Unscented Kalman Filter design for option pricing models

Piotr Orłowski

HEC Montréal

piotr.orlowski@hec.ca

July 13, 2018

Abstract

The option implied volatility surface can be succinctly summarised by the prices of certain option portfolios, which can be interpreted as the slope of the cumulant generating function of the underlying log return. In affine jump-diffusion models, the prices of some of such contracts are affine in the latent state variables. On the contrary, the prices of options are not. I show that state recovery with full option datasets, even when employing techniques that account for the non-linearity, is inferior to reduced-dimension filters which take the prices of CGF-slope replicating contracts.

Keywords: affine models, option pricing, filtering, unscented Kalman filter

---

1. Introduction

The most successful option valuation models capture the key features of the data by virtue of intricate specifications of the dynamics of volatility and jump factors which cannot be directly observed (Bates, 2000; Pan, 2002; Eraker et al., 2003; Andersen et al., 2015b; Calvet et al., 2015; Andersen et al., 2015c). Taking such models to the data requires three ingredients: a plausible specification of the statistical dynamics of index returns and latent factors, a valid specification of the corresponding risk premia, and a filtering method, which allows the econometrician or investor to form well-informed guesses about the values of the latent factors, and to ultimately make use of the model. While option prices are very sensitive to, and highly informative of, the values of the underlying factors, they are complicated, non-linear functions thereof. The underlying factor dynamics, to be plausible, necessarily becomes complicated and often includes features such as multiple stochastic volatilities and many types of jumps with time-varying occurrence probabilities. When filtering, one has to account for both the non-linearity (when describing the distribution of

the latent factors given observed data) and the statistical intricacies (when combining the former inference with predictions of future factor dynamics). In consequence, suitable filtering methods are typically complex and computationally expensive. One aspect of these methods that historically garnered less attention is the specification of the observation equations of the filter, i.e. the choice of what observable quantities should be used to uncover the latent factors. With this in mind, in this paper I show in a simulation study how a filtering approach that focuses on the prices of carefully selected option portfolios improves an agent's ability to correctly price options in an empirically relevant setting.

The natural point of reference is a filtering approach which extracts information from all option prices available at each point in time. Its evident implications are that the observed quantities are strongly correlated, and that the econometrician has to specify yet another large stochastic component which represents the error attached to option price observations, thus multiplying room for specification error. I propose the use of filtering equations based on the prices of option portfolios which replicate the risk-neutral variance and appropriately defined risk-neutral skew of the index return. In the class of affine jump diffusion models, their prices are affine functions of the underlying factors, and thus in principle the Kalman filter can be used for inference. Retrieving the risk-neutral moments exactly from the prices of traded options is not possible, however: only a limited and discrete range of strikes is traded. A standard solution to this problem is to inter- and extrapolate option prices appropriately and calculate the risk-neutral quantities, albeit for risk-neutral skewness and higher cumulants the results are very sensitive to modeling choices. I avoid such complications by formulating the observation equations in terms option portfolios which best approximate the risk-neutral quantities. Thus, I further reduce the potential for inadvertently introducing misspecification to the model. Furthermore, this construction of approximate risk-neutral moments retains their interpretation as trading strategies which are available to investors. Maintaining tradability and avoiding interpolation comes at a cost: even though the choses risk-neutral moments are affine in the latent factors, their tradable option approximations are only approximately affine, and the Kalman filter and its simple extension were shown to underperform in such situations (Christoffersen et al., 2014). I use the Unscented Kalman Filter (van der Merwe and Wan, 2001; Christoffersen et al., 2014) to sidestep this problem.

In the simulation study, I investigate the performance of three filtering setups applied to an affine jump diffusion model whose most important feature is a factor specific to the premium for negative jump risk, which is unrelated to variance risk. Such an unspanned skewness factor is necessary for any affine model to fit out-of-the-money put prices. I find that, compared to the reference approach, my specification of the observation equation brings important accuracy gains for the inference about the unspanned skew factor. The improvement is especially large when this unspanned factor takes high values, i.e. when the premium for negative jump risk is high, and put options are especially

expensive. I also show that when only variance swaps prices are used as observable inputs, the dynamics of the unspanned skew factor is recovered very badly. An investor who were to apply such a filter in order to recover state variable values and make predictions (price options) would base their decision on how well her model prices options, rather than on measures of raw filtering accuracy. To this end, I inspect the pricing errors that an investor would make when applying the filtering setups to a model with known parameters. Again, I find that significant reductions in pricing errors over the reference setup are available when using a filter with risk-neutral variance and skew observations. I conclude that augmenting the term structure of variance swaps with other affine contracts retains the simple form of the filtering problem and the information content of the complete implied volatility surface. Furthermore, as opposed to individual options, in the case of affine contracts it is straightforward to identify a contract's loadings on the factors.

While a number of studies limit the observed quantities to the prices of variance swaps (Li and Zinna, 2014; Ait-Sahalia et al., 2015; Fulop and Li, 2015), I show that whenever the specification posits a jump skewness component that is independent of variance dynamics (as in e.g. Andersen et al., 2015d,c, 2017), the quality of the state estimates deteriorates. The study of Feunou and Okou (2018) is the closest to the approach presented in this paper, taking risk-neutral cumulants, calculated from option prices, as observed inputs for model estimation. I too exploit the risk-neutral cumulant generating function, however there are important differences between the construction of the observation equation in my work and in Feunou and Okou. First, Feunou and Okou estimate risk-neutral cumulants by inter- and extrapolating option prices, while I work directly with tradable option portfolios which best approximate the risk-neutral moments with available option data. Second, my observation equations are cast in terms of the prices of the approximating option portfolios themselves, and not in terms of unavailable risk-neutral moments. Thus, while my approach is similar to Feunou and Okou, it has two features which give it potentially more robustness against data quality issues.

The importance of an unspanned skewness factor for modeling index option risk dynamics is underlined by the recent literature. Gruber et al. (2015) identify the unspanned risk factor with the dynamics of short-term implied volatility skew through a stochastic correlation factor between volatility processes and show that it is particularly important in the description of risk premia. Andersen et al. (2015c) find that an extremely $\mathbb{Q}$-persistent, self-exciting tail risk factor is necessary to describe risk premia implied by out-of-the-money put options. Importantly, the factor only drives premia, and is hardly detectable from return (volatility) dynamics. A high-frequency non-parametric analysis of option implied volatilities in Andersen et al. (2015a) finds that markedly different factors impact the pricing of low-struck puts as opposed to at-the-money calls: the former has a jump nature, while the latter is a Gaussian process and is associated with volatility. Calvet et al. (2015) link the jump factor to the dynamics of persistent changes in the level of volatility,

again decoupling the jump intensity from the evolution of spot volatility. Li and Zinna (2014) lean upon the term structure of variance swaps to find that the short end of the volatility risk premium responds strongly to the jump factor whose intensity is, again, independent of Brownian volatility dynamics.

The paper proceeds with a discussion of the properties and the interpretation of the forward-neutral cumulant generating function of log-returns in Section A, followed by the description and interpretation of CGF slope swaps 2. It presents results of simulation studies of various observation equation specifications in filtering in Section 3. Section 4 concludes.

2. Tradable risk-neutral moments

Breeden and Litzenberger (1978) showed that the forward risk-neutral density can be inferred directly from option prices. Equipped with that knowledge one can calculate the risk-neutral expectation – price – of an arbitrary function of the price of the underlying asset at option maturity, as later shown by Carr et al. (2001). The cumulant generating function of the returns is a natural target for such replication. In Appendix A I show that no-arbitrage restrictions on the cumulant generating function endow its first derivative – slope – with interesting properties, and that the slope is an affine function of the state variables in the class of affine jump diffusion models. In this section, I show that certain option-replicated contracts present in the literature correspond to the measurements of the slope of the CGF. Thus, the risk-neutral moments captured via the option-implied CGF can be understood as swap rates for trading strategies. The latter are listed below.

CGF slope swaps are financial contracts whose price (the fixed leg) is defined through equation (11) in Appendix A. A complete specification of the payoff requires recognizing that in Carr et al.'s result the payoff is delta-hedged, and that the frequency of delta-hedging can be increased arbitrarily, as its' a zero-cost operation under the risk-neutral measure. Sufficiently frequent delta hedging allows to interpret the strategy as a measure of realized return variation.

The payoff of a CGF slope swap can be constructed from (11) by taking the term inside the expectation, delta-hedging it, and adjusting the size of the position by the numerator:

$$\frac{\overbrace{\left(\frac{F_{t+\tau}}{F_t}\right)^p \ln\frac{F_{t+\tau}}{F_t} - \frac{(F_{t+\tau}-F_t)}{F_t}}^{\text{option position, non-linear payoff}} - \overbrace{\sum_{j=1}^{N}\left(\frac{1}{F_{t_{j-1}}} - \frac{1}{F_t}\right)\left(F_{t_j} - F_{t_{j-1}}\right)}^{\text{frequent delta-hedging}}}{e^{K_t'(p,\tau)}}. \tag{1}$$

By construction, the forward-neutral expectation of this expression is equal to (11).

Consider an implementation of the strategy with a single hedging step, i.e. with $t_N = t + \tau$ and $N = 1$. For $y := \ln F_{t+\tau}/F_t$ the non-linear component of the payoff in (1) can be written as $ye^{py} - (e^y - 1)$. A Taylor expansion of the payoff at $y = 0$ yields

$$ye^{py} - (e^y - 1) = \sum_{k=1}^{\infty} \left(\frac{p^k}{k!} - \frac{1}{(k+1)!} \right) y^{k+1}. \tag{2}$$

For all $p \neq 1/2$ the leading order of this expansion is y^2, while for $p = 1/2$ the first element vanishes, and the leading order becomes y^3. Whenever $N > 1$ and forward trading is more frequent, with $y_j = \ln F_{t_j}/F_{t_{j-1}}$, it is possible to write (1) in terms of sums of $\phi_p(y_j) - 1 - p(e^{y_j} - 1)$ and of $\phi_p'(y_j) - (e^{y_j} - 1)$, weighted by functions of $F_{t_{j-1}}/F_t$. This coincides with the notion of realized divergence, as defined in Schneider and Trojani (2015): CGF slope swaps are embedded in the family of power divergence swaps and their higher-order descendants. The total payoff in (1) has a well-defined limit for ultra-high frequency trading, that is for $\sup |t_j - t_{j-1}| \to 0$, for every $p \in [0, 1]$. I give the limits for selected ps below.

The contracts of interest, with $p \in \{0, 1/2, 1\}$ are already known in the literature, albeit to the best of my knowledge they have not been interpreted jointly in the context of the CGF. [1] In what follows, I give formulas of terminal payoffs and their continuous-time limits together with leading references.

2.1. $p = 0$: Variance swap (VIX)

The value of the square of the VIX index (CBOE, 2000; Jiang and Tian, 2007) corresponds to $-2\mathbb{E}_t^{\mathbb{Q}}[\ln F_{t+\tau}/F_t]$, i.e. the $p = 0$ slope swap is equivalent to the VIX^2 variance swap, after suitable rescaling and adjustment of the payoff function. The price and payoff of the variance swap are, respectively:

$$\begin{aligned} &-2\mathbb{E}_t^{\mathbb{Q}}[\ln F_{t+\tau}/F_t] \\ &-2\ln F_{t+\tau}/F_t + \sum_{j=1}^{N} \frac{F_{t_j} - F_{t_{j-1}}}{F_{t_{j-1}}} = -2\sum_{j=1}^{N} \ln \frac{F_{t_j}}{F_{t_{j-1}}} - \frac{F_{t_j} - F_{t_{j-1}}}{F_{t_{j-1}}}. \end{aligned}$$

Typically, this swap rate is used in transactions which exchange it for realized variance for returns, $RV_{t,\tau} = \sum_{j=1}^{N} \left(F_{t_j}/F_{t_{j-1}} - 1 \right)^2$. Consider the high frequency limits of the realized variance and the payoff (floating leg) of the $p = 0$ swap as defined above. In the absence of jumps, and after correcting for scale and sign, the two quantities would have exactly the same limit, integrated

[1] Orłowski et al. (2015a) in fact consider $K_t'(p)$ at $p = 0$ and $p = 1$ in the context of pricing kernel dispersion bounds.

variance: $\int_t^{t+\tau} \sigma_s^2 ds$. If jumps are present, this holds only approximately. For the details of the approximation, see Jiang and Tian (2005). The floating leg given in the equation above is the one correctly priced by the VIX^2 index (after rescaling), as shown by Schneider and Trojani (2015). Its continuous-time limit is:

$$\int_t^{t+\tau} \sigma_s^2 ds + 2 \sum_{t<s\leq t+\tau} \phi_0' \left(\frac{F_{t+s}}{F_{t+s-}} \right). \tag{3}$$

2.2. $p = 1$: *Gamma swap*

For $p = 1$ the CGF slope swap corresponds to the Gamma swap (Lee, 2010), also referred to as a Kullback-Leibler swap (Schneider and Trojani, 2015). The price and payoff of are, respectively:

$$2\mathbb{E}_t^{\mathbb{Q}} \left[\frac{F_{t+\tau}}{F_t} \ln F_{t+\tau}/F_t \right]$$
$$2F_{t+\tau}/F_t \ln F_{t+\tau}/F_t - 2\sum_{j=1}^{N} \frac{F_{t_j} - F_{t_{j-1}}}{F_{t_{j-1}}} = \sum_{j=1}^{N} F_{t_j}/F_t \ln F_{t_j}/F_{t_{j-1}} - \frac{F_{t_j} - F_{t_{j-1}}}{F_{t_{j-1}}}.$$

The payoff's continuous-time limit is, after appropriate scaling:

$$2\int_t^{t+\tau} \frac{F_{t+s}}{F_t} \sigma_s^2 ds + 2 \sum_{t<s\leq t+\tau} \frac{F_{t+s-}}{F_t} \phi_1' \left(\frac{F_{t+s}}{F_{t+s-}} \right), \tag{4}$$

that is, it accumulates quadratic variation weighted by the cumulative return on the asset.

2.3. $p = 1/2$: *Hellinger skew swap*

For $p = 1/2$, the CGF slope swap is a rescaled version of the Hellinger skew swap found in Schneider and Trojani (2015), with price and payoff:

$$-\frac{8}{e^{K_t(p,\tau)}} \mathbb{E}_t^{\mathbb{Q}} \left[\left(\frac{F_{t+\tau}}{F_t} \right)^{\frac{1}{2}} \ln F_{t+\tau}/F_t \right],$$
$$-\frac{8}{e^{K_t(p,\tau)}} \sqrt{\frac{F_{t+\tau}}{F_t}} \ln F_{t+\tau}/F_t + \frac{8}{e^{K_t(p,\tau)}} \sum_{j=1}^{N} \frac{F_{t_j} - F_{t_{j-1}}}{F_{t_{j-1}}}$$
$$= -\frac{8}{e^{K_t(p,\tau)}} \sum_{j=1}^{N} \phi_{\frac{1}{2}}' \left(\frac{F_{t_{j-1}}}{F_t} \right) \left(\phi_{\frac{1}{2}} \left(\frac{F_{t_{j-1}}}{F_t} \right) - \frac{F_{t_j} - F_{t_{j-1}}}{2F_{t_{j-1}}} \right) + \phi_{\frac{1}{2}} \left(\frac{F_{t_{j-1}}}{F_t} \right) \left(\phi_{\frac{1}{2}}' \left(\frac{F_{t_{j-1}}}{F_{t_j}} \right) - \frac{F_{t_j} - F_{t_{j-1}}}{2F_{t_{j-1}}} \right)$$

The continuous-time limit of the payoff is:

$$\int_t^{t+\tau} \frac{\phi'_{\frac{1}{2}}\left(\frac{F_{t+s-}}{F_t}\right)}{e^{K_t(p,\tau)}}\sigma_s^2 ds - \frac{8}{e^{K_t(p,\tau)}} \sum_{t\leq s\leq t+\tau} \left[\phi_{\frac{1}{2}}\left(\frac{F_{t+s-}}{F_t}\right)\phi'_{\frac{1}{2}}\left(\frac{F_{t+s-}}{F_{t+s}}\right) + \phi'_{\frac{1}{2}}\left(\frac{F_{t+s-}}{F_t}\right)\left(\phi_{\frac{1}{2}}\left(\frac{F_{t+s-}}{F_{t+s}}\right) - 1\right)\right] \tag{5}$$

The payoff accumulates – in the first term – continuous quadratic variation with a positive or negative weight, depending on whether $F_{t_{j-1}}$ is smaller or greater than the starting value F_t. Additionally, in the second term, it accumulates second-order jump variation, with a similar weighting rule.

2.4. CGF slope swaps in AJDs

Equation (15) is a counterpart of (11) in affine jump diffusion settings. It follows immediately that the prices of CGF slope contracts are affine in the state variables in this class of models.

Proposition 1 (affine contracts)**.** *The price of a CGF slope swap at p is equal to:*

$$\frac{\mathbb{E}_t^{\mathbb{Q}}\left[\phi'_p\left(\frac{F_{t+\tau}}{F_t}\right)\right]}{e^{K_t(p,\tau)}} = \alpha'(p,\tau) + \beta'(p,\tau)\cdot v_t,$$

where the coefficients α', β' are obtained by finding the transform in Proposition 3 in Duffie et al. (2000).

This fact has been known and exploited for long for $p = 0$ (the variance swap). Note that in Section 2 the payoffs were rescaled for an easier interpretation: the CGF slope at $p = 0$ ($p = 1/2$) is, to first order, exposed to the negative of the second (third) moment of returns.

As summarized in Section 1, unspanned skewness (i.e. stochastic skewness that is not driven by factors impacting stochastic volatility) is a primary feature of option implied volatility surfaces which are heavily skewed at the short end. In many dimension-reduced estimation attempts in the literature, the term structure of variance swaps is used, alongside the stock return, as an observable statistic summarising option information. In what follows, I focus on the term structures of the variance, Hellinger skew and Gamma swaps in example models, and analyse the information about state variable values contained therein.

I build example models starting from the SVJCD2 specification in Bates (2000). The base specification is denoted M1. The three remaining specifications modify the base model to introduce

various degrees of unspanned skewness. Model parameters are given in Table 1.

$$\begin{aligned}
\frac{dF_t}{F_{t-}} &= -\bar{\nu}_t dt + \phi_1\sqrt{v_{1t}}dW_{1t}^S + \phi_2\sqrt{v_{2t}}dW_{2t}^S + \int_{\mathbb{R}\setminus\{0\}} \left(e^{g(x,y)}-1\right)\nu_t(dx,dy)dt \\
dv_{1t} &= \kappa_1(\eta_1 - v_{1t})dt + \sigma_1\sqrt{v_{1t}}dW_{1t}^v + \int_{\mathbb{R}_+} y\nu_t(dx)dt \\
dv_{2t} &= \kappa_2(\eta_2 - v_{2t})dt + \sigma_2\sqrt{v_{2t}}dW_{2t}^v \\
dW_{it}^S dW_{it}^v &= \rho_i dt,
\end{aligned}$$

M1 through M3: $\frac{\nu_t(dx,dy)}{dx} = \frac{\lambda_1 v_{1t} + \lambda_2 v_{2t}}{\sqrt{2\pi}} e^{-\frac{(x-\mu_J)^2}{2\sigma_J^2}}$, while in M4: $\frac{\nu_t(dx,dy)}{dxdy} = \frac{\lambda_1 v_{1t}}{\rho_J \mu_J} e^{\frac{y}{\rho_J \mu_J}}$

name	M1	M2	M3	M4
η_1	0.0148	0.0148	0.0148	0.0148
κ_1	3.21	3.21	3.21	3.21
σ_1	0.24	0.24	0.24	0.24
ρ_1	-1	–	–	–
η_2	0.0195	0.0195	0.0195	0.0195
κ_2	0.85	0.85	0.85	0.85
σ_2	0.18	0.18	0.18	0.18
ρ_2	-0.314	0	-0.314	-0.314
λ_1	88.6	88.6	60.6	28.6
λ_2	–	–	15.2	–
μ_{JV}	–	–	–	0.05
μ_J	-0.054	-0.054	-0.054	–
ρ_{JV}	–	–	–	-0.99
σ_J	0.102	0.102	0.102	–
$\kappa_1^{\mathbb{Q}}$	2	2	2	2
$\kappa_2^{\mathbb{Q}}$	0.87	0.87	0.87	0.87
ϕ_1	1	0	0	0
ϕ_2	1	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$

Table 1: Example model parameters

In the base specification (M1), both variance factors impact the level of the Brownian volatility of the stock, but only v_1 drives the jump intensity. Simultaneously, the innovations to both factors are negatively correlated with corresponding innovations to the forward stock price. Under the parameter choices, in all models the (primary) jump-driving factor v_1 mean-reverts quicker than v_2. In specifications M1 through M3, the jump component only impacts returns, and apart from a negative mean, the jumps have symmetric Gaussian tails. In specification M4, jumps in factor v_1 are exponential and feed into the stock price process. In the supplementary specification M2:

- I remove the intensity driver's impact on Brownian asset volatility, setting $\lambda_1 = 0$;

- remove v_2's leverage effect, setting $\rho_2 = 0$;
- I increase the second factor's impact on Brownian volatility, setting $\phi_2 = \sqrt{2}$

. Specification M3 retains the changes from M2 and:

- I set $\lambda_2 = 15.2$ so that short-term skewness is influenced by both factors, but Brownian volatility only by v_2.

Specification M4 retains the features of M2, but alters the jump distribution:

- the asset price and v_1 co-jump, and the jumps are exponentially distributed self-exciting, i.e. $\lambda_1 = 28.6$, $\mu_{JV} = 0.05$, and the jump leverage parameter is set to $\rho_J = -0.99$;

Additionally, in specifications M2 through M4 I increase the remaining factor's contribution to Brownian volatility so as to achieve a similar level of total return variance as in M1.

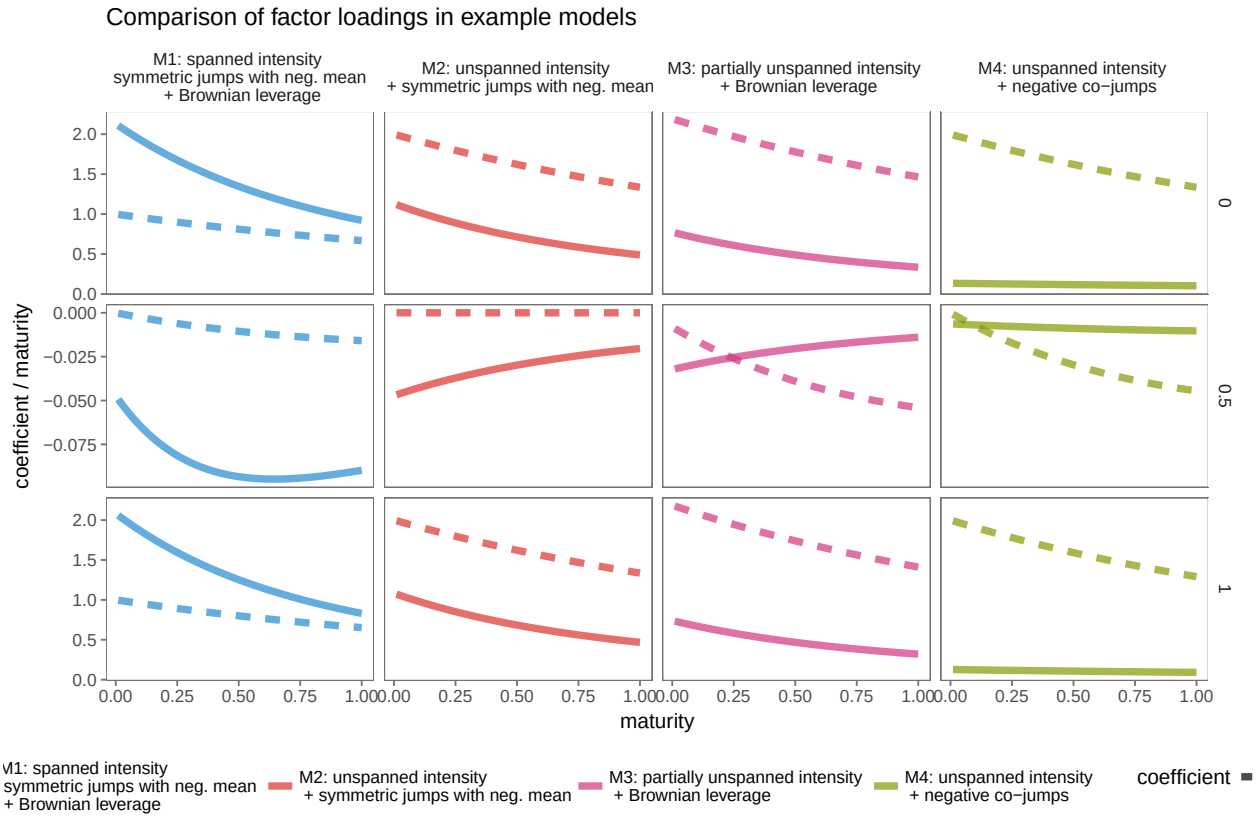

Fig. 1. Scaled factor loadings in affine contracts, $\beta_j'(p, \tau|v_t)/\tau$ against time to maturity in two-factor models.

Figure 1 contains plots of (maturity-scaled) factor loadings in affine contracts $\beta_j'(p, \tau)/\tau$, as given in Proposition 1. Loadings on the VIX variance swap ($p = 0$), the Hellinger skewness swap ($p = 1/2$), and the Gamma swap ($p = 1$) (top-to-bottom) in the aforementioned models (left-to-right) are

compared. Loadings on v_1 – unspanned by volatility dynamics in models M2 throgh M4 – are given in a solid line; on v_2 (volatility driver) in a dashed line.

While the plot is not a study of a general case, the changes in specification are rich enough to generate a diverse range of loading term structure behaviours. Note that the loadings of the variance and Gamma swap are almost identical in all models. This shifts the focus to the differences in variance and Hellinger swaps. The relative loadings of the variance swap rates vary little across the term structure, especially in models with unspanned intensity of symmetric jumps (M1 and M2), where the loadings are approximately collinear; their shapes are similar, too, across models. To the contrary, the value of loadings for the Hellingers skew swap vary with maturity in significantly different ways than loadings of variance swaps. Across models, the skew swap loadings also vary significantly more than the variance swap loadings. The Hellinger swap loadings' common feature is that at short maturites, the loadings on factor v_2 converge to 0 whenever it does not drive jump intensity. If additionally this factor does not generate a leverage effect, the price of the Hellinger swap does not depend on it at all. Finally, model M4 shows an example where the factor v_1 that drives asymmetric, self-exciting co-jumps, is absent from the term structure of variance swap prices.

Ex-ante, the econometrician does not know the characterization of factor loadings in the data. With recent empirical evidence pointing to the importance of unspanned skewness, the takeaway of this section is that using short-term Hellinger skew swaps in estimation data offers new statistics about latent volatility and skewness drivers, with richer information content than the term structure of variance swaps. Furthermore, in certain extreme situations, the variance swap term structure can have a loading of essentially 0 on the isolated skewness driver, and Helliger skew swaps become essential for uncovering its $\mathbb{Q}$-dynamics.

2.5. Spanning in option markets

Carr et al. (2001) showed that in the presence of complete option markets, that is if European options on the underlying are available for strikes $K \in [0, \infty]$, an arbitrary payoff function $\pi(x)$, non-linear in the forward return, is available for trading in the following sense:

$$\pi(F_{t+\tau}) - \pi(F_t) - \pi'(F_t)(F_{t+\tau} - F_t) = \int_0^\infty \pi''(K_t) O(K_t, F_{t+\tau}) dK_t. \tag{6}$$

In the equation above, $O(K_t, F_{t+\tau})$ is the payoff, at time $t+\tau$, of a maturing option. The option is a call (put) if at time t, K_t was greater (smaller) than F_t (i.e. it's an out-of-the-money option). The equation states that the delta-hedged increment [2] of function π can be obtained as a payoff of

[2] Bregman (1967) divergence, as noted by Schneider and Trojani (2015)

a portfolio of options at their maturity.

2.5.1. Incomplete markets

Equation (6) is an unobtainable idealization. Investors in real markets face two impediments: the option strike grid is discretized, and (more importantly), bounded. Their impact on the ability to hedge non-linear contracts is two-fold: discretization causes inaccuracies in the non-linear hedge if the final asset value stays within the strike grid, while boundedness limits the ability to hedge extreme events. Boundedness also implies that the replicating portfolio's price is only an approximation of the true contract price, and the error uncertainty is unaccounted for. Such extreme events, however, happen extremely rarely, as discussed in the appendix to Martin (2013) and by Orłowski et al. (2015b). Martin also discusses the underpricing for variance contracts in the context of an implementation of 6 by simple discretization. Orłowski et al. (2015b) shows that there are better alternatives, and optimal static replications can be chosen at any point in time, as to minimize expected hedging error. I use this method for calculating contract swap prices in the subsequent simulation exercises.

The model-based theoretical quantities are thus approximated as values of option portfolios):

$$K'_t(p,\tau) = \alpha'(p,\tau) + \beta'(p,\tau) \cdot v_t \approx \sum_{j=1}^{N} w_j^* O_\tau(K_{t,j}), \tag{7}$$

where w_j^* are portfolio weights that minimize expected hedging error. The choice of method for estimating $K'_t(p,\tau)$ from option data can have important implications for model estimation results. If a method such as (7) is chosen, the approximate contract prices are no longer linear in state variables. For certain ranges of factor values and availability of option contracts, the estimates can vary wildly from their theoretical CGF slope counterparts. While Martin (2013) discusses possible biases when estimating the prices of variance swaps and concludes they would be small, they might be much more significant for Hellinger skew swap contracts, which load much stronger on far out-of-the-money options than variance contracts. This is illustrated in Figure 2, which plots the relative approximation error of the (7) method. Other methods of estimating $K'_t(p,\tau)$ involve approximating the integral in (6) through estimating the implied volatility curve (and thus option prices), and extrapolating it both towards 0 and ∞. Such approaches introduce biases which are difficult to quantify and account for, as they are equivalent to making a statement about the specification of the jump distribution.

I account for potential biases in the filtering and estimation by constructing the filter directly on

replicating portfolios,[3] rather than on model-implied theoretcial prices of variance and Hellinger skew swaps. I find that the departures from linearity are small enough that the Unscented Kalman Filter (Wan and van der Merwe, 2000; van der Merwe and Wan, 2001) is a formidable tool for the task at hand (for applications in Finance, see for example **?** and Christoffersen et al. (2014)).

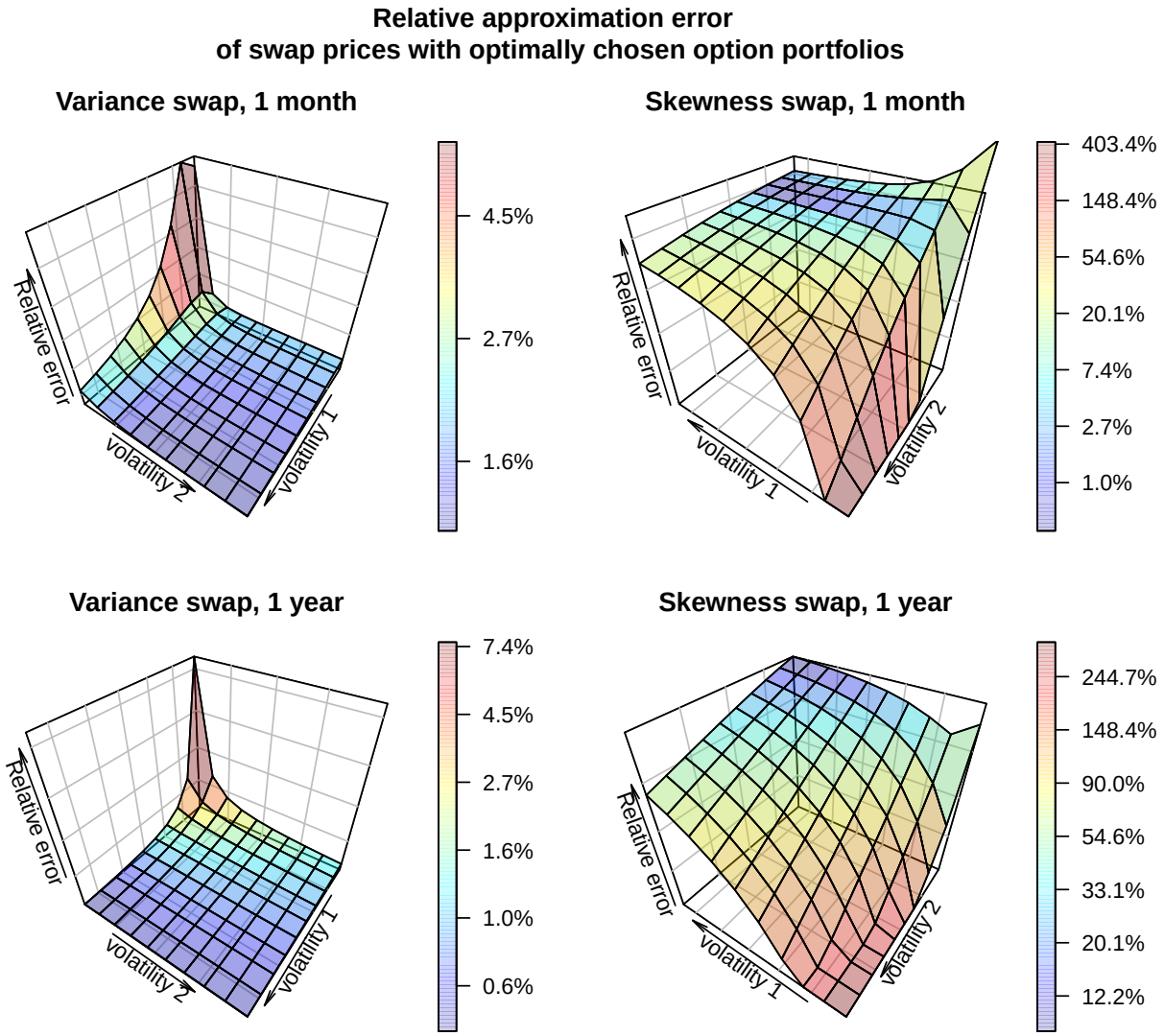

Fig. 2. Model M4: relative approximation error $100\%times\left(\frac{p_{\text{true}}-p_{\text{portfolio}}}{p_{\text{true}}}\right)$ of variance and skewness swap prices with option portfolios chosen with the method described in Orłowski et al. (2015b). Curvature in plotted surfaces represents departures from linearity of approximate contract price with respect to the factors. Extreme overpricing of skewness swaps is possible at very low values of factor v_1.

[3] I find the prices of the replicating portfolios in the model in the follwing sense: if in the option data, J strikes $\{K^*_{tj}\}_1^J$ are available, and the optimal replicating portfolio for a given swap rate at time t assigns weights $\{w^*_{tj}\}_1^J$ to the options, then in the filter I construct the model-implied price of the portfolio by finding the prices of options with the given strikes, and forming the portfolio with the optimal weights estimated in the data.

3. Simulation study

In the simulation study I focus on the consequences of estimation dataset choice for state variable recovery. I look at three aspects of the results. First, I examine the filtering error directly. The error feeds into further quantities of interest, such as factor premia and option prices. Thus, I translate the filtering error into option pricing error. This setup abstracts from parameter estimation error and allows to quantify signal fidelity losses that are solely due to factors under the econometrician's control, i.e. the composition of the dataset, while keeping the filtering method constant.

I scrutinize models M1 and M4, with specification as in Section 2.4 and parameters as in Table 1. The three datasets for filtering are:

- **opts**: for each date, the dataset contains the prices of 30 options per maturity, at 1- and 12-month maturities;
- **vs**: for each date, the dataset contains the prices of the 1- and 12-month variance swaps;
- **hs+vs**: for each date, the dataset contains the prices of the 1-month Hellinger skew swap and 12-month variance swap.

The prices of Hellinger skew and variance swaps in the two reduced-dimension datasets are calculated with the options available in dataset **opts** and in line with remarks in Section 2.5.1. In each dataset I assume that the signal-to-noise ratio is known, and I set the standard deviation of noise to 10% of a contract price's standard deviation. The noise is independent across contracts (be it swaps or options). Note that this assumption puts the reduced-dimension filters at a disadvantage, because independent noise in individual options would likely largely cancel out in portfolio formation.

3.1. State filtering

Two of the dataset setups – **vs** and **hs+vs** – reduce the option surface observed at each time point to a set of summary statistics. **Opts**, on the other hand, uses the complete price set for the same maturities. All datasets are noisy. I demonstrate, with the use of simulations, that the **hs+vs** dataset is no worse a base for the purpose of state recovery than the complete **opts** dataset.

Filtered factor values for models M1 and M4 are plotted in Figures 3 and 5, respectively. In M1, both factors are recovered with similar accuracy when using the **hs+vs** and **opts** datasets. With **vs**, recovery significantly deteriorates. A summary of filtering error distribution for model M1 is given in Figure 4. For factor v_1, which drives both volatility and jump intensity, recovery with

hs+vs and **opts** is approximately on par, whereas with **vs** the interquantile range of recovery error increases two-fold. For factor v_2, the **opts** dataset outperforms both **hs+vs** and **vs**, albeit **hs+vs** still reduces the interquantile range of recovery error by approximately 50% with respect to focusing on variance swaps. Model M4 is of particular interest – in this setup, v_1 drives strictly negative jumps the asset price, which correspond to jumps in volatility, as inspired by the recent empirical literature (**?**). For v_2 the results are similar to those in model M1. Filtering v_1, on the other hand, poses an insurmountable challenge to a filter based on **vs**. It also suffers with the **opts** dataset. The path of v_1 filtered from **vs** seems to bear little relation to the true path, with a correlation coefficient of approximately 30%. State recovery with this dataset is equally bad across all factor levels. The **opts** based filter correctly recovers the v_1 path when v_1 does not stray far from its long-run mean, however the filtered values are extremely noisy after v_1 jumps, when its value is very high. To the contrary, the filter based on **hs+vs** does not suffer from such extreme errors. Figure 6 plots filtering errors in model M4, stratified by the respective factor quantiles. When factor v_1 is below its 75th percentile, filtering with both the **opts** and **hs+vs** dataset yields reasonable results. When v_1 values rise, however, filtering with the whole option surface yields very noisy estimates – the interquantile range is equal to that for the variance swap dataset, and more extreme errors are recorded; the estimates seem nonetheless unbiased, contrary to those recovered with **vs**.

Across the whole paths of v_1 and v_2, using the **hs+vs** dataset reduces mean absolute filtering error by 57% and 44%, respectively, in model M1, and by 80% and 23% in model M4. In the absence of unspanned skewness (M1) the **hs+vs** dataset offers filtering performance equivalent to that of the full option panel in **opts**: it improves the recovery of v_1 by 20%, but hinders the recovery of v_2 by 23%. In the setting with significant unspanned skewness the **hs+vs** dataset, which compactifies the information in the volatility surface into two informative signals, significantly improves state recovery not only with respect to the usual **vs** approach, but also with respect to the whole implied volatility surface, the unfavourable observation noise assumption notwithstanding. Absolute filtering errors of v_1 and v_2 are reduced by 55% and 13%, respectively.

3.2. Option pricing

Section 3.1 gives a good indication of filtering performance gains when Hellinger skew swaps are introduced to the dataset. Filtering quality translates directly to a model's utility for inference about quantities such as option prices, and the importance of any state recovery improvements should be judged through the lens of the task at hand. From the point of view of an econometrician who estimates an asset pricing model, there are three sources of error in estimates of option prices: the filtering error, estimation error (understood as the distance between the estimated and (pseudo-

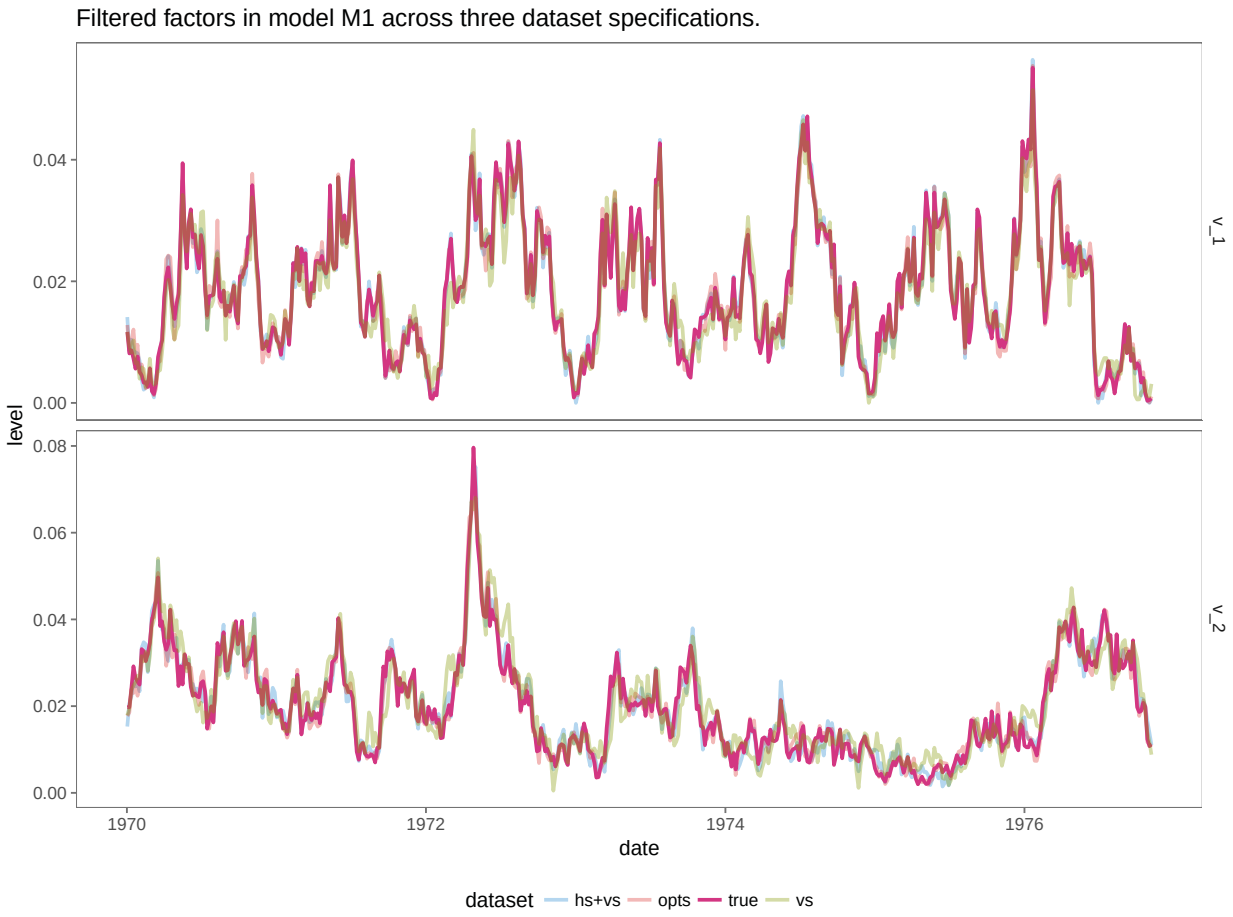

Fig. 3. Filtered factor paths in model M1 with three dataset setups: full option surface, **opts**; variance swaps, **vs**; Hellinger skew and variance swaps, **hs+vs**. For each factor, errors are grouped by the quartile of the factor's level.

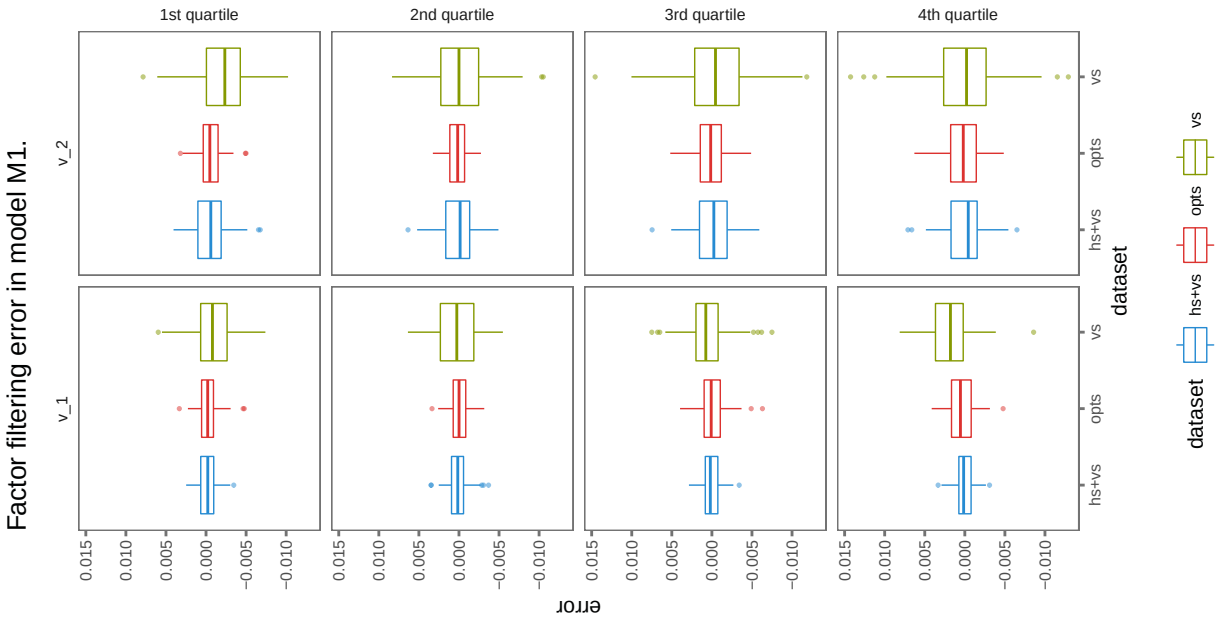

Fig. 4. Factor filtering error in model M1 with three three dataset setups: full option surface, **opts**; variance swaps, **vs**; Hellinger skew and variance swaps, **hs+vs**.

)true parameter values), and model specification error. I concentrate on the former: in models with known parameters, similarly to the previous section, I investigate the impact of the filtering dataset specification on estimated option prices. I judge the setups in terms of relative error of an option's implied volatility,

$$\epsilon_t(k,\tau) := \frac{IV_{\text{true}}(k,\tau) - IV_{\text{filtered}}(k,\tau)}{IV_{\text{true}}(k,\tau)}. \tag{8}$$

The option panels used for evaluation span maturities from 1 week to 1 year, with 30 options per maturity. The state variables are filtered from dataset which contain options with only two maturities: 1 month and 1 year. I include the single-week maturing options, because of the rising importance of the short-term options market, and because they are the most informative of jump risk determinants. Andersen et al. (2016) showed that the dynamics of "weeklies'" prices is indeed highly informative of tail risks prevailing in the index option market. I visualize the distributions of relative pricing error, and the values of mean absolute pricing error in Figures 7 through 10.

Distributions of option pricing errors are displayed in Figures 7 (model M1) and 8 (model M4). The options are split into categories across three dimensions: option maturity, option moneyness level (defined as $\frac{\ln K/F}{\sqrt{\tau} IV_{ATM}}$), and the quartiles of prevailing implied volatility skew for options of given maturity. Options are grouped in moneyness buckets in the following manner: the label denotes that the bucket contains options with moneyness levels smaller than the label, i.e. for 0 the bucket contains puts with moneyness levels $(-2, 0]$. The simulated option surfaces contain more puts than calls, as is typical in observed panels. Implied volatility skew is calculated as the difference between

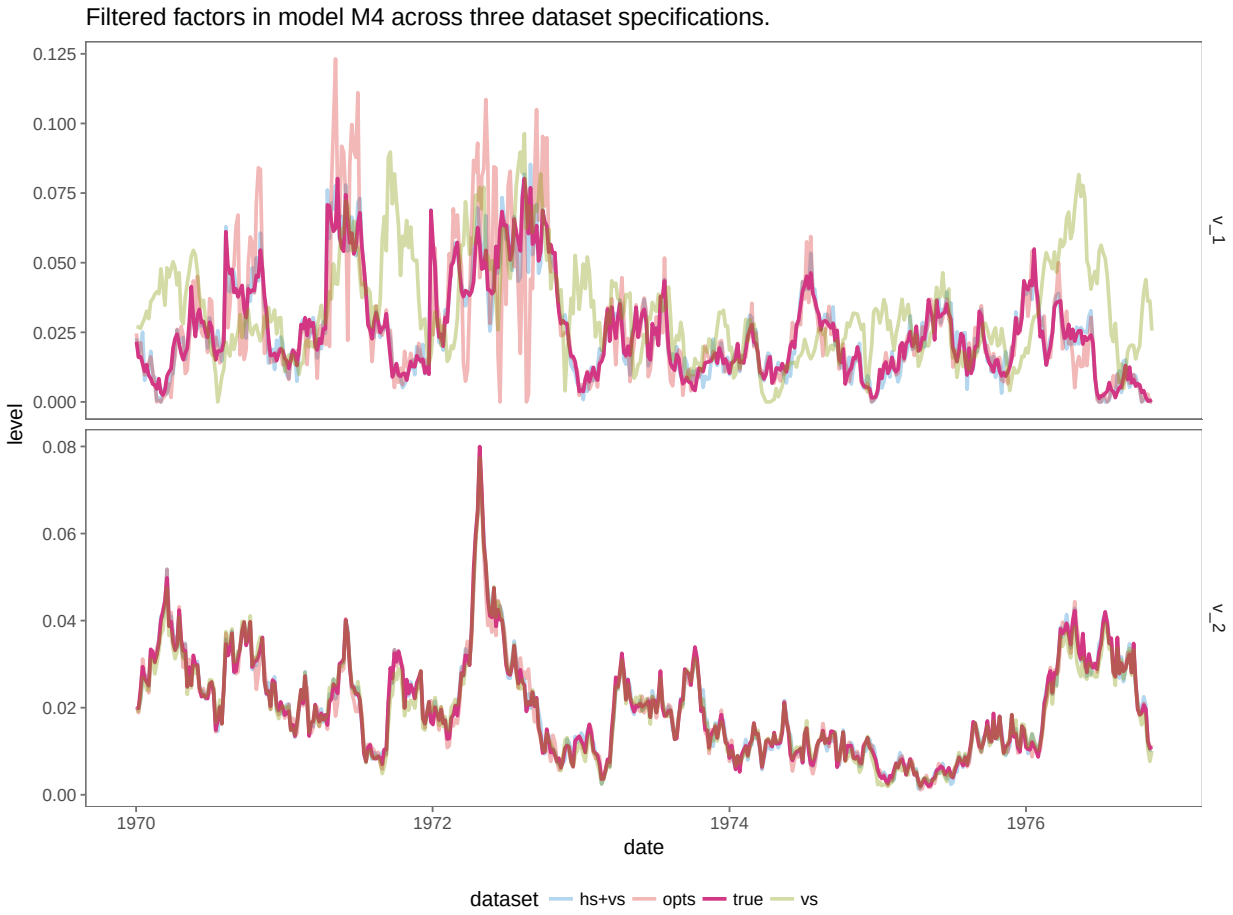

Fig. 5. Filtered factor paths in model M4 with three dataset setups: full option surface, **opts**; variance swaps, **vs**; Hellinger skew and variance swaps, **hs+vs**.

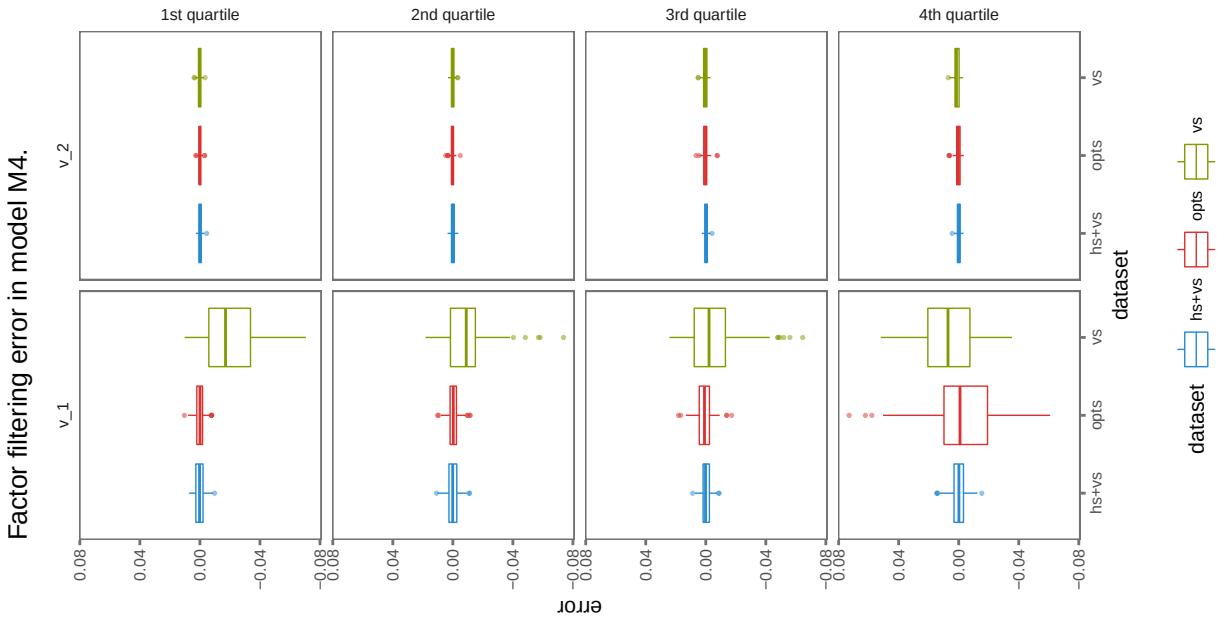

Fig. 6. Factor filtering error in model M4 with three three dataset setups: full option surface, **opts**; variance swaps, **vs**; Hellinger skew and variance swaps, **hs+vs**. For each factor, errors are grouped by the quartile of the factor's level.

median implied volatility of options in the $(-4, -2]$ moneyness bucket and at the money implied volatility, for every time stamp and maturity. In model M1, implied skew is strongly positively correlated with factor v_1 across all maturities and moneyness groups (correlations ranging from 0.8 to 0.95), and not correlated with factor v_2 (correlations ranging from -0.2 to 0.2). In model M4 the implied volatility skew is less correlated with factor v_1, even if this factor is the primary driver of time variation in the third moment of returns (correlations range between 0.5 and 0.8). Factor v_2, on the other hand, is strongly negative correlated with the IV skew (-0.8 to -0.5).

There are important differences in the range of pricing errors resulting from modifying the filtering dataset. Certain regularities are clearly visible for both model parametrizations: pricing errors decrease with option maturity (due to factor mean-reversion: the present factor value has little impact on long-maturity options), and for implied volatility skew levels above the median and maturities above 1 month, call options and at the money put options offer a greater challenge than out-of-the-money put options. Most importantly, for most maturities and implied volatility skew levels, the **vs** dataset yields the noisiest option price estimates. Judged in terms on mean absolute percentage pricing error (MAPE), the **vs** dataset is the worst perfomer across the board, as seen in Figures 9 and 10.

The performance differences of filters based on **hs+vs** and **opts** datasets are more nuanced. In terms of MAPE (Figures 9 and 10), in model M4 the **hs+vs** dataset offers an improvement over

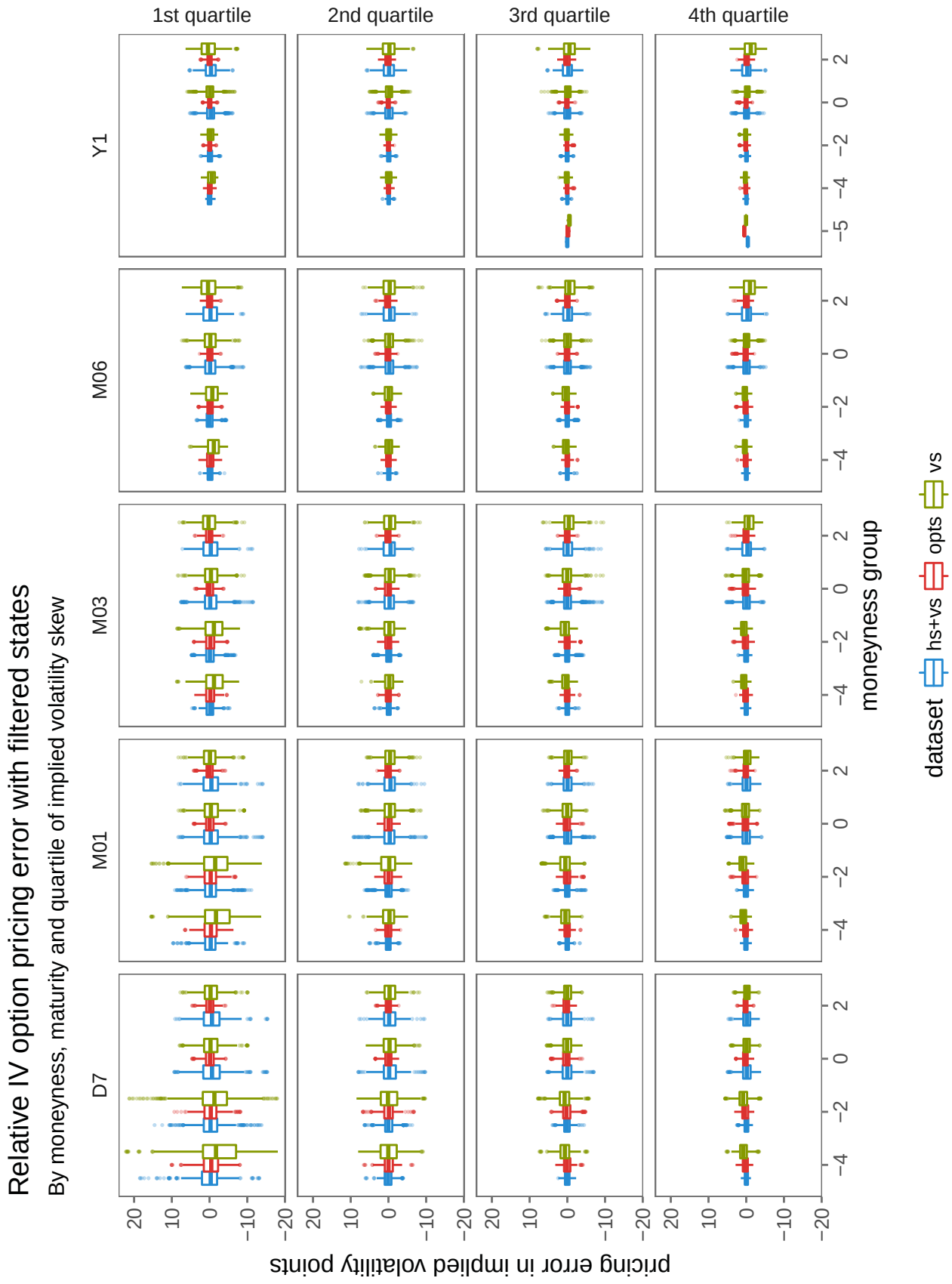

Fig. 7. Option pricing error in model M1 when parameters are known, but state variables are filtered. Error is expressed in percentage of a given option's true implied volatility, $\frac{IV_{\text{true}}(k,\tau)-IV_{\text{filtered}}(k,\tau)}{IV_{\text{true}}(k,\tau)}$. Pricing errors are grouped by option moneyness levels defined as $k = \frac{\ln K/F}{\sqrt{\tau}IV_{\text{ATM}}}$, and by the level of implied volatility skewness.

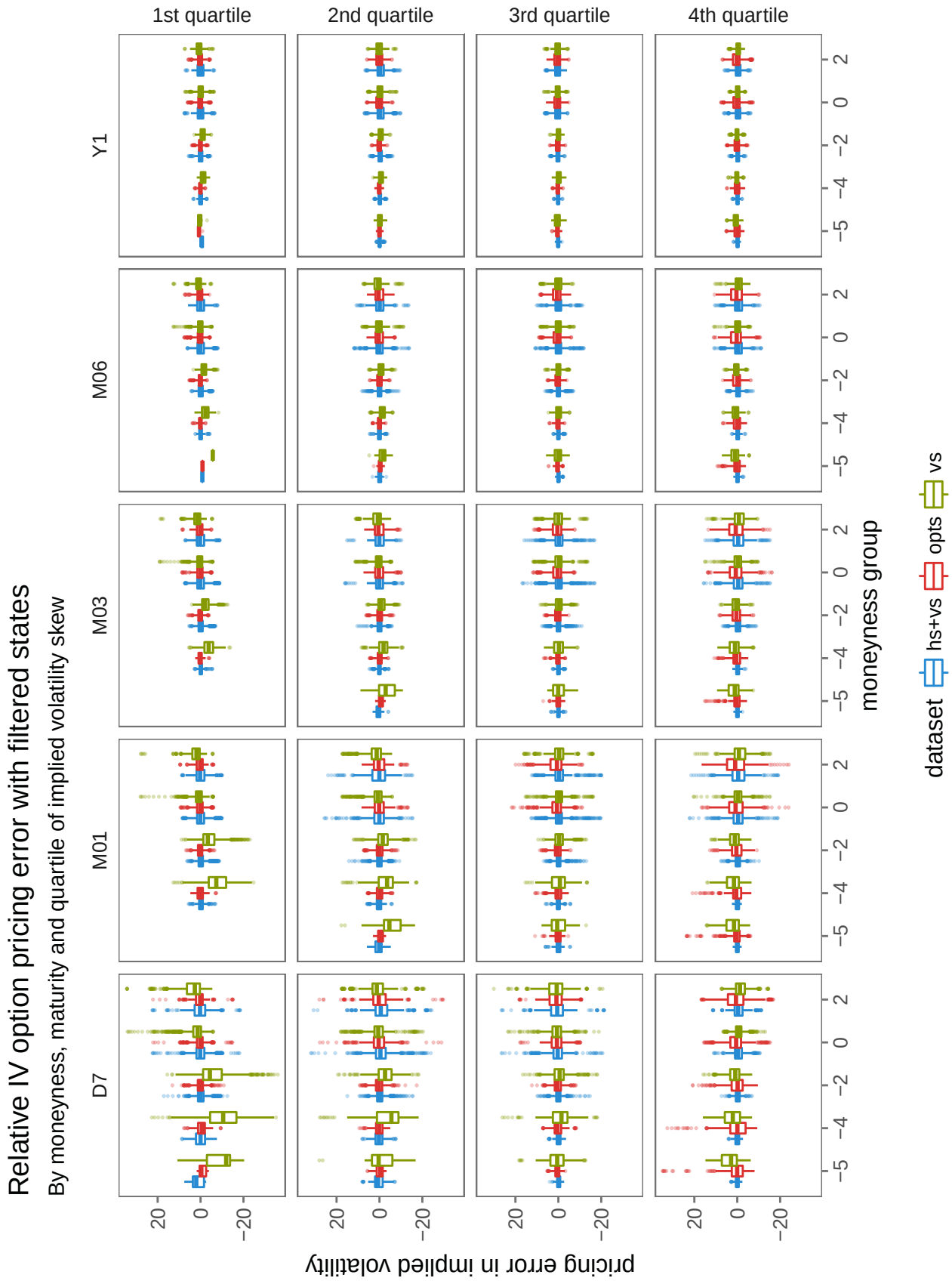

Fig. 8. Option pricing error in model M4 when parameters are known, but state variables are filtered. Error is expressed in percentage of a given option's true implied volatility, $\frac{IV_{\text{true}}(k,\tau)-IV_{\text{filtered}}(k,\tau)}{IV_{\text{true}}(k,\tau)}$. Pricing errors are grouped by option moneyness levels defined as $k = \frac{\ln K/F}{\sqrt{\tau}IV_{\text{ATM}}}$, and by the level of implied volatility skewness.

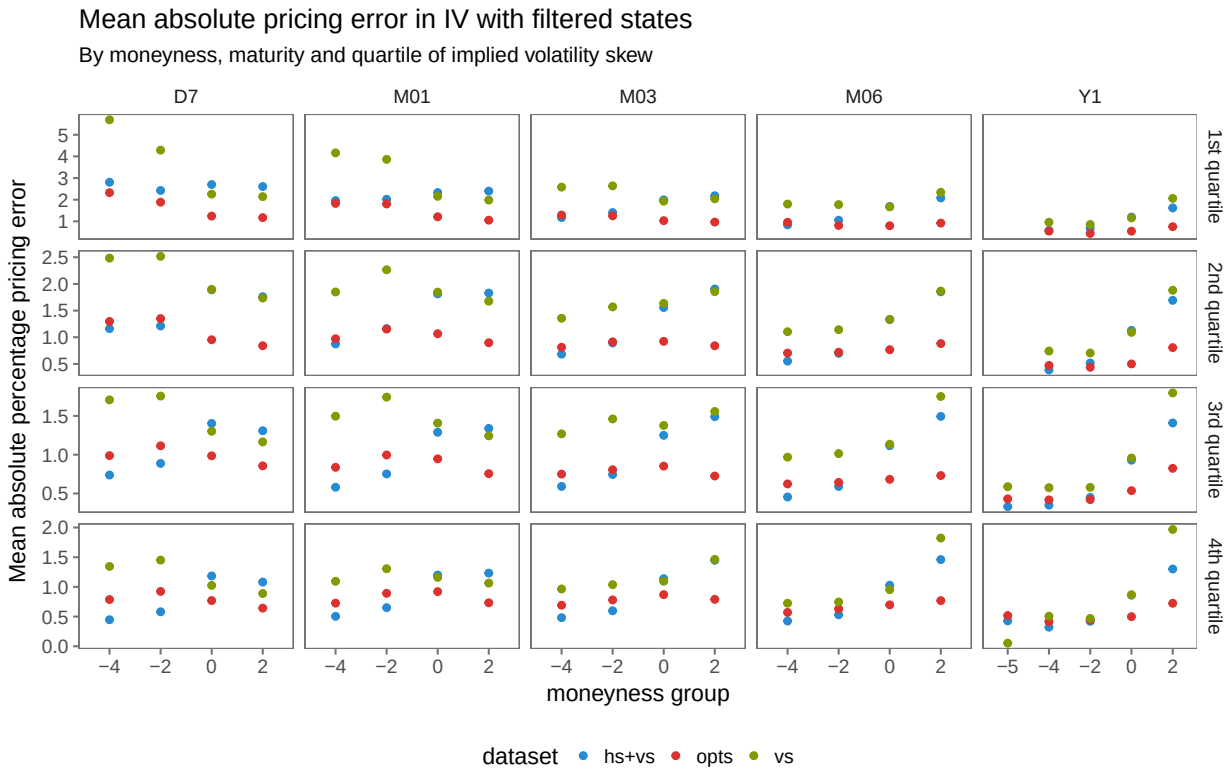

Fig. 9. Mean absolute percentage pricing error in model M1 when parameters are known, but state variables are filtered. Error is expressed in percentage of a given option's true implied volatility, $\frac{IV_{\text{true}}(k,\tau)-IV_{\text{filtered}}(k,\tau)}{IV_{\text{true}}(k,\tau)}$. Pricing errors are grouped by option moneyness levels defined as $k = \frac{\ln K/F}{\sqrt{\tau}IV_{\text{ATM}}}$, and by the level of implied volatility skewness.

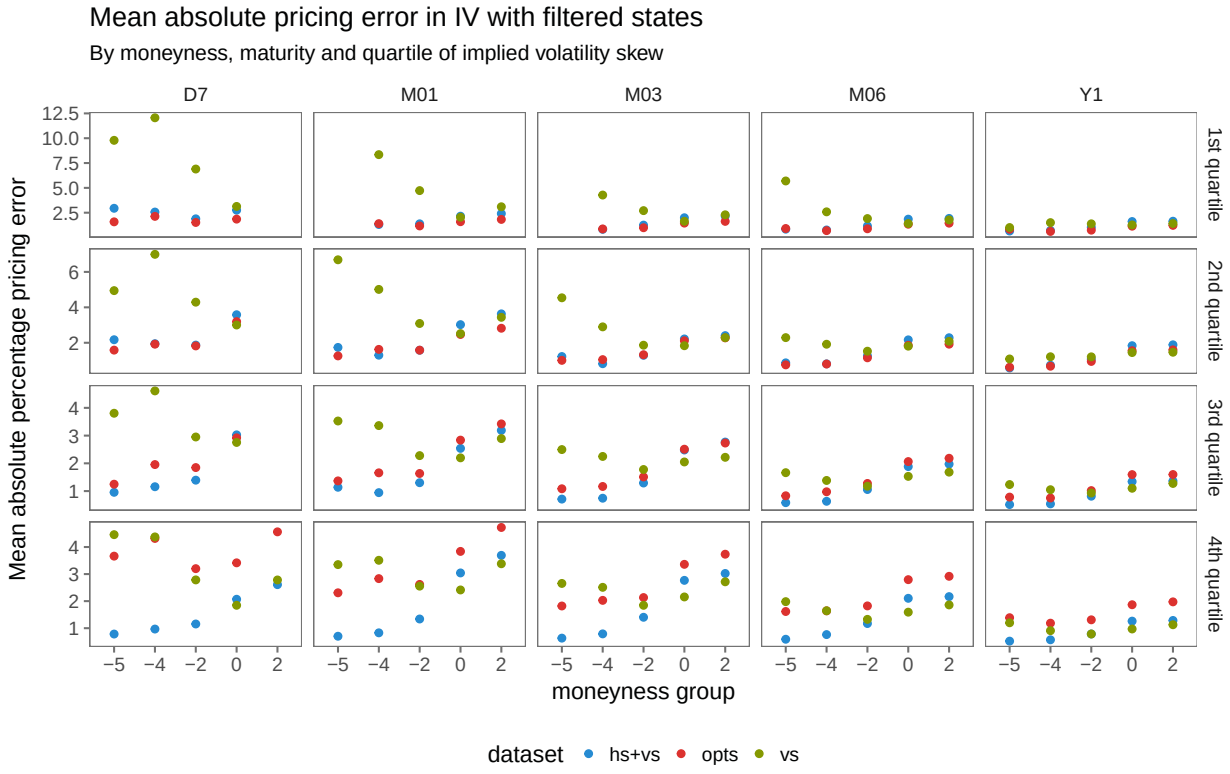

Fig. 10. Mean absolute percentage pricing error in model M4 when parameters are known, but state variables are filtered. Error is expressed in percentage of a given option's true implied volatility, $\frac{IV_{\text{true}}(k,\tau)-IV_{\text{filtered}}(k,\tau)}{IV_{\text{true}}(k,\tau)}$. Pricing errors are grouped by option moneyness levels defined as $k = \frac{\ln K/F}{\sqrt{\tau}IV_{\text{ATM}}}$, and by the level of implied volatility skewness.

opts, with MAPEs of 1.76% and 1.94%, respectively. In model M1, the **opts** dataset retains an advantage, with a MAPE of 0.87% vs **hs+vs**'s 1.22%. Performance gains in model M4 are most significant when the implied volatility skew is above its 75th percentile, for put options with maturities up to 3 months. Interquantile ranges of pricing errors are reduced by 50%, and outlying pricing errors are practically eliminated. The filter based on the **opts** dataset performs better when pricing call options, which are practically not impacted by the dynamics of factor v_1 in model M4.

As I argued in the introduction, recent literature implies that a model specification in the spirit of M4 is more relevant for empirical research than such as M1. A filter design which replaces short-maturity variance swaps with Hellinger skew swaps offers an improved identification of the path of a skewness-driving factor in a model such as M4, not only with respect to a variance-swap based filter, but also with respect to employing the whole option surface. The improvements in filtering translate into smaller option pricing errors when the model parameters are known. While a general recommendation for filter design is beyond the scope of this study, I present preliminary evidence that a dimension-reduced filter can perform on par or better than a more computationally involved alternative. The discussion in this paper also abstracts from the specifications of observation noise. It is self-evident that proposing an observation noise structure for an option panel with hundreds of observed assets opens more room for error than specifying one for a small number of swap rate contracts.

4. Conclusions

The econometrician who faces the task of estimating an option pricing model when state variables are unobservable, must specify the equations of the no-arbitrage pricing model, and the blueprint for the filter to uncover the latent variables. Assuming a given filtering method (a particle filter, or an UKF, as in this paper), the econometrician's task is limited to choosing the right observed quantities, which will yield good state recovery properties. This is much akin to the problem of sensor placement in the control literature, e.g. Tzoumas et al. (2016): using signals from more sensors is not always optimal; a sensor's signal to noise ratio changes over time, and when it is low, the sensor should be excluded from the observation set. Historically, the empirical option pricing literature resolved to use either all sensors – complete option panels, with multiple maturities and hundreds of options per each data point – or certain linear combinations of sensor signal – prices of variance swaps, which can be interpreted as portfolios of options.

I demonstrated that the variance swap observation design performs significantly worse in terms of both state filtering and, as a consequence, option pricing with filtered states, than the complete option panel dataset. In a next step, I showed how replacing short-term variance swaps with Hellinger

skew contracts in filter design makes the reduced-dimension filter competitive with one based on the whole option surface. In a simulated setting with rich volatility and jump intensity dynamics, which mimics features of factor dynamics uncovered by the recent literature, this augmented filtering system outperforms a filter based on the complete option surface. Furthermore, I characterized the variance swap and Hellinger skew contracts in the context of a broader class of CGF slope swaps. The swap rates in this class of contracts are linear in the unobserved latent states when affine jump diffusion dynamics is assumed.

The question of optimal filter design in option pricing model estimation remains open. I leave to extensions of this paper an empirical application of hereby developed methods and the problem of state-dependent observation equation design.

References

Ait-Sahalia, Y., M. Karaman, and L. Mancini

2015. The Term Structure of Variance Swaps and Risk Premia. *Ssrn*.

Andersen, T. G., O. Bondarenko, V. Todorov, and G. Tauchen

2015a. The fine structure of equity-index option dynamics. *Journal of Econometrics*, 187(2):532–546.

Andersen, T. G., N. Fusari, and V. Todorov

2015b. Parametric Inference and Dynamic State Recovery From Option Panels. *Econometrica*, 83(3):1081–1145.

Andersen, T. G., N. Fusari, and V. Todorov

2015c. Parametric Inference and Dynamic State Recovery From Option Panels. *Econometrica*, 83(3):1081–1145.

Andersen, T. G., N. Fusari, and V. Todorov

2016. The Pricing of Short-Term market Risk: Evidence from Weekly Options. *Journal of Finance*.

Andersen, T. G., N. Fusari, and V. Todorov

2017. Short-Term Market Risks Implied by Weekly Options. *Journal of Finance*.

Andersen, T. G., V. Todorov, and N. Fusari

2015d. The risk premia embedded in index options. *Journal of Financial Economics*, 117(3):558–584.

Bates, D.

1993. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in thePHLX Deutschemark Options. Technical Report 4596, National Bureau of Economic Research, Cambridge, MA.

Bates, D. S.

2000. Post-'87 Crash Fears in S&P500 Futures Options. *Journal of Econometrics*, 94(1-2):181–238.

Breeden, D. T. and R. H. Litzenberger

1978. Prices of State-contingent Claims Implicit in Option Prices. *The Journal of Bussiness*, 51(4):621–651.

Bregman, L. M.

1967. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *{USSR} Computational Mathematics and Mathematical Physics*, 7(3):200–217.

Calvet, L. E., M. Fearnley, A. J. Fisher, and M. Leippold

2015. What is beneath the surface? Option pricing with multifrequency latent states. *Journal of Econometrics*, 187(2):498–511.

Carr, P., X. Jin, and D. B. Madan

2001. Optimal investment in derivative securities. *Quantitative Finance*, 59(February 2000):33–59.

Carr, P. and R. Lee

2009. Put-call symmetry: Extensions and applications. *Mathematical Finance*, 19(4):523–560.

CBOE

2000. The CBOE Volatility Index – VIX. Technical report, CBOE.

Christoffersen, P., C. Dorion, K. Jacobs, and L. Karoui

2014. Nonlinear Kalman Filtering in Affine Term Structure Models. *Management Science*, 60(9):2248–2268.

Duffie, D., D. Filipović, and W. Schachermayer

2003. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053.

Duffie, D., J. Pan, and K. Singleton

2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68 (6):1343–1376.

Eraker, B., M. Johannes, and N. Polson

2003. The Impact of Jumps in Returns and Volatility. *Journal of Finance*, 58(3):1269–1300.

Feunou, B. and C. Okou

2018. Risk-neutral moment-based estimation of affine option pricing models*.

Fulop, A. and J. Li

2015. Inferring Volatility Dynamics and Variance Risk Premia: An Efficient Bayesian Approach. *Working P.*

Gruber, P. H., C. Tebaldi, and F. Trojani

2015. The Price of the Smile and Variance Risk Premia. *Working Paper.*

Jiang, G. and Y. S. Tian

2007. Extracting Model-Free Volatility from Option Prices: An Examination of the Vix Index. *Journal of Derivatives.*

Jiang, G. J. and Y. S. Tian

2005. The Model-Free Implied Volatility and Its Information Content. *Review of Financial Studies*, 18(4):1305–1342.

Lee, R.

2010. Weighted Variance Swap. In *Encyclopedia of Quantitative Finance.* John Wiley & Sons, Ltd.

Li, J. and G. Zinna

2014. The Variance Risk Premium: Components, Term Structures, and Stock Return Predictability.

Martin, I.

2013. Simple Variance Swaps. Working Paper 16884, National Bureau of Economic Research.

Orłowski, P., A. Sali, and F. Trojani

2015a. Arbitrage Free Dispersion.

Orłowski, P., P. Schneider, and F. Trojani

2015b. Big risk. *Working Paper.*

Pan, J.

2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3–50.

Schneider, P. G. and F. Trojani

2015. Divergence and The Price of Uncertainty.

Tzoumas, V., A. Jadbabaie, and G. J. Pappas

2016. Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms. In *2016 American Control Conference (ACC)*, Pp. 191–196.

van der Merwe, R. and E. A. Wan

2001. The square-root unscented Kalman filter for state and parameter-estimation. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '01)*, volume 6, Pp. 3461–3464.

Wan, E. A. and R. van der Merwe

2000. The unscented Kalman filter for nonlinear estimation. In *Proc. and Control Symp Adaptive Systems for Signal Processing, Communications 2000. AS-SPCC. The IEEE 2000*, Pp. 153–158.

Appendix A. The forward-neutral cumulant generating function

The CGF, through no-arbitrage restrictions, has several interesting properties under the forward-neutral probability measure. I show how under these restrictions the derivatives of the CGF can be interpreted as measures of expected return variation.

The cumulant generating function of the log-return, defined by

$$K(p,\tau|\mathcal{J}_t) := \ln \mathbb{E}\left[\left(F_{t+\tau}/F_t\right)^p |\mathcal{J}_t\right], \tag{9}$$

is a convex and continuous function of argument p. I also introduce the following notation, which will simplify many upcoming expressions:

$$\phi_p(x) := x^p. \tag{10}$$

The CGF uniquely identifies the forward-neutral probability distribution. Therefore, if it is possible to evaluate the CGF at arbitrary p, then it is possible to uniquely determine the state pricing measure. In empirical applications, the forward-neutral CGF has to be estimated from the prices of European option contracts, as in Section 2.5. For small and large values of p the estimates require the knowledge of prices deep out-of-the-money put and call contracts, respectively. If they are at all available, these prices are subject to high uncertainty through wide bid-ask spreads and potentially infrequent updating. As a consequence, estimates of CGF values for p outside of the $[0,2]$ interval are highly unreliable. This is because estimating the CGF (its slope) from option prices requires weighing the option prices by weights of order K^{p-2} $(\ln K K^{p-2})$. My experience

with option data shows that below $p = 0$ (*above* $= 2$) the put (call) option weights grow too fast to ensure decent swap rate approximations with the available option quotes. Fortunately, the CGF and especially its first derivative on the smaller interval $[0, 1]$ contain rich information about the second and third moments of simple returns.

An example CGF is plotted in Figure 11. Under the forward measure $\mathbb{Q}_{t,\tau}$ the time-t expectation of the forward return is $\mathbb{E}^{\mathbb{Q}_{t,\tau}}[F_{t+\tau}/F_t|\mathcal{J}_t] = 1$. [4] This property translates to $K_t(0,\tau) = K_t(1,\tau) = 0$ and has important consequences for the interpretation for the CGF. For instance, together with convexity and continuity, it implies that there exists a $p^* \in (0,1)$ which is a global minimum. It follows that the values of the CGF on the unit interval are bounded and easier to estimate.

The derivatives of the CGF evaluated at 0 convey information about the cumulants – and through them, moments – of the log-return. As a consequence of the particular features of the forward measure, and of the concavity of the logarithm function, the first derivative of the CGF evaluated at any p is informative of the second or third moment of *simple* returns $F_{t+\tau}/F_t$. To see that, note the form of $K'_t(p,\tau)$ in equation (11), and its expansion in simple return, around its forward-neutral mean, 1, in equation (12).

$$K'_t(p,\tau) = \mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{F_{t+\tau}}{F_t}\right)^p \ln\frac{F_{t+\tau}}{F_t}\right] / e^{K_t(p,\tau)} \tag{11}$$

$$= \frac{\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{F_{t+\tau}}{F_t}-1\right) + \left(p-\frac{1}{2}\right)\left(\frac{F_{t+\tau}}{F_t}-1\right)^2 + \left(\frac{p^2}{2}-p+\frac{1}{3}\right)\left(\frac{F_{t+\tau}}{F_t}-1\right)^3 + ...\right]}{e^{K_t(p,\tau)}}. \tag{12}$$

$K'_t(p,\tau)$ above is the slope of $K_t(p,\tau)$ plotted in the right panel of Figure 11. Under an arbitrary probability measure, $K'_t(p,\tau)$ would contain a first moment component. Under no-arbitrage restrictions of the probability measure, the first element of the expansion in (12) drops out, and the leading order of the expansion depends on p. While it is not immediately obvious from (11) what this quantity measures, the expansion (12) is rather instructive.

From expansion (12), the leading term in the slope of the CGF is determined for almost all ps by the second moment of the forward return. In the left panel of Figure 11 I plot the slope at three points of particular interest. The CGF's slope at 0 ($K'_t(0,\tau)$), marked with a green line, is the expectation of $\ln F_{t+\tau}/F_t$. Evaluated at 1 ($K'_t(1,\tau)$), marked with a red line, it is the expectation of $F_{t+\tau}/F_t \ln F_{t+\tau}/F_t$. The leading terms in these quantities are, to first order, determined by the forward-neutral variance of simple returns. They differ in how higher moments of the distribution of simple returns impact them. Importantly, to the left (right) of $p = 1/2$ the second moment of simple returns enters with a negative (positive) sign, therefore the steeper the CGF descends

[4] In order to de-clutter notation, I write $\mathbb{Q}$ for $\mathbb{Q}_{t,\tau}$, $K_t(p,\tau)$ for $K(p,\tau|\mathcal{J}_t)$, and $\mathbb{E}_t^{\mathbb{M}}[X_{t+\tau}]$ for $\mathbb{E}^{\mathbb{M}_{t,\tau}}[X_{t+\tau}|\mathcal{J}_t]$ whenever the contextual understanding does not suffer.

(ascends) at 0 (at 1), the higher the variance of simple returns. The slope at $p = 1/2$ ($K_t'(1/2, \tau)$), marked with a brown line, is a particularly interesting case. Indeed, in the expansion (12), the leading variance term drops out precisely at $p = 1/2$, which implies that the CGF slope at $p = 1/2$ is a measure of skewness of the simple return.

A precise notion of skewness can be assigned to the CGF slope at $p = 1/2$. Indeed, note that $(F_{t+\tau}/F_t)^p / e^{K_t(p,\tau)}$ can be interpreted as a Radon-Nikodym derivative, and (11) can be interpreted as the expectation of log return under a changed probability measure. The particular feature of this measure, as shown in Schneider and Trojani (2015), is that $K'(1/2, \tau) = 0$ if and only if the forward distribution is put-call symmetric (Carr and Lee, 2009). In the case of put-call symmetry, the minimum of the CGF falls at 1/2. Whenever the minimum falls to the left (right) of 1/2, the associated distribution is left-skewed (right-skewed).

In summary, in this section I showed how the CGF and its first derivatives are informative of the second and third moments of the forward-neutral distribution. If the forward measure allows no arbitrage, then prices of payoffs can be found by evaluating the relevant forward expectation. A natural consequence of this fact is that there exist trading strategies whose prices are expressed by $K'(p, \tau)$. I describe the strategies and their payoffs in the Section 2.

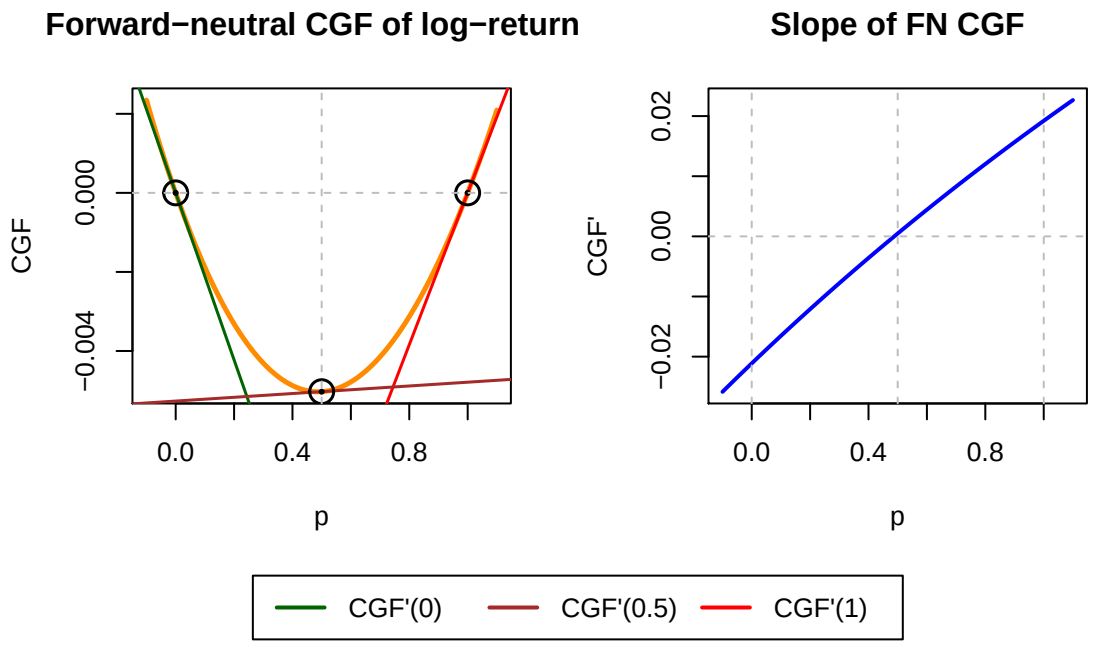

Fig. 11. Example cumulant generating function and its derivative. Derivatives of the CGF at three chosen points are illustrated with tangent lines in the left panel.

A.1. CGF in affine jump diffusions

I consider a general affine stock return dynamics, written as:

$$\begin{aligned} \frac{dF_t}{F_{t-}} &= \mu(V_{t-})dt + \Phi(V_{t-})dW_t^S + \int_{\mathbb{R}\setminus\{0\}} (e^{g(x,y)} - 1)\nu_t(dx,dy)dt \\ dV_t &= \kappa(V_{t-})dt + \Sigma(V_{t-})dW_t^V + \int_{\mathbb{R}_+^M} y\nu_t(dx,dy)dt \end{aligned} \tag{13}$$

with $\mu : \mathbb{R}^M \to \mathbb{R}$, $\kappa : \mathbb{R}^M \to \mathbb{R}^M$, $\Phi'\Phi$, and the diagonal of $\Sigma'\Sigma$ (the off-diagonal elements are 0) affine in V. ν_t is a counting measure whose compensator is affine in V. All components of dW^S are independent, and so are those of dW^V. It is, however, possible to allow for correlation between dW^S and dW^V to model the "leverage effect", that is to reflect the fact that negative returns often accompany icreases in volatility. This notation covers basically all models existing in the literature on standard state spaces.

Duffie et al. (2000) and Duffie et al. (2003) give the general form of the joint cumulant generating function of the log-return and variance factors in settings like (13) and impose necessary technical conditions on the jump measure ν. Borrowing from the intuitions of Section A, I focus on the marginal CGF of the log-return and its derivative. For $p \in \mathbb{R}$, write

$$K_t(p,\tau) = \alpha(p,\tau) + \beta(p,\tau) \cdot v_t \tag{14}$$

$$K_t'(p,\tau) = \alpha'(p,\tau) + \beta'(p,\tau) \cdot v_t. \tag{15}$$

The coefficients $\alpha, \alpha', \beta, \beta'$ can be found by solving systems of ordinary differential equations, as stated in Propositions 1 and 3 in Duffie et al. (2000), respectively for the CGF and its derivative.

It is instructive to investigate the sensitivity of the shape of the CGF and its derivative to the parameter values of a simple affine jump diffusion model. From changes in the shape and values of the CGF and its derivative, induced by parameter changes, one can infer *a*) which risk-neutral moments of log-returns the parameter changes impact, and *b*) how contract prices might reflect higher-order moment information.

The Bates (1993) specification is the simplest "laboratory" which allows for a rich comparative statics analysis. The model allows for jumps in the underlying (with a constant jump intensity)

and a Brownian leverage effect.

$$\begin{aligned}
\frac{dF_t}{F_{t-}} &= -\lambda_0 \mathbb{E}[e^x - 1]dt + \sqrt{v_t} dW_t^S + \int_{\mathbb{R}\setminus\{0\}} (e^x - 1)\,\nu(dx)dt \qquad (16)\\
dv_t &= \kappa(\eta - v_t)dt + \sigma\sqrt{v_t} dW_t^v \\
dW_t^S dW_t^v &= \rho dt, \text{ and} \qquad \frac{\nu(dx)}{dx} = \frac{\lambda_0}{\sqrt{2\pi}} e^{-\frac{(x-\mu_J)^2}{2\sigma_J^2}}
\end{aligned}$$

I modify the parameters from those contained in the paper so that they better reflect the characteristics of stock returns, rather than foreign currency returns, which served as the estimation sample. I set a negative jump mean and add a substantial leverage effect. The exact values are not of central interest, albeit they are reported in Figure 12.

Consider the CGF and its derivative as mappings over the underlying model parameter space, Θ, for fixed state variable value, and for a range of values of p. I adopt the notation $K_t(\theta; p, \tau)$ and $K_t'(\theta; p, \tau)$ for the CGF and its derivative, respectively. Both functions map from Θ to $\mathbb{R}$. A direct graphical representation of both functions for multiple parameter value sets is not informative, as it is hard to visually discern the subtle shape changes induced even by substantial parameter modifications. Therefore, in Figure 12 I plot the mappings $\frac{\partial}{\partial\theta_j} K_t(\theta; p, \tau)$ (in orange) and $\frac{\partial}{\partial\theta_j} K_t'(\theta; p, \tau)$ (in blue) for all forward-neutral parameters of the model, while holding the $v_t = 0.02$ constant. These curves, and particularly their values at $p \in \{0, 1/2, 1\}$, illustrate how the forward-neutral moments change with model parameters.

The interpretation of these curves requires some prudence. $\frac{\partial}{\partial\theta_j} K_t'(\theta; p, \tau)$, plotted in blue in Figures 12, gives precisely the effect of a parameter change on the risk-neutral moments of returns. Focus on $p \in \{0, 1/2, 1\}$. In the beginning of section A I showed that at $p = 0$ and $p = 1$ ($p = 1/2$), the leading term in $K_t'(p, \tau)$ is determined by the variance (skew) of returns. Recall, however, that $K_t'(0, \tau) \approx -1/2\mathbb{V}_t[y_{t+\tau}]$, $K_t'(1, \tau) \approx 1/2\mathbb{V}_t[y_{t+\tau}]$, and $K_t'(1/2, \tau) \approx -1/24\mathbb{E}_t[y_{t+\tau} - 1]^3$ for $y_{t,\tau} = F_{t+\tau}/F_t$. It follows that the lower (higher) the blue curve at $p = 0$ ($p = 1$) in Figure 12, the more the forward-neutral variance increases with an increase in the parameter. Similarly, the higher the blue curve at $p = 1/2$, the more negative skewness there is in the model. The parameters $(\lambda_0, \mu_J, \sigma_J, \eta, \kappa)$ have an almost linear impact on the shape of skewness, and the curves pass close to 0 at $p = 1/2$ in Figure 12. The values of the blue curves for at 0 and 1 give directly their marginal impact on signed return variance. Thus $(\lambda_0, \mu_J, \sigma_J, \eta, \kappa)$ are the key determinants of return variance in this model. For the remaining parameters, from the magnitudes of the response of the CGF slope I can presuppose whether the perturbation of the parameter changes the third or fourth risk-neutral moment. An increase in the leverage parameter from the initial value of $\rho = -1$ or from $\sigma = 0.38$ has no impact on the expected log-return, and thus on return variance, because the blue curves pass through 0 in Figure 12. This implies that the price of the variance swap

does not reflect Brownian skewness, and it is also not affected by structural changes in volatility of volatility. The slope of the CGF at $p = 1/2$ decreases in ρ at the starting parameter value, that is an inrease in ρ decreases negative skewness. The two remaining parameters impact the slope of the CGF at $p = 1/2$ via the fourth moment of returns rather than the third. An increase in the volatility-of-volatility parameter σ increases the fourth moment of returns, and the impact is visible in the CGF slope at $p = 1/2$.

Before considering $\frac{\partial}{\partial \theta_j} K_t(\theta; p, \tau)$, note that the CGF values at $p = 0$ and $p = 1$ are anchored at 0 by no-arbitrage. It follows that if the level of the CGF changes at some point inside $p \in [0, 1]$, it must, due to convexity and continuity, entail changes in slope and/or curvature of the CGF at some points. Curvature is directly related to the second derivative of the CGF with respect to the argument p, $K_t''(p, \tau)$, and thus to the second cumulant (variance) of the log-return, which translates to the third or fourth moment of the simple return, depending on p. The shapes of $\frac{\partial}{\partial \theta_j} K_t(\theta; p, \tau)$ in Figure 12 tells us whether a change in a parameter value predominantly impacts CGF's curvature or symmetry around 1/2. Note that $\frac{\partial}{\partial \theta_j} K_t(\theta; p, \tau)$ curves are almost symmetric around $p = 1/2$ for λ_0, μ_J, σ^J, κ and η, as opposed to other parameters. Imagine, for each parameter, adding the orange curves in Figure 12 to the CGF plot in the left panel of Figure 11. As values of these parameters increase, the CGF becomes flatter (more curved) for μ_J (λ_0, σ_J and η), for κ the sign of the parameter derivative would be opposite were $v_t > \eta$; these parameters predominantly govern return variance. For the remaining parameters, the effect is not symmetric around $p = 1/2$. Again, add the orange curve in the ρ panel of Figure 12 to the CGF plot in Figure 11. This moves the minimum to the right, in this case decreasing the distance between $p* = \text{argmin} K_t(p, \tau)$ and 1/2, which is a measure of skewness. In this way, the curve $\frac{\partial}{\partial \theta_j} K_t(\theta; p, \tau)$ shows which parameters predominantly govern return variance, and which govern higher moments, but does not show the magnitude of the change.

The static analysis above, conducted in an example model, shows that, in isolation, the price of a variance swaps conveys little information about leverage or volatility-of-volatility parameters, while the price of a skew swap is not informative of jump intensity, or of the jump variance parameter. By no means is that a general rule, but rather a demonstration of how our intuitions about model parameters translate to the properties of the CGF. The following intuition should be gleaned from this section: in an estimation exercise certain parameters can be better identified from the dynamics of some contracts, than others. In the example model, one would rightly expect that a single observation of the term structure of variance swaps is completely uninformative of the values of ρ and σ. While in an estimation exercise with variance swap and return, these parameters can be identified from, respectively, the dynamics of returns, and the dynamics of variance swaps together with returns, one can strengthen their identification by adding the term structure of skewness swaps to the observation equation, as estimating skewness features from returns is notoriously difficult.

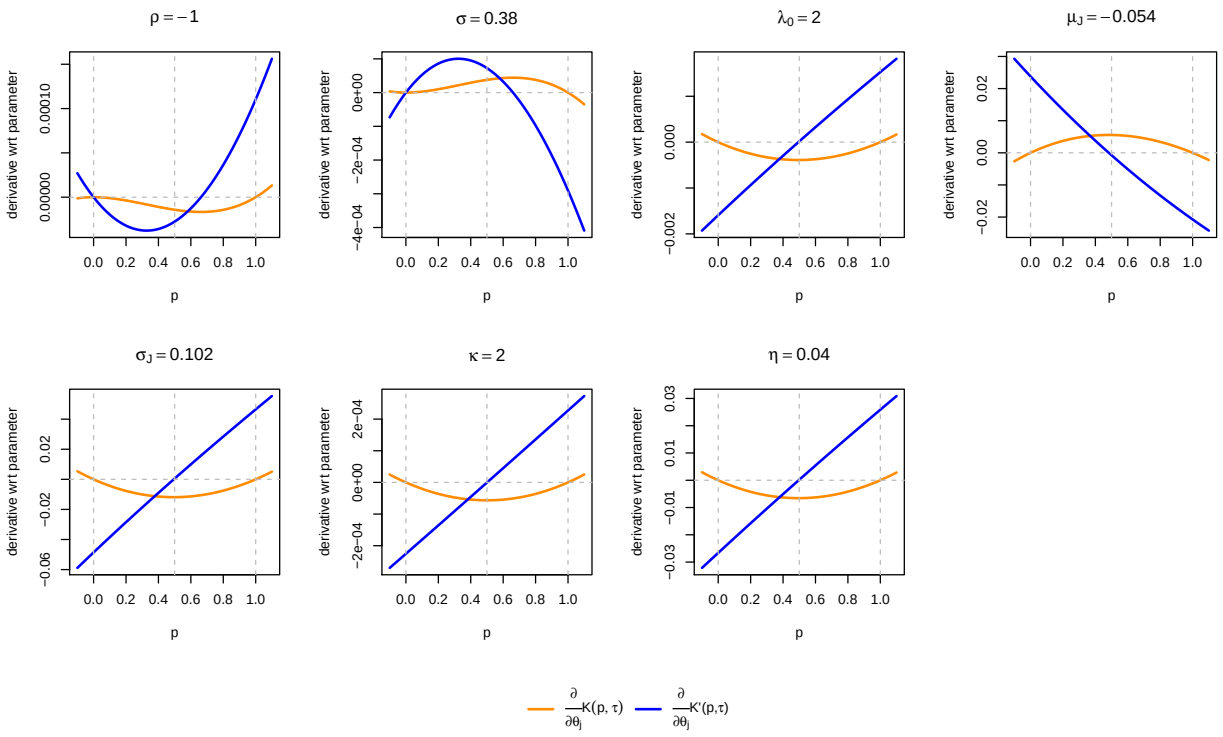

Fig. 12. Comparative statics: sensitivity of cumulant generating function of the log-return, and of its first derivative to changes in model parameters.