastic volatility mode to capture the dynamics of variance swap rates in the sample period. We first present the model specification, and formulate the variance swap rates under this model. At the end, we provide the solutions to the optimal variance swap allocation problem under this self-exciting stochastic volatility model.

3.1 Model Specification and Properties

For analytic tractability, we adopt the popular double jump model used by Aït-Sahalia, Karaman and Mancini (2015), apart from the specification for the intensity of the counting process. That is, we assume that the stock price, volatility and its long-run mean under a risk-neutral measure Q are given as follows:

$$\begin{aligned}
\frac{dS_t}{S_{t-}} &= (r-\delta)dt + \sqrt{(1-\rho^2)v_t}dB_{1t}^Q + \rho\sqrt{v_t}dB_{2t}^Q + (\exp(J_t^{s,Q})-1)dN_t - g^Q\lambda_t dt, \\
dv_t &= \kappa_v^Q(m_t - v_t)dt + \sigma_v\sqrt{v_t}dB_{2t}^Q + J_t^{v,Q}dN_t \\
dm_t &= \kappa_m^Q(\theta_m^Q - m_t)dt + \sigma_m\sqrt{m_t}dB_{3t}^Q.
\end{aligned} \tag{4}$$

As in Aït-Sahalia, Karaman and Mancini (2015), we specify the market price of risks for the Brownian motions by γ_i $(i = s, v, m)$ in the following way:

$$\Lambda_t^\top = [\gamma_s\sqrt{(1-\rho^2)v_t}, \gamma_v\sqrt{v_t}, \gamma_m\sqrt{m_t}],$$

and then, under the objective probability P, the stock price and variance dynamics can be represented as follows:

$$\begin{aligned}
\frac{dS_t}{S_{t-}} &= \mu_t dt + \sqrt{(1-\rho^2)v_t} dB_{1t}^P + \rho\sqrt{v_t} dB_{2t}^P + (\exp(J_t^{s,P}) - 1)dN_t - g^P \lambda_t dt, \\
dv_t &= \kappa_v^P (m_t \kappa_v^Q / \kappa_v^P - v_t)dt + \sigma_v \sqrt{v_t} dB_{2t}^P + J_t^{v,P} dN_t \\
dm_t &= \kappa_m^P (\theta_m^P - m_t)dt + \sigma_m \sqrt{m_t} dB_{3t}^P,
\end{aligned} \tag{5}$$

where $\mu_t = r - \delta + \gamma_s(1-\rho^2)v_t + \gamma_v \rho v_t + (g^P - g^Q)\lambda_t$, $\kappa_v^P = \kappa_v^Q - \gamma_v \sigma_v$, $\kappa_m^P = \kappa_m^Q - \gamma_m \sigma_m$, and $\theta_m^P = \theta_m^Q \kappa_m^Q / \kappa_m^P$, while r is the risk free rate, and δ is the dividend yield, both taken to be constant for simplicity. The correlation parameter ρ is used to capture the so-called leverage effect between stock returns and variance changes. The three Brownian motions, $B_{it}^Q, i = 1, 2, 3$, are uncorrelated.[8]

The dynamics of the spot variance of the price, v_t, is driven by a two-factor model, while the speed of mean revision is κ^P under P (κ^Q under Q accordingly). The long-term mean of the variance is governed by the pure-diffusion process m_t that has a similar specification with v_t but equipped with a parameter triple of κ_m^P (κ_m^Q), θ_m^P (θ_m^Q) and σ_m, respectively. As a result, the process v_t presents the fast mean reverting and volatile pattern and captures sudden movements in variance with the jump process, while the process m_t has no jump and is less volatile and persistent and characterizes the central tendency of variance.

Meanwhile, the jump size in the stock price, $J^{s,Q}$, is independent of both Brownian and jump components, and is assumed to follow a normal distribution with mean μ_j^Q and variance σ_j^2 so that $g^Q = \exp(\mu_j^Q + \sigma_j^2/2) - 1$. Similarly, we may have $g^P = \exp(\mu_j^P + \sigma_j^2/2) - 1$ under the objective probability measure P. However, the jump size in the spot variance, $J^{v,Q}$, is positive. It is independent of Brownian motions and the jump component in the stock price, and follows an exponential distribution with a parameter μ_v^Q, i.e., $E^Q[J^{v,Q}] = \mu_v^Q$, and so is the jump size $J^{v,P}$. This specification thus captures sudden upward movement of v_t. Furthermore, the two-factor model studied by Pan (2002) that allows for jumps only in stock

[8]The variant of the specifications in Model (29) and (30) is widely used in the literature (see Bakshi, Cao and Chen (1997), Chernov and Ghysels (2000), Bates (2000,2006), Pan (2002), Eraker, Johannes and Polson (2003), Broadie, Chernov and Johannes (2007), Egloff, Leippold and Wu (2010) and Todorov (2009) and references therein).

price can be obtained if both μ_v^P and μ_v^Q are set as zero (e.g., $\mu_v^P = \mu_v^Q \equiv 0$).

We now turn to modeling jump intensity under the measure P and Q, respectively. Empirical studies suggest that the jump intensity of asset prices is stochastic and clustered in time (see Bates (2006) and Aït-Sahalia, Cacho-Diaz and Laeven (2015)). We then assume that the jump intensity λ_t of the counting process N_t under the measure Q follows a self-exciting process as follows:

$$d\lambda_t = \alpha(\lambda_\infty - \lambda_t)dt + \beta_0 J_t^{v,Q} dN_t, \tag{6}$$

where α, λ_∞ and $\beta_0 > 0$.[9] Unlike the the specification of λ_t in Aït-Sahalia, Karaman and Mancini (2015) in the form of $\lambda_t = \lambda_0 + \lambda_1 v_t$, implying that the jump intensity is uniquely determined by volatility, Equation (31) suggests that a jump in either price or variance may cause the intensities to jump up, governing by β_0 and the jump intensity decays exponentially back towards a level λ_∞ at speed α. It partially disentangles the jump intensity from volatility in the sense that λ_t is proportional to v_t when v_t has large movements driven by the jump instead of small one caused by the diffusion.

More importantly, our model is especially tractable in that we solve the optimal portfolio choice problem with variance swaps in closed form[10]. Although the specification in Aït-Sahalia, Karaman and Mancini (2015) allows for more jumps to occur during volatile periods with the intensity bounded by a positive constant ($\lambda_0 > 0$), it is subject to the underestimation of volatility in the long run. That is, the mean-reverting nature of volatility suggests that the long-term expected volatility declines over time, implying that the intensity of jumps is also a decreasing function of time. This in fact imposes an unnecessary restriction on the dynamics of jump intensity λ_t.

[9] Accordingly, its dynamics under the measure P can be represented as $d\lambda_t = \alpha(\lambda_\infty - \lambda_t)dt + \beta_0 J_t^{v,P} dN_t$.

[10] Our model is also tractable for pricing European options as it is one of affine models developed by Duffie, Pan and Singleton (2000). More recently, Fulop, Li and Yu (2015) propose a self-exciting asset pricing model that takes into account co-jumps between prices and volatility and self-exciting jump clustering. They find that the self-exciting jump intensity has become more important since the onset of the 2008 global financial crisis and illustrate good model performance for the S&P 500 index option data.

3.2 Term Structure of Variance Swap Rates

Similar to Equation (9) in Aït-Sahalia, Karaman and Mancini (2015), the variance swap rate is given by

$$VS_{t,t+\tau} = E_t^Q\left[\frac{1}{\tau}\int_t^{t+\tau} v_u du + \frac{1}{\tau}\sum_{u=N_t}^{N_{t+\tau}} (J_u^s)^2\right] = \bar{v}_{t,t+\tau}^Q + E_t^Q[(J^s)^2]\bar{\lambda}_{t,t+\tau}^Q,$$

where $E_t^Q[(J^{s,Q})^2] = (\mu_j^Q)^2 + \sigma_j^2$, and $E_t^Q[\lambda_s] = \frac{\alpha\lambda_\infty}{(\alpha-\beta^Q)}\left[1 - e^{-(\alpha-\beta^Q)(s-t)}\right] + \lambda_t e^{-(\alpha-\beta^Q)(s-t)}$, with $\beta^Q = \beta_0 E_t^Q[J_t^{v,Q}]$, associated with

$$\bar{\lambda}_{t,t+\tau}^Q = \frac{1}{\tau}\int_t^{t+\tau} E_t^Q[\lambda_s]ds = \frac{\alpha\lambda_\infty}{(\alpha-\beta^Q)}\left[1 - \frac{(1-e^{-(\alpha-\beta^Q)\tau})}{\tau(\alpha-\beta^Q)}\right] \frac{\lambda_t}{\tau(\alpha-\beta^Q)}(1-e^{-(\alpha-\beta^Q)\tau}).$$

Likewise, we can also obtain the following results:

$$\begin{aligned}
E_t^Q[m_s] &= \theta_m^Q\left[1 - e^{-\kappa_m^Q(s-t)}\right] + m_t e^{-\kappa_m^Q(s-t)}, \\
E_t^Q[v_s] &= \theta_m^Q\left[1 + \frac{\kappa_m^Q}{\kappa_v^Q - \kappa_m^Q}e^{-\kappa_v^Q(s-t)} - \frac{\kappa_v^Q}{\kappa_v^Q - \kappa_m^Q}e^{-\kappa_m^Q(s-t)}\right] \\
&+ \frac{\alpha\mu_v^Q\lambda_\infty}{\kappa_v^Q(\alpha-\beta^Q)}\left[1 + \frac{\alpha-\beta^Q}{\kappa_v^Q - (\alpha-\beta^Q)}e^{-\kappa_v^Q(s-t)} - \frac{\kappa_v^Q}{\kappa_v^Q - (\alpha-\beta^Q)}e^{-(\alpha-\beta^Q)(s-t)}\right] \\
&+ \frac{\kappa_v^Q}{\kappa_v^Q - \kappa_m^Q}\left[e^{-\kappa_m^Q(s-t)} - e^{-\kappa_v^Q(s-t)}\right] m_t \\
&+ \frac{\mu_v^Q}{\kappa_v^Q - (\alpha-\beta^Q)}\left[e^{-(\alpha-\beta)(s-t)} - e^{-\kappa_v^Q(s-t)}\right]\lambda_t + v_t e^{-\kappa_v^Q(s-t)}.
\end{aligned}$$

Thus, under the risk neutral probability Q, the rate of a variance swap contract with the life time of τ, staring from time t, can be specified as follows:

$$\begin{aligned}
VS_{t,t+\tau} &= \frac{1}{\tau}\int_t^{t+\tau} E_t^Q[v_s]ds + \frac{1}{\tau}E_t^Q[(J^s)^2]\int_t^{t+\tau} E_t^Q[\lambda_s]ds \\
&= \phi_\theta(\tau)\theta_m^Q + \phi_\lambda^0(\tau)\lambda_\infty + \phi_v(\tau)v_t + \phi_m(\tau)m_t + \phi_\lambda(\tau)\lambda_t,
\end{aligned}$$

where

$$
\begin{aligned}
\phi_\theta(\tau) &= 1 + \frac{\kappa_m^Q}{\kappa_v^Q\tau(\kappa_v^Q-\kappa_m^Q)}\left[1-e^{-\kappa_v^Q\tau}\right] - \frac{\kappa_v^Q}{\kappa_m^Q\tau(\kappa_v^Q-\kappa_m^Q)}\left[1-e^{-\kappa_m^Q\tau}\right],\\
\phi_\lambda^0(\tau) &= \frac{\alpha\mu_v^Q}{\kappa_v^Q(\alpha-\beta^Q)}\left[1+\frac{(\alpha-\beta^Q)(1-e^{-\kappa_v^Q\tau})}{\kappa_v^Q\tau(\kappa_v^Q-(\alpha-\beta^Q))} - \frac{\kappa_v^Q(1-e^{-(\alpha-\beta^Q)\tau})}{(\alpha-\beta^Q)\tau(\kappa_v^Q-(\alpha-\beta^Q))}\right]\\
&\quad + \frac{\alpha E_t^Q[(J^{s,Q})^2]}{(\alpha-\beta^Q)}\left[1-\frac{1}{\tau(\alpha-\beta^Q)}(1-e^{-(\alpha-\beta^Q)\tau})\right],\\
\phi_v(\tau) &= \frac{1}{\kappa_v^Q\tau}(1-e^{-\kappa_v^Q\tau}),\\
\phi_m(\tau) &= \frac{\kappa_v^Q}{\tau(\kappa_v^Q-\kappa_m^Q)}\left[\frac{1-e^{-\kappa_m^Q\tau}}{\kappa_m^Q} - \frac{1-e^{-\kappa_v^Q\tau}}{\kappa_v^Q}\right],\\
\phi_\lambda(\tau) &= \frac{\mu_v^Q}{\tau(\kappa_v^Q-(\alpha-\beta^Q))}\left[\frac{1-e^{-(\alpha-\beta^Q)\tau}}{\alpha-\beta^Q} - \frac{1-e^{-\kappa_v^Q\tau}}{\kappa_v^Q}\right] + \frac{E_t^Q[(J^{s,Q})^2](1-e^{-(\alpha-\beta^Q)\tau})}{\tau(\alpha-\beta^Q)}.
\end{aligned}
\tag{7}
$$

Note that for a given τ, the variance swap rate $VS_{t,t+\tau}$ is a martingale under Q-measure. Hence, under the objective probability P, $VS_{t,t+\tau}$ follows the equation below

$$
\begin{aligned}
dVS_{t,t+\tau} &= [\phi_v(\tau)\sigma_v\gamma_2 v_t + \phi_m(\tau)\sigma_m\gamma_3 m_t - (\phi_v(\tau)+\beta_0\phi_\lambda(\tau))\mu_v^Q\lambda_t]dt\\
&\quad + \phi_v(\tau)\sigma_v\sqrt{v_t}dB_{2t}^P + \phi_m(\tau)\sigma_m\sqrt{m_t}dB_{3t}^P + (\phi_v(\tau)+\beta_0\phi_\lambda(\tau))J_t^{v,P}dN_t.
\end{aligned}
\tag{8}
$$

And as such, the difference of the drift components in (8) under the measure Q and P characterizes the term structure of risk premia across the time-to-maturity τ in the variance swap market, in spirit, which is similar to the definition of the variance swap premium in Aït-Sahalia, Karaman and Mancini (2015).

3.3 MCMC Estimation for Variance Swap Rates

Before analyzing the decision of variance swap investments, we first discuss the choice of the variance swap contracts possessing rich information for predictability on the stock returns, and then study the empirical performance of four models, including the model investigated by Pan (2002) (termed the "JP" model hereafter), the one examined by Egloff, Leippold and Wu (2010) (or the "ELW" model) and the one studied by Aït-Sahalia, Karaman and

Mancini (2015) (or the "AKM" model) as well as our model with a self-exciting process for the jump intensity (or the "HJ" model). After calibrating the models to both the stock returns and the variance swap rate, we examine their goodness-of-fit performance and the dynamics of the filtered latent state variables.

To highlight the economic role of jumps in either price or volatility within the framework of the two-factor pure-diffusion model in Egloff, Leippold and Wu (2010), we calibrate all the four models to the empirical term structure of variance swap rates reported in Table 3, and then refer to the ELW model as the benchmark.

The estimation procedure is implemented by the Markov Chain Monte Carlo (MCMC) method. MCMC estimation is a Bayesian inference technique. Specifically, the MCMC method obtains the point estimator by sampling from a posterior distribution. Usually, the difficulty of MCMC comes from deriving a simple and easy posterior distribution. The basic idea of MCMC is that we assume some prior distribution on the parameter that needs to be estimated. Then, we find the posterior distribution of the parameter and draw samples from the posterior distribution. According to Bayesian Theorem, the posterior distribution summarizes the sample information regarding the parameters and the latent variables:

$$p(\Theta, V, J, G|Y) \propto p(Y|\Theta, V, J)p(\Theta, V, J) \tag{9}$$

where Θ is the set of model parameters, V is the latent volatility variable, J is the latent jump variable and Y is the sample data observable in the market. In particular, the latter three variables are vectors containing the time series of these latent variables. This posterior then combines the likelihood, $p(Y|\Theta, V, J)$ and the prior, $p(\Theta, V, J)$.

If we can directly sample for the distribution (9), the estimation procedure will be fairly easy. However, the posterior in (9) is usually not known in closed form, and so direct sampling is not feasible. Alternatively, we can sample by constructing a Markov Chain over the parameters and latent variables whose equilibrium transition density converges to the desired posterior distribution. The sampling procedure is done by iterations. The best estimate is the empirical mean of samples simulated. Before taking the mean value, we also

need to cut off some samples from the beginning as it takes time to reach the stationary state. This refers to the 'burn-in' sample. Appendix C provides the details about the MCMC implementations of all the four models, including the ELW, JP, AKM and HJ model.[11]

We conduct the model calibration that matches up the dynamics of the S&P500 index and the quoted variance swap rates in the market, reported in Table 3 by minimizing the root mean-squared errors (RMSEs):

$$\text{RMSE}(\Theta; i) = \sqrt{\frac{\sum_{j=1}^{N}(E_i^P[VS_{t,t+\tau_j}|\Theta] - E_{Market}^P[VS_{t,t+\tau_j}|\Theta])^2}{N}}, \tag{10}$$

for $i \in \{ELW, JP, AKM, HJ\}$, and Θ denotes the set of model parameters, and $N = 5$ indicates the total number of the time-to-maturities of variance swap contracts. Following such a procedure, we finally obtain all the required parameters, as reported in Table 4. It shows that the market prices of both the instantaneous variance and the central tendency factor, γ_v and γ_m are negative, despite moderate differences in absolute magnitude. In particular, the highly negative market price of variance risk (γ_v) is confirmed by several studies (see Bakshi and Kapadia (2003), Bondarenko (2004), Carr and Wu (2009) and Todorov (2009) among others).

The negative market prices make the statistical mean-reverting speeds (κ^P) larger and the statistical long-run means (θ_m^P) smaller than their risk-neutral counterparts (κ^Q and θ_m^Q respectively) in all the four models. Linking back to Equation (11), the three long-term means show the order of $\theta_m^Q > \theta_m^P > \theta_v^P$ in each model, while each of them presents a declining pattern across models duo to the presence of jumps in variance, as suggested in Table 4. Moreover, since the risk neutral mean of the variance jump size is larger than the statistical mean in the AKM and HJ model, this indicates a negative jump risk premium in variance swap rates defined in Equation (8).

We now examine the performance of the four models in capturing the physical dynamics of the S&P500 index in the sample period. Based on estimated model parameters reported

[11]For each model, we run 250,000 simulations and use the final 100,000 simulation paths for the purposes of the parameter estimation and pricing performance analysis.

Parameters	ELW		JP		AKM		HJ	
	Estim.	S.E.	Estim.	S.E.	Estim.	S.E.	Estim.	S.E.
κ_v^P	1.997	0.001	2.022	0.005	3.377	0.000	2.793	0.003
σ_v	0.521	0.000	0.698	0.000	0.648	0.000	0.613	0.000
κ_m^P	0.348	0.000	0.524	0.000	0.566	0.000	0.451	0.000
σ_m	0.330	0.000	0.350	0.000	0.329	0.000	0.292	0.000
θ_m^P	0.054	0.000	0.053	0.000	0.051	0.000	0.053	0.000
ρ	-0.794	0.010	-0.806	0.011	-0.772	0.011	-0.775	0.011
γ_s	-0.024	0.943	-0.065	0.954	-0.612	0.950	-0.590	0.944
γ_v	-0.743	0.002	-0.091	0.007	-0.417	0.003	-0.438	0.005
γ_m	-0.629	0.001	-0.347	0.001	-0.547	0.000	-0.454	0.001
λ_0	-	-	0.064	0.000	0.075	0.000	-	-
λ_1	-	-	9.230	0.007	1.720	0.003	-	-
μ_j^P	-	-	0.004	0.012	0.001	0.023	-0.013	0.079
μ_j^Q	-	-	-0.001	0.010	-0.005	0.021	-0.027	0.074
σ_j	-	-	0.111	0.000	0.167	0.000	0.208	0.002
μ_v^P	-	-	-	-	0.067	0.022	0.066	0.022
μ_v^Q	-	-	-	-	0.071	0.000	0.243	0.000
α	-	-	-	-	-	-	184.312	71.747
λ_∞	-	-	-	-	-	-	0.059	0.000
β_0	-	-	-	-	-	-	2.214	1.979
σ_{e_1}	0.008	0.000	0.008	0.000	0.007	0.000	0.008	0.000
σ_{e_2}	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000
σ_{e_3}	0.004	0.000	0.003	0.000	0.003	0.000	0.004	0.000
σ_{e_4}	0.006	0.000	0.005	0.000	0.004	0.000	0.004	0.000
σ_{e_5}	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4: **Model Parameters**. All the parameters are based on the data set reported in Table 3. The variance swap rates with five typical maturities (including 1-, 2-, 3-, 6- and 24-month time to maturity). All the four models, including ELW, JP, AKM and HJ Model, are calibrated to the data set composed of both variance swap rates and the S&P500 index log-returns in the sample period from November 4, 2008 to October 13, 2014.

in Table 4 and latent volatility/jump variables, we calculate the standardized residuals for both return, volatility and long-term tendency variables, e.g., ϵ^{SP}_{t+1}, ϵ^{v}_{t+1} and ϵ^{m}_{t+1}. If a given model is correctly specified, the distribution of these three residuals should be close to $N(0,1)$. Figure 2 plots the kernel density estimators of $\epsilon^{SP}, \epsilon^{v}$ and ϵ^{m}, separately. It clearly shows that the incorporation of jump in the state variables can significantly improve the model performance in capturing the dynamics of the S&P500 index. More importantly, the distributions of the kernel densities among all the four models suggest that a flexible specification of jump intensity in the HJ model can facilitate the model calibration, evident by the relatively small distance towards the standard normal distribution.

In addition to graphical illustrations, we further conduct tests on the performance of the standard residuals related to the state variables and of pricing VS contracts by using the Kolmogov-Smirnov (KS hereafter) test and the Bayes Factor. Table 5 reports the statistics of these tests for all the four models. The KS test results reject the null hypothesis that the model residuals follow $N(0,1)$, given the relatively large distance between the empirical cumulative distribution function (CDF) and the theoretical one, reported by the KS statistics, showing that none of these models can perfectly capture the dynamics of the S&P500 index and latent variables. But the KS statistics suggest that the calibration performance is improving from the ELW model to the HJ model. This is also confirmed by the Bayes Factors reported in Panel E. It shows that the JP, AKM and HJ model clearly outperform the ELW model in capturing the dynamics of the S&P5 index in the sample period from November 4, 2008 to October 13, 2014, demonstrating the importance of incorporating jumps in either returns or latent variables. The further comparison among the JP, AKM and HJ model suggests that jumps in variance may make relatively small contributions to the improvement in model calibration and pricing VS contracts in the sample period, supported by their small Bayes Factors. On the other hand, the p-values suggest that the 2- and 24-m VS contracts can be priced precisely in all the four models with the pricing errors following $N(0,1)$, which will be further discussed in Section 3.3.

Finally, we examine the filtered latent volatility and jump variables. Figure 3 plots the dynamics of the state variables, v and m in the four models. Historically speaking,

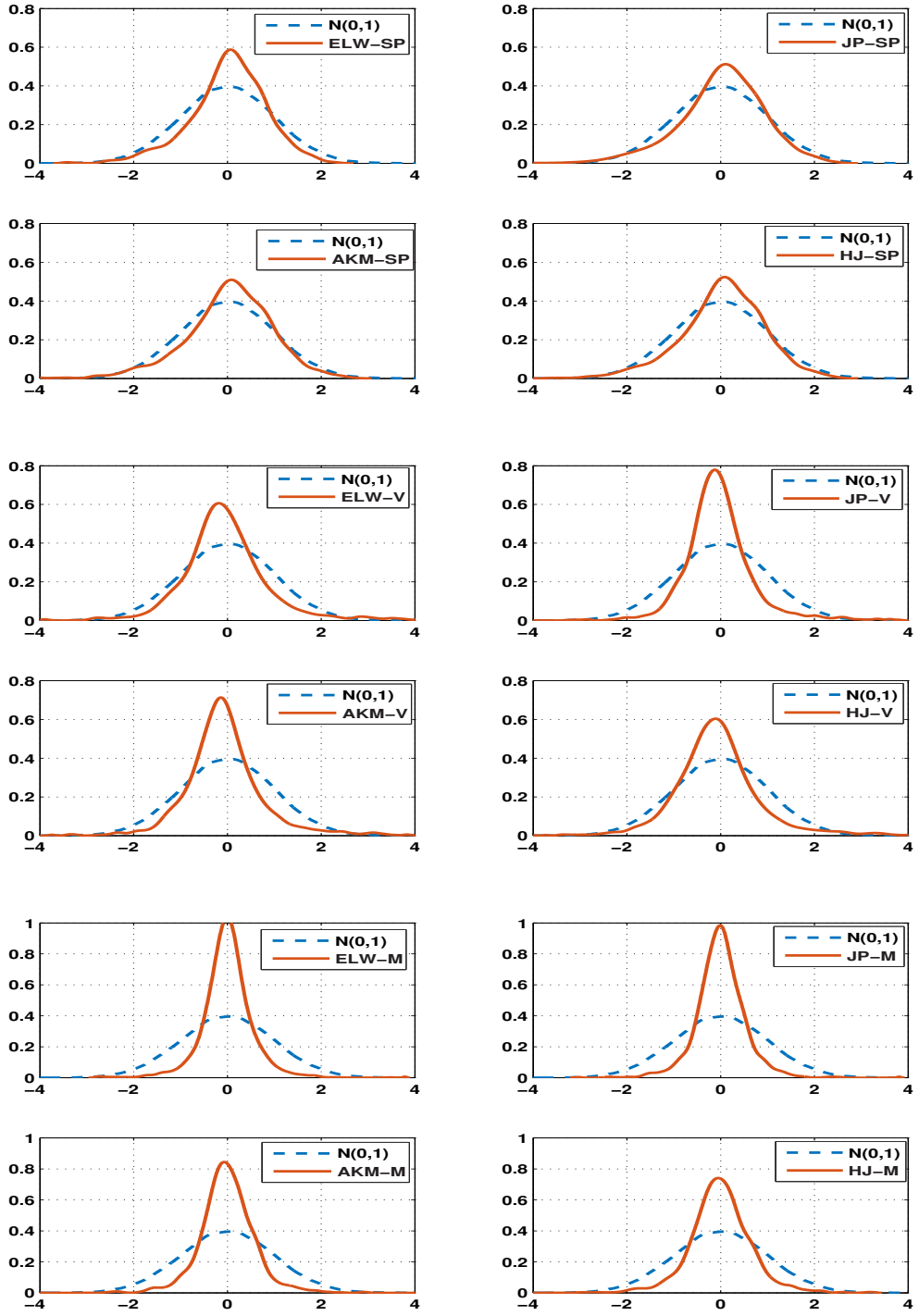

Figure 2: **Kernel Densities of Standardized Model Residuals**. The kernel densities of standardized residuals of returns (denoted by SP), variance (V) and long-term variance (M) in ELW, JP, AKM and HJ Model. They are estimated using daily S&P500 log-returns and the variance swap rates between November 4, 2008 to October 13, 2014. There are 1,495 daily observations for each time-to-maturity.

Panel A: Performance of ELW Model in Capturing S&P500 Index Returns and Pricing VS Contracts								
	State Variables			Time to Maturity				
	S&P500	V	M	1	2	3	6	24
KS statistics	0.101	0.111	0.199	0.303	0.038	0.325	0.333	0.032
P-values	0.000	0.000	0.000	0.000	0.330	0.000	0.000	0.492
Panel B: Performance of JP Model in Capturing S&P500 Index Returns and Pricing VS Contracts								
KS statistics	0.089	0.160	0.197	0.265	0.032	0.258	0.192	.031
P-values	0.000	0.000	0.000	0.000	0.490	0.000	0.000	0.530
Panel C: Performance of AKM Model in Capturing S&P500 Index Returns and Pricing VS Contracts								
KS statistics	0.087	0.151	0.174	0.224	0.031	0.188	0.138	0.031
P-values	0.000	0.000	0.000	0.000	0.524	0.000	0.000	0.529
Panel D: Performance of HJ Model in Capturing S&P500 Index Returns and Pricing VS Contracts								
KS statistics	0.088	0.137	0.146	0.254	0.032	0.241	0.125	0.030
P-values	0.000	0.000	0.000	0.000	0.498	0.000	0.000	0.549
Panel E: Relative Performance of Four Models in Capturing S&P Index Returns								
	JP-ELW	AKM-ELW	HJ-ELW	AKM-JP	HJ-JP	HJ-AKM		
Bayes Factor (BF)	131.515	139.154	143.421	5.926	11.906	5.980		

Table 5: **Kolmogorov-Smirnov (KS) Test and Bayes Factors in ELW, JP, AKM and HJ Model**. The table provides Kolmogorov-Smirnov (KS) tests of the hypotheses that both the standardized residuals related to the index return, variance and long-term tendency factor of each of the four models (e.g., the ELW, JP, AKM and HJ model) and pricing errors of the 1-, 2-,3-, 6- and 24-m VS contracts follow $N(0,1)$. The KS statistics and p-values for model residuals and pricing errors are reported at the significance level of 5%. The residuals and pricing errors are collected in the sample period (11/04/2008-10/13/2014). The bayes factors (BF) (see Eraker, Johannes and Polson (2003) for more details) are calculated in the log-ratio differences, while the second model is the base. The first model is preferred if the Bayes Factor is larger than 10, e.g., BF $>$ 10.

the volatility variable tended to be relatively stable, following highly volatile states in the financial crisis in 2008, and bounded up again in the European debt crisis from 2010 to 2012, and showed a small spike in 2015 when Chinese financial market experienced a significant turmoil in the summer. Compared with the JP model, the incorporation of jumps in volatility in the AKM and HJ model clearly widens the range of volatility states, which may better capture the dramatic changes in volatile market. In contrast, the long-term tendency variable m evolved gently over the sample period, and gradually turns down to the relative low level (with the annual rate of about 5%) in recent years, reflecting the clam sentiments among investors.

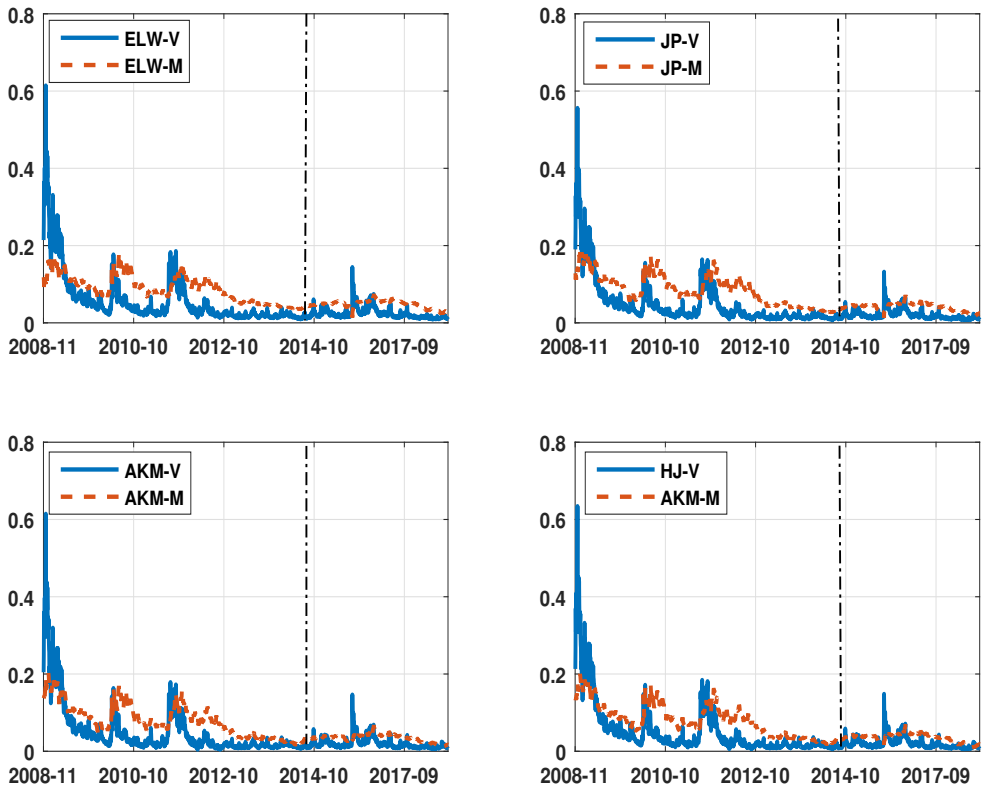

Figure 3: **Dynamics of Filtered Volatility Variables (v and m) in ELW, JP, AKM and HJ Model**. The dynamics of the filtered volatility variables V and M in ELW, JP, AKM and HJ Model are reported. The time series of v (m) is plotted by the solid (dashed) line. The dash-dotted line splits the full sample into the in-the-sample period and the out-of-sample one. There are 2,242 daily observations in the full sample with 1,495 observations in the-sample period.

Figure 4 further plots the dynamics of the filtered jump variables over the sample period. First, the paths of the jump intensity λ_t in the AKM (JP as well) model and HJ model

are in sharp contrast due to their distinct specifications. Second, it seems that more jumps in volatility rather than the S&P500 index are captured, consistent with the dynamics of the index plotted in Figure 2. More interestingly, the AKM model tends to capture those positive jumps in the S&P500 index, while the negative jumps in the index are usually detected in the HJ model. Partially, this difference is rooted in the specification of the jump intensity. This further suggests that the latter specification for the jump intensity (e.g., $d\lambda_t = \alpha(\lambda_\infty - \lambda_t)dt + \beta_0 J_t^v dN_t$ in the HJ model) can explain the style facts better in the empirical literature that market volatility is usually driven by unexpected price drops (see Li and Zinna (2017)).

In Section 2.1, we have identified the number of the common factor with $m = 3$ over the full sample of the VS contracts. Accordingly, their factor scores, the projects of the 10 VS contracts on these factors can be further worked out from the factor loadings. To investigate the linkage between the common factors and the risk factors (v_t, m_t, and λ_t) specified in the four models, we further examine the correlations between the estimated factor scores and state variables, as reported in Table 6. The results show that the estimated risk factors v_t and m_t are highly related to the first two common factors in all the models, respectively. In the HJ model, the role of the third common factor is decomposed into the variance variable v (with the positive correlation) and the jump intensity variable λ_t (with the negative correlation). All these observations suggest that our model specification has its solid root and that the self-exciting jump intensity process indeed improve the valuation performance for the VS contracts.

3.4 Model Performance

This section examines the performance of calibration to the term structure of variance swap rates. First, the risk loadings on the components that make the contributions to the variance swap rates are analyzed. They present distinct behavior across time-to-maturities, showing the time-varying role of the state variables (v_t, m_t, λ_t) in shaping the term structure of the variance swap rates. We then estimate the mean term structure of variance swap rates

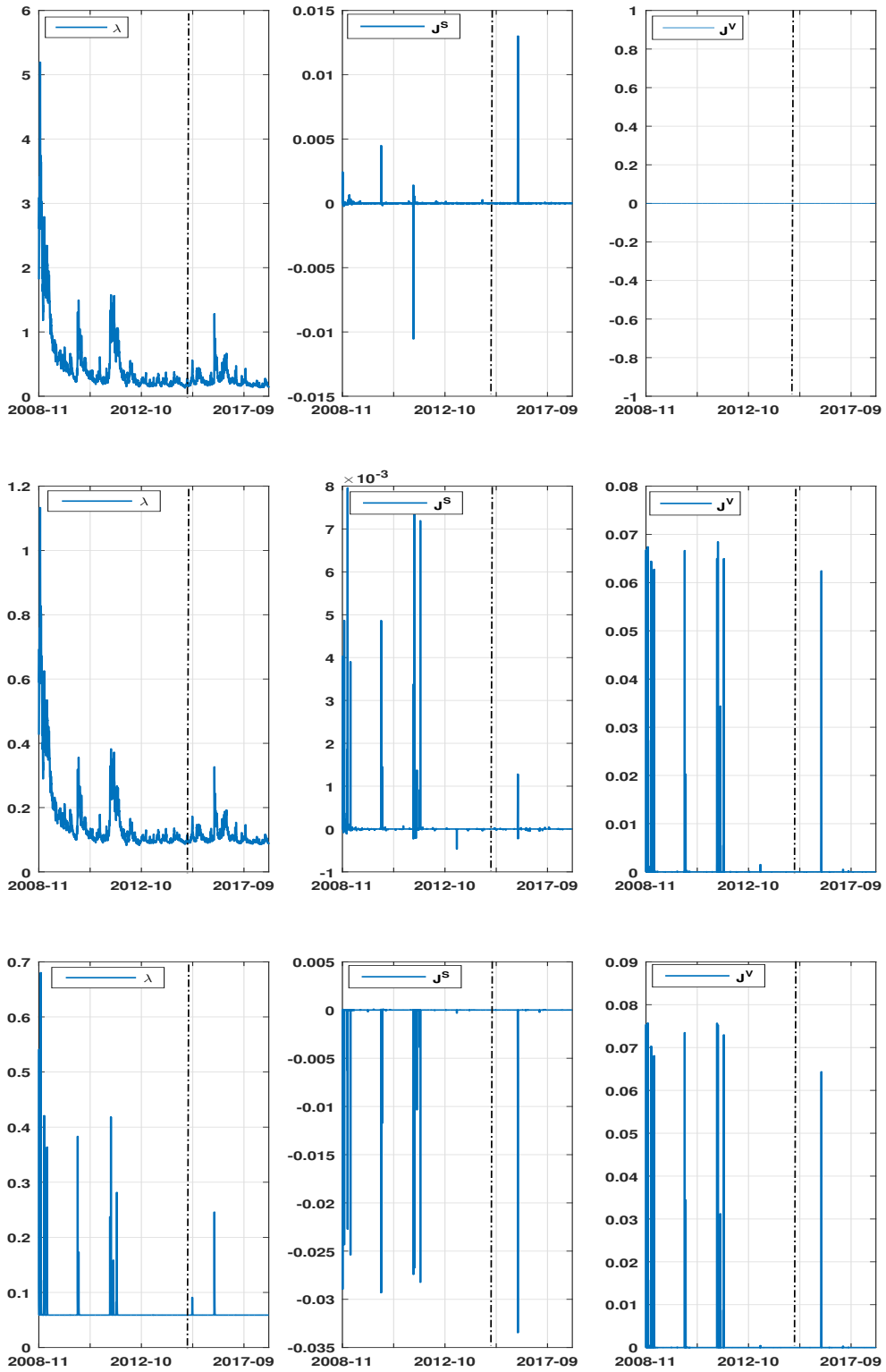

Figure 4: **Dynamics of Filtered Jump Variables (λ, J^S, and J^V) in ELW, JP, AKM and HJ Model**. The dynamics of the filtered jump variables λ, J^S and J^V in ELW, JP, AKM and HJ Model are reported. The dash-dotted line splits the full sample into the in-the-sample period and the out-of-sample one. There are 2,242 daily observations in the full sample with 1,495 observations in the-sample period.

	ELW			JP			AKM			HJ		
	v_t	m_t	λ_t	v_t	m_t	λ_t	v_t	m_t	λ_t	v_t	m_t	λ_t
$\mathbf{F}_1$ Score	0.858	0.120	-	0.871	0.478	-	0.887	0.449	-	0.875	0.471	0.027
$\mathbf{F}_2$ Score	0.500	0.989	-	0.478	0.974	-	0.448	0.941	-	0.288	0.954	-0.040
$\mathbf{F}_3$ Score	-	-	-	-	-	-	-	-	-	0.334	0.017	-0.227

Table 6: **Correlation between Factor Scores and Estimated State Variables in ELW, JP, AKM and HJ Model**. The correlations between the factor scores estimated from the factor analysis and the time series of state variables from the MCMC procedure are calculated during the full sample period (11/04/2008-09/29/2017). For the ELW, JP and AKM model, we set $m = 2$ to work out the factor scores to reflect two risk factors in the pricing formula for a VS contract, while m is equal to 3 in the HJ model accordingly.

to verify the performance of model calibration. We further gauge the model stability by investigating the out-of-sample pricing performance over all the VS contracts. In addition to the mean term structure of the variance swap rates, we further follow the practice in the literature to formulate the model-based variance swap premiums (VRPs) which show the magnitude of VRPs highly depends the time-to-maturity of the contracts during the sample period.

3.4.1 Mean Term Structure of Variance Swap Rates

We now specify the mean term structure of variance swap rates in all the four models to investigate the contribution of each component to the determination of the rate. More specifically, we first obtain the mean term structure of variance swap rates under the objective probability measure P by taking the unconditional expectation as follows:

$$
\begin{aligned}
&E^P_{ELW}[VS_{t,t+\tau}] = (1 - \phi_v^{ELW}(\tau) - \phi_m^{ELW}(\tau))\theta_m^Q + \phi_v^{ELW}(\tau)\theta_v^P + \phi_m^{ELW}(\tau)\theta_m^P, \\
&E^P_{AKM}[VS_{t,t+\tau}] = (1 + \lambda_1 E_t^Q[(J^{s,Q})^2])(1 - \phi_v^{AKM}(\tau) - \phi_m^{AKM}(\tau))\tilde{\theta}_m^Q + E_t^Q[(J^{s,Q})^2]\lambda_0 \\
&+ (1 + \lambda_1 E_t^Q[(J^{s,Q})^2])\phi_v^{AKM}(\tau)\theta_v^P + (1 + \lambda_1 E_t^Q[(J^{s,Q})^2])\phi_m^{AKM}(\tau)\theta_{\tilde{m}}^P, \\
&E^P_{HJ}[VS_{t,t+\tau}] = (1 - \phi_v^{HJ}(\tau) - \phi_m^{HJ}(\tau))\theta_m^Q + \phi_\lambda^0(\tau)\lambda_\infty + \phi_v^{HJ}(\tau)\theta_v^P + \phi_m^{HJ}(\tau)\theta_m^P \\
&+ \phi_\lambda^{HJ}(\tau)\theta_\lambda^P,
\end{aligned}
\tag{11}
$$

where the mean term structure in the ELW model is a weighted average of the statistical mean of the instantaneous variance rate, θ_v^P, the statistical mean of the central tendency factor, θ_m^P ($\theta_{\tilde{m}}^P$ instead in the AKM model where $\tilde{\theta}_m^Q = (\kappa_v^Q \theta_m^Q + \mu_v^Q \lambda_0)/\tilde{\kappa}_v^Q$, and $\tilde{m} = (\kappa_v^Q m_t + \mu_v^Q \lambda_0)/\tilde{\kappa}_v^Q$ with $\tilde{\kappa}_v^Q = \kappa_v^Q - \mu_v^Q \lambda_1$), and and the common risk-neutral (unconditional) long-term mean for both the variance rate v_t and the central tendency m_t, θ_m^Q. In addition to these three factors, the constant jump intensity λ_0 is counted in the AKM model, while both the risk-neutral long-term mean of the jump intensity, λ_∞ and the statistical intensity mean of the jump factor, θ_λ^P are taken into account in the HJ model. In particular, if $\lambda_0 = \lambda_1 = \alpha = \lambda_\infty = \beta_0 \equiv 0$, the jumps in both models vanish, and both models then converge to the ELW model (with no jump components). Also, as suggested by Aït-Sahalia, Karaman and Mancini (2015), the JP model can be regarded as a special case of the AKM model when both μ_v^Q and μ_v^P are set as zero (i.e., $\mu_v^Q = \mu_v^P \equiv 0$), and the intensity of jumps may employ the same specification that is a function of variance.

As suggested in Equation (11), the loading coefficients measure the magnitude of the contemporaneous responses of the variance swap term structure towards unit shocks on risk components (e.g., variance, central tendency and jump). Panel A of Figure 5 plots the term structure of all risk responses. In all the four models, the monotonically decreasing function of risk loading coefficient over time shows that the variance risk factor v_t has a transient and dominant contribution on the mean term structure of variance swap rates at short maturities, and such influence gradually declines over maturities. In particular, the AKM model amongst all the models puts the highest weights on the volatility factor in the short term. This results from the contributions made by jumps to capture price variance driven by the large estimator λ_1, and such contributions are eventually reflected by the jump-adjusted weights on the volatility factor due to its specification of the jump intensity that is a function of variance. Accordingly, the similar pattern of the factor loading on v_t can be observed in the JP model.

In contrast, the impact of the central tendency factor m_t is mainly governed by ϕ_m which is persistent and substantial over time in order to construct an upward-sloping mean term structure. The increasing coefficients of m_t in the ELW and JP model, associated with their

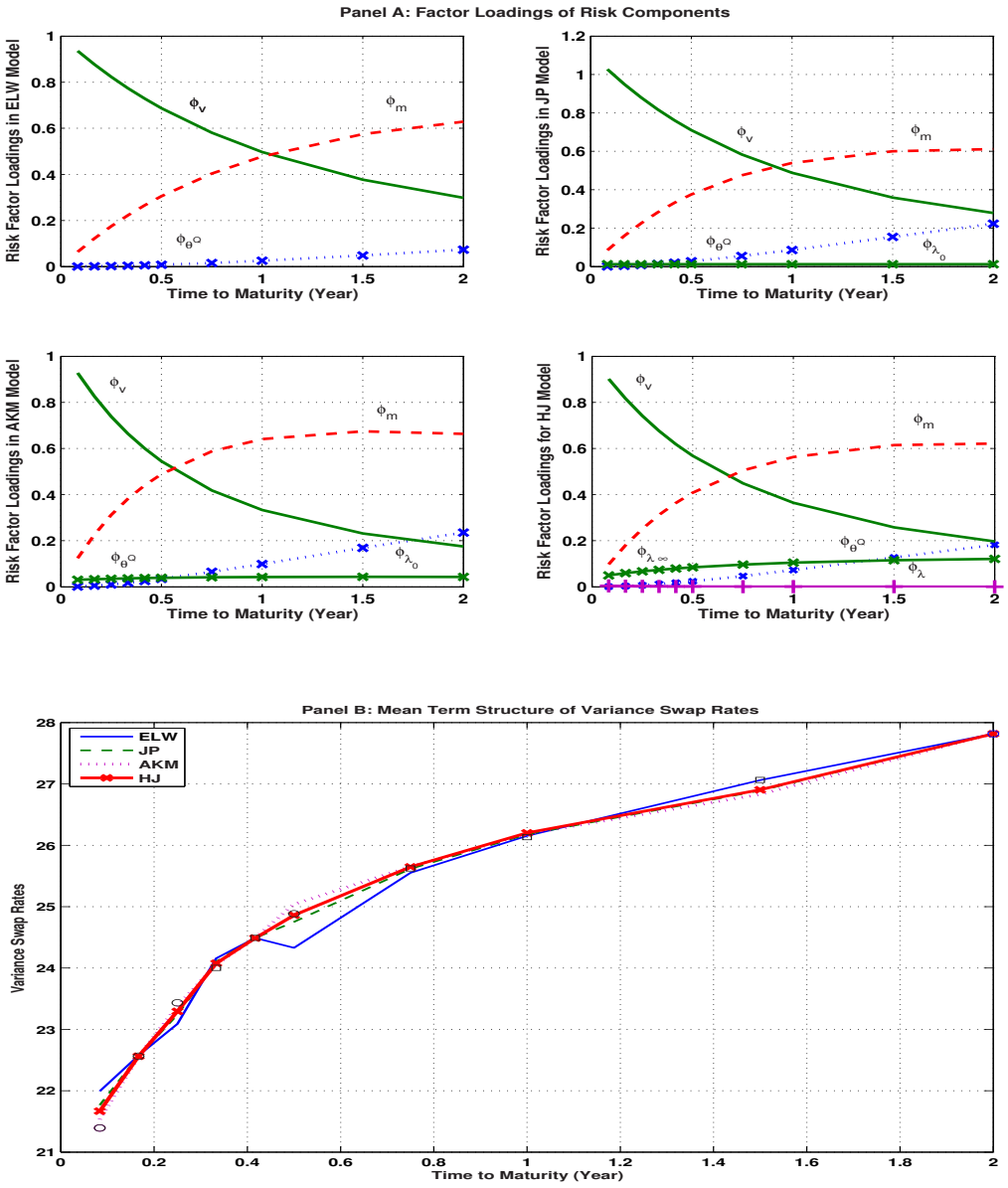

Figure 5: **Risk Loadings and Mean Term Structure of Variance Swap Rates.** In Panel A, the contemporaneous responses of the variance swap term structure to unit shocks on the instantaneous variance rate v_t (denoted by the solid line), the central tendency factor m_t (e.g., ϕ_λ denoted by the dashed line) and the jump risk factor λ_t (denoted by the solid line with "+" in HJ Model) are plotted, while the loading on the long-term mean of the central tendency (θ^Q) is represented by the dotted line with "x". The responses of other risk factors are denoted accordingly (e.g., ϕ_{λ_0} in JP and AKM Model and ϕ_{λ_∞} in HJ Model), but ϕ_{λ_∞} collapses to the horizontal axis due to their small values. In Panel B, the mean term structure of variance swap rates produced by ELW,JP, AKM and HJ Model is denoted by the dashed line, the dotted line, the solid line and the solid line with "x", respectively, while the empirical means of the VS rates with the five time-to-maturities reported in Table 3 are represented by "$\circ$", while the other five VS contracts are represented by "$\square$" (e.g., $\tau = 4, 5, 9, 12$ and 18 months). The in-the-sample period is from November 4, 2008 to October 13, 2014.

relatively large magnitudes, show that this factor's influence intensifies progressively with the increasing maturity since inception, especially in the presence of jumps in price. Furthermore, if jumps in variance are allowed in both the AKM and HJ model, the coefficients of m_t increase steadily till the medium term and then turn down gradually afterwards, suggesting the declining contemporaneous contributions towards the term structure of variance swap rates. Moreover, the weight of the risk-neutral long-term mean θ^Q (a constant) monotonically grows as the maturity of variance swap increases in all models as an additional adjustment, which is helpful to mitigate the derivation of variance rates in the long term.

It is distinguishable in both the AKM and HJ model about the manner that the jump risk factor (λ_t) contributes to the responses of the variance swap term structure. Compared to the ELW model, the specification of the jump intensity in the AKM model suggests that the contribution of λ_t is decomposed into two components: the weight function ϕ_{λ_0} and the weight adjustments on other factors (e.g., v_t, m_t and θ^Q) that result in the substantial upward shifts in the coefficients of v_t, m_t and θ^Q. On the other hand, the self-exciting specification of the jump intensity in the HJ model suggests that the contribution of the jump component is characterized only by two factors: λ_∞ with the weight function ϕ_{λ_∞} and λ_t with ϕ_λ, independent of those risk factors associated with diffusion components. Nevertheless, the values of these two weight functions are relatively small due to the nature of jumps estimated in the present paper.

Based on those response functions of the risk factors discussed above, we estimate the mean term structure of variance swap rates with a set of time-to-maturities from one month up to two years. As reported in Panel B of Figure 5, the variance swap rates with maturity up to two years range from 21% to 28% in terms of volatility percentage units. This upward-sloping term structure is certainly consistent with the negative market prices (γ_v and γ_m) in Table 4. As shown in Table 7, it seems that apart from the ELW model, the rest three models are calibrated to the empirical term structure of variance swap rates quite well during the in-the-sample period, and that both the AKM and HJ model achieve the very close RMSE of about 0.60 volatility units. Among all these models, in particular, the HJ model well captures the dynamics of the variance swap rate means across all the time-to maturities in

the 6-year sample period from November 4, 2008 to October 13, 2014. This indicates that jumps in price or variance or both do play a crucial role in pricing variance swaps, and then reminds the importance of the jump risk (λ_t) as the third pricing factor, apart from the variance factor (v_t) and the central tendency factor (m_t) in the variance swap market, as suggested in Section 2.1.

3.4.2 Out-of-Sample Pricing Performance

We further gauge the stability of all the four models by investigating their out-of-sample performance of pricing the VS contracts. The out-of-sample period is from October 13, 2014 to September 29, 2017. This sample contains about 3 years of data, 748 daily observations for both the S&P500 index returns and the variance swap rates.

To corroborate the estimators obtained from the MCMC approach, we first analyze the variance swap pricing errors for the four models obtained in both the in-the-sample and out-of-sample analysis. Table 7 summarizes the pricing errors in both the in-the-sample and out-of-sample periods, which are defined as model-based minus actual variance swap rates, both in volatility terms.

First, the in-the-sample pricing errors of variance swaps suggest that the diffusion-jump models (e.g., the JP, AKM and HJ model) perform much better than the diffusion model (the ELW model) across all the time-to-maturities with $\tau = 1, 2, 3, 6$, and 24 months, while the 2- and 24-m VS contracts can be priced with much smaller errors.[12] Second, when the rich information in the S&P500 index is captured by the jump component, apart from a diffusion one, this greatly reduces pricing errors from 0.760 volatility units to 0.641, equivalently, a 15.78% decrease in pricing errors. Moreover, the introduction of jumps in volatility clearly can further enhance the pricing performance, evident by the decrease in the overall RMSE from 0.641 in the JP mdoel to 0.605 in the AKM model (namely, a 5.62% decrease in pricing errors) and 0.609 in the HJ model (e.g., 4.99%). As a result, jump in both the index and volatility can improve the pricing performance by over 20% in total. Third, the over-all

[12] As shown in Panel B of Figure 5, the price errors for the 4-, 5-, 9-, 12, and 18-m VS contracts during the in-the-sample period are also very small.

performance of the HJ model is competitive with the AKM model, implying that the self-exciting specification for the jump intensity is a good alternative for the one of the linear function of the volatility (e.g., $\lambda_t = \lambda_0 + \lambda_1 v_t$).[13] Interestingly, the AKM model apparently has the better pricing performance in those short-term contracts (e.g., the 1-,2- and 3-m VS contracts), while those medium- and long-term contracts (e.g., the 6- and 24-m ones) can be priced with relatively smaller errors.

The out-of-sample pricing performance in Panel B of Table 7 confirms the stability of these models. Clearly, the incorporation of jumps in both the index and volatility (e.g., the AKM and HJ model) can improve the performance of model calibration to the quoted variance swap rates in terms of root mean square error (RMSE) and bias. Also, the 2- and 24-m VS contracts can be priced with high precision, consistent with their performance in the sample period. Meanwhile, the AKM and HJ model perform similarly to the one presented in Panel A, e.g., the very close overall RMSEs in both models and the relative better performance of those short-term contracts in the AKM model, compared with the improvement of pricing errors of the medium- and long-term contracts in the HJ model.

Furthermore, Panel C of Table 7 reports the pricing errors of the 4-, 5-, 9-, 12- and 18-m VS contracts which are not used for model calibration in the sampler period. Surprisedly, the overall performance of the four models is substantially improved across all the time-to-maturities, with the overall RMSE of 0.22 volatility units, compared with the performance in the 1-, 2-, 3-, 6- and 24-m VS contracts, with the overall RMSE ranging from 0.611 units in the ELW model to 0.494 units in the HJ model. This then indicates the less information in these contracts, consistent with our analysis in Section 3.1. These results then confirm that the 1-, 2-, 3-, 6- and 24-m VS contracts indeed possess the richest information in the variance swap market underlying the S&P500 index. Equipped with these observations, we further investigate the power of these VS contracts towards return predictability in the out-of-sample period.

[13]Note that the self-exciting specification for the jump intensity proposed in this article has another advantage, compared with the linear specification. That's, this specification may provide an analytical solution to the dynamic asset allocation problem in the variance swap market, while the latter one does not. This further motivates us to investigate the role of jump in variance swap investments.

Panel A: In-the-Sample Pricing Errors of Variance Swaps							
τ		1	2	3	6	12	Overall
ELW	RMSE	1.228	0.207	0.635	0.965	0.027	0.760
	Bias	0.599	0.013	-0.344	-0.555	0.000	-
JP	RMSE	1.100	0.198	0.540	0.715	0.029	0.641
	Bias	0.372	0.003	-0.192	-0.132	0.000	-
AKM	RMSE	1.047	0.079	0.498	0.690	0.029	0.605
	Bias	0.140	-0.001	-0.057	0.155	0.000	-
HJ	RMSE	1.067	0.050	0.517	0.669	0.017	0.609
	Bias	0.273	0.000	-0.139	-0.024	0.000	-
Panel B: Out-of-Sample Pricing Errors of Variance Swaps							
τ		1	2	3	6	12	Overall
ELW	RMSE	0.930	0.081	0.483	0.873	0.049	0.611
	Bias	0.653	0.009	-0.385	-0.753	0.002	-
JP	RMSE	0.870	0.064	0.439	0.753	0.039	0.552
	Bias	0.576	0.002	-0.338	-0.615	0.000	-
AKM	RMSE	0.795	0.041	0.388	0.666	0.028	0.496
	Bias	0.485	0.002	-0.283	-0.504	.001	-
HJ	RMSE	0.801	0.031	0.392	0.649	0.025	0.494
	Bias	0.500	0.001	-0.290	-0.506	0.000	-
τ		4	5	9	12	18	Overall
ELW	RMSE	0.186	0.060	0.205	0.118	0.495	0.262
	Bias	0.106	-0.012	-0.086	0.005	0.316	-
JP	RMSE	0.178	0.030	0.187	0.131	0.433	0.233
	Bias	0.111	-0.005	-0.070	0.015	0.281	-
AKM	RMSE	0.165	0.029	0.171	0.157	0.406	0.222
	Bias	0.075	-0.003	-0.042	0.014	.211	-
HJ	RMSE	0.168	0.021	0.164	0.163	0.400	0.220
	Bias	0.079	-0.001	-0.041	0.016	0.207	-

Table 7: **Pricing Errors of Variance Swap Rates in ELW, JP, AKM and HJ Model**. The pricing errors is defined as the model-based minus observed variance swap rates, both in volatility percentage units, e.g., $(\sqrt{VS_{t,t+\tau}^{Model}} - \sqrt{VS_{t,t+\tau}^{Market}}) \times 100$. Both the mean (Bias) and RMSE of pricing errors for variance swap rates under all the four models both in the sample period (11/04/2008-10/13/2014) and out of sample period (10/14/2014-09/29/2017).

Finally, we conduct the Diebold-Mariano(DM) test (Diebold and Mariano (1995)) which is indeed a t-statistic test to the pricing errors for all the four models. Table 8 reports the statistics for both the in-the-sample and out-of-sample pricing errors. The results show that in bother period, the AKM and HJ model outperform other two models, while they achieve the very close pricing performance for the 1-, 3- and 6-m VS contracts. For the the 4-, 5-, 9- and 12-m VS contracts, none of these model can persistently outperform the other across all the time-to-maturities, partially rooted in the relatively small RMSEs in all the models, while the HJ model outperforms the others for the 18-m VS contract.

3.4.3 Term Structure of Variance Risk Premiums (VRPs)

After examining the calibration of the mean term structure of the variance swap rates and the out-of-sample pricing performance, we further investigate the dynamics of the variance risk premiums (VRPs) which reflect how much investors are willing to pay against variance risk. Following Aït-Sahalia, Karaman and Mancini (2015), the annualized time-t variance risk premium (VRP) is defined as the difference between the objective and risk-neutral quadratic variations at time t as follows:

$$\begin{aligned} VRP(t,\tau) &= E_t^P[QV_{t,t+\tau}] - E_t^Q[QV_{t,t+\tau}] \\ &= E_t^P\left[\frac{1}{\tau}\int_t^{t+\tau} v_u du + \frac{1}{\tau}\sum_{u=N_t}^{N_{t+\tau}}(J_u^s)^2\right] - E_t^Q\left[\frac{1}{\tau}\int_t^{t+\tau} v_u du + \frac{1}{\tau}\sum_{u=N_t}^{N_{t+\tau}}(J_u^s)^2\right]. \end{aligned} \tag{12}$$

It represents the expected payoff for a long position in a VS contract over the period from t to $t+\tau$. Also, this measure provides rich information of aggregate risk aversion amongst investors toward economic uncertainty (Bollerslev, Tauchen and Zhou (2009)). The definition of the VRP in Equation (12) implies the contribution of the jump component in the current setup can be formulated as

$$VRP_J(t,\tau) = E_t^P\left[\frac{1}{\tau}\sum_{u=N_t}^{N_{t+\tau}}(J_u^s)^2\right] - E_t^Q\left[\frac{1}{\tau}\sum_{u=N_t}^{N_{t+\tau}}(J_u^s)^2\right].$$

The upper panel in Figure 6 plots the term structure of the VRPs over time. Over the

Panel A: DM Test for In-the-Sample Pricing Errors of Variance Swaps						
τ	ELW-JP	ELW-AKM	ELW-HJ	JP-AKM	JP-HJ	AKM-HJ
1	3.446	2.658	3.287	1.364	1.379	-0.864
2	6.617	9.387	9.205	8.885	8.763	7.839
3	4.143	3.653	4.322	2.202	1.938	-1.380
6	3.215	2.437	3.275	0.426	1.781	0.518
24	-11.092	-11.107	9.819	10.274	10.422	10.313
Panel B: DM Test for Out-of-Sample Pricing Errors of Variance Swaps						
1	2.870	3.352	5.357	3.819	9.421	-0.389
2	8.267	9.008	9.167	9.738	9.790	9.860
3	2.863	3.362	5.321	3.946	12.328	-0.319
6	2.692	2.712	4.570	2.664	6.537	0.449
24	16.389	16.369	16.464	15.869	16.152	15.622
4	3.256	4.741	3.755	4.369	3.546	-1.702
5	19.899	19.055	18.819	6.026	12.320	14.325
9	2.734	2.569	3.014	2.217	2.906	2.580
12	-2.104	-4.557	-5.141	-6.126	-6.404	-2.017
18	2.603	2.561	3.075	2.048	3.443	0.859

Table 8: **Diebold-Mariano(DM) Test for Pricing Errors of Variance Swap Rates in ELW, JP, AKM and HJ Model**. The DM test conducts a one-sided t-statistic test for the time series of the squared pricing errors, e.g., $\{e_t^2\}_{t=1}^T$, where the pricing error is defined in Table 7. The test statistics measure whether the first model in each pair has significantly smaller squared pricing errors than the second model, as indicated by the positive sign and the negative one otherwise. The critical value at the 5% level for the one-sided test (e.g., the left- or right-tailed one) is ± 1.76. The full sample splits into two periods: the in-the-sample one (11/04/2008-10/13/2014) and the out-of-sample one (10/14/2014-09/29/2017).

full sample period, the VRPs are negative across to the time-to-maturities, and start to converge when the market regains the recovery from the 2008 financial crisis and the 2010-2102 European debt crisis. In particular, dynamics of the 1-m VRPs suggest that as the response to the Lehman Brothers' bankruptcy on September 15, 2008, the 1-m VRPs present the highest level in absolute magnitude in the early sample period, increase moderately during the European crisis, and then stay at the low level in the rest sample period, showing that the short-term economic uncertainty among investors has been substantially mitigated especially in the recent years. In contrast, the long-term (24-m) VRPs are much larger than the others in absolute magnitude in all the models, showing that investors are willing to pay more premiums to hedge unexpected increases in variance in the long run. More specifically, investors have started to pay less (about 1% on average) when facing tranquil markets (e.g., since 2012 after the European debt crisis), compared with the more premiums paid in the volatile market due to the "search-for-yield" effect, for example, ranging from 1% to 5% in the ELW model during the financial crisis in 2008. It seems that the JP model delivers the relatively small VRPs across all maturities, while the VRPs produced by the AKM and HJ model exhibit very similar behavior. In sum, these plots complement the empirical results reported in Aït-Sahalia, Karaman and Mancini (2015) after the final crisis in 2008.

Similarly, the term structure of the jump VRPs implied in the JP, AKM and HJ model are plotted in the low panel in Figure 6. Compared with the VRPs reported in the upper panel, the jump VRPs are much smaller in each time-to-maturity category. Since the jump intensity depends on the variance state variable in both the JP and AKM model, e.g., $\lambda_t = \lambda_0 + \lambda_1 v_t$, the jump VRPs behave similarly to those VRPs in the upper panel. On the other hand, the behavior of the jump VRPs in the HJ model are quite different from those in the JP and AKM model, mainly because of its distinct specification of the jump intensity. Meanwhile, the HJ model produces much larger premiums in absolute magnitude, on average, ranging from 0.045% for $\tau = 1$ month to 0.345% for $\tau = 24$ months, compared with the corresponding premium of 0.003% and 0.047% in the AKM model. These gaps in the jump VRPs suggest that the value of jump risk in the JP and AKM model might be underestimated, which may further result in a substantial loss in investors' economic benefits, as discussed in Section 5.

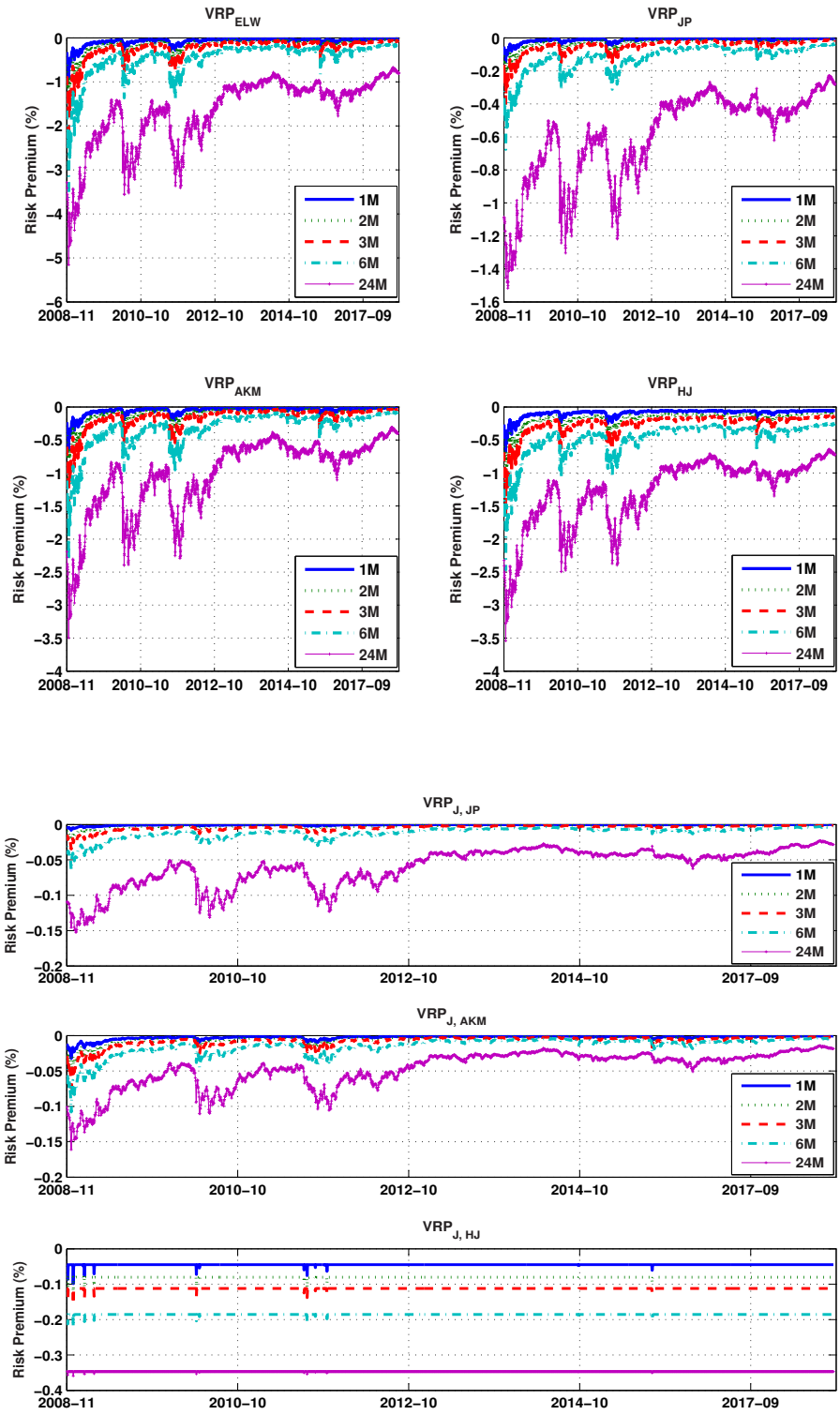

Figure 6: **Term Structure of Variance Risk Premiums (VRPs).** In the top panel, the term structures of variance swap premiums produced by ELW,JP, AKM and HJ Model across the 1-, 2-, 3-, 6- and 24-m time-to-maturities are denoted by the solid line, the dotted line, the dashed line, the dash-dotted line and the solid line with "+", respectively, while the term structure of the jump risk premiums in these models are plotted similarly. The sample period ranges from November 4, 2008 to September 27, 2017.

4 Predictability of Stock Returns

To understand the nature of these pricing models better, we now further exploit the predictability of stock returns using the information extracted from the out-of-sample VS contracts under these models. The underlying logic is rooted in the observation that since the variance swap rate can be formulated with both model parameters and state variables in each model, this then builds up a direct channel between the variance swap rates and the state variables in a model-based setup. On the one hand, a number of recent empirical studies document that the information embedded in the VS contracts (for example, the variance risk premium (VRP)) has the power to predict stock returns in short horizons (e.g, Bollerslev, Tauchen and Zhou (2009) and Li and Zinna (2017) among others). This consensus in the empirical literature is also supported by the analysis in Section 3.1. On the other hand, Adrian and Rosenberg (2008) evidence that both the long- and short-run components of market volatility are significantly priced towards the attribution of stock returns. As a result, by virtue of this channel, we can measure the efficiency of a pricing model that extracts the information from the VS contracts in terms of by its predictability of stock returns, apart from the performance of pricing VS contracts.

Following the definition of the k-period return $R_{t+k,t}$ in Equation (2), we consider the regressions of the S&P500 index returns on the state variables and on the VRPs estimated from the VS contracts as follows:

$$\begin{aligned} R_{t+k,t} &= b_0(k) + b_v(k)v_t + b_m m_t + b_\lambda(k)\lambda_t \mathbf{1}_{\mathcal{M}_i=HJ} + \epsilon_t(k), \\ R_{t+k,t} &= c_0(k) + c_v(k)VRP_\tau(t) + \varepsilon_t(k) \end{aligned} \tag{13}$$

where the state variables v_t, m_t and λ_t are estimated by calibrating the models to the VS contracts in the out-of-sample period, and $\mathbf{1}_{(\cdot)}$ is equal to one if the HJ model is considered and zero for the ELW, JP and AKM model, while VRP_τ indicates the estimated VRPs from the VS contracts in the model. In both regressions, we use the adjusted R^2 to measure the prediction performance in each model.

We further define an out-of-sample R^2 measure as follows:

$$R^2_{OS,\mathcal{M}_i-\mathcal{M}_j} = 1 - \frac{\sum_{t,k}^{\mathcal{M}_i}(R_{t+k,t} - \hat{R}_{t+k,t})^2}{\sum_{t,k}^{\mathcal{M}_j}(R_{t+k,t} - \hat{R}_{t+k,t})^2}, \tag{14}$$

where $\hat{R}_{t+k,t}$ indicates the predicated cumulated returns under Model $\mathcal{M}$. In each model, we estimate $\hat{R}_{t+k,t}$ as follows:

- We choose a set of the VS contracts during the out-of-sample period, and run the minimization at time t:

$$\min_{v_t>0,m_t>0,l_t>0} \Sigma_n(VS_{n,t} - \beta_0 - \beta_v v_t - \beta m_t - \beta_\lambda \mathbf{1}_{\mathcal{M}_i=HJ}\lambda_t)^2,$$

  where the coefficients $\beta_0, \beta_v, \beta_m, \beta_\lambda$ can be backed up from the pricing formula for the variance swap rates in each model, and n indicates the number of selected contracts.

- We then use the estimated state variables $\hat{v}, \hat{m}$ and $\hat{\lambda}$ to run the first regression in Equation (13) to collect the coefficient estimators, $\hat{b}_0, \hat{b}_v, \hat{b}_m$ and $\hat{b}_\lambda$ for each horizon k.

- The estimated cumulated return $\hat{R}_{t+k,t}$ is then given by

$$\hat{R}_{t+k,t} = \hat{b}_0(k) + \hat{b}_v(k)\hat{v}_t + \hat{b}_m\hat{m}_t + \hat{b}_\lambda(k)\hat{\lambda}_t\mathbf{1}_{\mathcal{M}_i=HJ}.$$

For a comprehensive comparison, we finally conduct a regression of the k-period return $R_{t+k,t}$ on all the 10 VS contracts, and follow the PCA analysis used in Section 2.2 to extract the full information embedded in the sample.

Table 9 reports the regression statistics to measure the return predictability of the state variables estimated from the out-of-sample VS contracts with the time-to-maturities of $\tau = 1, 2, 3, 6$, and 24 months.[14] The resulted adjusted-R^2 suggests that the estimated volatility and jump variables do have the power to predict the S&P500 index returns in all the models,

[14]We check the correlations of the latent variables used in the regressions, and find that in all the models, the correlation coefficients among the state variables (v, m, λ) and $\lambda \equiv 0$ in the ELW model are lower than 0.50 in absolute magnitude with all the forecasting horizons.

Panel A: Regressions on State Variables with Calibration of $VS_{t,t+\tau(\tau=1,2,3,6,24)}$						
Model	R^2	Forecasting Horizon (k months)				
		1	3	6	12	24
ELW	Adj-R^2	0.073	0.144	0.163	0.387	0.384
JP	Adj-R^2	0.073	0.142	0.164	0.387	0.383
AKM	Adj-R^2	0.073	0.140	0.165	0.378	0.380
HJ	Adj-R^2	0.071	0.147	0.169	0.415	0.433
JP-ELW	R^2_{OS}	0.000	-0.002	0.001	0.000	-0.002
AKM-ELW	R^2_{OS}	0.000	-0.004	0.002	-0.014	-0.007
HJ-ELW	R^2_{OS}	0.000	0.005	0.009	0.049	0.084
HJ-JP	R^2_{OS}	0.000	0.007	0.008	0.048	0.085
HJ-AKM	R^2_{OS}	0.000	0.009	0.007	0.062	0.090
Panel B: Regression on State Variables with Calibration of All VS Contracts						
ELW	Adj-R^2	0.074	0.146	0.189	0.435	0.383
JP	Adj-R^2	0.073	0.144	0.191	0.435	0.382
AKM	Adj-R^2	0.073	0.142	0.191	0.423	0.379
HJ	Adj-R^2	0.072	0.146	0.200	0.480	0.416
JP-ELW	R^2_{OS}	0.000	-0.002	0.002	0.000	-0.002
AKM-ELW	R^2_{OS}	0.000	-0.005	0.003	-0.021	-0.005
HJ-ELW	R^2_{OS}	0.000	0.002	0.015	0.081	0.058
HJ-JP	R^2_{OS}	0.000	0.004	0.012	0.081	0.059
HJ-AKM	R^2_{OS}	0.000	0.007	0.012	0.100	0.063
Panel C: Regression on VRPs with Calibration of $VS_{t,t+\tau(\tau=1,2,3,6,24)}$						
VRP^{ELW}_{1m}	Adj-R^2	0.071	0.140	0.070	0.112	0.392
VRP^{ELW}_{2m}	Adj-R^2	0.071	0.137	0.075	0.124	0.389
VRP^{ELW}_{3m}	Adj-R^2	0.071	0.134	0.081	0.134	0.387
VRP^{ELW}_{6m}	Adj-R^2	0.069	0.121	0.099	0.175	0.376
VRP^{ELW}_{24m}	Adj-R^2	0.033	0.031	0.157	0.3791	0.087

Table 9: **Out-of-Sample Return Predictability in ELW, JP, AKM and HJ Model**. The stock return predictability of the state variables is examined in the out-of-sample period (10/14/2014-09/29/2017) with the 1-, 2-, 3-, 6-, 24-m VS contracts (in Panel A) and all the ten VS contracts with $\tau = 1, 2, 3, 4, 5, 6, 9, 12, 18, 24$ (in Panel B), while Panel C reports the regression statistics on VRPs estimated in Section 5.2 under ELW Model.

Panel A: Regressions on All VS Contracts using PCA						
PCA Factors	R^2	Forecasting Horizon (k months)				
		1	3	6	12	24
2	Adj-R^2	0.073	0.140	0.186	0.411	0.379
3	Adj-R^2	0.072	0.181	0.185	0.427	0.385
4	Adj-R^2	0.082	0.188	0.276	0.565	0.397
5	Adj-R^2	0.140	0.219	0.359	0.565	0.458
6	Adj-R^2	0.172	0.218	0.408	0.565	0.489
7	Adj-R^2	0.192	0.226	0.411	0.568	0.572
Panel B: Akaike information criterion (AIC)						
2	-	-7.053	-6.385	-5.846	-5.493	-6.216
3	-	-7.051	-6.424	-5.845	-5.527	-6.217
4	-	-7.091	-6.465	-6.048	-5.683	-6.218
5	-	-7.088	-6.462	-6.055	-5.724	-6.366
6	-	-7.180	-6.536	-6.181	-5.783	-6.392
7	-	-7.209	-6.534	-6.202	-5.780	-6.634
Panel C: Bayesian information criterion (BIC)						
2	-	-7.034	-6.365	-5.824	-5.467	-6.173
3	-	-7.025	-6.397	-5.816	-5.493	-6.160
4	-	-7.059	-6.432	-6.012	-5.641	-6.146
5	-	-7.050	-6.422	-6.013	-5.673	-6.280
6	-	-7.136	-6.490	-6.131	-5.723	-6.291
7	-	-7.158	-6.481	-6.145	-5.712	-6.519

Table 10: **Regression Statistics of the Out-of-Sample Return Predictability on All VS Contracts**. In Panel A, the stock return predictability of all the VS contracts is examined in the out-of-sample period (10/14/2014-09/29/2017) by applying the PCA decomposition with 7 PCA factors (< 10). In Panel B and C, the AICs and BICs obtained from the regressions are reported.

while R^2_{OS}s suggest that the HJ model can capture more information from the VS contracts to have better predication on the underlying stock return in the long run. Moreover, it seems that the incorporation of jumps in volatility, associated with a flexible specification of jump intensity, may improve the predictability of the index returns in terms of adjusted-R^2s, but its magnitude is partially impacted by less jumps in the out-of-sample period, as shown in Figure 4. More interestingly, Panel C in Table 9 shows that the VRPs with various time-to-maturities have the close performance in predicting stock return over all the horizons, which somehow devalues the importance of the VRPs in stock return predictability widely studied in the literature.[15]

Finally, the PCA analysis on the regression of $R_{t+k,t}$ on all the VS contracts shows that only about 7 VS contracts of 10 are sufficient to achieve the much high prediction power in the future stock returns, as implied by the marginal changes in both the AICs and BICS. As a result, we do not report the regression statistics with more factors. In sum, Table 9 and 10 suggest that using the variance swap rates (equivalently, the state variable v, m and λ), rather than the VRPs, can achieve better predictability of the underlying stock returns.

5 Empirical Analysis of Variance Swap Investments

The preceding analysis suggests that the JP, AKM and HJ model can be calibrated to the empirical mean term structure of variance swap rates quite well. This indicates that these two models can serve as good alternatives for the AKM model in terms of valuing variance swap rate contracts. However, it is unclear how jumps in variance may affect investor's decisions on variance swap investment. Since the AKM model seems hard to provide an analytical solution to the problem of variance swap allocation, we mainly consider the other three models (namely, the ELW, JP and HJ model). Specifically, we analyze the optimal allocations to variance swap contracts, study the role of jumps in variance swap portfolios, and quantify the cost of economic welfare in investing variance swaps in the presence of both model and parameter mis-specification.

[15]We also check the prediction performance of the VRPs obtained in the JP, AKM and HJ model, and have the very similar results observed in the ELW model.

5.1 Optimal Variance Swap Allocation: Theoretical Results

We first solve the optimal portfolio choice problem with variance swaps before stepping into an empirical analysis. Specifically, to better understand a risk-averse investor's behavior in variance investments in a stochastic environment, particularly in comparison with the short-long strategy in the pure-diffusion model of Egloff, Leippold and Wu (2010), we consider a model where this investor can trade three variance swaps and a money market account.[16] As in Egloff, Leippold and Wu (2010) and Jin and Zhang (2012), we assume that at time t the investor initiates three new variance swap contracts with the delivery prices equal to the prevailing variance swap rates $K_1 = VS_{t,t+\tau_1}$, $K_2 = VS_{t,t+\tau_2}$ and $K_3 = VS_{t,t+\tau_3}$. Thus her wealth W_t can be written as

$$W_t = W_t^M + W_{1t}(VS_{t,t+\tau_1} - K_1) + W_{2t}(VS_{t,t+\tau_2} - K_2) + W_{3t}(VS_{t,t+\tau_3} - K_3),$$

where W_t^M denotes the amount of money invested in the money market account, and W_{1t}, W_{2t} and W_{3t} denote the dollar notional amount invested in the three variance swaps, respectively.

[16] It is worth mentioning that incorporating the stock will introduce two more sources of risk: the diffusion B_{1t}^P and the jump N_t with jump size $J_t^{s,P}$. In essence, the jump in stock price and the jump in volatility are considered as two different jumps although they occur simultaneously because the two jumps have different random jump sizes. As a result, in order to deliver closed-form solution to the optimal portfolio choice problem when the investor can invest in the stock, we need to incorporate a new asset in addition to the stock. The price of the new asset is driven by the diffusion B_{1t}^P. For this, we can extend the stock price model in Section 2.3 of Liu (2007) by incorporating jump in the stock price. Specifically, the stock price is

$$\frac{dS_t}{S_{t-}} = (r_t + \gamma_v \rho v_t + \gamma_s r_t)dt + \sigma_r \sqrt{r_t} dB_{1t}^P + \rho \sqrt{v_t} dB_{2t}^P + (\exp(J_t^{s,P}) - 1)dN_t - g^P \lambda_t dt,$$

where r_t is the short rate. And then the investor is allowed to trade a zero-coupon bond, the stock and three variance swaps. In this model, the variance swap rate includes a new term $R_t = E^Q\left[\exp\left(-\int_t^T r_s ds\right)\int_t^T r_s ds\right]$. The dynamics of the expectation can be explicitly derived by using the methods in Duffie, Pan and Singleton (2000) for an affine short rate process. In particular, by adopting the Vasicek model for the short rate r_t, we can solve the optimal portfolio choice problem in the ODE-based closed form. If r_t is modeled by the CIR process, then, unlike the previous case, the optimal portfolio choice problem can be solved by combining the simulation-based method in Jin and Zhang (2012) and the ODE-based approach in the present paper. A noteworthy feature of the new model is that we can study how the interest rate affect the variance swap rate due to the presence of the new term R_t and the investor's demands for the stock, the bond and variance swap. Allowing the investor to access both stock and bond in addition to variance swap will certainly enrich the analysis. We leave this extension as future research. Also, we will show that a third variance swap is redundant in the model of Aït-Sahalia, Karaman and Mancini (2015).

As a result, we can write the wealth dynamics as

$$\frac{dW_t}{W_t} = r_t dt + w_{1t} dVS_{t,t+\tau_1} + w_{2t} dVS_{t,t+\tau_2} + w_{3t} dVS_{t,t+\tau_3} \tag{15}$$

where w_{1t}, w_{2t} and w_{2t} denote the fractions of wealth in the three variance swaps, respectively. Plugging the equations for variance swaps into Equation (15), we can recast the wealth dynamics as

$$\begin{aligned}
\frac{dW_t}{W_t} &= r_t dt + w_{1t}[\phi_v(\tau_1)\sigma_v\gamma_v v_t + \phi_m(\tau_1)\sigma_m\gamma_m m_t - (\phi_v(\tau_1) + \beta_0\phi_\lambda(\tau_1))\mu_v^Q\lambda_t]dt \\
&\quad + w_{2t}[\phi_v(\tau_2)\sigma_v\gamma_v v_t + \phi_m(\tau_2)\sigma_m\gamma_m m_t - (\phi_v(\tau_2) + \beta_0\phi_\lambda(\tau_2))\mu_v^Q\lambda_t]dt \\
&\quad + w_{3t}[\phi_v(\tau_3)\sigma_v\gamma_v v_t + \phi_m(\tau_3)\sigma_m\gamma_m m_t - (\phi_v(\tau_3) + \beta_0\phi_\lambda(\tau_3))\mu_v^Q\lambda_t]dt \\
&\quad + w_{1t}[\phi_v(\tau_1)\sigma_v\sqrt{v_t}dB_{2t}^P + \phi_m(\tau_1)\sigma_m\sqrt{m_t}dB_{3t}^P + (\phi_v(\tau_1) + \beta_0\phi_\lambda(\tau_1))J_t^{v,P}dN_t] \\
&\quad + w_{2t}[\phi_v(\tau_2)\sigma_v\sqrt{v_t}dB_{2t}^P + \phi_m(\tau_2)\sigma_m\sqrt{m_t}dB_{3t}^P + (\phi_v(\tau_2) + \beta_0\phi_\lambda(\tau_2))J_t^{v,P}dN_t] \\
&\quad + w_{3t}[\phi_v(\tau_3)\sigma_v\sqrt{v_t}dB_{2t}^P + \phi_m(\tau_3)\sigma_m\sqrt{m_t}dB_{3t}^P + (\phi_v(\tau_3) + \beta_0\phi_\lambda(\tau_3))J_t^{v,P}dN_t].
\end{aligned}$$

The next result gives the indirect value function.

Proposition 1 *Under the above assumptions, we have the following result:*

$$J(t, W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma}[f(t, X_t)]^\gamma = \frac{W^{1-\gamma}}{1-\gamma}\left[e^{A(t)+B_1(t)v_t+B_2(t)m_t+B_3(t)\lambda_t}\right]^\gamma \tag{16}$$

where the functions $A(t)$, $B(t) = (B_1(t), B_2(t))^\top$ and $B_3(t)$ satisfy the following equations:

$$\begin{aligned}
\frac{dA}{dt} &+ \kappa_m^P\theta_m^P B_2 + \alpha\lambda_\infty B_3 + \frac{1-\gamma}{\gamma}r = 0, \\
\frac{dB_1}{dt} &- \left(\kappa_v^P - \frac{1-\gamma}{\gamma}\sigma_v\gamma_v\right)B_1 + \frac{1}{2}\sigma_v^2 B_1^2 + \frac{1-\gamma}{2\gamma^2}\gamma_v^2 = 0, \\
\frac{dB_2}{dt} &+ \kappa_v^Q B_1 - \left(\kappa_m^P - \frac{1-\gamma}{\gamma}\sigma_m\gamma_m\right)B_2 + \frac{1}{2}\sigma_m^2 B_2^2 + \frac{1-\gamma}{2\gamma^2}\gamma_m^2 = 0, \\
\frac{dB_3}{dt} &- \alpha B_3 + \frac{\gamma-1}{\gamma}\widetilde{\pi}_{q1}^* E^Q[J^{v,Q}] + \frac{1}{\gamma}E^P\left[\left(\widetilde{\pi}_{q1}^* J^{v,P} + 1\right)^{1-\gamma} e^{\gamma(B_1+B_3\beta_0)J^{v,P}} - 1\right] = 0,
\end{aligned}$$

with $A(T) = B_1(T) = B_2(T) = B_3(T) = 0$ and $\widetilde{\pi}_{q1}^*$ given in Proposition 4 below.

Proof. See Appendix A. ■

It is interesting to note that the indirect value function $J(t, W_t, X_t)$ is independent of the maturities τ_1, τ_2 and τ_3 of the three variance swaps because the functions $A(t)$, $B(t) = (B_1(t), B_2(t))^\top$ and $B_3(t)$ do not depend on the three maturities. This can be seen from the above ordinary differential equations satisfied by the four functions. In other words, for a CRRA investor, as long as there are three different variance swaps for trading, the maturities of variance swaps are irrelevant for her investment performance measured by the indirect value function, $J(\cdot)$. This conclusion also holds true in Egloff, Leippold and Wu (2010) as shown in their Proposition 3. In essence, any two variance swaps with different maturities can span two diffusions in the two-factor pure-diffusion model of Egloff, Leippold and Wu (2010) while, as indicated by the nonsingular matrix Σ below, any three variance swaps with different maturities can span three sources of risk in the variance swap market. In contrast, the optimal portfolio weights depend on the maturities of three variance swaps shown below. To this purpose, from the equation (8), we let

$$\Sigma = \begin{pmatrix} \phi_v(\tau_1)\sigma_v\sqrt{v_t} & \phi_m(\tau_1)\sigma_m\sqrt{m_t} & \phi_v(\tau_1) + \beta_0\phi_\lambda(\tau_1) \\ \phi_v(\tau_2)\sigma_v\sqrt{v_t} & \phi_m(\tau_2)\sigma_m\sqrt{m_t} & \phi_v(\tau_2) + \beta_0\phi_\lambda(\tau_2) \\ \phi_v(\tau_3)\sigma_v\sqrt{v_t} & \phi_m(\tau_3)\sigma_m\sqrt{m_t} & \phi_v(\tau_3) + \beta_0\phi_\lambda(\tau_3) \end{pmatrix} \tag{17}$$

In general, the above matrix is nonsingular and thus a third variance swap is not redundant in our model since we disconnect the jump intensity from the variance. In the model of Aït-Sahalia, Karaman and Mancini (2015), according to their equations (8) and (9), the corresponding matrix can be represented as

$$\Sigma_1 = \begin{pmatrix} \phi_v(\tau_1)\sigma_v\sqrt{v_t} & \phi_m(\tau_1)\sigma_m\sqrt{m_t} & \lambda_1[(\mu_j^Q)^2 + \sigma_j^2]\phi_v(\tau_1) \\ \phi_v(\tau_2)\sigma_v\sqrt{v_t} & \phi_m(\tau_2)\sigma_m\sqrt{m_t} & \lambda_1[(\mu_j^Q)^2 + \sigma_j^2]\phi_v(\tau_2) \\ \phi_v(\tau_3)\sigma_v\sqrt{v_t} & \phi_m(\tau_3)\sigma_m\sqrt{m_t} & \lambda_1[(\mu_j^Q)^2 + \sigma_j^2]\phi_v(\tau_3) \end{pmatrix}$$

Clearly, the the matrix Σ_1 is singular because its first and third columns are proportional and this, in turn, implies that the third variance swap is redundant. Thus, variance swaps cannot provide independent exposures to two diffusions and one jump, making it difficult to

obtain analytic solution to optimal portfolio choice problem involving variance swap. The next result presents the optimal solution for w_{1t}, w_{2t} and w_{3t} in our model.

Proposition 2 *The optimal portfolio weight* $w^* = (w_{1t}^*, w_{2t}^*, w_{3t}^*)$ *is given by*

$$w^* = \left(\widetilde{\pi}_{b1}^*, \widetilde{\pi}_{b2}^*, \widetilde{\pi}_{q1}^*\right) \Sigma^{-1} \tag{18}$$

where

$$\widetilde{\pi}_{b1}^* = \frac{\gamma_v \sqrt{v_t}}{\gamma} + \sigma_v \sqrt{v_t} B_1(t), \widetilde{\pi}_{b2}^* = \frac{\gamma_m \sqrt{m_t}}{\gamma} + \sigma_m \sqrt{m_t} B_2(t),$$

and $\widetilde{\pi}_{q1}^*$ *solves the following optimization problem:*

$$\sup_{\widetilde{\pi}_{q1} \in [0,\infty)} -\widetilde{\pi}_{q1} E^Q[J^{v,Q}] + \frac{1}{1-\gamma} E^P\left[\left(1 + \widetilde{\pi}_{q1} J^{v,P}\right)^{1-\gamma} e^{\gamma(B_1 + B_3\beta_0) J^{v,P}} - 1\right]. \tag{19}$$

Furthermore, if $\gamma_v < 0, \gamma_m < 0, E^Q[J^{v,Q}] > E^P[J^{v,P}], B_1(t) < 0, B_2(t) < 0$ *and* $B_3(t) < 0$*, then*

$$\widetilde{\pi}_{b1}^* < 0, \widetilde{\pi}_{b2}^* < 0, \widetilde{\pi}_{q1}^* = 0. \tag{20}$$

Proof. See Appendix A. ■

Our calibration exercise shows that conditions of results (20) hold true. Intuitively, the first two results in (20) say that the investor has negative diffusion exposures to profit from the negative market prices of risk γ_v and γ_m. Interestingly, the third result in (20) suggests that the investor takes zero instead of a negative jump exposure despite the negative market prices of risk $\lambda_t(E^P[J^{v,P}] - E^Q[J^{v,Q}])$. The two reasons account for this result. On the one hand, the investor is prohibited from having a negative jump exposure by the no-bankruptcy constraint: $\widetilde{\pi}_{q1} \in [0, \infty)$. On the other hand, a positive jump exposure means losing the negative jump risk premium. As a result, the investor optimally takes zero jump exposure. We now apply these theoretical results to the sample of the variance swap contracts.

5.2 Optimal Variance Swap Allocations: A Short-Long-Short Strategy

Unlike the stream of literature that attempts to rationalize the magnitude of the risk premium based on various economic issues, we instead study how a trader allocates her wealth to variance swap contracts in order to benefit from the risk premium dynamics and further investigate the impact of jumps in variance on her asset allocations. We now assume that the trader allocates the initial wealth W_t at time t between the money market and the variance swap market. The trader has access to the money market account to balance out the investments by earning a risk-free interest rate. The variance swap contracts are initialized with zero costs and so they have zero initial values.

As aforementioned before, our closed-form solution presents a new way for understanding the short-long strategy studied in Egloff, Leippold and Wu (2010). For this, by turning off jumps in variance, we have their two-factor variance risk structure so that the variance swap rates across all maturities are determined by two sources of variations. Accordingly, the trader just need choose any two variance swap contracts with distinct time to maturities (say $0 < \tau_1 < \tau_2 < \infty$ without loss of generality), which could sufficiently span all the sources of risks in the variance swap market. When the investor only invests in the money market and variance swap contracts, the optimal portfolio weight of her wealth invested in these two contracts, $w^* = (w^*_{1t}, w^*_{2t})$ is equal to $w^* = (\tilde{\pi}^*_{b1}, \tilde{\pi}^*_{b2})\Sigma^{-1}$, as given in Equation (18) with $\tilde{\pi}^*_{q1} \equiv 0$ (and $B_3 \equiv 0$ in Proposition 1 as well), while the 2×2 matrix Σ is given as follows:

$$\Sigma = \begin{pmatrix} \phi_v(\tau_1)\sigma_v\sqrt{v_t} & \phi_m(\tau_1)\sigma_m\sqrt{m_t} \\ \phi_v(\tau_2)\sigma_v\sqrt{v_t} & \phi_m(\tau_2)\sigma_m\sqrt{m_t} \end{pmatrix}$$

and ϕ_v and ϕ_m are given in Equation (7), which is equivalent to the formulas in Egloff, Leippold and Wu (2010)(see Equation (43) and (44) in page 1298). It is clear that these results are valid in both the ELW and JP model but with distinct parameter sets reported in Table 4. Note that for the JP model, the exposures of the two contracts to the sources of risk, ϕ_v and ϕ_m in the matrix Σ are obtained by setting $\mu^Q_v = \mu^P_v = 0$ in the AKM model.

Inspecting the optimal investment decisions in the variance swap contracts under the two-factor variance risk specification (e.g., the ELW and JP model)[17], Egloff, Leippold and Wu (2010) suggest that the optimal allocations in the two variance swap contracts at short investment horizons depend not only on the market prices of both the variance risk (γ_v) and the central tendency risk (γ_m), but also on the exposures of the two contracts towards the risk factors, ϕ_v and ϕ_m. This can be seen clearly from the portfolio weights $w^* = (\tilde{\pi}^*_{b1}, \tilde{\pi}^*_{b2})\Sigma^{-1}$.

More specifically, investment in the short-term contract is more sensitive to the market price of the variance risk, while investment in the long-term contract depends more on the market price of the central tendency risk, owing to the distinct patterns of risk loadings of these two factors over time. To see this, considering the example where $\tau_1 = 2$ months and $\tau_2 = 2$ years and the ELW model's parameters are given in Table 1, the sensitive matrix Σ^{-1} is given by

$$\Sigma^{-1} = \begin{pmatrix} 13.1570 & -4.7567 \\ -5.9110 & 36.9530 \end{pmatrix}.$$

And thus,

$$w^* = (13.1570\tilde{\pi}^*_{b1} - 5.9110\tilde{\pi}^*_{b2}, -4.7567\tilde{\pi}^*_{b1} + 36.9530\tilde{\pi}^*_{b2}).$$

In this case, the investor uses short-term and long-term variance swaps to exploit the risk premia $\gamma_v = -1.229$ through $\tilde{\pi}^*_{b1}$ and $\gamma_m = -0.704$ through $\tilde{\pi}^*_{b2}$, separately, as indicated by the positive numbers 13.1570 and 36.9530. In the meantime, the investor takes positive exposure to the second Brownian motion in the short-term contract (-5.9110) to offset the negative exposure of the long-term contract to the second Brownian motion. For the same reason, the investor takes positive exposure to the first Brownian motion in the long-term contract (-4.7567) to offset the negative exposure of the short-term contract to the first Brownian motion. In fact, negative market prices of the two risk factors result in short

[17]Note that the investor's optimal investment problem in the presence of jumps in stock price (e.g., the JP model) can be solved in the ELW model. We have examined the impact of jumps in stock price on variance swap investments, and found that the differences in the holdings of either the short-term contracts or the long-term contracts are less than seven in absolute magnitude value. In this sense, jumps in price do not cause substantial impact on the allocations decisions to variance swap contracts, as the variance market is still completed and can be spanned by the combinations of any two contracts, apart from the minor differences in magnitude.

positions -0.5271 and -0.9634 in both contracts. In contrast to the short-long strategy in Egloff, Leippold and Wu (2010), this example implies that the investor is still able to hedge via the special structure[18] of Σ^{-1} without taking a short position in one variance swap and a long position in the other. In other words, each variance swap play dual roles: exploiting risk premium and hedging. Moreover, the positions in two variance swap contracts depend on the relative magnitude of the market prices of the two sources of risk, γ_v and γ_m when their maturity gap tends to be relatively moderate, which affects the structure of Σ^{-1}. This dependence partially supports the institution that the two variance swap contracts are insufficient to chase the dynamics of variance swap risk premia, although they can complete the market in the assumed two-factor model. Then the optimal allocations could involve short positions in short-term contracts, but long positions in long-term contracts, if the variance price term γ_v is sufficiently larger than the central tendency price term γ_m in absolute value, as demonstrated in Figure 7.[19]

We now turn to our model. Jumps in variance make the variance swap market incomplete. The dynamics of variance swap rates are now driven by a three-factor variance risk structure, and hence another variance swap contract is required to span all the sources of variations. Figure 8 plots the allocations to variance swap contracts in the presence of jumps in variance for a 2-month investment horizon. Due to the small estimators for jump sizes in variance in both the measure P and Q, as reported in Table 4, these values cause a linearity problem to the 3×3 matrix, Σ in Equation (18). This makes its inverse matrix large and in turn the holdings of each contract are very large, as shown in Figure 8.

More importantly, in stark contrast to the portfolio weights in the two-factor model, Figure 8 suggests that the sign of optimal position on each variance swap is irrelevant to the relative magnitudes of the market prices of the two sources of risk, γ_v and γ_m. More specifically, in the presence of jumps in variance, it is always optimal to take long positions in medium-term variance swap contracts and short positions in both short-term and long-term

[18]We find this structure does not change by varying the parameters.

[19]When the market prices of the risk factors vary, unlike the way in Egloff, Leippold and Wu (2010), we adjust all the estimators under the measure P by fixing the estimators in the measure Q in order to reflect their impact on the intertemporal hedging demand.

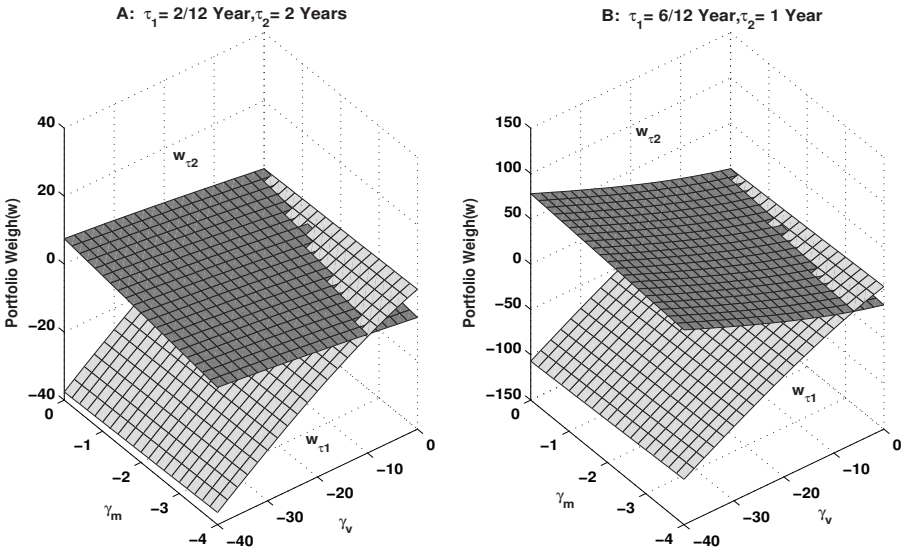

Figure 7: **Optimal Investments in Variance Swap Contracts in ELW Model.** The optimal investments in variance swap contracts (as the fractions of total wealth in notional) in ELW Model (with $\gamma = 5$) is plotted as a function of the market price of the variance risk (γ_v) and the market price of the central tendency factor (γ_m).Panel A shows the optimal investments in 2-month and 2-year variance swap contracts, while Panel B shows the investment in 6-month and 1-year contracts. The surface on the top in each panel denotes the holdings of the long-term contract with time-to-maturity τ_2, while the surface on the below shows the holdings of the short-term contract with time-to-maturity τ_1. The investment horizon is set as two months (e.g., $T = 2$ months).

contracts, that is, a "short-long-short" strategy. Also, the trader should sell even more long-term contracts than short-term ones. For a better understanding, we rewrite the portfolio weights $w^* = (\tilde{\pi}^*_{b1}, \tilde{\pi}^*_{b2}, \widetilde{\pi}^*_{q1})\Sigma^{-1}$ in the example where $\tau_1 = 2$ months, $\tau_2 = 1$ year and $\tau_3 = 2$ years, the matrix Σ^{-1} for the parameters in Table 2 is given by

$$\Sigma^{-1} = \begin{pmatrix} 77.4910 & -359.3994 & 346.9892 \\ 13.8305 & -114.1198 & 197.7839 \\ -1.9143 & 12.2678 & -12.1893 \end{pmatrix}. \tag{21}$$

And thus, we have the holdings of each contract as follows:

$$\begin{aligned} w^*_1 &= 77.4910\tilde{\pi}^*_{b1} + 13.8305\tilde{\pi}^*_{b2} - 1.9143\widetilde{\pi}^*_{q1} = -6.6042, \\ w^*_2 &= -359.3994\tilde{\pi}^*_{b1} - 114.1198\tilde{\pi}^*_{b2} + 12.2678\widetilde{\pi}^*_{q1} = 31.4253, \\ w^*_3 &= 346.9892\tilde{\pi}^*_{b1} + 197.7839\tilde{\pi}^*_{b2} - 12.1893\widetilde{\pi}^*_{q1} = -30.9234. \end{aligned} \tag{22}$$

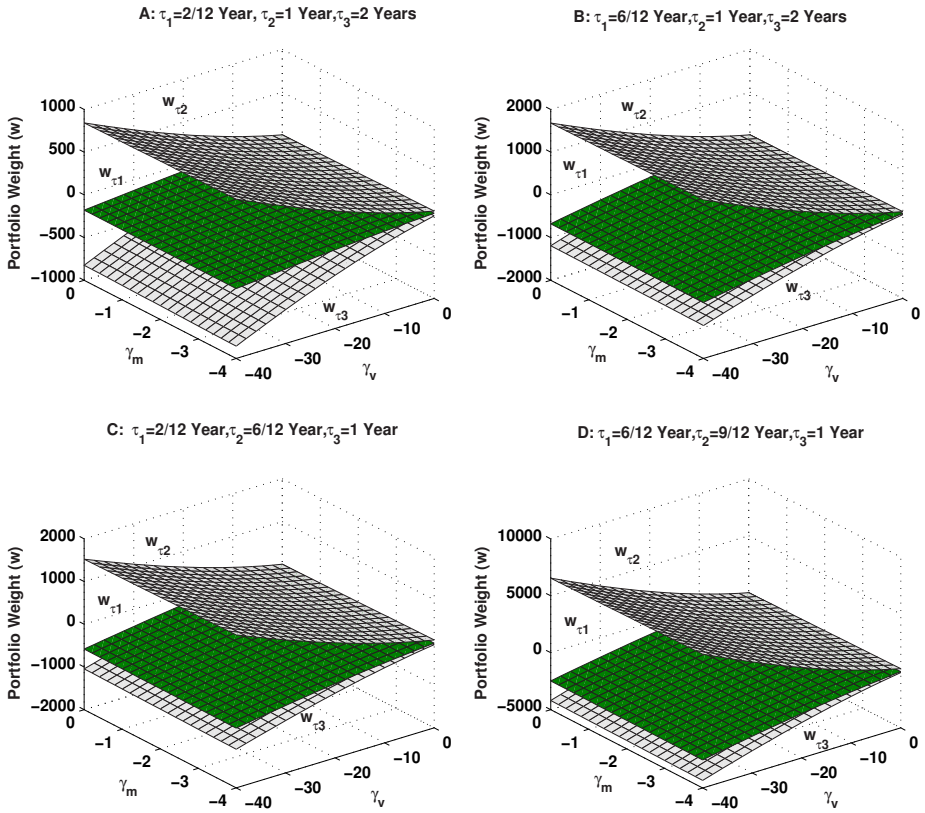

Figure 8: **Optimal Investments in Variance Swap Contracts in HJ Model.** The optimal investments in variance swap contracts (as the fractions of total wealth in notional) in HJ Model (with $\gamma = 5$ and $T = 2$ months) are plotted as a function of the market price of the variance risk (γ_v) and the market price of the central tendency factor (γ_m). When the maturities of the first two contracts, including the 2-month and 2-year variance swap contract in Panel A and the 6-month and 1-year contract in Panel B, C and D, are specified, the four panels show the optimal investments in three contracts by positioning the third one with the different time to maturity. In each panel, the surface on the top denotes the holdings of the medium-term contract, and the surface on the below shows the holdings of the long-term contract, while the surface in the middle presents the holdings of the short-term contract.

Unlike the case in Egloff, Leippold and Wu (2010) discussed above, adding a third variance swap disentangles the dual roles played by each of two variance swaps in their two-factor model in that each of three variance swaps in our double-jump model plays a single role: either gaining risk premium or hedging. Specifically, the shortest-term swap is used to exploit the risk premia γ_v through $\tilde{\pi}^*_{b1}$ and γ_m through $\tilde{\pi}^*_{b2}$, respectively, as indicated by their positive loadings of $\tilde{\pi}^*_{b1}$ and $\tilde{\pi}^*_{b2}$: 77.4910 and 13.8305. For the same reason, the longest-term variance swap is used to gain the risk premia too. In the meantime, the investor uses the medium-term variance swap for hedging due to its negative coefficients of $\tilde{\pi}^*_{b1}$ and $\tilde{\pi}^*_{b2}$: -359.3994 and -114.1198. Moreover, in our experiments, we find that the functions $B_1(t), B_2(t)$ and $B_3(t)$ are all negative and $\widetilde{\pi}^*_{q1} = 0$, implying, by Proposition 4, that $\tilde{\pi}^*_{b1} < 0$ and $\tilde{\pi}^*_{b2} < 0$. That is, the investor has negative exposures to the diffusion risks to gain significant variance risk premia. As a result, given their positive coefficients of $\tilde{\pi}^*_{b1}$ and $\tilde{\pi}^*_{b2}$, w^*_1 and w^*_3 are always negative regardless of the relative magnitudes of the market prices of the two sources of risk, γ_v and γ_m. Similarly, w^*_2 is always positive. In short, the investor uses two variance swaps to exploit the significant volatility risk premium and one swap to hedge the large positions in the other two swaps. Intuitively, the reason for choosing the medium-term variance swap as a hedging asset is that this swap lies in the middle in terms of maturity and thus is simultaneously closest to each of other two variance swaps, implying the simultaneously highest correlations with each of other two variance swaps, making this variance swap the most effective asset for hedging relative to the other two swaps.

In addition, we can make further observations from the above optimal portfolio weights. First, for each variance swap, the sensitivity to jump is much smaller in magnitude compared with the sensitivities to the two diffusion risks, indicated by those relatively small holdings of $\widetilde{\pi}^*_{q1}$. The reason for this is that the value of ϕ_λ is very small due to small jump size in variance. Second, for each variance swap, the sensitivity to the variance exposure via $\tilde{\pi}^*_{b1}$ (the first row of Σ^{-1}) is much larger than those to other two exposures (the second and third rows of Σ^{-1}) in magnitude. Third, the sensitivities to all the risk factors in the long-term contracts through the optimal weights w^*_2 and w^*_3 given by Equation (22) (the second and third columns of Σ^{-1}) are much larger than those in the short-term one (the first column of

Σ^{-1}) in magnitude. This is also consistent with an empirical finding in Aït-Sahalia, Karaman and Mancini (2015) and Filipović, Gourier and Mancini (2015) that long term variance swaps carry more (cumulative) volatility risk premiums than short-term contracts. Finally, given the term structure of variance swap risk premia in Panel B of Figure 6, the trader may gain from the unexpected high frequency jumps in variance (governed by $\beta_0 = 2.214$) such that the premiums for medium-term contracts are paid out by the compensations from short positions in both short-term and long-term contracts.

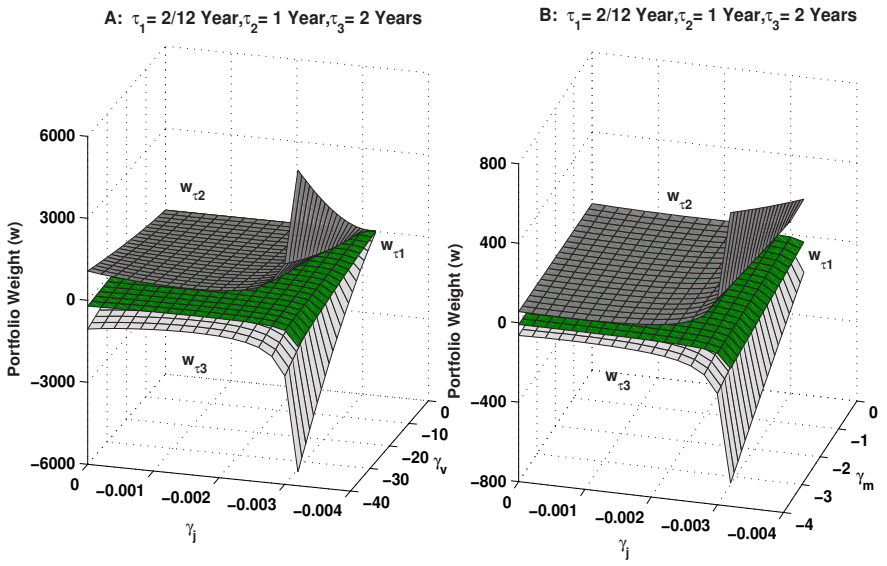

Figure 9: **Sensitivity of Optimal Investments in Variance Swap Contracts to Jump Risk in HJ Model.** The optimal investment in variance swap contracts (as the fractions of total wealth in notional) in HJ Model (with $\gamma = 5$ and $T = 2$ months) is plotted as a function of the market price of the variance risk (γ_v), the market price of the central tendency factor (γ_m) and the market price of the jump risk (γ_j). Three contracts are traded, including the 2-month, 1-year and 2-year variance swap contract (denoted by τ_1, τ_2 and τ_3 respectively). In each panel, the surface on the top denotes the holdings of the tim-to-maturity-τ_2 contract (the medium-term contract), and the surface on the below shows the holdings of the long-term contract with time-to-maturity τ_3, while the surface in the middle presents the holdings of the short-term contract with time-to-maturity τ_1.

To examine the empirical properties of the portfolio strategies discussed above, we conduct robustness tests below. Due to the small market price compensated for jump risk, as reported in Table 4, we then adjust μ_v^Q by fixing μ_v^P in order to reflect the sensitivity of optimal investments in variance swap contracts towards jump risk, i.e., $\mu_v^Q = \mu_v^P - \gamma_j$ where

γ_j is chosen to ensure the positivity of the long-term jump intensity λ_t, e.g., $\mu_v^{\mathcal{Q}} \leq \alpha/\beta_0$. Similar to the results in Figure 8, Figure 9 suggests that it is still optimal to take long positions in medium-term contracts, and short positions in both short-term contracts and long-term contracts. Interestingly, the optimal investments are insensitive to small changes in the market price of jump risk when it is relatively high, and then turns to be very sensitive when the market price of jump risk increases up to 0.4%. In particular, when jump risk in variance is compensated by a low market price, the relatively flat surfaces of the holdings in each contract suggest that jump risk has minor impact on the allocation to variance swap contracts, and the trader's investment decision is mainly affected by variance risk and central tendency risk. However, jump risk plays a substantial role in investment decision when its mark price turns to be high. In order to maximize her expected utility in the investment horizon $T = 2$ months, for example, the trader exploits opportunities by rapidly increasing holdings (in magnitude) of both medium- and long-term contracts while only slightly increasing the position (in magnitude) in the short-term contract, as demonstrated in Figure 9. This further suggests that long-term contracts are more profitable than short-term contracts.

5.3 Optimal Hedging Demands

In this section, we further empirically examine the roles of the three variance swaps in hedging the two state variables v_t and m_t. Proposition 2 shows that the optimal allocation to the variance swap contracts consists of a myopic component that is the optimal portfolio with a constant opportunity set, and an intertemporal hedging demand that a trader may ask for to reduce the impact of shocks to the indirect utility of wealth when facing stochastic opportunities. As a result, the portfolio rule $w^* = (w_{1t}^*, w_{2t}^*, w_{3t}^*)$ is the sum of the myopic demand and the intertemporal hedging demand. In this context, the stochastic variance risk (including the central tendency risk) represents those stochastic investment opportunities, which induces an intertemporal hedging demand when we invest in the variance swap contracts alone.

In the literature, the hedging demand for volatility is not significant in a realistic portfolio problem within a stochastic variance environment excluding variance swap, as discussed by Buraschi, Porchia and Trojani (2010). It is necessary to investigate whether this empirical observation still holds in variance swap investments as the contracts provide direct exposure to volatility risk. And moreover, it is interesting to examine the role of each variance swap in hedging both the transitory volatility risk v and the persistent central tendency risk m. For this reason, we set $\sum^{-1} = (\tilde{\sigma}_{i,j})_{3\times 3}$ for each variance swap j ($j = 1, 2, 3$) based on Proposition 2, and obtain the following demands on this contract at time t:

$$\begin{aligned} &\text{Total myopic demand:} M = \frac{1}{\gamma}(\tilde{\sigma}_{1,j}\gamma_v\sqrt{v_t} + \tilde{\sigma}_{2,j}\gamma_m\sqrt{m_t}); \\ &\text{Hedging demand for } v_t\text{:} H_v = \tilde{\sigma}_{1,j}\sigma_v\sqrt{v_t}B_1(t); \\ &\text{Hedging demand for } m_t\text{:} H_m = \tilde{\sigma}_{2,j}\sigma_m\sqrt{m_t}B_2(t). \end{aligned} \tag{23}$$

We now can work out the hedging ratios for both the volatility risk v and the central tendency risk m, reported as the percentages of the myopic portfolio in Table 11 separately. From the results in Table 11, we make the following observations. First, these ratios clearly show that the intertemporal hedging demands, for example, for the volatility risk and central tendency risk, are indeed significant in the context of variance swap investments. In particular, when the trader tends to be more risk averse, indicated by the increasing degree of risk aversion from $\gamma = 2$ to $\gamma = 40$, the hedging demands for volatility vary from 0.040 to 0.088 within the investment horizon of $T = 20$ years, while the changes in the hedging demands for the central tendency risk can be more substantial, ranging from 0.005 to 0.228. The large hedging demand for the central tendency relative to the one for volatility risk is primarily because the state variable m_t is more strongly persistent than the state variable v_t as suggested by $\kappa_m^P = 0.491$ and $\kappa_v^P = 5.340$ in Table 4.

Second, the total hedging demands, as a percentage of the myopic portfolio (e.g., $(H_v + H_m)/M$), are highly significant as opposed to the empirical results in Buraschi, Porchia and Trojani (2010) in magnitude. For example, given the availability of the variance swap contracts in Table 11, the total hedging demands generated by a trader with $\gamma = 5$ and

2-year variance swap can be as high as 24% for the investment horizon of $T = 10$ years. Moreover, the total hedging demands are even more sensitive to the degree of relative risk aversion (γ), e.g., approaching to 31% for $\gamma = 40$. In contrast, the empirical results in Table II of Buraschi, Porchia and Trojani (2010) show that the largest volatility hedging demand is around 13% (15%) for $\gamma = 6$ ($\gamma = 41$) and the investment horizon of $T = 10$ years. This is caused by the persistent central tendency variable m_t in the present model.[20]

Third, for the same reason as above, the hedging demands for the central tendency risk exhibit strong horizon effects. Specifically, considering the 2-year variance swap with $\gamma = 5$, the hedging ratio is 0.017 for the investment horizon of $T = 6$ months while the hedging ratio is 0.112 for the investment horizon of $T = 5$ years. In contrast, it is evident that the hedging demands for the volatility risk show much weaker horizon effects.

Fourth and more interestingly, the hedging demands also show maturity effects. More specifically, the shortest-term variance swap is mainly used to hedge the variance risk (v_t) while the medium-term and the longest-term variance swaps are mainly used to hedge the central tendency risk m_t. For example, given the relative risk aversion coefficient $\gamma = 5$ and the investment horizon $T = 5$ years, the hedging ratio of the shortest-term variance swap for volatility equals 0.073 as opposed to the hedging ratio of 0.051 for the central tendency. In the meantime, we can observe from Panel I of Table 11 that the variance hedging ratio of the shortest-term variance swap is larger than those of other two variance swaps. In contrast, the hedging ratio of the longest-term variance swap for volatility is 0.070 in comparison with a much higher hedging ratio of 0.157 for the central tendency. And furthermore, Panel II of Table 11 indicates that the central tendency hedging ratios of the medium-term and the longest-term variance swaps are much larger than the one of the shortest-term variance swap. The reason for the above results is that, as can be seen from Figure 1, ϕ_v is much larger than ϕ_m for short maturities while ϕ_m is much larger than ϕ_v for long maturities, and thus short-term variance swaps are more sensitive to v_t than to m_t while long-term variance swaps are more sensitive to m_t than to v_t.

[20]We also re-exam the sensitivities of the results in Table 11 by varying the parameter estimators in Table 4 and find these results are robust regarding the changes in parameters.

Panel I: Hedging Demands for Volatility (H_v/M)						
RRA	Investment Horizon (T) (Year)					Myopic Portfolio
	0.5	1	5	10	20	(M)
2	0.040	0.043	0.043	0.043	0.043	-15.862
	0.039	0.042	0.042	0.042	0.042	75.431
	0.038	0.041	0.041	0.041	0.041	-74.193
5	0.066	0.073	0.073	0.073	0.073	-6.345
	0.064	0.071	0.071	0.071	0.071	30.172
	0.063	0.069	0.070	0.070	0.070	-29.677
40	0.083	0.091	0.092	0.092	0.092	-0.793
	0.081	0.089	0.090	0.090	0.090	3.772
	0.079	0.087	0.088	0.088	0.088	-3.710
Panel II: Hedging Demands for Central Tendency (H_m/M)						
2	0.005	0.010	0.030	0.032	0.033	-15.862
	0.010	0.023	0.065	0.071	0.072	75.431
	0.014	0.032	0.091	0.099	0.100	-74.193
5	0.008	0.018	0.051	0.057	0.057	-6.345
	0.017	0.039	0.112	0.124	0.125	30.172
	0.023	0.054	0.157	0.173	0.175	-29.677
40	0.009	0.022	0.066	0.074	0.075	-0.793
	0.021	0.048	0.145	0.162	0.164	3.772
	0.029	0.067	0.202	0.226	0.228	-3.710

Table 11: **Hedging Ratios for Volatility (v) and Central Tendency (m)**. The hedging ratios for volatility v and central tendency m are calculated using the hedging demands in Equation (23) with various risk aversions of the trader: the less risk aversion ($\gamma = 2$), the moderate risk aversion ($\gamma = 5$) and the extreme risk aversion ($\gamma = 40$). The variance swap contracts with typical time-to-maturities are used, including a set of the 2-month, 1-year and 2-year variance swap contracts. Each entry of the array in both panels consists of three components: the first of which is the hedging ratio for the 2-month contract (τ_1), the second one for the 1-year contract (τ_2) and the third one for the 2-year contract (τ_3), respectively.

5.4 Economic Welfare in Variance Swap Investments

Previously, we analyze the optimal allocations to the variance swap contracts when the variance market is driven by the 2- or 3- factor structural models, respectively. As demonstrated before, the two contracts suggested by a two-factor model are insufficient to span the variance swap market and consequently the trading strategies may sub-optimally utilize the information in the term structure of variance risk premia, for example, shown in Figure 6. Hence it is interesting to further investigate the economic costs for the trader who invests heavily in this market if the dynamics of the variance swap is mis-specified by the two-factor model (e.g., the ELW model or the JP model), while the HJ model is assumed to be the true specification for its dynamics. In particular, the investor follows the strategy $w = (n_{1t}, n_{2t})^\top$ given by Equation (43) and (44) in Egloff, Leippold and Wu (2010).

We follow the literature to evaluate the economic costs by a certainty equivalent loss (CE) defined by: The utility cost, CE, of following the suboptimal strategy $w = (n_{1t}, n_{2t})^\top$ satisfies the equation below:

$$J(t, W_t(1 - CE), X_t) = J^{(1)}(t, W_t, X_t),$$

where $J(t, W_t, X_t)$ is the indirect value of the portfolio choice problem in the HJ model given by Proposition 4 with three traded variance swaps and $J^{(1)}(t, W_t, X_t)$ is the value function provided in Proposition 5 below corresponding to the the suboptimal strategy $w = (n_{1t}, n_{2t})^\top$ in HJ model with two traded variance swaps. Intuitively, CE is the percentage of initial wealth an investor is willing to pay to switch from the suboptimal strategy w to the optimal strategy w^*. The following result presents the calculation of CE.

Proposition 3 *Under the above assumptions, we have the following result:*

$$J^{(1)}(t, W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma}[f(t, X_t)]^\gamma = \frac{W^{1-\gamma}}{1-\gamma}\left[e^{A^{(1)}(t)+B_1^{(1)}(t)v_t+B_2^{(1)}(t)m_t+B_3^{(1)}(t)\lambda_t}\right]^\gamma \tag{24}$$

where the functions $A^{(1)}(t)$, $B^{(1)}(t) = (B_1^{(1)}(t), B_2^{(1)}(t))^\top$ *and* $B_3^{(1)}(t)$ *satisfy the following*

equations:

$$\begin{aligned}
\frac{dA^{(1)}}{dt} &+ \kappa_m^P \theta_m^P B_2^{(1)} + \alpha\lambda_\infty B_3^{(1)} + \frac{1-\gamma}{\gamma} r = 0, \\
\frac{dB_1^{(1)}}{dt} &- [\kappa_v^P - (1-\gamma)\psi_v\sigma_v^2] B_1^{(1)} + \frac{1}{2}\gamma\sigma_v^2 \left(B_1^{(1)}\right)^2 + \frac{1-\gamma}{\gamma}\sigma_v\gamma_v\psi_v + \frac{\gamma-1}{2}\psi_v^2\sigma_v^2 = 0, \\
\frac{dB_2^{(1)}}{dt} &+ \kappa_v^Q B_1^{(1)} - [\kappa_m^P - (1-\gamma)\psi_m\sigma_m^2] B_2^{(1)} + \frac{1}{2}\gamma\sigma_m^2 \left(B_2^{(1)}\right)^2 + \frac{1-\gamma}{\gamma}\sigma_m\gamma_m\psi_m \\
&+ \frac{\gamma-1}{2}\psi_m^2\sigma_m^2 = 0, \\
\frac{dB_3^{(1)}}{dt} &- \alpha B_3^{(1)} + \frac{\gamma-1}{\gamma}\widetilde{\pi}_{q1}\mu_v^Q + \frac{1}{\gamma}E^P\left[\left(\widetilde{\pi}_{q1}J^{v,P}+1\right)^{1-\gamma} e^{\gamma(B_1^{(1)}+B_3^{(1)}\beta_0)J^{v,P}} - 1\right] = 0,
\end{aligned}$$

with $A^{(1)}(T) = B_1^{(1)}(T) = B_2^{(1)}(T) = B_3^{(1)}(T) = 0$. *Here* $\psi_v = n_{1t}\phi_v(\tau_1) + n_{2t}\phi_v(\tau_2)$ *and* $\psi_m = n_{1t}\phi_m(\tau_1) + n_{2t}\phi_m(\tau_2)$. $\widetilde{\pi}_{q1}$ *is the jump exposure generated by the suboptimal strategy* $w = (n_{1t}, n_{2t})^\top$ *and represented by*

$$\widetilde{\pi}_{q1} = n_{1t}(\phi_v(\tau_1) + \beta_0\phi_\lambda(\tau_1)) + n_{2t}(\phi_v(\tau_2) + \beta_0\phi_\lambda(\tau_2)).$$

The utility cost, CE, *of following the suboptimal strategy* $w = (n_{1t}, n_{2t})^\top$ *satisfies the equation below:*

$$J(t, W_t(1-CE), X_t) = J^{(1)}(t, W_t, X_t),$$

and thus

$$CE = 1 - \left[e^{A^{(1)}(t) - A(t) + (B_1^{(1)}(t) - B_1(t))v_t + (B_2^{(1)}(t) - B_2(t))m_t + (B_3^{(1)}(t) - B_3(t))\lambda_t}\right]^{\frac{\gamma}{1-\gamma}}$$

Proof. See Appendix B. ■

Unlike the indirect value function $J(t, W_t, X_t)$, the above result suggests that the value function $J^{(1)}(t, W_t, X_t)$ depends on the maturities of the two variance swaps through the functions $A^{(1)}(t), B_1^{(1)}(t), B_2^{(1)}(t)$ and $B_3^{(1)}(t)$. The utility cost, CE, thus depends on the two maturities. Also, Proposition 3 suggests that the jump risk exposure $\tilde{\pi}_{q1}$ play a crucial role

in determining the magnitude of CE. It is then necessary to evaluate $\tilde{\pi}_{q1}$ before numerically assessing the utility cost of following a suboptimal strategy w.

As well-understood in the literature of portfolio choice problem in jump-diffusion model, see, e.g., Proposition 1 of Liu, Longstaff and Pan (2003), the investor must restrict her jump exposure $\tilde{\pi}_{q1}$ to guarantee that her wealth remains positive when jump occurs. In particular, $\tilde{\pi}_{q1}$ satisfies: $\tilde{\pi}_{q1} \geq 0$ since the support of the variance jump size J_v^P is $[0, \infty)$. As a result, before presenting the empirical results for CE given by Proposition 3, it is interesting to examine whether or not the restriction on π_{q1} is violated by the suboptimal strategy w. The reason for this is that, as mentioned in Egloff, Leippold and Wu (2010), an investor often takes an extreme short position in a variance swap due to significantly negative variance risk premium in variance swap rate, and thus this is very likely to lead to violation of the restriction. Table 12 shows that all the jump exposures $\tilde{\pi}_{q1}$s caused by the ELW model are negative across the various investment horizons, suggesting that the trader is subject to the substantial jump risk exposure by ignoring jumps in volatility.[21] This table then confirms our concern, that is, the restriction is overwhelmingly violated by the the suboptimal strategy w, implying a 100 percent loss.

Also, it is very interesting to see whether trading stocks in addition to variance swap portfolios can mitigate the risk of bankruptcy in the two cases of stock being tradable i) in the ELW model and further ii) in the HJ model. Intuitively, the investor in the former case, according to Egloff, Leippold and Wu (2010), could take short positions in stock in order to hedge variance swaps given the market price of variance risk being more negative. The investor hence holds more extreme positions in variance swaps, compared with those portfolios without stock, suggesting that the values of $\widetilde{\pi}_{q1}$ turn to be more negative. In the latter case, the investor would take another risk exposure to jumps in stock prices and

[21] We also examine $\tilde{\pi}_{q1}$ in the case of $n_{1t} \times n_{2t} < 0$ and find that the values of $\tilde{\pi}_{q1}$ are negative due to the dominance of γ_v. Moreover, the jump exposures caused by the JP model are still negative with the larger absolute magnitude, and thus they are not reported. Meanwhile, we exam the jump exposure $\tilde{\pi}_{q1}$ of both the ELW and JP model if the AKM model is assumed to be the true model, and find that all the jump exposures are negative but with the relatively smaller absolute values, implying the underestimation of the unhedged jump risk. It is expected that this situation will become much more severe when the expected jump size μ_v^P is much bigger than 0.001 estimated in Aït-Sahalia, Karaman and Mancini (2015) and used here.

RRA	Maturity Pair (Year)	Investment Horizon (T) (Year)					
		0.5	1	5	10	20	30
$\gamma = 5$	$(\tau_1 = 2/12, \tau_2 = 2)$	-1.32	-1.35	-1.47	-1.54	-1.60	-1.61
	$(\tau_1 = 6/12, \tau_2 = 1)$	-1.50	-1.53	-1.69	-1.79	-1.86	-1.88
$\gamma = 40$	$(\tau_1 = 2/12, \tau_2 = 2)$	-0.17	-0.17	-0.19	-0.21	-0.22	-0.22
	$(\tau_1 = 6/12, \tau_2 = 1)$	-0.19	-0.19	-0.22	-0.24	-0.26	-0.26

Table 12: **Jump Exposure $\widetilde{\pi}_{q1}$ in ELW Model with Different Risk Aversions (γs)**. The jump exposures $\widetilde{\pi}_{q1}$ caused by ELW Model due to model mis-specification are presented with the two risk aversions: the moderate one ($\gamma = 5$) and the extreme one ($\gamma = 40$), given HJ Model being the true model. The typical maturity pairs of the variance swap contracts are used, as in Figure 7, including the pair with a long maturity gap (i.e., $\tau_1 = 2/12, \tau_2 = 2$) and the one with a moderate gap (i.e., $\tau_1 = 6/12, \tau_2 = 1$).

the investor must take positive positions in stock to avoid bankruptcy. This indicates that the short positions in the ELW model violates this non-bankruptcy condition. Therefore, trading stocks in variance swap portfolios in both cases makes the situation even worse. This indeed suggests that under this circumstance the investor is certainly subject to the risk of bankruptcy.

As observed in our tests in Table 4, $\widetilde{\pi}_{q1}$s generated by suboptimal strategies in the ELW model fall in the range of $(-2, 0)$ for $\gamma = 5$ and of $(-1, 0)$ for $\gamma = 40$. Thus, to further evaluate the utility costs in the case where the suboptimal strategy w is feasible in the HJ model, we intentionally and conservatively change the support $[0, \infty)$ of the variance jump size J_v^P into $[0, k]$ $(k > 0)$ with $k = 0.5$ here. This means that the variance can jump at most $k \times 100$ percent, and then the restriction on $\tilde{\pi}_{q1}^*$ in the optimization problem in Proposition 2 becomes $\tilde{\pi}_{q1}^* > -1/k$ in order to work out the finite values of CE in $(0, 1)$.[22] In the meantime, this change does not affect the variance swap rate due to the negligible probability $P(J_v^P > k)$. The investor, however, still incurs significant costs as reported in Table 5.

[22]The current restriction $\tilde{\pi}_{q1}^* \geq 0$ easily results in the bankruptcy of the investor in her trading positions, e.g., $CE \to 1$. This assumption is also supported by the dynamics of the CBOE Volatility Index (or VIX) over the past twenty-five years at http://www.cboe.com/delayedquote/advchart.aspx?ticker=VIX from 1990 to 2015. Its historical close quotes reached the highest level of 80.96% on November 20, 2008, which suggests a value of $k = 0.81$ in our setup.

Panel I: Utility Costs (CEs) by ELW Model and JP Model for $\gamma = 5$							
Model	Maturity Pair (Year)	Investment Horizon (T) (Year)					
		0.5	1	5	10	20	30
ELW	$(\tau_1 = 2/12, \tau_2 = 2)$	0.0065	0.0129	0.0593	0.1123	0.2101	0.2967
	$(\tau_1 = 6/12, \tau_2 = 1)$	0.0056	0.0111	0.0510	0.0973	0.1852	0.2639
JP	$(\tau_1 = 2/12, \tau_2 = 2)$	0.0043	0.0086	0.0423	0.0862	0.1665	0.2355
	$(\tau_1 = 6/12, \tau_2 = 1)$	0.0036	0.0072	0.0370	0.0816	0.1647	0.2333
Panel II: Utility Costs (CEs) by ELW Model and JP Model for $\gamma = 40$							
ELW	$(\tau_1 = 2/12, \tau_2 = 2)$	0.0054	0.0104	0.0476	0.0922	0.1764	0.2532
	$(\tau_1 = 6/12, \tau_2 = 1)$	0.0053	0.0101	0.0466	0.0906	0.1753	0.2531
JP	$(\tau_1 = 2/12, \tau_2 = 2)$	0.0041	0.0078	0.0356	0.0724	0.1464	0.2124
	$(\tau_1 = 6/12, \tau_2 = 1)$	0.0022	0.0042	0.0197	0.0455	0.1144	0.1929

Table 13: **Utility Costs by Model Mis-specification**. The utility costs caused by ELW and JP Model due to model mis-specification are presented, given HJ Model being the true model, associated with two risk aversions of the trader: the moderate risk aversion ($\gamma = 5$) and the extreme risk aversion ($\gamma = 40$). The typical time-to-maturity pairs of the variance swap contracts are used, as in Figure 7, including the pair with a long maturity gap (i.e., $\tau_1 = 2/12, \tau_2 = 2$) and the one with a moderate gap (i.e., $\tau_1 = 6/12, \tau_2 = 1$).

Rooted in the results in Table 12, Table 13 report the utility costs of switching from a suboptimal strategy w generated in the ELW or JP model to the optimal strategy w^* generated in the HJ model. Under a two-factor variance structure, two variance swap contracts would be sufficient to span the market. As shown in Figure 7, two typical combinations of contracts, one with the 2-month and 2-year contracts ($\tau_1 = 2/12$ and $\tau_2 = 2$) and the other with the 6-month and 1-year contracts ($\tau_1 = 6/12$ and $\tau_2 = 1$), are used to quantify the utility (economic) costs that the trader has to bear due to model mis-specification. The first panel in Table 13 reports the utility costs for the trader with $\gamma = 5$ in the various investment horizons (T) by assuming the rolling-over of specific variance swap contracts. Specifically, the utility costs in two portfolios are very close and steadily increasing with the growth of T, suggesting that the utility costs in the ELW and JP model are less dependent on the maturity gap, but more sensitive to the length of investment horizon (T). Compared with the JP model, the ignorance of jumps in both price and variance can make the ELW model produce

relatively higher economic costs when the investment horizon increases from 6 months to 30 years, implying that incorporating jumps into stock price is of the first-order importance for variance swap investments. The reason for this is that the JP model is much closer to the HJ model in terms of fitting variance swap rate as observed in Panel B of Figure 5. Nevertheless, the utility costs in the JP model are still significant despite the small expected volatility jump size. This further underscores the importance of taking into account jumps in variance.

The second panel in Table 13 further reports the utility costs for the trader with $\gamma = 40$. Similar to the case of $\gamma = 5$, the utility costs of the contract combinations in both models are still increasing significantly with the length of the investment horizon (T). Also, the magnitudes of utility costs shift down overwhelmingly across all investment horizons when the degree of risk aversion turns to be much higher from $\gamma = 5$ to $\gamma = 40$. In particular, the utility cost of each contract combination in the JP model is persistently reduced for each investment horizon (T), compared with those costs reported in the first panel. Similar to the first pannel, this seems to suggest that for the extremely risk-averse trader, the incorporation of jumps into price returns has large impact on utility costs in the context of asset allocation to variance swap contracts, compared with those produced by the ELW model with the moderate risk aversion (e.g., $\gamma = 5$). This is caused partially because the extremely high risk aversion ($\gamma = 40$) enforces the trader to make much smaller investments in variance swaps and thus the impact of variance jump components in model mis-specification is mitigated to certain extent but remains significant.

Table 13 has demonstrated the impact of jumps in prices on utility costs caused by model mis-specification. It is interesting to further investigate the impact of jumps on variance on economic costs, given the small estimators of μ_v^P and μ_v^Q in Table 4 that capture jump size in variance under the measure P and Q. Table 14 reports the utility costs caused by the ELW model with a range of jump sizes in variance for a portfolio of the 2-month and 2-year variance swap contracts (i.e., $\tau_1 = 2/12, \tau_2 = 2$).[23] It clearly shows that apart from

[23]Since the jump size $J^{v,P}$ follows an exponential distribution, we then adjust the value of $J^{v,P}$ to ensure its mean equal to μ_v^P under a truncated exponential distribution, namely, $E[J^{v,P}\mu_v^P/m_0] = \mu_v^P$ for $J^{v,P} \in [0,1]$, by scaling up a constant μ_v^P/m_0 with $m_0 = \mu_v^P - exp(-1/\mu_v^P)(1+\mu_v^P)$. We also exam the utility costs

T (Year)	Jump Size in Variance (μ_v^P)					
	0.001	0.015	0.055	0.095	0.135	0.175
0.5	0.0065	0.0067	0.0085	0.0115	0.0164	0.0229
1	0.0129	0.0142	0.0174	0.0233	0.0331	0.0468
5	0.0593	0.0756	0.0896	0.1157	0.1635	0.2387
10	0.1123	0.1484	0.1741	0.2212	0.3080	0.4406
20	0.2101	0.2772	0.3206	0.3965	0.5280	0.7022
30	0.2967	0.3859	0.4404	0.5317	0.6771	0.8403

Table 14: **Utility Costs with Different Jump Sizes in Variance**. The utility costs that a trader with $\gamma = 5$ may suffer in ELW Model due to model mis-specification are reported with a range of jump size in variance by fixing $\mu_v^Q - \mu_v^P = 0.001$. The typical time-to-maturity pair of the variance swap contracts is used, including the 2-month and 2-year variance swap contracts (i.e., $\tau_1 = 2/12, \tau_2 = 2$). Note that for each pair of (μ_v^P, μ_v^Q), HJ Model is re-calibrated to the empirical mean term structure of variance swap rates reported in Table 3, which results in the RMSEs with mean 2.31% and standard deviation 0.15%.

the investment horizon (T), jumps in variance do have substantial impact on utility costs when the dynamics of variance is improperly specified. That is, large jumps in variance can result in high utility costs, which again emphasizes the importance of incorporating jumps in variance in the context of variance swap investments.

We further investigate the sensitivity of utility costs towards parameter mis-specification in the HJ model. Suppose that one parameter in Table 4 is mis-specified. We then use Proposition 1 and 2 to obtain the optimal portfolio wight $w = (n_{1t}, n_{2t}, n_{3t})$. Then, the corresponding CE is still calculated by Proposition 3 with the following modifications:

$$\begin{aligned} \psi_v &= n_{1t}\phi_v(\tau_1) + n_{2t}\phi_v(\tau_2) + n_{3t}\phi_v(\tau_3), \\ \psi_m &= n_{1t}\phi_m(\tau_1) + n_{2t}\phi_m(\tau_2) + n_{3t}\phi_m(\tau_3), \\ \tilde{\pi}_{q1} &= \sum_{i=1}^{3} n_{it}(\phi_v(\tau_i) + \beta_0\phi_\lambda(\tau_i)). \end{aligned} \tag{25}$$

In the literature, the CE caused by parameter mis-specification is much smaller than one

caused by the ELW and JP model for a portfolio of the 6-month and 1-year variance swap contracts (i.e., $\tau_1 = 6/12, \tau_2 = 1$), and obtain the consistent results with those reported in Table 13 and 14.

caused by model mis-specification (see, for example, Table 5 and Table 6 in Zhou and Zhu (2012) in a pure-diffusion model). But in variance swap investments, it may be different because $\tilde{\pi}_{q1}$ can be easily become negative for incorrect parameters, as implied by the preceding analysis primarily due to the investor's extreme positions in variance swaps.[24] Table 15 confirms our concern. In this table, we take 5.34 as the true value of κ_v^P. It clearly shows that the trader may be easily bankrupt when the parameter κ_v^P is underestimated, and may suffer utility costs when κ_v^P is overestimated. Intuitively, when κ_v^P is underestimated by a standard deviation of 0.4, for example, $\kappa_v^P = 4.94$, the state variable v_t is more persistent in the wrong model with $\kappa_v^P = 4.94$ than in the true model with $\kappa_v^P = 5.34$. Then the investor over hedges the risk stemming from v_t in the wrong model than she would do in the true model.[25] In other words, the investor takes more extreme positions in the wrong model than in the true model, leading to bankruptcy in the true model. Conversely, when κ_v^P is overestimated, the investor takes less extreme positions in the wrong model than in the true model. As a result, this make the suboptimal strategy less likely to go bankrupt in the true model. In the meantime, the utility costs tend to be large when the investment horizon T increases from $T = 6$ months to $T = 20$ years. Recall in the previous section that the hedging ratios for volatility are significant in the context of variance swap investments. These results then suggest that the proper estimation strategy of parameters (e.g., κ_v^P) that capture the dynamics of variance, may significantly mitigate the impact of parameter mis-specification so that the trader may suffer less utility costs when entering into the volatility market.

6 Conclusion

In the present paper, we propose a tractable three-factor model to extend the existing two-factor term structure models for variance swap rates. The empirical results document that this new model outperforms the two-factor jump-diffusion models widely studied in the

[24] We mainly report the utility costs by mis-specifying κ_v^P, given both its relative large magnitude and the importance of measuring the mean-reversion speed for the instantaneous variance rate, after varying the estimators of the parameters reported in Table 4. To conduct the analysis, we restrict $\tilde{\pi}_{q1}^* \geq 0$ in the optimization problem in Proposition 2.

[25] We test this by varying the value κ_v^P from 4.14 to 6.54, including the true estimator of 5.34, and find the value of $B_1(t)$ increaases monotonically from -0.1093 to -0.0664.

RRA	T	4.14	4.54	4.94	**5.34**	5.74	6.14	6.54
	0.5	-	-	-	0	0.0017	0.0031	0.0041
	1.0	-	-	-	0	0.0035	0.0061	0.0081
$\gamma = 2$	5.0	-	-	-	0	0.0179	0.0309	0.0404
	10.0	-	-	-	0	0.0341	0.0591	0.0771
	20.0	-	-	-	0	0.0643	0.1113	0.1445
	0.5	-	-	-	0	0.0007	0.0013	0.0017
	1.0	-	-	-	0	0.0015	0.0026	0.0035
$\gamma = 5$	5.0	-	-	-	0	0.0090	0.0149	0.0192
	10.0	-	-	-	0	0.0167	0.0281	0.0364
	20.0	-	-	-	0	0.0299	0.0516	0.0674
	0.5	-	-	-	0	0.0001	0.0002	0.0002
	1.0	-	-	-	0	0.0002	0.0003	0.0005
$\gamma = 40$	5.0	-	-	-	0	0.0014	0.0022	0.0028
	10.0	-	-	-	0	0.0025	0.0041	0.0053
	20.0	-	-	-	0	0.0043	0.0073	0.0096

Table 15: **Utility Costs by Mis-specifying κ_v^P in HJ Model**. The utility costs by mis-specifying κ_v^P in HJ Model are calculated using the notation in Equation (25) with various risk aversions of the trader: the less risk aversion ($\gamma = 2$), the moderate risk aversion ($\gamma = 5$) and the extreme risk aversion ($\gamma = 40$). The variance swap contracts with typical time-to-maturities are used, including a set of the 2-month (τ_1), 1-year (τ_2) and 2-year (τ_3) variance swap contracts. The bold number in the top row is the true estimator for κ_v^P reported in Table 4. The symbol of "-" denotes the bankruptcy of the trader's trading position due to the negative value of $\tilde{\pi}_{q1}$.

literature over the out-of-sample period in terms of pricing variance swaps across all time-to-maturities and predicting stock returns with various time horizons ranging from one month to two years. And more importantly, we find that variance swap rates have even more predictive power for predicting market returns.

We further explicitly solve the optimal investment problem in variance swaps, and find that both jumps in stock price and variance are important for variance swap investments, which leads to a "short-long-short" trading strategy. This further enhances the understanding of the optimal trading strategies in the variance swap market in the literature. Also, the trader may suffer from substantial economic losses caused by both model and parameter mis-specification.

References

Adrian, T. and J. Rosenberg (2008). Stock returns and volatilities: Pricing the long-run and short-run components of market risk. Journal of Finance, 63(6), 2997-3030.

Aït-Sahalia, Y., Cacho-Diaz,J., Laeven, R.J.A., 2015. Modeling financial cotagion: Using mutually exciting jump processes. *Journal of Financial Economics*, 117(3), 585-606.

Aït-Sahalia, Y., Karaman, M., Mancini, L., 2015. The term structure of variance swap and risk premia, *working paper*, Princeton University.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *Journal of Finance*, 52, 2003-2049.

Bakshi, G., N. Kapadia, 2003. Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies*, 16, 527-566.

Bandi, F., Reno, R., 2015. Price and volatility co-jumps. Forthcoming in *Journal of Financial Economics*.

Bates, D.S., 2000. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94, 181-238.

Bates, D.S., 2006. Maximum likelihood estimation of latent affine process. *Review of Financial Studies*, 19, 909-965.

Bondarenko, O., 2004. Market price of variance risk and performance of hedge funds. *Working Paper*, University of Illinois at Chicago.

Bollerslev, T., G. Tauchen and H. Zhou (2009). Expected stock returns and variance risk premia. Review of Financial Studies, 22(11), 4463-4492.

Bollerslev, T., V. Todorov and Xu (2015). Tail Risk Premia and Return Predictability. Journal of Financial Economics. 118, 113-134.

Britten-Jones, M., Neuberger, A. and Nolte, I. (2011). Improved Inference in Regression with Overlapping Observations. Journal of Business Finance & Accounting, 38(5-6), 657-683.

Broadie, M., Chernov, M. Johannes, M., 2007. Model specification and risk premia: Evidence from futures options. *Journal of Finance*, 62, 1453-1490.

Buraschi, A., P. Porchia and F. Trojani (2010). Correlation Risk and Optimal Portfolio Choice, Journal of Finance, LXV(1), 393-420.

Carr, P., Madan, D.B. (1998). *Towards a theory of volatility trading*. In: Jarrow, R.(Ed.), Risk Book on Volatility. New York: Risk, pp.417-427.

Carr, P., L. Wu (2009). Variance risk premiums. *Review of Financial Studies*, 22, 1311-1341.

Chernov, M., Gallant, A. R., Ghysels, E., Tauchen, G. T. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116, 225-257.

Chernov, M., Ghysels, E., 2000. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of option valuation. *Journal of Financial Economics*, 57, 407-458.

Damine, P., J. Wakefield and S. Walker (1999). Gibbs Sampling for Bayesian Non-conjugate and heirarchical models by using auxiliary variables. J.R..Stat.Soc. 61, 331-344.

Diebold, F. and R. Mariano (1995). Comparing predicative accuaracy. Journal of Business and Economic Statistics, 13, 253-265.

Duffie, D., J. Pan, K. Singleton, 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68, 1343-1376.

Egloff, D., Leippold, M., Wu, L., 2010. The term structure of variance swap rates and optimal variance swap investments. *Journal of Financial and Quantitative Analysis*, 45, 1279-1310.

Eraker, B., Johannes, M.S., Polson, N., 2003. The impact of jumps in equity index volatility volatility and returns. *Journal of Finance*, 58, 1269-1300.

Eraker, B., 2004. Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *Journal of Finance*, 59, 1367-1403.

Filipović, D., Gourier, E., Mancini, L., 2015. Quadratic variance swap models. *Journal of Financial Economics*, forthcoming.

Fulop, A., Li, J., Yu, J., 2015. Self-Exciting Jumps, Learning, and Asset Pricing Implications. *Review of Financial Studies*, 28(3), 876-912.

Hodrick, R.(1992). Dividend yields and expected stock returns alternative procedures for inference and measurement. Review of Financial Studies, 5(3), 357-386.

Hong, Y. and X. Jin (2018). Semi-Analytical Solutions for Dynamic Portfolio Choice in Jump-Diffusion Models and the Optimal Bond-Stock Mix. European Journal of Operational Research, 265(1), 389-398.

Garlappi, L., R. Uppal and T. Wang (2007). Portfolio selection with parameter and model Uncertainty: A multi-prior approach. *Review of Financial Studies*, 20, 41-81.

Jiang, G.J., Oomen, R.C., 2008. Testing for jumps when asset price are observed with noise - A "swap variance" approach. *Journal of Econometrics*, 144, 352-370.

Jin, X., A. Zhang, 2012. Decomposition of optimal portfolio weights in a jump-diffusion model and its applications. *Review of Financial Studies*, 25, 2877-2919.

Johannes, Michael, Nicholas Polson and Jonathan Stroud (2009). Optimal filtering of jumpdiffusions: extracting jumps and volatility from prices, Review of Financial Studies, 22(7), 2759-2799.

Li, J.Y. and G. Zinna (2017). The Variance Risk Premium: Components, Term Structures, and Stock Return Predictability. Journal of Business & Economic Statistics, available at https://doi.org/10.1080/07350015.2016.1191502.

Liu, J., 2007. Portfolio selection in stochastic environments. *Review of Financial Studies*, 20, 1-39.

Liu, J., F. Longstaff, J. Pan, 2003. Dynamic asset allocation with event risk. *Journal of Finance*, 58, 231-259.

Liu, J., J. Pan, 2003. Dynamic derivative strategies. *Journal of Financial Economics*, 69, 401-430.

Neuberger, A., 1994. The log contract. *Journal of Portfolio Management*, 20, 74-80.

Pan, J., 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, 63, 3-50.

Santa-Clara, P. and S. Yan (2010). Crashes, volatility, and the equity premium: Lessons from S$P500 options. Review of Economics and Statistics, 92(2), 435C451.

Tobin, J., 1958. Liquidity preference as behavior towards risk, *Review of Economic Studies*, 25, 65-86.

Todorov, V., 2009. Variance risk premium dynamics: The role of jumps. *Review of Financial Studies*, 23, 345-383.

Wu, L., 2008, *Modeling financial security returns using Levy processes*, Handbooks in Operations Research and Management Science: Financial Engineering, 15, Eds. John Birge and Vadim Linetsky, Elsevier, 2008.

Yu, C.L., H.T. Li and M.T. Wells (2011). MCMC Estimation of Levy Jump models Using Stock and Option Prices. Mathematical Finance, 21(3), 383-422.

Zhou, G., Zhu, Y., 2012. Volatility trading: What is the role of the long-run volatility component? *Journal of Financial and Quantitative Analysis*, 47, 273-307.

Appendix A-C

A Proof of Proposition 1 and Proposition 2

For an investor with the CRRA utility, we apply Proposition 1 and 2 in Hong and Jin (2018) for the proofs by setting $\eta = 0$. The state variables in our model are v_t, m_t and λ_t. Note that in this case

$$b - r\mathbf{1}_3 = \begin{pmatrix} \phi_v(\tau_1)\sigma_v\gamma_v v_t + \phi_m(\tau_1)\sigma_m\gamma_m m_t - (\phi_v(\tau_1) + \beta_0\phi_\lambda(\tau_1))\mu_v^Q\lambda_t \\ \phi_v(\tau_2)\sigma_v\gamma_v v_t + \phi_m(\tau_2)\sigma_m\gamma_m m_t - (\phi_v(\tau_2) + \beta_0\phi_\lambda(\tau_2))\mu_v^Q\lambda_t \\ \phi_v(\tau_3)\sigma_v\gamma_v v_t + \phi_m(\tau_3)\sigma_m\gamma_m m_t - (\phi_v(\tau_3) + \beta_0\phi_\lambda(\tau_3))\mu_v^Q\lambda_t \end{pmatrix} = \Sigma \begin{pmatrix} \gamma_v\sqrt{v_t} \\ \gamma_m\sqrt{m_t} \\ -\mu_v^Q\lambda_t \end{pmatrix},$$

where Σ is given by (17) in Section 2.2. Hence

$$\theta = \begin{pmatrix} \theta_1^b \\ \theta_2^b \\ \theta_1^q \end{pmatrix} = \Sigma^{-1}(b - r\mathbf{1}_3 + \Sigma_q E^P[J_t^{v,P}]\lambda_t) = \begin{pmatrix} \gamma_v\sqrt{v_t} \\ \gamma_m\sqrt{m_t} \\ -(E^Q[J_t^{v,P}] - E^P[J_t^{v,P}])\lambda_t \end{pmatrix}.$$

By noticing that the state variable λ_t is a pure-jump process, applying Proposition 1 of Hong and Jin (2018) gives the following indirect value function:

$$J(t, W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma}[f(t, X_t)]^\gamma = \frac{W^{1-\gamma}}{1-\gamma}\left[e^{A(t)+B_1(t)v_t+B_2(t)m_t+B_3(t)\lambda_t}\right]^\gamma \tag{26}$$

where the functions $A(t)$, $B(t) = (B_1(t), B_2(t))^\top$ and $B_3(t)$ satisfy the following equations:

$$\begin{aligned}
\frac{dA}{dt} &+ \left(k + \frac{1-\gamma}{\gamma}g_0\right)^\top B^\top + \frac{1}{2}B[h_0 + (1-\gamma)l_0]B^\top \\
&+ \frac{1-\gamma}{2\gamma^2}H_0 + \frac{1-\gamma}{\gamma}\delta_0 = 0, \\
\frac{dB}{dt} &+ \left(-K + \frac{1-\gamma}{\gamma}g_1\right)^\top B^\top + \frac{1}{2}B[h_1 + (1-\gamma)l_1]B^\top \\
&+ \frac{1-\gamma}{2\gamma^2}H_1 + \frac{1-\gamma}{\gamma}\delta_1 = 0 \\
\frac{dB_3}{dt} &- \alpha B_3 + (\gamma-1)\widetilde{\pi}_{q1}^* E^Q[J^{v,Q}] + E^P\left[\left(\widetilde{\pi}_{q1}^* J^{v,P} + 1\right)^{1-\gamma} e^{\gamma(B_1+B_3\beta_0)J^{v,P}} - 1\right] = 0,
\end{aligned}$$

Furthermore, by noticing that

$$\sigma^x(X_t) = \begin{pmatrix} \sigma_v\sqrt{v_t} & 0 \\ 0 & \sigma_m\sqrt{m_t} \end{pmatrix},$$

we can get the following parameters: $k = (0, \kappa_m^P\theta_m^P)^\top$, $h_{111} = (\sigma_v^2, 0)$, $h_{112} = h_{121} = (0,0)$, $h_{122} = (0, \sigma_m^2)$, $\delta_0 = r$, $\delta_1 = 0$, $H_0 = 0$, $H_1 = (\gamma_v^2, \gamma_m^2)$, $g_0 = 0$, $l_0 = 0, l_1 = 0$.

$$K = \begin{pmatrix} \kappa_v^P & -\kappa_v^Q \\ 0 & \kappa_m^P \end{pmatrix},\ g_1 = \begin{pmatrix} \sigma_v\gamma_v & 0 \\ 0 & \sigma_m\gamma_m \end{pmatrix},$$

Proposition 1 and Proposition 2 except (20) will follow from results in Proposition 1 and Proposition 2 of Hong and Jin (2018). We now turn to the proof of (20) in Proposition 2. From the conditions of (20), it is easy to see that $\widetilde{\pi}_{b1}^* < 0$ and $\widetilde{\pi}_{b2}^* < 0$. To prove that $\widetilde{\pi}_{q1}^* = 0$, we let

$$f(\widetilde{\pi}_{q1}) = -\widetilde{\pi}_{q1}E^Q[J^{v,Q}] + \frac{1}{1-\gamma}E^P\left[\left(1+\widetilde{\pi}_{q1}J^{v,P}\right)^{1-\gamma}e^{\gamma(B_1+B_3\beta_0)J^{v,P}} - 1\right].$$

Then, we have

$$\frac{df(\widetilde{\pi}_{q1})}{d\widetilde{\pi}_{q1}} = -E^Q[J^{v,Q}] + E^P\left[\left(1+\widetilde{\pi}_{q1}J^{v,P}\right)^{-\gamma}J^{v,P}e^{\gamma(B_1+B_3\beta_0)J^{v,P}}\right].$$

Furthermore, by noticing that $\widetilde{\pi}_{q1} \geq 0, B_1(t) < 0$ and $B_3(t) < 0$, we obtain

$$E^P\left[\left(1+\widetilde{\pi}_{q1}J^{v,P}\right)^{-\gamma}J^{v,P}e^{\gamma(B_1+B_3\beta_0)J^{v,P}}\right] \leq E^P\left[J^{v,P}\right],$$

implying that $\frac{df(\widetilde{\pi}_{q1})}{d\widetilde{\pi}_{q1}} < 0$ since $E^Q[J^{v,Q}] > E^P[J^{v,P}]$. Consequently, $\widetilde{\pi}_{q1}^* = 0$. ■

B Proof of Proposition 3

We assume that the model "SV2F-PJ-VJ" in Aït-Sahalia, Karaman and Mancini (2015) is the true model and then evaluate the utility costs of suboptimal strategies based on the models "SV2F" and "SV2F-PJ", respectively. The indirect value function J of the true model "SV2F-PJ-VJ" is given in Proposition 3. We now first derive the indirect utility corresponding to the model "SV2F".

Suppose that the two variance swaps have maturities of τ_1 and τ_2. Let n_{1t} and n_{2t} denote the optimal portfolio strategy in the model "SV2F" which are given by (43) and (44) in Proposition 4 of Egloff, Leippold and Wu (2010). Consider the corresponding strategy $w = (n_{1t}, n_{2t})^\top$ in the the model "SV2F-PJ-VJ" and let $J^{(1)}$ denote its indirect utility. Then $J^{(1)}$ satisfies the the partial differential equation below

$$J_t^{(1)} + \mathcal{L}J^{(1)} = 0,$$

where $\mathcal{L}J^{(1)}$ is the infinitesimal generator of $J^{(1)}$. Thus, we have

$$\begin{aligned} 0 &= J_t^{(1)} + \frac{1}{2}W^2 w^\top \Sigma_b^{(1)} (\Sigma_b^{(1)})^\top w J_{WW}^{(1)} + W(w^\top(b - r\mathbf{1}_2) + r)J_W^{(1)} \\ &+ (b^x)^\top J_X^{(1)} + \alpha(\lambda_\infty - \lambda_t)J_\lambda^{(1)} + W w^\top \Sigma_b^{(1)} \sigma^{x\top} J_{WX}^{(1)} + \frac{1}{2}Tr(\sigma^x \sigma^{x\top} J_{XX^\top}^{(1)}) \\ &+ \lambda_t E^P \left[J^{(1)}(W\left(\widetilde{\pi}_{q1} J^{v,P} + 1\right), v_t + J^{v,P}, m_t, \lambda_t + \beta_0 J^{v,P}) - J^{(1)}(W, v_t, m_t, \lambda_t) \right], \quad (27) \end{aligned}$$

where $X_t = (v_t, m_t)^\top$,

$$\widetilde{\pi}_{q1} = n_{1t}(\phi_v(\tau_1) + \beta_0\phi_\lambda(\tau_1)) + n_{2t}(\phi_v(\tau_2) + \beta_0\phi_\lambda(\tau_2)),$$

$$\Sigma_b^{(1)} = \begin{pmatrix} \phi_v(\tau_1)\sigma_v\sqrt{v_t} & \phi_m(\tau_1)\sigma_m\sqrt{m_t} \\ \phi_v(\tau_2)\sigma_v\sqrt{v_t} & \phi_m(\tau_2)\sigma_m\sqrt{m_t} \end{pmatrix}$$

$$b - r\mathbf{1}_2 = \begin{pmatrix} \phi_v(\tau_1)\sigma_v\gamma_v v_t + \phi_m(\tau_1)\sigma_m\gamma_m m_t - (\phi_v(\tau_1) + \beta_0\phi_\lambda(\tau_1))\mu_v^Q\lambda_t \\ \phi_v(\tau_2)\sigma_v\gamma_v v_t + \phi_m(\tau_2)\sigma_m\gamma_m m_t - (\phi_v(\tau_2) + \beta_0\phi_\lambda(\tau_2))\mu_v^Q\lambda_t \end{pmatrix}.$$

Note that according to the formulas (43) and (44) in Egloff, Leippold and Wu (2010)the jump exposure $\widetilde{\pi}_{q1}$ is deterministic. We guess the following indirect utility function:

$$J^{(1)}(t, W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma}[f^{(1)}(t, X_t)]^\gamma = \frac{W^{1-\gamma}}{1-\gamma}\left[e^{A^{(1)}(t) + B_1^{(1)}(t)v_t + B_2^{(1)}(t)m_t + B_3^{(1)}(t)\lambda_t}\right]^\gamma \quad (28)$$

By using the same method as in the proof of Proposition 1, substituting the above function $J^{(1)}$ into the equation (27) gives the ODEs for the functions $A^{(1)}(t)$, $B^{(1)}(t) = (B_1^{(1)}(t), B_2^{(1)}(t))^\top$ and $B_3^{(1)}(t)$ in Proposition 3. Furthermore, the utility cost, CE, of following the suboptimal strategy

$w = (n_{1t}, n_{2t})^\top$ is obtained by applying the formulas of $J(t, W_t(1-CE), X_t)$ and $J^{(1)}(t, W_t, X_t)$.

■

C MCMC Estimation for Variance Swap Rates

C.1 Model Discretization

we assume that the stock price, volatility and its long-run mean under a risk-neutral measure Q are given as follows:

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= (r-\delta)dt + \sqrt{v_t}dB_{1t}^Q + (\exp(J_t^{s,Q}) - 1)dN_t - g^Q\lambda_t dt, \\ dv_t &= \kappa_v^Q(m_t - v_t)dt + \sigma_v\sqrt{v_t}dB_{2t}^Q + J_t^{v,Q}dN_t \\ dm_t &= \kappa_m^Q(\theta_m^Q - m_t)dt + \sigma_m\sqrt{m_t}dB_{3t}^Q, \end{aligned} \tag{29}$$

where $corr(dB_{1t}^Q, dB_{2t}^Q) = \rho dt$ and $g^Q = e^{\mu_j^Q + \sigma_j^2/2} - 1$. We specify the market price of risks for the Brownian motions by γ_i $(i = s, v, m)$ in the following way:

$$\Lambda_t^\top = [\gamma_s\sqrt{(1-\rho^2)v_t}, \gamma_v\sqrt{v_t}, \gamma_m\sqrt{m_t}],$$

and then, under the objective probability P, the stock price and variance dynamics can be represented as follows:

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= \mu_t dt + \sqrt{v_t}dB_{1t}^P + (\exp(J_t^{s,P}) - 1)dN_t - g^P\lambda_t dt, \\ dv_t &= \kappa_v^P(\frac{\kappa_v^Q}{\kappa_v^P}m_t - v_t)dt + \sigma_v\sqrt{v_t}dB_{2t}^P + J_t^{v,P}dN_t \\ dm_t &= \kappa_m^P(\theta_m^P - m_t)dt + \sigma_m\sqrt{m_t}dB_{3t}^P, \end{aligned} \tag{30}$$

where $\mu_t = r_f - \delta + \gamma_s(1-\rho^2)v_t + \gamma_v\rho v_t + (g^P - g^Q)\lambda_t$, $\kappa_v^P = \kappa_v^Q - \gamma_v\sigma_v$, $\kappa_m^P = \kappa_m^Q - \gamma_m\sigma_m$, $g^P = e^{\mu_j^P + \sigma_j^2/2} - 1$ and $\theta_m^P = \theta_m^Q\kappa_m^Q/\kappa_m^P$, while r_f is the risk free rate, and δ is the dividend yield, both taken to be constant for simplicity, and $corr(dB_{1t}^P, dB_{2t}^P) = \rho dt$. The correlation parameter ρ is used to capture the so-called leverage effect between stock returns and variance changes. The two Brownian motions, B_{3t}^Q, B_{3t}^P, are uncorrelated with other Brownian motions.

We then assume that the jump intensity λ_t of the counting process N_t under the measure Q

follows a self-exciting process as follows:

$$d\lambda_t = \alpha(\lambda_\infty - \lambda_t)dt + \beta_0 J_t^{v,Q} dN_t, \tag{31}$$

where α, λ_∞ and $\beta_0 > 0$. Accordingly, its dynamics under the measure P can be represented as

$$d\lambda_t = \alpha(\lambda_\infty - \lambda_t)dt + \beta_0 J_t^{v,P} dN_t. \tag{32}$$

Now let $Y_t = \ln(S_t)$ be the natural log of the stock price, S_t at time $t > 0$. Applying an Euler discretization to the stock price in (30) equally with a time interval $\triangle$:

$$\begin{aligned}
&Y_{t+1} - Y_t = (r + \gamma_s(1-\rho^2)v_t + \gamma_v \rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle + \sqrt{v_t}\triangle B_{1,t}^P + J_{t+1}^{s,P}\triangle N_t;\\
&v_{t+1} - v_t = [\kappa_v^P(m_t - v_t) + m_t\gamma_v\sigma_v]\triangle + \sigma_v\sqrt{v_t}\triangle B_{2,t}^P + J_{t+1}^{v,P}\triangle N_t;\\
&m_{t+1} - m_t = \kappa_m^P(\theta_m^P - m_t)\triangle + \sigma_m\sqrt{m_t}\triangle B_{3,t}^P;\\
&\lambda_t - \lambda_{t-1} = \alpha(\lambda_\infty - \lambda_{t-1})\triangle + \beta_0 J_t^{v,P}\triangle N_{t-1},
\end{aligned} \tag{33}$$

where $r = r_f - \delta$ and $\triangle B_{i,t}^P = B_{i,t+1}^P - B_{i,t}^P$ $(i = 1, 2, 3)$, and so $\triangle B_{i,t}^P \sim N(0, \triangle)$, and $\triangle N_t = N_{t+1} - N_t$ and the latent occurrence of jumps in both price and volatility on day are expressed as

$$\triangle N_t \sim Bernoulli(\lambda_{t,\triangle}), \tag{34}$$

suggesting $prob(\triangle N_t = 1) = \lambda_t\triangle$. In particular, the discretized jump intensity, λ_t, is conditionally deterministic and thus possesses a structure that is analogous to the structure prescribed by a GARCH model for latent volatility, with the lagged jump occurrences playing a similar role to the lagged (squared) returns in a GARCH model. Without a jump in either price or volatility, λ_t will revert back to its long run, or steady-state value λ_0. Assuming stationarity of the (discretized) jump intensity process, it is known that its long-term run under the time interval of $\triangle$ is given by

$$\lambda_0 = \frac{\alpha}{\alpha - \beta_0\mu_v^P}\lambda_\infty, \tag{35}$$

which implies that $0 < \beta_0 \max(\mu_v^P, \mu_Q^P) < \alpha$ and $\beta_0 > 0$ and $\lambda_\infty > 0$.[26]

We further assume that variance swap rates are collected with measurement errors. Let n denote the number of contract maturities. We may obtain n observation equations:

$$VS_{t,\tau_i} = VS^M_{t,\tau_i}(v_t, m_t, \lambda_t; \mathbf{\Theta}) + \epsilon_{t,\tau_i}, \tag{36}$$

where $VS_{t,\tau_i,i\in\{1,2,\cdots,n\}}$ presents the time-t market quote of a variance swap rate with time-to-maturity τ, and $VS^M_{t,\tau_i}(v_t, m_t, \lambda_t; \mathbf{\Theta})$ is the theoretical rate for this variance swap contract, conditional on the latten variables v_t, m_t, λ_t and the model parameters $\mathbf{\Theta}$. The observation errors, ϵ_{t,τ_i}, are assumed to follow a normal distribution with zero mean and a variance of $\sigma_{\epsilon,i}$, independent of the state process and across time, e.g.,[27]

$$\epsilon_{t,\tau} = (\epsilon_{t,\tau_1}, \cdots, \epsilon_{t,\tau_n}) \sim N(0, \boldsymbol{\sigma}_\epsilon^2 I_n), \tag{37}$$

where $\boldsymbol{\sigma}_\epsilon = (\sigma_{\epsilon,1}, \cdots, \sigma_{\epsilon,n})$ and I_n is an $n \times n$ identity matrix. As a result, the joint likelihood of the observation errors is represented by an n-dimensional normal density function, while the pricing error is defined as the model-based VS rate minus VS rate in volatility percentage units in terms of the root mean-squared errors (RMSEs):

$$\text{RMSE}(\Theta; i) = \sqrt{\frac{\sum_{t=1}^T (\sqrt{E^Q[VS_{t,\tau_i}|\Theta]} \times 100 - \sqrt{VS_{t,\tau_i}} \times 100])^2}{T}}, \tag{38}$$

In sum, we have observations $(Y_t, VS_{t,\tau})_{t=0}^T$; latent variables including variances $(v_t)_{t=0}^T$, the central-tendency $(m_t)_{t=0}^T$, jump times $(N_t)_{t=1}^T$, price/volatility jump sizes $(J^s, J^v)_{t=1}^T$, the jump intensity $(\lambda_t)_{t=1}^T$; and parameters:

$$\mathbf{\Theta} = \{(\kappa_v^P, \sigma_v, \rho, \mu_v^P, \kappa_m^P, \theta_m^P, \sigma_m, \mu_j^P, \sigma_j), (\mu_j^Q, \mu_v^Q), (\gamma_s, \gamma_v, \gamma_m), (\alpha, \lambda_\infty, \beta_0), (\rho_\epsilon, \sigma_\epsilon)\} \tag{39}$$

[26] Alternatively, it can be derived from the final equation in (33):

$$\lambda_0 = E[\lambda_t] = E[\lambda_{t-1}] = \frac{1}{\triangle} E[\triangle N_{t-1}]$$

which implies $\alpha\lambda_\infty - (\alpha - \beta_0 \mu_v^P)\lambda_0 = 0$.

[27] Aït-Sahalia, Karaman and Mancini (2015) document that the correlations of the observation errors in variance swap prices are close to zero across maturities.

where the first group of parameters is unique to the physical measure P, the second one is unique to the risk-neutral measure Q, the third one represents the market prices of return, volatility and the central tendency, the fourth one reports all parameters involving the jump intensity process, and the final one presents variance swap pricing errors.

C.2 Prior Distributions of Model Parameters

To simplify notation, we denote the price returns as $\mathbf{Y} = \{Y_t\}_{t=0}^T$, the variance swap rates as $\mathbf{VS} = \{VS_{t,\tau}\}_{t=0}^T$, the volatility variables as $\mathbf{V} = \{v_t\}_{t=0}^T$, the central tendency as $\mathbf{M} = \{m_t\}_{t=0}^T$, the price/volatility jump times/sizes as $\mathbf{J}^s = \{N_t, J^s\}_{t=1}^T$ and $\mathbf{J}^v = \{N_t, J^v\}_{t=1}^T$ (so $\mathbf{J} = \{\mathbf{J}^s, \mathbf{J}^v\}$) and $\mathbf{N} = \{N_t\}_{t=1}^T$, the jump intensity as $\boldsymbol{\lambda} = \{\lambda_t\}_{t=1}^T$.

Given the assumed dynamics of variance swap rates, we have

$$P(\mathbf{VS}|\mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}, \boldsymbol{\Theta}) = \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \tag{40}$$

where $\Sigma = \boldsymbol{\sigma}_\epsilon^2 I_n$ and $\boldsymbol{\epsilon}_{t+1,\tau} = VS_{t+1,\tau} - VS^M_{t+1,\tau}$. Also, the full-information likelihood of the price returns is a product of bivariate normals as follows:

$$P(\mathbf{Y}|\mathbf{V}, \mathbf{M}, \mathbf{J}, \boldsymbol{\Theta}) = \prod_{t=0}^{T-1} \frac{\exp(-\frac{(\epsilon^y_{t+1})^2 - 2\rho\epsilon^y_{t+1}\epsilon^v_{t+1} + (\epsilon^v_{t+1})^2}{2(1-\rho^2)})}{2\pi\sigma_v v_t \triangle\sqrt{1-\rho^2}}, \tag{41}$$

where

$$\epsilon^y_{t+1} = (Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v\rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle - J^{s,P}_{t+1}\triangle N_t)/\sqrt{v_t\triangle}$$

and

$$\epsilon^v_{t+1} = (v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - m_t\gamma_v\sigma_v\triangle - J^{v,P}_{t+1}\triangle N_t)/(\sigma_v\sqrt{v_t\triangle}).$$

We then consider the priors for parameters of all four models. To simplify the simulation procedure, we use standard conjugate priors as follows:

- **Priors for All Models**. We consider the following priors distributions:
  - **Parameters under P:** $\kappa_v^P \sim N(0,1)1_{\kappa_v^P>0}, \sigma_v \sim \frac{1}{\sigma_v}$, $\kappa_m^P \sim N(0,1)1_{\kappa_m^P>0}$, $\theta_m^P \sim N(0,1)1_{\theta_m^P>0}, \sigma_m \sim \frac{1}{\sigma_m}$, $\rho \sim$ Uniform(0,1), and $\sigma_\epsilon \sim \frac{1}{\sigma_\epsilon}$;

– **Parameters under Q:** $\gamma_s \sim N(0,1)$, $\gamma_v \sim N(0,1)$, $\gamma_m \sim N(0,1)$.

- **Specific Priors for Jumps in AKM, JP and HJ Model**. We further consider the following specific priors distributions:

  – **Parameters under P:** $\mu_j^P \sim N(0,1)$, $\sigma_j \sim \frac{1}{\sigma_j}$, $\mu_v^P \sim \frac{1}{\mu_v^P}$;

  – **Parameters under Q:** $\mu_j^Q \sim N(0,1)$, $\mu_v^Q \sim N(0,1)1_{\mu_v^Q>0}$.

  Moreover, the prior for jump intensity (λ_t) in AKM and JP Model is specified as follows:

  – **Parameters for $\boldsymbol{\lambda}$:** $\lambda_0 \sim N(0,5)1_{\lambda_0>0}$ and $\lambda_1 \sim N(0,5)1_{\lambda_1>0}$,

  while it is given in HJ Model as follows:

  – **Parameters for $\boldsymbol{\lambda}$:** $\alpha \sim \text{Uniform}(0, 1/\triangle)$, $\beta_0 \sim N(0,5)1_{\beta_0>0}$ and $\lambda_\infty \sim N(0,5)1_{\lambda_\infty>0}$.

C.3 Posterior Distributions of Model Parameters and Variables

This section discusses the posterior distributions of model parameters and latent variables for all the four models. Compared with the case of stock prices only, the posterior likelihood in the presence of volatility derivatives prices has additional component of the likelihood of variance swap pricing errors. In particular, those parameters that appear in the variance swap pricing formula usually do not have known posterior distributions. We follow the method proposed by Damine, Wakefield and Walker (1999) (hereafter DWW) and adopted in Yu, Li and Wells (2011)) to update them. Otherwise, we may follow the standard way to draw samples from the posterior distributions.

More specifically, we consider the posterior distributions of model parameters and state variables in ELW AKM, JP and HJ Model. By setting $J^s = J^v = 0$ and $\lambda = 0$, we then conduct the MCMC estimation in ELW model, while JP Model can be estimated by setting $J^v = 0$.

C.3.1 MCMC Methods for Diffusion Components

We first derive the posteriors for all the relevant parameters.

- **Posterior for κ_v^P.** Conditional on the prior $\kappa_v^P \sim N(0,1)$, the posterior of κ_v^P can be expressed as follows:

$$p(\kappa_v^P|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\boldsymbol{\Theta}) \propto p(\mathbf{VS}|\mathbf{V},\mathbf{M},\mathbf{J},\boldsymbol{\Theta}) \times p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\boldsymbol{\Theta}) \times p(\kappa_v^P), \qquad (42)$$

where

$$
\begin{aligned}
&p(\mathbf{VS}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) = \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \\
&p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\kappa_v^P) \propto \\
&\prod_{t=0}^{T-1} \frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{C_{t+1}^2}{v_t\triangle} - \frac{2\rho C_{t+1}(D_{t+1}-\kappa_v^P(m_t-v_t)\triangle)}{v_t\sigma_v\triangle} + \frac{(D_{t+1}-\kappa_v^P(m_t-v_t)\triangle)^2}{\sigma_v^2 v_t\triangle})}}{2\pi\sigma_v v_t\triangle\sqrt{1-\rho^2}} \times \frac{e^{-\frac{(\kappa_v^P)^2}{2}}1_{\kappa_v^P>0}}{\sqrt{2\pi}} \\
&\propto N(\frac{S}{W},\frac{1}{W})1_{\kappa_v^P>0},
\end{aligned}
\tag{43}
$$

with $\Sigma = \sigma_\epsilon^2 I_n$ and so we have

$$
p(\kappa_v^P|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto \underbrace{\prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}}_{\ell(\kappa_v^P)} \times N(\frac{S}{W},\frac{1}{W})1_{\kappa_v^P>0}, \tag{44}
$$

where

$$
\begin{aligned}
C_{t+1} &= Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v\rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle - J_{t+1}^{s,P}\triangle N_t; \\
D_{t+1} &= v_{t+1} - v_t - m_t\gamma_v\sigma_v\triangle - J_{t+1}^{v,P}\triangle N_t; \\
S &= \frac{1}{\sigma_v(1-\rho^2)}\sum_{t=0}^{T-1}\frac{(m_t-v_t)}{v_t}(\frac{D_{t+1}}{\sigma_v} - \rho C_{t+1}) \\
W &= 1 + \frac{\triangle}{\sigma_v^2(1-\rho^2)}\sum_{t=0}^{T-1}\frac{(m_t-v_t)^2}{v_t};
\end{aligned}
\tag{45}
$$

Conditional on the g-th sample $\kappa_v^{P,g}$, we follow the DWW method to draw a sample for $\kappa_v^{P,g+1}$ as follows:

1. Draw $\kappa_v^{P,*}$ from $N(\frac{S}{W},\frac{1}{W})1_{\kappa_v^P>0}$;
2. Draw an auxiliary variable u from Uniform$(0,\ell(\kappa_v^{P,g}))$;
3. Accept $\kappa_v^{P,g+1} = \kappa_v^{P,*}$ if $\ell(\kappa_v^{P,*}) > u$, and otherwise set $\kappa_v^{P,g+1} = \kappa_v^{P,g}$.

- **Posterior for** σ_v. Conditional on the prior $\sigma_v \sim 1/\sigma_v$, the posterior of σ_v can be expressed as follows:

$$
p(\sigma_v|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\sigma_v), \tag{46}
$$

where

$$p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\sigma_v) \propto \prod_{t=0}^{T-1} \frac{e^{-\frac{1}{2(1-\rho^2)}(C_{t+1}^2 - 2\rho C_{t+1}(\frac{D_{t+1}}{\sigma_v} - \frac{m_t\gamma_v\sqrt{\triangle}}{\sqrt{v_t}}) + (\frac{D_{t+1}}{\sigma_v} - \frac{m_t\gamma_v\sqrt{\triangle}}{\sqrt{v_t}})^2)}}{2\pi\sigma_v v_t \triangle\sqrt{1-\rho^2}} \times \frac{1}{\sigma_v}$$
$$\propto (\frac{1}{\sigma_v^2})^{\frac{T}{2}+\frac{1}{2}} \times \exp(-\frac{\sum_{t=0}^{T-1} D_{t+1}^2}{2(1-\rho^2)}\frac{1}{\sigma_v^2} + \frac{\rho\sum_{t=0}^{T-1} C_{t+1}D_{t+1} + \gamma_v\sqrt{\triangle}\sum_{t=0}^{T-1}\frac{D_{t+1}m_t}{\sqrt{v_t}}}{1-\rho^2}\frac{1}{\sigma_v}), \tag{47}$$

where

$$\begin{aligned} C_{t+1} &= \frac{Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v\rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle - J_{t+\triangle}^{s,P}\triangle N_t}{\sqrt{v_t\triangle}}; \\ D_{t+1} &= \frac{v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - J_{t+\triangle}^{v,P}\triangle N_t}{\sqrt{v_t\triangle}}. \end{aligned} \tag{48}$$

Define

$$\ell(\sigma_v) \triangleq \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}_{t+1,\tau}'\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \times \exp(\frac{\rho\sum_{t=0}^{T-1} C_{t+1}D_{t+1} + \gamma_v\sqrt{\triangle}\sum_{t=0}^{T-1}\frac{D_{t+1}m_t}{\sqrt{v_t}}}{1-\rho^2}\frac{1}{\sigma_v}).$$

We then have

$$p(\sigma_v|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto \ell(\sigma_v) \times (\frac{1}{\sigma_v^2})^{\frac{T}{2}+\frac{1}{2}} \times \exp(-\frac{\sum_{t=0}^{T-1} D_{t+1}^2}{2(1-\rho^2)}\frac{1}{\sigma_v^2}) = \ell(\sigma_v) \times f(\frac{1}{\sigma_v^2}),$$

where $f(x) \sim Gamma(x; a, b)$ with

$$\begin{aligned} a &= \frac{T}{2} + \frac{3}{2}; \\ b &= (\frac{\sum_{t=0}^{T} D_{t+1}^2}{2(1-\rho^2)})^{-1}; \end{aligned} \tag{49}$$

Given the sample of σ_v^g, the sample of σ_v^{g+1} is drawn as follows:

1. Draw σ_v^* from $Gamma(\frac{1}{\sigma_v^2}; a, b)$;
2. Draw an auxiliary variable u from $Uniform(0, \ell(\sigma_v^g))$;
3. Accept $\sigma_v^{g+1} = \sigma_v^*$ if $\ell(\sigma_v^*) > u$, and otherwise set $\sigma_v^{g+1} = \sigma_v^g$.

- **Posterior for** ρ. Conditional on the prior $\rho \sim Uniform(0, 1)$, the posterior of ρ can be

expressed as follows:

$$
\begin{aligned}
& p(\rho|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\rho), \\
& \propto \prod_{t=0}^{T-1} \frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{(C_{t+1}-\rho(\gamma v-\gamma s\rho)v_t\triangle)^2}{v_t\triangle}-2\rho\frac{(C_{t+1}-\rho(\gamma v-\gamma s\rho)v_t\triangle)D_{t+1}}{\sigma v v_t\triangle}+\frac{D_{t+1}^2}{\sigma_v^2 v_t\triangle})}}{2\pi\sigma_v v_t\triangle\sqrt{1-\rho^2}} \times 1 \\
& \propto (1-\rho^2)^{-\frac{T-1}{2}} \exp(-\frac{1}{2(1-\rho^2)}\sum_{t=0}^{T-1}\frac{1}{v_t\triangle}(C_{t+1}^2+\frac{D_{t+1}^2}{\sigma_v^2})) \\
& \times \exp(-\frac{\rho^2(\gamma_v-\gamma_s\rho)}{2(1-\rho^2)}[(\gamma_v-\gamma_s\rho)\sum_{t=0}^{T-1}v_t\triangle+2\sum_{t=0}^{T-1}\frac{D_{t+1}}{\sigma_v}]) \\
& \times \exp(\frac{\rho}{(1-\rho^2)}[(\gamma_v-\gamma_s\rho)\sum_{t=0}^{T-1}C_{t+1}+\sum_{t=0}^{T-1}\frac{C_{t+1}D_{t+1}}{\sigma_v v_t\triangle}]) \\
& \triangleq \ell(\rho)
\end{aligned}
\tag{50}
$$

where

$$
\begin{aligned}
C_{t+1} &= Y_{t+1}-Y_t-(r+\gamma_s v_t-(e^{\mu_j^Q+\frac{\sigma_j^2}{2}}-1)\lambda_t-\frac{1}{2}v_t)\triangle-J_{t+\triangle}^{s,P}\triangle N_t; \\
D_{t+1} &= v_{t+1}-v_t-\kappa_v^P(m_t-v_t)\triangle-m_t\gamma_v\sigma_v\triangle-J_{t+\triangle}^{v,P}\triangle N_t.
\end{aligned}
$$

We then follow the approach used in Yu, Li and Wells (2011) to draw the sample for ρ. Given the sample of ρ^g, the sample of ρ_v^{g+1} is drawn as follows:

1. Draw $\frac{1}{2}\log(\frac{1+\rho^*}{1-\rho^*})$ from $N(\frac{1}{2}\log(\frac{1+\rho_r}{1-\rho_r}),\frac{1}{T-3})$ to generate a sample ρ^*, where $\rho_r = Corr(\mathbf{C},\mathbf{D})$, with $\mathbf{C}=\{C_{t+1}\}_{t=0}^{T-1}, \mathbf{D}=\{D_{t+1}\}_{t=0}^{T-1}$, and define $z(\rho_r)=\frac{1}{2}\log(\frac{1+\rho_r}{1-\rho_r})$ which is the formula of Fisher's Z transformation[28];

2. Accept $\rho^{g+1}=\rho^*$ with probability $\alpha(\rho^g,\rho^*)$, where

$$
\alpha(\rho^g,\rho^*) = \min(\frac{\ell(\rho^*)}{\ell(\rho^g)} \times \frac{\exp(-\frac{(z(\rho^g)-z(\rho^r))^2}{\frac{2}{T-3}})}{\exp(-\frac{(z(\rho^*)-z(\rho^r))^2}{\frac{2}{T-3}})}, 1) \tag{51}
$$

Yu, Li and Wells (2011) suggest that this approach converges more quickly than the one without the transformation by removing the negative skewness of the sampling distribution for Pearson's correlation ρ.

[28]The Fisher transformation implies that $z(\rho)$ is approximately normally distributed with mean $\frac{1}{2}\log(\frac{1+\rho}{1-\rho})$ and variance $\frac{1}{T-3}$, where T indicates the observation number.

- **Posterior for** κ_m^P. Conditional on the prior $\kappa_m^P \sim N(0,1)1_{\kappa_m^P>0}$, the posterior of κ_m^P can be expressed as follows:

$$p(\kappa_m^P|\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\mathbf{M}|\mathbf{V},\mathbf{J},\mathbf{\Theta}) \times p(\kappa_v^P), \tag{52}$$

where

$$\begin{aligned}
&p(\mathbf{VS}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}_{t+1,\tau}'\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \triangleq \ell(\kappa_m^P) \\
&p(\mathbf{M}|\mathbf{V},\mathbf{J},\mathbf{\Theta}) \times p(\kappa_m^P) \propto \\
&\prod_{t=0}^{T-1} \frac{1}{\sqrt{2\pi m_t \triangle}\sigma_m} \exp(-\frac{(E_{t+1}-\kappa_m^P(\theta_m^P-m_t)\triangle))^2}{2\sigma_m^2 m_t \triangle}) \times \frac{e^{-\frac{(\kappa_m^P)^2}{2}}1_{\kappa_m^P>0}}{\sqrt{2\pi}} \\
&\propto N(\frac{S}{W},\frac{1}{W})1_{\kappa_m^P>0},
\end{aligned} \tag{53}$$

where

$$\begin{aligned}
E_{t+1} &= m_{t+1} - m_t; \\
S &= \frac{1}{\sigma_m^2}\sum_{t=0}^{T-1}\frac{(\theta_m^P - m_t)E_{t+1}}{m_t} \\
W &= 1 + \frac{\triangle}{\sigma_m^2}\sum_{t=0}^{T-1}\frac{(\theta_m^P - m_t)^2}{m_t};
\end{aligned} \tag{54}$$

and so the posterior for κ_m^P is given by

$$p(\kappa_m^P|\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto \ell(\kappa_m^P) \times N(\frac{S}{W},\frac{1}{W})1_{\kappa_m^P>0}. \tag{55}$$

We follow the DWW method to draw sample for $\kappa_m^{P,g+1}$ as follows, conditional on the g-th sample $\kappa_m^{P,g}$:

1. Draw $\kappa_m^{P,*}$ from $N(\frac{S}{W},\frac{1}{W})1_{\kappa_m^P>0}$;
2. Draw an auxiliary variable u from $Uniform(0,\ell(\kappa_m^{P,g}))$;
3. Accept $\kappa_m^{P,g+1} = \kappa_m^{P,*}$ if $\ell(\kappa_m^{P,*}) > u$, and otherwise set $\kappa_m^{P,g+1} = \kappa_m^{P,g}$.

- **Posterior for** θ_m^P. Conditional on the prior $\theta_m^P \sim N(0,1)1_{\theta_m^P>0}$, the posterior of θ_m^P can be

expressed as follows:

$$p(\theta_m^P|\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\mathbf{M}|\mathbf{V},\mathbf{J},\mathbf{\Theta}) \times p(\kappa_v^P), \tag{56}$$

where

$$\begin{aligned}
&p(\mathbf{VS}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}_{t+1,\tau}^{'}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \triangleq \ell(\theta_m^P) \\
&p(\mathbf{M}|\mathbf{V},\mathbf{J},\mathbf{\Theta}) \times p(\theta_m^P) \propto \\
&\prod_{t=0}^{T-1} \frac{1}{\sqrt{2\pi m_t \triangle}\sigma_m} \exp(-\frac{(E_{t+1}-\kappa_m^P\theta_m^P\triangle)^2}{2\sigma_m^2 m_t \triangle}) \times \frac{e^{-\frac{(\theta_m^P)^2}{2}}1_{\theta_m^P>0}}{\sqrt{2\pi}} \\
&\propto N(\frac{S}{W},\frac{1}{W})1_{\theta_m^P>0},
\end{aligned} \tag{57}$$

where

$$\begin{aligned}
E_{t+1} &= m_{t+1} - (1-\kappa_m^P\triangle)m_t; \\
S &= \frac{\kappa_m^P}{\sigma_m^2}\sum_{t=0}^{T-1}\frac{E_{t+1}}{m_t} \\
W &= 1 + \frac{(\kappa_m^P)^2\triangle}{\sigma_m^2}\sum_{t=0}^{T-1}\frac{1}{m_t};
\end{aligned} \tag{58}$$

and so the posterior for θ_m^P is given by

$$p(\theta_m^P|\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto \ell(\theta_m^P) \times N(\frac{S}{W},\frac{1}{W})1_{\theta_m^P>0}. \tag{59}$$

We follow the DWW method to draw sample for $\theta_m^{P,g+1}$ as follows, conditional on the g-th sample $\theta_m^{P,g}$:

1. Draw $\theta_m^{P,*}$ from $N(\frac{S}{W},\frac{1}{W})1_{\theta_m^P>0}$;
2. Draw an auxiliary variable u from $Uniform(0,\ell(\theta_m^{P,g}))$;
3. Accept $\theta_m^{P,g+1} = \theta_m^{P,*}$ if $\ell(\theta_m^{P,*}) > u$, and otherwise set $\theta_m^{P,g+1} = \theta_m^{P,g}$.

- **Posterior for** σ_m. Conditional on the prior $\sigma_m \sim 1/\sigma_m$, the posterior of σ_m can be expressed

as follows:

$$p(\sigma_m|\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\mathbf{M}|\mathbf{V},\mathbf{J},\mathbf{\Theta}) \times p(\sigma_m), \tag{60}$$

where

$$\begin{aligned} p(\mathbf{VS}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) &= \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \triangleq \ell(\sigma_m) \\ p(\mathbf{M}|\mathbf{V},\mathbf{J},\mathbf{\Theta}) \times p(\sigma_m) &\propto \prod_{t=0}^{T-1} \frac{1}{\sqrt{2\pi m_t \triangle}\sigma_m} \exp(-\frac{(E_{t+1})^2}{2\sigma_m^2 m_t \triangle}) \times \frac{1}{\sigma_m} \\ &\propto (\frac{1}{\sigma_m^2})^{\frac{T}{2}+\frac{1}{2}} \exp(-\frac{1}{2\triangle}\sum_{t=0}^{T-1}\frac{E_{t+1}^2}{m_t} \times \frac{1}{\sigma_m^2}), \end{aligned} \tag{61}$$

where

$$E_{t+1} = m_{t+1} - m_t - \kappa_m^P(\theta_m^P - m_t)\triangle.$$

We then have

$$p(\sigma_m|\mathbf{VS},\mathbf{V},\mathbf{MJ},,\mathbf{\Theta}) \propto \ell(\sigma_m) \times (\frac{1}{\sigma_m^2})^{\frac{T}{2}+\frac{1}{2}} \exp(-\frac{1}{2\triangle}\sum_{t=0}^{T-1}\frac{E_{t+1}^2}{m_t} \times \frac{1}{\sigma_m^2}) = \ell(\sigma_m) \times f(\frac{1}{\sigma_m^2}),$$

where $f(x) \sim Gamma(x;a,b)$ with

$$\begin{aligned} a &= \frac{T}{2} + \frac{3}{2}; \\ b &= (\frac{1}{2\triangle}\sum_{t=0}^{T-1}\frac{E_{t+1}^2}{m_t})^{-1}; \end{aligned}$$

We now follow the DWW method to draw the sample for σ_m. Given the sample of σ_m^g, the sample of σ_m^{g+1} is drawn as follows:

1. Draw $1/(\sigma_m^*)^2$ from $Gamma(\frac{1}{\sigma_m^2};a,b)$ to generate a sample σ_m^*;
2. Draw an auxiliary variable u from $Uniform(0,\ell(\sigma_m^g))$;
3. Accept $\sigma_m^{g+1} = \sigma_m^*$ if $\ell(\sigma_m^*) > u$, and otherwise set $\sigma_m^{g+1} = \sigma_m^g$.

- **Posterior for γ_s.** Conditional on the prior $\gamma_s \sim N(0,1)$, the posterior of γ_s can be expressed

as follows:

$$p(\gamma_s|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\gamma_s), \tag{62}$$

where

$$\begin{aligned}
&p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\gamma_s) \propto \\
&\prod_{t=0}^{T-1} \frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{(C_{t+1}-(1-\rho^2)v_t\triangle\gamma s)^2}{v_t\triangle} - \frac{2\rho(C_{t+1}-(1-\rho^2)v_t\triangle\gamma s)D_{t+1}}{v_t\sigma v\triangle} + \frac{D_{t+1}^2}{\sigma_v^2 v_t\triangle})}}{2\pi\sigma_v v_t\triangle\sqrt{1-\rho^2}} \times e^{-\frac{\gamma_s^2}{2}} \\
&\propto \prod_{t=0}^{T-1} e^{-\frac{1}{2}[(1-\rho^2)v_t\triangle\gamma_s^2 - 2(C_{t+1}-\frac{\rho}{\sigma v}D_{t+1})\gamma_s]} \times e^{-\frac{\gamma_s^2}{2}} \\
&\propto N(\frac{S}{W}, \frac{1}{W}),
\end{aligned} \tag{63}$$

where

$$\begin{aligned}
C_{t+1} &= Y_{t+1} - Y_t - (r + \gamma_v\rho v_t - \frac{1}{2}v_t - (e^{\mu_j^Q + \sigma_j^2/2} - 1)\lambda_t)\triangle - J_{t+1}^{s,P}\triangle N_t; \\
D_{t+1} &= v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - m_t\gamma_v\sigma_v\triangle - J_{t+1}^{v,P}\triangle N_t; \\
S &= \sum_{t=0}^{T-1}(C_{t+1} - \frac{\rho}{\sigma_v}D_{t+1}) \\
W &= 1 + (1-\rho^2)\triangle\sum_{t=0}^{T-1} v_t;
\end{aligned} \tag{64}$$

- **Posterior for** γ_v. Conditional on the prior $\gamma_v \sim N(0,1)$, the posterior of γ_v can be expressed as follows:

$$p(\gamma_v|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\gamma_v), \tag{65}$$

where

$$
\begin{aligned}
&p(\mathbf{VS}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta})\prod_{t=0}^{T-1}\frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\triangleq\ell(\gamma_v)\\
&p(\mathbf{Y}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta})\times p(\gamma_v)\propto\\
&\prod_{t=0}^{T-1}\frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{(C_{t+1}-\rho v_t\triangle\gamma v)^2}{v_t\triangle}-\frac{2\rho(C_{t+1}-\rho v_t\triangle\gamma v)(D_{t+1}-m_t\gamma v\sigma v\triangle)}{v_t\sigma v\triangle}+\frac{(D_{t+1}-m_t\gamma v\sigma v\triangle)^2}{\sigma_v^2 v_t\triangle})}}{2\pi\sigma_v v_t\triangle\sqrt{1-\rho^2}}\times e^{-\frac{\gamma_v^2}{2}}\\
&\propto\prod_{t=0}^{T-1}e^{-\frac{1}{2(1-\rho^2)}[(\rho^2\triangle(v_t-2m_t)+\frac{m_t^2\triangle}{v_t})\gamma_v^2-2(\rho C_{t+1}(1-\frac{m_t}{v_t})-\frac{D_{t+1}}{\sigma v}(\rho^2-\frac{m_t}{v_t}))\gamma_v]}\times e^{-\frac{\gamma_v^2}{2}}\\
&e^{-\frac{1}{2}[(\frac{\rho^2\triangle\sum_{t=0}^{T-1}(v_t-2m_t)+\triangle\sum_{t=0}^{T-1}\frac{m_t^2}{v_t}}{1-\rho^2}+1)\gamma_v^2-2(\frac{\rho\sum_{t=0}^{T-1}C_{t+1}(1-\frac{m_t}{v_t})}{1-\rho^2}-\frac{\sum_{t=0}^{T-1}D_{t+1}(\rho^2-\frac{m_t}{v_t})}{(1-\rho^2)\sigma v})\gamma_v]+\cdots}\\
&\propto N(\frac{S}{W},\frac{1}{W}),
\end{aligned}
\tag{66}
$$

where

$$
\begin{aligned}
C_{t+1}&=Y_{t+1}-Y_t-(r+\gamma_s(1-\rho^2)v_t-\frac{1}{2}v_t-(e^{\mu_j^Q+\sigma_j^2/2}-1)\lambda_t)\triangle-J_{t+\triangle}^{s,P}\triangle N_t;\\
D_{t+1}&=v_{t+1}-v_t-\kappa_v^P(m_t-v_t)\triangle-J_{t+\triangle}^{v,P}\triangle N_t;\\
S&=\frac{1}{1-\rho^2}(\rho\sum_{t=0}^{T-1}C_{t+1}(1-\frac{m_t}{v_t})-\frac{1}{\sigma_v}\sum_{t=0}^{T-1}D_{t+1}(\rho^2-\frac{m_t}{v_t}));\\
W&=1+\frac{\rho^2\triangle\sum_{t=0}^{T-1}(v_t-2m_t)+\triangle\sum_{t=0}^{T-1}\frac{m_t^2}{v_t}}{1-\rho^2},
\end{aligned}
\tag{67}
$$

and so the posterior for γ_v is given by

$$
p(\gamma_v|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta})\propto\ell(\gamma_v)\times N(\frac{S}{W},\frac{1}{W}). \tag{68}
$$

Conditional on the sample γ_v^g, the sample of γ_v^{g+1} is drawn by applying the DWW method:

1. Draw a sample γ_v^* from $N(\frac{S}{W},\frac{1}{W})$;
2. Draw an auxiliary variable u from $Uniform(0,\ell(\gamma_v^g))$;
3. Accept $\gamma_v^{g+1}=\gamma_v^*$ if $\ell(\gamma_v^*)>u$, and otherwise set $\gamma_v^{g+1}=\gamma_v^g$.

- **Posterior for** γ_m. Conditional on the prior $\gamma_m\sim N(0,1)$, the posterior of γ_m can be

expressed as follows:

$$
\begin{aligned}
p(\gamma_m|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) &\propto p(\mathbf{VS}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\gamma_m) \\
&\propto \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \times \exp(-\frac{\gamma_m^2}{2}) \triangleq \ell(\gamma_m).
\end{aligned} \tag{69}
$$

Conditional on the sample γ_m^g, the sample of γ_m^{g+1} is drawn by applying the DWW method:

1. Draw a sample γ_m^* from $N(\gamma_m^g, 1)$;
2. Draw an auxiliary variable u from $Uniform(0,1)$;
3. Accept $\gamma_m^{g+1} = \gamma_m^*$ if $\min(\frac{\ell(\gamma_m^*)}{\ell(\gamma_m^g)}, 1) > u$, and otherwise set $\gamma_m^{g+1} = \gamma_m^g$.

- **Posterior for $\boldsymbol{\sigma}_\epsilon$.** Conditional on the prior $\boldsymbol{\sigma}_\epsilon = (\sigma_{\epsilon,1}, \cdots, \sigma_{\epsilon,n})$ with $\sigma_{\epsilon,i} \sim 1/\sigma_{\epsilon,i}$, the posterior of $\boldsymbol{\sigma}_\epsilon$ can be expressed as follows:

$$
\begin{aligned}
& p(\boldsymbol{\sigma}_\epsilon|\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\mathbf{M}|\mathbf{V},\mathbf{J},\mathbf{\Theta}) \times p(\boldsymbol{\sigma}_\epsilon) \\
& \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \times \prod_{i=1}^{n} \frac{1}{\sigma_{\epsilon,i}} \\
& \propto \prod_{i=1}^{n} \{ \prod_{t=0}^{T-1} (\frac{1}{\sigma_{\epsilon,i}^2})^{\frac{1}{2}} \exp(-\frac{1}{2}\epsilon_{t+1,\tau_i}^2) \times \frac{1}{\sigma_{\epsilon,i}^2}) \times (\frac{1}{\sigma_{\epsilon,i}^2})^{\frac{1}{2}} \} \\
& \propto \prod_{i=1}^{n} (\frac{1}{\sigma_{\epsilon,i}^2})^{\frac{T}{2}+\frac{1}{2}} \exp(-\frac{1}{2}(\sum_{t=0}^{T-1} \epsilon_{t+1,\tau_i}^2) \times \frac{1}{\sigma_{\epsilon,i}^2}) \\
& \propto \prod_{i=1}^{n} Gamma(\frac{1}{\sigma_{\epsilon,i}^2}; a, b_i),
\end{aligned} \tag{70}
$$

where $\Sigma = \boldsymbol{\sigma}_\epsilon^2 \times I_n$, and

$$
\begin{aligned}
a &= \frac{T}{2} + \frac{3}{2}, \\
b_i &= \frac{1}{\frac{1}{2}\sum_{t=0}^{T-1} \epsilon_{t+1,\tau_i}^2}.
\end{aligned}
$$

Then the sample of $\boldsymbol{\sigma}_m^{g+1}$ is drawn from $Gamma(\frac{1}{\sigma_{\epsilon,i}^2}; a, b_i)$ for $i = 1, \cdots, n$.

C.3.2 MCMC Methods for Jump Components

- **Posterior for μ_j^P and μ_j^Q.** Conditional on the prior $\mu_j^P \sim N(m, M^2)$, the posterior of μ_j^P can be expressed as follows:

$$\begin{aligned} p(\mu_j^P|\mathbf{J}^s) &\propto p(\mathbf{J}^s|\mu_j^P) \times p(\mu_j^P) \\ &\propto \prod_{t=0}^{T-1} \frac{1}{\sqrt{2\pi}\sigma_j} \exp(-\frac{(J_{t+1}^{s,P} - \mu_j^P)^2}{2\sigma_j^2}) \times \exp(-\frac{(\mu_j^P - m)^2}{2M^2}) \\ &\propto \exp(-\frac{1}{2}[(\frac{T}{\sigma_j^2} + \frac{1}{M^2})(\mu_j^P)^2 - 2(\frac{m}{M^2} + \frac{\sum_{t=0}^{T-1} J_{t+1}^{s,P}}{\sigma_j^2})\mu_j^P]) \\ &\propto N(\frac{S}{W}, \frac{1}{W}), \end{aligned} \tag{71}$$

where $W = \frac{T}{\sigma_j^2} + \frac{1}{M^2}$ and $S = \frac{m}{M^2} + \frac{\sum_{t=0}^{T-1} J_{t+1}^{s,P}}{\sigma_j^2}$. Then the sample of μ_j^P can be drawn directly. Similarly, the posterior for μ_j^Q is given as follows:

$$p(\mu_j^Q|\mathbf{Y}, \mathbf{VS}, \mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta} \times p(\mu_j^Q|\mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}) \times p(\mu_j^Q) \tag{72}$$

where [29]

$$\begin{aligned} & p(\mathbf{VS}|\mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta} \propto \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}_{t+1,\tau}^{'} \Sigma^{-1} \boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \triangleq \ell(\mu_j^Q) \\ & p(\mu_j^Q|\mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}) \times p(\mu_j^Q) \\ & \propto \prod_{t=0}^{T-1} \frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{(C_{t+1}+\mu_j^Q \lambda_t \triangle)^2}{v_t \triangle} - 2\rho \frac{(C_{t+1}+\mu_j^Q \lambda_t \triangle) D_{t+1}}{\sigma v v_t \triangle} + \frac{D_{t+1}^2}{\sigma_v^2 v_t \triangle})}}{2\pi\sigma_v v_t \triangle \sqrt{1-\rho^2}} \times e^{-\frac{(\mu_j^Q - m)^2}{2M^2}} \\ & \propto e^{-\frac{1}{2(1-\rho^2)}(\triangle \sum_{t=0}^{T-1} \frac{\lambda_t^2}{v_t}(\mu_j^Q)^2) - 2\sum_{t=0}^{T-1} \frac{\lambda_t}{v_t}(\frac{\rho D_{t+1}}{\sigma v} - C_{t+1})\mu_j^Q) + \cdots} \times e^{-\frac{(\mu_j^Q - m)^2}{2M^2}} \\ & \propto N(\frac{S}{W}, \frac{1}{W}), \end{aligned} \tag{73}$$

[29] To sample μ_j^Q, we follow the Taylor expansion of e^x for $(e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)$.

where

$$C_{t+1} = Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v \rho v_t - \frac{1}{2}v_t - \frac{\sigma_j^2}{2}\lambda_t)\triangle - J_{t+\triangle}^{s,P}\triangle N_t;$$

$$D_{t+1} = v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - m_t\gamma_v\sigma_v\triangle - J_{t+\triangle}^{v,P}\triangle N_t;$$

$$S = \frac{1}{1-\rho^2}\sum_{t=0}^{T-1}\frac{\lambda_t}{v_t}(\frac{\rho D_{t+1}}{\sigma_v} - C_{t+1}) + \frac{m}{M^2}$$

$$W = \frac{1}{M^2} + \frac{\triangle}{1-\rho^2}\sum_{t=0}^{T-1}\frac{\lambda_t^2}{v_t}.$$

Conditional on the sample $\mu_j^{Q,g}$, the sample of $\mu_j^{Q,g+1}$ is drawn by applying the Metropolis-Hastings algorithm:

1. Draw a sample $\mu_j^{Q,*}$ from $N(\frac{S}{W}, \frac{1}{W})$;
2. Draw an auxiliary variable u from $Uniform(0, \ell(\mu_j^{Q,g}))$;
3. Accept $\mu_j^{Q,g+1} = \mu_j^{Q,*}$ if $\ell(\mu_j^{Q,*}) > u$, and otherwise set $\mu_j^{Q,g+1} = \mu_j^{Q,g}$.

- **Posterior for** σ_j. Conditional on the prior $\sigma_j^2 \sim IG(m, M)$, the posterior of σ_j^2 can be expressed as follows:

$$p(\sigma_j|\mathbf{VS}, \mathbf{V}, \mathbf{MJ}, , \mathbf{\Theta}) \propto p(\mathbf{VS}|\mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta}) \times p(\sigma_j|\mathbf{Y}, \mathbf{V}, \mathbf{MJ}, \mathbf{\Theta}) \times p(\mathbf{J}^s|\sigma_j) \times p(\sigma_j)$$

where [30]

$$\begin{aligned}
p(\mathbf{VS}|\mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta}) &\propto \prod_{t=0}^{T-1}\frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \\
p(\sigma_j^2|\mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta}) &\propto \prod_{t=0}^{T-1}\frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{(C_{t+1}+\sigma_j^2\lambda_t\triangle/2)^2}{v_t\triangle} - 2\rho\frac{(C_{t+1}+\sigma_j^2\lambda_t\triangle/2)D_{t+1}}{\sigma_v v_t\triangle} + \frac{D_{t+1}^2}{\sigma_v^2 v_t\triangle})}}{2\pi\sigma_v v_t\triangle\sqrt{1-\rho^2}} \\
&\propto e^{-\frac{(\sigma_j^2/2)^2\triangle}{2(1-\rho^2)}\sum_{t=0}^{T-1}\frac{\lambda_t^2}{v_t} + \frac{\sigma_j^2/2}{1-\rho^2}\sum_{t=0}^{T-1}\frac{\lambda_t}{v_t}(\frac{\rho D_{t+1}}{\sigma_v} - C_{t+1})} \qquad (74) \\
p(\mathbf{J}^s|\sigma_j^2) \times p(\sigma_j) &\propto \prod_{t=0}^{T-1}\frac{1}{\sqrt{2\pi}\sigma_j}\exp(-\frac{(J_{t+1}^{s,P} - \mu_j^P)^2}{2\sigma_j^2}) \times \frac{M^m(\sigma_j^2)^{-m-1}e^{-\frac{M}{\sigma_j^2}}}{\Gamma(m)} \\
&\propto (\frac{1}{\sigma_j^2})^{\frac{T}{2}+m+1}\exp(-(M + \frac{1}{2}\sum_{t=0}^{T-1}(J_{t+1}^{s,P} - \mu_j^P)^2)\frac{1}{\sigma_j^2}),
\end{aligned}$$

[30] To sample σ_j, we follow the same way used in sampling μ_j^Q.

where

$$C_{t+1} = Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v \rho v_t - \frac{1}{2}v_t - \mu_j^Q \lambda_t)\triangle - J_{t+\triangle}^{s,P}\triangle N_t;$$
$$D_{t+1} = v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - m_t\gamma_v\sigma_v\triangle - J_{t+\triangle}^{v,P}\triangle N_t.$$

Now we define

$$\ell(\sigma_j^2) \triangleq p(\mathbf{VS}|\mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta}) \times p(\sigma_j^2|\mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}, \mathbf{\Theta}). \tag{75}$$

We then have

$$p(\sigma_j^2|\mathbf{VS}, \mathbf{V}, \mathbf{MJ}, , \mathbf{\Theta}) \propto \ell(\sigma_j^2) \times f(\frac{1}{\sigma_j^2}),$$

where $f(x) \sim Gamma(x; a, b)$ with

$$\begin{aligned} a &= \frac{T}{2} + m; \\ b &= \frac{1}{M + \frac{1}{2}\sum_{t=0}^{T-1}(J_{t+1}^{s,P} - \mu_j^P)^2}; \end{aligned} \tag{76}$$

We now follow the DWW method to draw the sample for σ_j. Given the sample of σ_j^g, the sample of σ_j^{g+1} is drawn as follows:

1. Draw $1/(\sigma_j^*)^2$ from $Gamma(\frac{1}{\sigma_j^2}; a, b)$ to generate a sample σ_j^*;
2. Draw an auxiliary variable u from $Uniform(0, \ell(\sigma_j^g))$;
3. Accept $\sigma_j^{g+1} = \sigma_j^*$ if $\ell(\sigma_j^*) > u$, and otherwise set $\sigma_j^{g+1} = \sigma_j^g$.

- **Posterior for μ_v^P and μ_v^Q.** Conditional on the prior $\mu_v^P \sim IG(m, M)$, the posterior of μ_v^P can be expressed as follows:

$$\begin{aligned} p(\mu_v^P|\mathbf{J}^v) &\propto p(\mathbf{J}^v|\mu_v^P) \times p(\mu_v^P) \\ &\propto \prod_{t=0}^{T-1} \frac{1}{\mu_v^P} \exp(-\frac{J_{t+1}^{v,P}}{\mu_v^P}) \times \frac{M^m(\mu_v^P)^{-m-1}e^{-\frac{M}{\mu_v^P}}}{\Gamma(m)} \\ &\propto (\frac{1}{\mu_v^P})^{T+m+1} \exp(-\frac{1}{\mu_v^P}(M + \sum_{t=0}^{T-1} J_{t+1}^{v,P})) \\ &= Gamma(\frac{1}{\mu_v^P}; a, b) \end{aligned} \tag{77}$$

where $a = T+m$ and $b = \frac{1}{M+\sum_{t=0}^{T-1} J_{t+1}^{v,P}}$. Then the sample of μ_v^P is drawn from $Gamma(\frac{1}{\mu_v^P}; a, b)$.

Similarly, the posterior for μ_v^Q is given as follows:

$$\begin{aligned} p(\mu_v^Q|\mathbf{Y},\mathbf{VS},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) &\propto p(\mathbf{VS}|\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) \times p(\mu_v^Q) \\ &\propto \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \times e^{-\frac{(\mu_v^Q)^2}{2}} 1_{\mu_v^Q>0} \triangleq \ell(\mu_v^Q). \end{aligned} \tag{78}$$

Then, the sample of $\mu_v^{Q,g+1}$ is drawn by applying the Metropolis-Hastings algorithm with a proposal density of the Gamma distribution.

We now present the posteriors for the parameters that characterize the process of the jump density (λ_t) in AKM, JP and HJ Model. Given the priors for λ_i ($i = 0, 1$), $\lambda_i \sim N(m, M)1_{\lambda_i>0}$:

$$p(\lambda_i|VS,\mathbf{\Theta}) \propto \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \times \frac{e^{-\frac{(\lambda_i-m)^2}{2}}1_{\lambda_i>0}}{\sqrt{2\pi}M^2}, \tag{79}$$

with $m = 0$ and $M = 5$. We use a normal distribution centered at the previous draw with constant variance M to generate a sample for λ_i, and then follow the DWW procedure to update the sampling through the Metropolis-Hasting algorithm. For HJ Model, the sampling of λ_∞ and β_0 can be done similarly, while the posterior for α is given by:

$$p(\alpha|VS,\mathbf{\Theta}) \propto \prod_{t=0}^{T-1} \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+1,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+1,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \times p(\alpha), \tag{80}$$

where $\alpha \sim U(\beta_0\mu_v^Q, 1/\triangle)$, and the sampling of α is drawn from its posterior through the Metropolis-Hasting algorithm.

C.3.3 Updating State Variables

We now derive the posteriors of those state variables, including $\mathbf{N}, \mathbf{J}, \mathbf{V}, \mathbf{M}$.

- **Posterior for** $\triangle N_t$. For the jump times, the conditional posterior is Bernoulli, since $\triangle N_t$ can only take two values between 0 and 1 in each time interval. Accordingly, the Bernoulli

probability is given by

$$\begin{aligned}
&p(\triangle N_t = 1|\mathbf{\Theta}, \mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}) \\
&\propto p(Y_{t+1}, v_{t+1}|\Theta, \mathbf{V}, \mathbf{M}, \mathbf{J}, \triangle N_t = 1) \times p(\triangle N_t = 1|\Theta) \\
&\propto (\lambda_t) \frac{e^{-\frac{1}{2(1-\rho^2)}(C_{t+1}^2 - 2\rho C_{t+1} D_{t+1} + D_{t+1}^2)}}{2\pi\sigma_v v_t \triangle \sqrt{1-\rho^2}}
\end{aligned}$$

where

$$\begin{aligned}
C_{t+1} &= \frac{Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v \rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle - J_{t+1}^{s,P}}{\sqrt{v_t \triangle}}; \\
D_{t+1} &= \frac{v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - m_t \gamma_v \sigma_v \triangle - J_{t+1}^{v,P}}{\sigma_v \sqrt{v_t \triangle}};
\end{aligned}$$

Then $p(\triangle N_t = 0|\mathbf{\Theta}, \mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J})$ is given by a Bernoulli probability as follows.

$$\begin{aligned}
&p(\triangle N_t = 0|\mathbf{\Theta}, \mathbf{Y}, \mathbf{V}, \mathbf{M}, \mathbf{J}) \\
&\propto p(Y_{t+1}, v_{t+1}|\mathbf{\Theta}, \mathbf{V}, \mathbf{M}, \mathbf{J}, N_{t+1} = 0) \times p(\triangle N_t = 0|\mathbf{\Theta}) \\
&\propto (1 - \lambda_t) \frac{e^{-\frac{1}{2(1-\rho^2)}(\tilde{C}_{t+1}^2 - 2\rho \tilde{C}_{t+1} \tilde{D}_{t+1} + \tilde{D}_{t+1}^2)}}{2\pi\sigma_v v_t \triangle \sqrt{1-\rho^2}}
\end{aligned}$$

where

$$\begin{aligned}
\tilde{C}_{t+1} &= \frac{Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v \rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle}{\sqrt{v_t \triangle}}; \\
\tilde{D}_{t+1} &= \frac{v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - m_t \gamma_v \sigma_v \triangle}{\sigma_v \sqrt{v_t \triangle}};
\end{aligned}$$

In sum, we may have the posterior distribution for N_{t+1} as follows:[31]

$$\triangle N_t \sim Bernoulli(\frac{\alpha_1}{\alpha_1 + \alpha_2}), \tag{81}$$

[31]To generate a random variable x that follows a $Bernoulli(p)$ distribution, we follow the procedure:1) Generate u from a uniform distribution between 0 and 1; 2) Get $x = \mathbf{1}_{u \leq p}$ where $\mathbf{1}(\cdot)$ is an indicator function.

where

$$\alpha_1 = \lambda_t \times e^{-\frac{1}{2(1-\rho^2)}(C_{t+1}^2 - 2\rho C_{t+1}D_{t+1} + D_{t+1}^2)},$$
$$\alpha_2 = (1-\lambda_t) \times e^{-\frac{1}{2(1-\rho^2)}(\tilde{C}_{t+1}^2 - 2\rho \tilde{C}_{t+1}\tilde{D}_{t+1} + \tilde{D}_{t+1}^2)}.$$

- **Posteriors for** $J^{s,P}$ and $J^{v,P}$. Conditional on the prior $J^{s,P} \sim N(\mu_j^P, \sigma_j)$, the posterior of $J_{t+1}^{s,P}$ can be expressed as follows:

$$\begin{aligned}
&p(J_{t+1}^{s,P}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{N},\mathbf{\Theta}) \propto p(\mathbf{Y},\mathbf{V}|J_{t+1}^{s,P}, J_{t+1}^{v,P}, \triangle N_t = 1, \mathbf{M},\mathbf{N},\mathbf{\Theta}) \times p(J_{t+1}^{s,P}),\\
&\propto \frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{(C_{t+1}-J_{t+1}^{s,P}\triangle N_t)^2}{v_t\triangle} - \frac{2\rho(C_{t+1}-J_{t+1}^{s,P}\triangle N_t)(D_{t+1}-J_{t+1}^{v,P}\triangle N_t)}{v_t\sigma_v\triangle} + \frac{(D_{t+1}-J_{t+1}^{v,P}\triangle N_t)^2}{\sigma_v^2 v_t\triangle})}}{2\pi\sigma_v v_t\triangle\sqrt{1-\rho^2}} \times e^{-\frac{(J_{t+1}^{s,P}-\mu_j^P)^2}{2\sigma_j^2}}\\
&\propto N(\frac{S_{t+1}}{W_{t+1}}, \frac{1}{W_{t+1}}),
\end{aligned} \tag{82}$$

where

$$\begin{aligned}
C_{t+1} &= Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v\rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle;\\
D_{t+1} &= v_{t+1} - v_t - \kappa_v^P(m_t - v_t)\triangle - m_t\gamma_v\sigma_v\triangle;\\
S_{1,t+1} &= \frac{\triangle N_t}{(1-\rho^2)v_t\triangle}(C_{t+1} - \frac{\rho}{\sigma_v}D_{t+1}) + \frac{\mu_j^P}{\sigma_j^2} + \frac{\rho(\triangle N_t)^2}{(1-\rho^2)\sigma_v v_t\triangle}J_{t+1}^{v,P};\\
W_{1,t+1} &= \frac{(\triangle N_t)^2}{(1-\rho^2)v_t\triangle} + \frac{1}{\sigma_j^2};
\end{aligned} \tag{83}$$

We now turn to derive the posterior for $J^{v,P}$. In order to derive its posterior, we first obtain the joint posterior as follows:

$$\begin{aligned}
&p(J_{t+1}^{s,P}, J_{t+1}^{v,P}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{N},\mathbf{\Theta}) \propto p(J_{t+1}^{s,P}, J_{t+1}^{v,P}|\triangle N_t = 1, \mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{\Theta})\\
&\propto \frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{(C_{t+1}-J_{t+1}^{s,P}\triangle N_t)^2}{v_t\triangle} - \frac{2\rho(C_{t+1}-J_{t+1}^{s,P}\triangle N_t)(D_{t+1}-J_{t+1}^{v,P}\triangle N_t)}{v_t\sigma_v\triangle} + \frac{(D_{t+1}-J_{t+1}^{v,P}\triangle N_t)^2}{\sigma_v^2 v_t\triangle})}}{2\pi\sigma_v v_t\triangle\sqrt{1-\rho^2}}\\
&\times e^{-\frac{(J_{t+1}^{s,P}-\mu_j^P)^2}{2\sigma_j^2}} \times \frac{1}{\mu_v^P}e^{-\frac{J_{t+1}^{v,P}}{\mu_v^P}}1_{J^{v,P}>0}\\
&\propto e^{-\frac{W_{1,t+1}(J_{t+1}^{s,P})^2 - 2A_{t+1}J_{t+1}^{s,P} + G_{t+1}(J_{t+1}^{v,P})^2 - 2F_{t+1}J_{t+1}^{v,P} - 2B_{t+1}J_{t+1}^{s,P}J_{t+1}^{v,P}}{2}}
\end{aligned} \tag{84}$$

where

$$
\begin{aligned}
A_{t+1} &= \frac{\triangle N_t}{(1-\rho^2)v_t\triangle}(C_{t+1} - \frac{\rho}{\sigma_v}D_{t+1}) + \frac{\mu_j^P}{\sigma_j^2};\\
B_{t+1} &= \frac{\rho(\triangle N_t)^2}{(1-\rho^2)\sigma_v v_t\triangle};\\
G_{t+1} &= \frac{(\triangle N_t)^2}{(1-\rho^2)\sigma_v^2 v_t\triangle};\\
F_{t+1} &= \frac{\triangle N_t}{(1-\rho^2)\sigma_v v_t\triangle}(\frac{D_{t+1}}{\sigma_v} - \rho C_{t+1}) - \frac{1}{\mu_v^P}.
\end{aligned}
\tag{85}
$$

Conditional on the prior $J^{v,P} \sim \frac{1}{\mu_v^P}e^{-\frac{J^{v,P}}{\mu_v^P}}$, the posterior of $J_{t+1}^{v,P}$ can be expressed as follows:

$$
\begin{aligned}
p(J_{t+1}^{v,P}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{\Theta}) &\propto \frac{p(J_{t+1}^{s,P},J_{t+1}^{v,P}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{N},\mathbf{\Theta})}{p(J_{t+1}^{s,P}|\mathbf{Y},\mathbf{V},\mathbf{M},\mathbf{J},\mathbf{N},\mathbf{\Theta})}\\
&\propto \frac{e^{-\frac{W_{t+1}(J_{t+1}^{s,P})^2 - 2A_{t+1}J_{t+1}^{s,P} + G_{t+1}(J_{t+1}^{v,P})^2 - 2F_{t+1}J_{t+1}^{v,P} - 2\frac{\rho(\triangle N_t)^2}{(1-\rho^2)\sigma_v v_t\triangle}J_{t+1}^{s,P}J_{t+1}^{v,P}}{2}}}{e^{-\frac{W_{1,t+1}(J_{t+1}^{s,P} - S_{1,t+1}/W_{1,t+1})^2}{2}}}\\
&\propto e^{-\frac{1}{2}[G_{t+1}(J_{t+1}^{v,P})^2 - 2F_{t+1}J_{t+1}^{v,P} - \frac{(A_{t+1}+B_{t+1}J_{t+1}^{v,P})^2}{W_{1,t+1}}]}\\
&\propto N(\frac{S_{2,t+1}}{W_{2,t+1}}, \frac{1}{W_{2,t+1}})1_{J^{v,P}>0},
\end{aligned}
\tag{86}
$$

where

$$
\begin{aligned}
S_{2,t+1} &= \frac{\triangle N_t}{(1-\rho^2)\sigma_v v_t\triangle}(\frac{D_{t+1}}{\sigma_v} - \rho C_{t+1}) - \frac{1}{\mu_v^P} + \frac{A_{t+1}B_{t+1}}{W_{1,t+1}};\\
W_{2,t+1} &= \frac{(\triangle N_t)^2}{(1-\rho^2)\sigma_v^2 v_t\triangle} - \frac{B_{t+1}^2}{W_{1,t+1}};
\end{aligned}
\tag{87}
$$

- **Posterior for** v_{t+1}. Given the information of variance swap rates, the conditional posterior

for v_{t+1} $(0 < t+1 < T)$ is given by

$$
\begin{aligned}
&p(v_{t+1}|\mathbf{\Theta}, \mathbf{Y}, \mathbf{VS}, \mathbf{J}, \mathbf{M}) \\
&\propto p(VS_{t+2}|v_{t+2}, m_{t+2}, \lambda_{t+2}, J^{v,P}_{t+2}, \triangle N_{t+1}, \mathbf{\Theta}) \\
&\times p(Y_{t+2}, v_{t+2}|\Theta, v_{t+1}, m_{t+1}, J^{s,P}_{t+2}, J^{v,P}_{t+2}, \triangle N_{t+1}) \\
&\times p(Y_{t+1}, v_{t+1}|\Theta, v_t, m_t, J^{s,P}_{t+1}, J^{v,P}_{t+1}, \triangle N_t) \\
&\propto \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+2,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+2,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \\
&\times \frac{1}{v_{t+1}} \exp\{-\frac{[(C_{t+2})^2 - 2\rho C_{t+2}D_{t+2} + (D_{t+2})^2]}{2(1-\rho^2)}\} \times \exp\{-\frac{[-2\rho C_{t+1}D_{t+1} + (D_{t+1})^2]}{2(1-\rho^2)}\},
\end{aligned}
$$

where

$$
\begin{aligned}
C_{t+1} &= \frac{Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v\rho v_t - (e^{\mu_j^Q + \frac{\sigma_j^2}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle - J^{s,P}_{t+1}\triangle N_t}{\sqrt{v_t\triangle}}, \\
D_{t+1} &= \frac{v_{t+1} - v_t - \kappa^P_v(m_t - v_t)\triangle - m_t\gamma_v\sigma_v\triangle - J^{v,P}_{t+1}\triangle N_t}{\sigma_v\sqrt{v_t\triangle}}.
\end{aligned}
$$

In particular, the above posterior only has the second exponential part for $t+1 = T$ because v_T depends only on v_{T-1}, while the posterior of v_0 depends on both v_0 and the first exponential part.

- **Posterior for** m_{t+1}. Given the information of variance swap rates and $\mathbf{V}$, the conditional posterior for m_{t+1} $(0 < t+1 < T)$ is given by

$$
\begin{aligned}
&p(m_{t+1}|\mathbf{\Theta}, \mathbf{Y}, \mathbf{V}, \mathbf{VS}, \mathbf{J}) \\
&\propto p(VS_{t+2}|v_{t+2}, m_{t+2}, \lambda_{t+2}, J^{v,P}_{t+2}, \triangle N_{t+1}, \mathbf{\Theta}) \\
&\times p(Y_{t+2}, v_{t+2}|\Theta, v_{t+1}, m_{t+1}, J^{s,P}_{t+2}, J^{v,P}_{t+2}, \triangle N_{t+1}) \times p(m_{t+2}|\Theta, m_{t+1}) \times p(m_{t+1}|\Theta, m_t)
\end{aligned}
$$

where

$$
p(VS_{t+2}|v_{t+2}, m_{t+2}, \lambda_{t+2}, J^{v,P}_{t+2}, \triangle N_{t+1}, \mathbf{\Theta}) \propto \frac{\exp(-\frac{1}{2}\boldsymbol{\epsilon}'_{t+2,\tau}\Sigma^{-1}\boldsymbol{\epsilon}_{t+2,\tau})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \triangleq \ell(m_{t+1})
$$

and

$$
\begin{aligned}
&p(Y_{t+2}, v_{t+2}|\Theta, v_{t+1}, m_{t+1}, J^{s,P}_{t+2}, J^{v,P}_{t+2}, \triangle N_{t+1}) \times p(m_{t+2}|\Theta, m_{t+1}) \times p(m_{t+1}|\Theta, m_t) \\
&\propto \frac{e^{-\frac{1}{2(1-\rho^2)}(\frac{C^2_{t+2}}{v_{t+1}\triangle} - \frac{2\rho C_{t+2}(D_{t+2}-\kappa^Q_v \triangle m_{t+1})}{\sigma_v v_{t+1}\triangle} + \frac{(D_{t+2}-\kappa^Q_v\triangle m_{t+1})^2}{\sigma^2_v v_{t+1}\triangle})}}{2\pi\sigma_v v_{t+1}\triangle\sqrt{1-\rho^2}} \times \frac{e^{-\frac{[m_{t+2}-(1-\kappa^P_m\triangle)m_{t+1}-\kappa^P_m\theta^P_m\triangle]^2}{2\sigma^2_m m_{t+1}\triangle}}}{\sigma_m\sqrt{m_{t+1}\triangle}} \\
&\times \frac{e^{-\frac{[m_{t+1}-(m_t+\kappa^P_m(\theta^P_m-m_t)\triangle)]^2}{2\sigma^2_m m_t\triangle}}}{\sigma_m\sqrt{m_t\triangle}} 1_{m_{t+1}>0} \\
&\propto e^{-\frac{[\frac{(\kappa^Q_v)^2\triangle}{\sigma^2_v v_{t+1}} m^2_{t+1} - 2(\frac{\kappa^Q_v D_{t+2}}{\sigma^2_v v_{t+1}} - \rho\frac{\kappa^Q_v C_{t+2}}{\sigma_v v_{t+1}})m_{t+1}]}{2(1-\rho^2)}} \times \frac{e^{-\frac{E^2_{t+2}}{2}}}{\sqrt{m_{t+1}}} \times e^{-\frac{m^2_{t+1}-2((m_t+\kappa^P_m(\theta^P_m-m_t)\triangle))m_{t+1}}{\sigma^2_m m_t\triangle}} 1_{m_{t+1}>0} \\
&\propto e^{-\frac{[\frac{(\kappa^Q_v)^2\triangle}{\sigma^2_v v_{t+1}} m^2_{t+1} - 2\frac{\kappa^Q_v}{\sigma_v v_{t+1}}(\frac{D_{t+2}}{\sigma_v} - \rho C_{t+2})m_{t+1}]}{2(1-\rho^2)}} \times \frac{e^{-\frac{E^2_{t+2}}{2}}}{\sqrt{m_{t+1}}} \times e^{-\frac{m^2_{t+1}-2((m_t+\kappa^P_m(\theta^P_m-m_t)\triangle))m_{t+1}}{\sigma^2_m m_t\triangle}} 1_{m_{t+1}>0}
\end{aligned}
$$

where

$$
\begin{aligned}
C_{t+1} &= Y_{t+1} - Y_t - (r + \gamma_s(1-\rho^2)v_t + \gamma_v\rho v_t - (e^{\mu^Q_j + \frac{\sigma^2_j}{2}} - 1)\lambda_t - \frac{1}{2}v_t)\triangle - J^{s,P}_{t+1}\triangle N_t, \\
D_{t+1} &= v_{t+1} - (1-\kappa^P_v\triangle)v_t - J^{v,P}_{t+1}\triangle N_t, \\
E_{t+1} &= \frac{m_{t+1} - (1-\kappa^P_m\triangle)m_t - \kappa^P_m\theta^P_m\triangle}{\sigma_m\sqrt{m_t\triangle}} \\
\kappa^Q_v &= \kappa^P_v + \gamma_v\sigma_v.
\end{aligned}
$$

We then can follow a Metropolis-Hastings algorithm to draw the sample for m_{t+1}.