Abstract

Extensive research illustrates the jump and discretisation errors that affect the valuation of standard swap contracts. We introduce a vector space of price and return characteristics that allow to define swaps which can be valued exactly, assuming only that the market is free of arbitrage. Although fair-value swap rates are independent of monitoring frequency, the associated risk premiums are not. A historical analysis based on 16 years of S&P500 data demonstrates the diversity of the risk exposures attainable through trading these swaps, as well as floating-floating swaps that trade differential risk premiums and maturities.

Keywords: Aggregation property, calendar swaps, frequency swaps, fourth moment trading, realised skewness, risk premium, straddle swaps, variance swaps

1 Introduction

Variance swaps are popular over-the-counter instruments for trading variance risk premiums by exchanging a floating realised variance with a fixed swap rate, based on some notional amount. A risk-neutral market participant can offer this premium to speculators or risk-averse investors who hedge their exposure to realised variance. When a bank issues a variance swap that pays realised variance, with payment settled at maturity, the rate it charges should be determined so that it expects a small profit after hedging its exposure to realised variance. A theoretical, fair-value variance swap rate provides an indicative quote for the rate actually charged. Variance swap rates have been available from broker dealers for many years and Hafner and Wallmeier [2008] demonstrate that fair-value rates are normally within the bid-ask spread of market rates, indicating an active market where banks may not be hedging all their exposures in order to charge competitive rates. By contrast, during the turbulent year
surrounding the Lehman Brothers collapse in September 2008 market rates were frequently more than 5% greater than their fair-values – see Ait-Sahalia et al. [2012].

The fair-value swap rate is set so that the swap’s expected pay-off at inception is zero. Since the seminal paper of Demeterfi et al. [1999], the computation of a theoretical fair value for a variance swap rate, i.e. the expected realised variance, has been the subject of extensive academic research. The standard fair-value variance swap rate calculation supposes three things, namely: (1) monitoring of the realised leg happens continuously; (2) the discounted underlying price follows a martingale diffusion process; and (3) vanilla options on the underlying with the same maturity as the swap are traded at a continuum of strikes. As long as these assumptions hold the fair value of realised variance – which, under assumption (1), becomes the quadratic variation of the log price – can be derived precisely from the market prices of these vanilla options by applying the replication theorem of Carr and Madan [2001]. See Jiang and Tian [2005] for further details. However, in the real world none of these assumptions hold. Consequently, a considerable body of literature has developed on quantifying the errors associated with these assumptions.

In practice the realised leg of the swap must be monitored discretely and the theoretical fair-value derived under assumption (1) is subject to a bias, which decreases with the term of the swap. As shown by Carr and Lee [2009] this discretisation bias is dominated by the third moment of returns, so it is largest during excessively volatile periods. Moreover, the continuous fair-value, i.e. the quadratic variation of the log price, under (over) estimates the correct (discretely-monitored) fair value when the third moment is negative (positive).

Jarrow et al. [2013] show that, for some otherwise reasonable price processes, a discretely-monitored swap rate need not exist. They derive discretisation error bounds that get tighter as the monitoring frequency increases and prove several results on the convergence of the discretely-monitored swap rate to its continuously-monitored counterpart, assuming some specific stochastic volatility diffusion processes. Bernard and Cui [2014] generalize some of the results in Jarrow et al. [2013] and give conditions for the discretisation bias to be positive.
or negative, and Hobson and Klimmek 2012 derive model-free discretisation error bounds and super- and sub-replication strategies for hedging variance swaps.

Relaxing assumption (1), Broadie and Jain 2008 derive analytic fair-value variance (and volatility) swap rates based on the discrete-monitoring assumption. However, these swap rates are not model-free. They consider several stochastic volatility diffusion and jump models, claiming that for most realistic contract specifications the discrete monitoring error is actually smaller than the error due to the violation of assumption (2). Bernard and Cui 2014 extend their analysis to include a much wider variety of processes, but their results are still not entirely model free. Ignoring the jump component in an underlying process induces another bias to the fair-value swap rate, whose sign and size depend on the direction and magnitude of the jumps. Rompolis and Tzavalis 2013 derive bounds for this jump bias and demonstrate, via simulations and an empirical study, that price jumps induce a systematic negative bias which is particularly apparent during excessively volatile periods.

Thus, the jump bias and the discretisation bias work in the same direction, to substantially under-estimate the fair-value swap rate when the term of the swap includes a particularly volatile period. This is potentially good news for the investor. If the issuer omits to add on to his premium to cover these two known biases then the investor could pay a much lower fixed rate than he would if the swap were priced under more realistic assumptions. By contrast, the issuer of the swap could find himself under-hedged during exactly those excessive volatile periods that drive the variance swap market. Hence, there are good reasons for issuer’s premiums to contain substantial add-ons to cover the uncertainty in the size and sign of jump and discretisation biases. Indeed, Ait-Sahalia et al. 2012 demonstrate that traded swap rates can deviate significantly from fair-value quoted rates, especially for longer-term variance swaps, during periods where there are large jumps in the price of the underlying. Jiang and Tian 2005 also address the problems attendant to assumption (3), i.e. impossibility of exact replication due to the availability of only a finite set of strikes for traded vanilla options. They derive upper bounds for the truncation error (i.e. the error due to the
use of a finite range of strikes) and use a model-dependent simulation to illustrate that truncation errors are negligible if the strike range is more than two standard deviations from the separation strike. Model-free error bounds are also provided, but they are not as tight as the model dependent ones. Based on a finite number of traded strikes Davis et al. [forthcoming] derive much tighter model-free arbitrage bounds for continuously-monitored variance swap rates. They also find that market rates are surprisingly close to their lower bound yet remain consistent with the absence of arbitrage.

The terms and conditions of a standard variance swap define the realised variance as the average squared daily log return on some underlying, commonly an equity index. However, a different definition for the realised characteristic could result in fair values that are easier to price and hedge. Martin [2013] advocates the use of a sum of squared ‘simple’ returns, rather than log returns, to define the realised variance. With this modification both jump and discretisation biases are minimised. Likewise, the gamma swaps described by Lee [2010] weight the realised variance characteristic in such a way that replication and valuation are relatively straightforward under the continuous martingale assumption.

The idea to change the definition of the realised characteristic leads us to the path-breaking work of Neuberger [2012], which lays the foundation for our research. Neuberger demonstrates that a different modification of the definition of realised variance results in a swap which is completely free from any errors arising from assumptions (1) and (2) above. In other words, an exact theoretical fair-value swap rate can be derived for any price process without assuming continuous monitoring. Indeed, the expected realised variance does not depend on the frequency of monitoring and the same fair-value swap rate applies whether the floating leg is based on intra-day, daily, weekly, or monthly returns. In fact, the monitoring of the floating leg does not even have to be regular; any time-partition of the term of the swap can be defined without affecting the fair-value swap rate. The only assumption made about the underlying is that the market is free from arbitrage opportunities. The crucial condition is that the function used for defining the floating leg (the realised characteristic)
must satisfy an aggregation property. Using this property Neuberger introduces a particular third-moment swap where the floating leg is based on a new realised skewness characteristic.

The starting point of our research is an extension of Neuberger’s ideas to other characteristics of a price (or a log return) distribution. We introduce a vector space (over \( \mathbb{R} \)) of time-discretisation invariant (TDI) swaps which contains an infinite variety of second, third and higher-moment swaps, and swaps that are not even associated with moments of the underlying distribution. The realised leg of a TDI swap is defined in such a way that the discrete monitoring error is zero, which implies that the jump error is also zero, and that the aggregation property is satisfied. We show that, provided some technical conditions, the absence of a discrete monitoring error as well as the aggregation property are equivalent to a second order system of partial differential equations and provide elegant analytic solutions for the characteristics. The practical importance of our work is that, based only on the no-arbitrage assumption and without requiring any further model specifications, the theoretical fair-value rate for a TDI swap can be derived exactly from vanilla option prices, i.e. both the discretisation bias and the jump bias are zero. An infinite variety of TDI swaps can be defined. For brevity we focus on just a few interesting examples, including third-moment and fourth-moment swaps for which the associated risk premiums have relatively low correlation. In all these examples the theoretical fair-value swap rate may be expressed in terms of prices of certain synthetic fundamental contracts, of similar ilk to the log and entropy contracts introduced by Neuberger [1994] and Neuberger [2012].

Since the same fair-value rate applies to a realised characteristic irrespective of the monitoring frequency, the fair value of a floating-floating swap which exchanges a daily-monitored for a weekly-monitored realised characteristic is zero. However, we demonstrate both theoretically and empirically that the risk premium on a TDI swap is non-zero, in general, and that it will depend on the monitoring frequency. This leads to the definition of frequency swaps, which exchange one monitoring frequency for another. Similarly, the TDI risk premium depends on the maturity of the swap. This observation motivates the introduction of
calendar swaps, which exchange two TDI characteristics that refer to different maturities, and which provide direct access to the forward variance, third moment or fourth moment risk premiums.

While it is reasonable in practice to assume an arbitrage-free market, so that the underlying is a martingale, the supposition that options on a continuum of strikes are traded is certainly not valid. Indeed, as previously discussed, when this assumption is needed an approximation error appears in the actual calculation of the theoretical fair-value swap rate. This error can be relatively large, especially during volatile markets when the volatility skew is pronounced – see [Alexander and Leontsinis 2011]. Motivated by this observation we introduce a subspace of swaps whose fair value may be computed exactly, without the need for numerical integration, as a discrete weighted sum over vanilla option prices at the available, traded strikes. These ‘strike-discretisation invariant’ swaps also preserve the TDI property, so we refer to them using the more general term discretisation invariant (DI) swaps. A particular subset of DI swaps is called straddle swaps because their theoretical swap rates are simply the product of the market prices of a call and a put options with the same strike. The returns on these swaps have very low correlation with returns on variance swaps.

In the following: Section 2 sets the background by briefly describing the errors that enter the calculation of theoretical rates for standard variance swaps and relating them to the aggregation property introduced by [Neuberger 2012]; Section 3 presents the aforementioned second order system of partial differential equations, derives the vector space of realised characteristics for TDI swaps, develops examples of variance and higher-moment swaps for which the fair-value swap rates and replication portfolios may be expressed in terms of fundamental contracts, and focuses on a particular subset of DI swaps called ‘straddle swaps’ for which replication and valuation are particularly simple; Section 4 analyses the risk premiums on TDI swaps under the geometric Brownian motion assumption and introduces ‘frequency swaps’ and ‘calendar swaps’ based on TDI characteristics; The empirical results are presented in Section 5 and Section 6 concludes. All proofs are in the Appendix.
2 Background and Motivation

The distinction between martingale and non-martingale processes will be central to our arguments so it helps to distinguish them in our notation. Univariate martingale processes are here denoted using upper-case letters; and non-martingales with lower-case. We can simplify notation because we need to consider only one maturity date, T, but various partitions of the interval \( \Pi := [0, T] \) are also needed, in particular the regular partition \( \Pi_D := \{0, 1, \ldots, T\} \) which we term the ‘daily’ partition for brevity. The increments along a partition are denoted using a ‘carat’ (\(^\hat{\text{a}}\)), e.g. for some process \( y \) we set \( \hat{y}_t := y_t - y_{t-1} \) for increments under the daily partition. We use \( \mathbb{E}_t[.] := \mathbb{E}[.|F_t] \) to denote the expectation conditional on a filtration \( F_t \) at time \( t \), under the risk-neutral measure (unless otherwise stated), and write \( \mathbb{E}[.] := \mathbb{E}_0[.] \).

Let \( s := \{s_t\}_{t \in \Pi} \) be the price process underlying a variance swap of maturity \( T \); let \( F := \{F_t\}_{t \in \Pi} \) be the fair-value price process of a futures contract on \( s \) with maturity \( T \), i.e. \( F_t := \mathbb{E}_t[s_T] \); and let \( x := \{x_t\}_{t \in \Pi} \) be the log futures price process, i.e. \( x_t := \ln F_t \). Using our notation the standard, daily, realised variance may be written:

\[
\sum_{\Pi_D} \hat{x}_t^2 := \sum_{t=1}^{T} (x_t - x_{t-1})^2, \tag{1}
\]

where \( \hat{x}_t := x_t - x_{t-1} \) denotes the daily log returns. In practice, the floating leg of a variance swap is set equal to the average realised variance \( \bar{T}^{-1} \sum_{\Pi_D} \hat{x}_t^2 \), where \( \bar{T} \) denotes the number of trading days during the lifespan \( \Pi \), rather than the total variance as in (1). However, including this level of detail would only add an unnecessary level of complexity to our analysis. Carr and Wu [2009] discuss the idealised case where continuous monitoring is possible. In other words, the realised variance (1) becomes the quadratic variation of \( x \), denoted \( \langle x \rangle_T \). Then, under the assumption of a generic decomposition of the underlying process into a pure jump and a pure geometric diffusion component, they apply the replication theorem of Carr.
and Madan 2001 to derive the variance swap rate\(^1\)

\[
\mathbb{E}[\langle x \rangle_T] = 2 \int_{\mathbb{R}^+} k^{-2} q(k) dk + \eta,
\]

where \(q(k)\) denotes the forward price of a vanilla out-of-the-money (OTM) option with strike \(k\) and maturity \(T\).\(^2\) The jump error, \(\eta\), is zero when \(F\) follows a pure diffusion.

The second main source of error in the theoretical fair-value swap rate stems from the fact that the realised leg of the swap is monitored only at discrete points in time. For instance, when monitoring is based on the daily partition the discrete monitoring error may be written

\[
\varepsilon := \mathbb{E} \left[ \sum_{\Pi_D} \hat{x}^2_t - \langle x \rangle_T \right].
\]

(2)

Both these errors affect the theoretical price of the fixed leg. For instance, with the realised variance \(\text{(1)}\) the fair-value variance swap rate may be written

\[
\mathbb{E} \left[ \sum_{\Pi_D} \hat{x}^2_t \right] = 2 \int_{\mathbb{R}^+} k^{-2} q(k) dk + \eta + \varepsilon.
\]

(3)

By ignoring these errors the risk-neutral expectation of the pay-off becomes \(\eta + \varepsilon\) rather than zero and therefore the estimator for the variance risk premium is biased under the standard variance swap pricing formula.

In practice, the integral in \(\text{(3)}\) that is used to approximate the fair-value swap rate must be estimated using the prices of vanilla options that are actually traded. So there is a third, estimation bias affecting the actual computation of the fair-value rate approximation. Typically we use a fairly restricted range of quoted strikes, because deep-OTM options lack

\(^1\)In an arbitrage-free market, as introduced by Harrison and Kreps 1979, the bank issuing a variance swap to a representative investor will compute this expected pay-off under a risk-neutral measure \(Q\). In a complete market the risk-neutral measure for a representative investor corresponds to the market implied measure \(M\) (see Breeden and Litzenberger 1978). In this case a unique fair value for the variance swap rate may be derived as the expectation of realised variance under \(M\).

\(^2\)When \(k \leq F_0\) the option is a put and when \(k > F_0\) the option is a call. This choice of separation strike is standard in the variance swap literature, e.g. in Bakshi et al. 2003.
sufficient liquidity to have reliable prices. Indeed the estimation error can be quite significant, especially during volatile periods, as shown by Alexander and Leontsinis [2011]. The second aim of our research is to define characteristics that are also free from this estimation error because their fair-value can be derived without using the replication theorem of Carr and Madan [2001].

Let \( \{\Pi_N\}_{N=1,2,...} \) be a sequence of partitions \( \Pi_N = \{t_i\}_{i=0,...,N} \) of the lifespan \( \Pi \), having the properties \( 0 = t_0 \leq t_1 \leq \ldots \leq t_N = T \) and \( \max_{i\in\{1,...,N\}} [t_i - t_{i-1}] \to 0 \) as \( N \to \infty \). This second property is written \( \Pi_N \to \Pi \) for brevity. Clearly, the daily partition \( \Pi_D \) is a special case of \( \Pi_N \) where \( t_i = i \) and \( T = N \). To be completely general we allow the realised characteristics to describe distributional properties of an \( n \)-dimensional stochastic process \( z := \{z_t\}_{t\in\Pi} \in \mathbb{R}^n \). Then, given a continuous function \( f : \mathbb{R}^n \to \mathbb{R} \), the realised \( f \)-characteristic of \( z \) w.r.t. \( \Pi_N \) is defined as

\[
\sum_{\Pi_N} f(\hat{z}_i) := \sum_{i=1}^{N} f(z_{t_i} - z_{t_{i-1}}), \quad (4)
\]

where \( \hat{z}_i := z_{t_i} - z_{t_{i-1}} \) denote the increments in \( z \) along \( \Pi_N \). If it exists (and this depends on the choice of \( f \) and \( z \) we define the \( f \)-variation of \( z \) as the continuously monitored limit of the realised characteristic, i.e.

\[
\langle z \rangle_f^z := \lim_{\Pi_N \to \Pi} \sum_{\Pi_N} f(\hat{z}_i). \quad (5)
\]

In the following we only consider characteristics \( f \) with \( f(0) = 0 \), otherwise the limit in [5] would not be finite.\(^3\) The \( f \)-variation is a purely theoretical construct that, if it exists, can be used to derive a theoretical fair-value swap rate by taking its expected value based on some assumed process for the underlying. This is the approach taken by Jarrow et al. (2013) and many other papers that analyse the discrete monitoring error for variance swaps. As long

\(^3\) However, we do not need to assume the existence of the \( f \)-variation because it does not preclude the definition of an \( f \)-swap as a financial contract that exchanges a realised \( f \)-characteristic [4] with a fixed value, called the \( f \)-swap rate.
as the $f$-variation exists and is finite the discrete monitoring error for an $f$-swap under the partition $\Pi_N$ may be written

$$\varepsilon_N(f, z) := \mathbb{E} \left[ \sum_{\Pi_N} f(\hat{z}_i) - \langle z \rangle_T^f \right]. \quad (6)$$

For instance, with $z = x$ and $f(\hat{x}) = \hat{x}^2$ the definition (5) corresponds to the quadratic variation of the log price and the discrete monitoring error is given by (2).

For a given $f$ and $z$ there may be zero, one or more partitions $\Pi_N$ of $\Pi$ for which $\varepsilon_N(f, z) = 0$. Our focus is on those combinations of $f$ and $z$ for which

$$\mathbb{E} \left[ \sum_{\Pi_N} f(\hat{z}_i) \right] = \mathbb{E} \left[ \langle z \rangle_T^f \right], \quad \text{for all } \Pi_N. \quad (7)$$

If (7) holds $\forall \Pi_N$ then it holds for the trivial partition $\Pi_1 = [0, T]$, for which the above becomes: $\mathbb{E} [f(z_T - z_0)] = \mathbb{E} [\langle z \rangle_T^f]$. But $f(z_T - z_0) = f \left( \sum_{\Pi_N} \hat{z}_i \right)$, so (7) implies

$$\mathbb{E} \left[ \sum_{\Pi_N} f(\hat{z}_i) \right] = \mathbb{E} \left[ f \left( \sum_{\Pi_N} \hat{z}_i \right) \right] = \mathbb{E} [f(z_T - z_0)], \quad \forall \Pi_N. \quad (8)$$

Note that the lack of path-dependence of this expectation also implies that the jump error $\eta$ must be zero. In other words, when the discrete monitoring error is zero under all partitions then, even if investors differ in their views about jump risk in an incomplete market, they would still agree on the fair-value $f$-swap rate as long as they agree on the measure for $\mathbb{E}[^\cdot^\cdot]$.

The aggregation property (8) was introduced by Neuberger [2012], albeit with different notation and motivation. He also noted that exact pricing of a discretely monitored swap is possible when the aggregation property (AP) holds because the replication theorem of Carr and Madan [2001] is applicable to the right hand side of (8), which he called the

---

4Although the absence of a discrete monitoring error (7) is a stronger assumption than the aggregation property (8), we show in Corollary 1 that the two properties are equivalent as long as $f$ is twice continuously differentiable and the $f$-variation of $z$ is finite. In other words, if the $f$-variation does exist (and we do not need to assume this for the main results in our paper) then (8) holds for a pair $(f, z)$ iff $\varepsilon_N(f, z) = 0$. 

implied characteristic. The AP does not hold for \((f, x)\) when \(f(\hat{x}) = \hat{x}^2\) is the standard variance characteristic, but Neuberger finds two alternative variance characteristics: The log characteristic \(l(\hat{x}) := 2(e^{\hat{x}} - 1 - \hat{x})\) for which the AP holds when \(x\) is the log of any martingale; and the entropy characteristic \(h(\hat{x}) := 2(e^{\hat{x}e^\hat{x}} - e^{\hat{x}} + 1)\) which satisfies the AP under the additional assumption of independent increments \(\hat{x}_i, i \in \{1, \ldots, N\}\). Using Taylor expansion about the origin, one can see that both \(g \equiv l\) and \(g \equiv h\) may be associated with the second moment of the distribution of \(\hat{x}\), because they satisfy \(\lim_{\hat{x} \to 0} \frac{g(\hat{x})}{\hat{x}^2} = 1\).

Neuberger also classifies the functions \(f\) which satisfy the AP when: (i) \(z = \{F, \psi\}\) and \(\psi := \{\psi_t\}_{t \in \Pi}\) is the conditional variance process defined as \(\psi_t := \mathbb{E}_t [(F_T - F_t)^2]\); and (ii) \(z = (x, v_g)'\) with \(v_g := \{v_{gt}\}_{t \in \Pi}\) being a generalised variance process, i.e. \(v_{gt} := \mathbb{E}_t [g(x_T - x_t)]\).

Two particular subsets of this second class of AP-characteristics correspond to \(v_{gt} := v_{lt} = \mathbb{E}_t [l(x_T - x_t)]\) and \(v_{gt} := v_{ht} = \mathbb{E}_t [h(x_T - x_t)]\). These are called the log and entropy variance processes because they are closely related to the log contract, which pays \(x_T\), and the entropy contract, which pays \(F_Tx_T\) at maturity, respectively. Because the log characteristic is an AP-characteristic w.r.t. the log of any martingale, \(x\), one can change the definition of the floating leg of a variance swap from \(\Pi\) to \(\sum_{\Pi_D} l(\hat{x}_t)\), and the result will be a log variance swap that has no discrete monitoring error and no jump error. In this case \((3)\) becomes

\[
\mathbb{E} \left[ l \left( \sum_{\Pi_D} \hat{x}_t \right) \right] = 2 \int_{\mathbb{R}^+} k^{-2} q(k) dk.
\]

Neuberger also values and replicates a third-moment AP characteristic whose risk premium is strongly correlated with the variance risk premium (see Kozhan et al. [2013]). This further motivates our search for a more general class of \((f, z)\) for which the AP holds and through which more diverse risk premiums can be accessed.

---

5See Neuberger [2012], p.7: “If the measure is a pricing measure, it says that the fair price of a one-month variance swap computed daily (a swap that pays the realized daily variance over a month) is the same as the price of a contingent claim that pays \((F_T - F_0)^2\). Indeed, because the relationship holds under any pricing measure (because the process is a martingale under any pricing measure), it also implies that a variance swap can be perfectly replicated if the contingent claim exists (or can be synthesized from other contingent claims) and the underlying asset is traded.”
3 Discretisation Invariant Swaps

Consider a multivariate stochastic process \( z \in \mathbb{R}^n \) which contains only deterministic functions of futures prices \( F := \{ F_t \}_{t \in \Pi} \in \mathbb{R}^d \) of \( d \) tradable assets, or derivatives on these assets, in an arbitrage-free market. The process \( z \) may e.g. contain futures prices or the logs of these prices and we make the minimal, arbitrage-free assumption only to ensure that the futures prices follow a multivariate \( Q \)-martingale. Let \( \Delta \in \mathbb{R}^{n \times d} \) and \( \Gamma \in \mathbb{R}^{n \times d \times d} \) denote the first and second partial derivatives of \( z \) w.r.t. \( F \).

Now let \( f : \mathbb{R}^n \to \mathbb{R} \) be some twice-differentiable deterministic function and denote by \( J(\hat{z}) \in \mathbb{R}^n \) the Jacobian vector and \( H(\hat{z}) \in \mathbb{R}^{n \times n} \) the Hessian matrix of first and second partial derivatives of \( f \) w.r.t. \( \hat{z} \). A time-discretisation invariant (TDI) swap is any \( f \)-swap on \( z \) for which the discrete monitoring error (7) is zero and consequently the aggregation property (8) holds. Two trivial TDI swaps follow immediately from the definition: (a) if \( f \) is linear, say \( f(\hat{z}) = a^T \hat{z} \) for some \( a \in \mathbb{R}^n \), then (8) holds for any process \( z \) because \( \sum_{\Pi_N} \hat{z}_i = z_T - z_0 \); (b) if \( z \) contains only constant processes then \( \hat{z}_i = 0 \) \( \forall i \in \{1, \ldots, N\} \), so (8) holds for any function with \( f(0) = 0 \). Note that (7) also holds in both cases: in (a) because \( \langle z \rangle^f_T = z_T - z_0 \) and in case (b) because \( \langle z \rangle^f_T = 0 \), provided \( f(0) = 0 \). Therefore, in order to find the most general set of characteristics that define TDI swaps we shall assume that \( z \) can be stochastic and we consider only those characteristics \( f \in C^2 \) for which \( f(0) = 0 \).

Theorem 1: Equivalence of the Aggregation Property

If \((f, z)\) is such that either the AP (8) holds, or the \( f \)-variation of \( z \) exists and (7) holds, then the following second-order system of partial differential equations holds:

\[
[J(\hat{z}) - J(0)]' \Gamma + \Delta' [H(\hat{z}) - H(0)] \Delta = 0. \tag{9}
\]

Moreover, if \( F \) follows a diffusion with finite \( f \)-variation, then (7), (8) and (9) are equivalent.
The above system can be solved numerically to yield, for given \( z \), the characteristics that define a TDI swap on \( z \). However, we are only interested in the analytic solutions of (9), for which we can value and replicate realised TDI characteristics, as in the following:

**Theorem 2: Time-Discretisation Invariant Characteristics**

Let \( F \) follow any \( d \)-dimensional martingale process and set \( z = (F, x)' \) with \( x = \ln F \). Then the solutions to (9) form vector space over \( \mathbb{R} \), defined by:

\[
\mathcal{V} := \left\{ f : \mathbb{R}^{2d} \to \mathbb{R} \mid f(\hat{z}) = \alpha' \hat{z} + \hat{F}' \Omega \hat{F} + \beta' (e^{\hat{x}} - 1), \; \alpha \in \mathbb{R}^{2d}, \; \Omega \in \mathbb{R}^{d \times d}, \; \beta \in \mathbb{R}^d \right\}.
\]

The idea of the proof is to find candidates for \( \mathcal{V} \) by solving (9) for this particular \( z \) and then show, by straight forward evaluation of (7), that the necessary condition is sufficient.

The generality of the conditions for this theorem allows TDI swaps to be defined for a wide variety of underlying variables. For instance, we can include the log contract \( \Lambda_t := \mathbb{E}_t [x_T] \) in \( F_t \) and more generally set \( F_t = \mathbb{E}_t [m(F_T)] \) for some European payoff functions \( m : \mathbb{R} \to \mathbb{R}^d \). Note that with \( z = (F, \Lambda, x)' \) we can relate the characteristics introduced by Neuberger [2012] to specific characteristics in \( \mathcal{V} \).

Before examining some specific examples of TDI swaps on \( z = (F, x)' \) we prove one further result which provides a characterisation of the value process for a general TDI swap in a form that clarifies how to replicate them. From henceforth we assume that replication is under the risk-neutral measure \( Q \). Also, for ease of exposition we use the daily partition \( \Pi_D \) in the text, although all proofs in the Appendix are for general \( \Pi_N \).

Since the fair-value \( f \)-swap rate equals the risk-neutral expectation of the realised \( f \)-
characteristic, the value process \( V := \{ V_t \}_{t \in \Pi} \) of the \( f \)-swap is given by

\[
V_t := E_t \left[ \sum_{D \in \Pi} f(\tilde{z}_t) \right] - E \left[ \sum_{D \in \Pi} f(\tilde{z}_t) \right].
\]

**Theorem 3: Replicating TDI Swaps**

The increments along \( \Pi_D \) of the value process for a TDI \( f \)-swap may be written

\[
\hat{V}_t = \hat{U}_t - F'_{t-1} (\Omega' + \Omega) \hat{F}_t + \beta' (e^{\hat{x}_t} - 1),
\]

where \( U_t = E_t [\alpha' z_T + F'_T \Omega F_T] \).

Theorem 3 characterises the unrealised profit and loss (P&L) which accrues to the issuer of a TDI swap who pays the fixed swap rate \( E \left[ \sum_{D \in \Pi} f(\tilde{z}_t) \right] \) and receives the floating leg defined by the realised characteristic. When the elements of \( F \) and the contract \( U := \{ U_t \}_{t \in \Pi} \) are both tradable the swap can not only be priced exactly, but is also replicable in discrete time by means of a dynamic trading strategy. For instance, the value increments of the swap on the log characteristic \( l(\hat{x}) := 2(e^{\hat{x}} - 1 - \hat{x}) \) are \( \hat{V}_{\Lambda t} = 2(e^{\hat{x}_t} - 1 - \hat{\Lambda}_t) \). Hence this variance swap can be dynamically hedged by selling \( 2F_{t-1}^{-1} \) futures contracts and buying two log contracts.

We now describe some simple canonical examples of TDI swaps which relate to just one underlying futures contract with price \( F_t \). For each swap we derive the fair-value formula, an expression for the value increments in terms of certain ‘fundamental contracts’ and, assuming as in Neuberger [2012] that such contracts can be traded, we also define their dynamic replication.

---

8We follow Neuberger [2012]’s assumption here: of a complete market in which the log contract – and the other fundamental contracts to be defined presently – are tradable. By Carr and Madan [2001] the replication value of the log contract is \( \Lambda_t = x_t - \int_{\mathbb{R}^+} k^2 q_t(k) \, dk \), where \( q_t(k) \) is the time-\( t \) forward price of a vanilla OTM option with strike \( k \) and maturity \( T \). Although neither \( x_t \) nor the options portfolio are tradable (the latter due to the stochastic nature of the separation strike) \( \Lambda_t \) is a \( \mathbb{Q} \)-martingale and so \( \Lambda_t = 0 \) under \( \mathbb{Q} \). However, this is not the case under \( \mathbb{P} \), in general.
Example 1: Variance Swap. Let \( z = F \) and consider the characteristic \( f(\hat{z}) = \hat{F}^2 \). As opposed to squared log returns, squared futures price changes do satisfy the AP. The fair-value swap rate is \( E\left[(F_T - F_0)^2\right] = \Sigma_0 - F_0^2 \) and the value increments of this swap may be written \( \hat{V}_{\Sigma t} := \hat{F}_t^2 + \hat{\psi}_t = E_t \left[(F_T - F_t)^2\right] \). However, it is more convenient to follow Theorem 3 and use the square contract to write \( \hat{V}_{\Sigma t} = \hat{\Sigma}_t - 2F_{t-1}\hat{F}_t \). Hence, this swap can be hedged by selling a square contract and dynamically holding \( 2F_{t-1} \) futures from time \( t-1 \) to \( t \).

Example 2: Covariance Swap. Let \( z = (F, \Sigma)' \) and \( f(\hat{z}) = \hat{F}\hat{\Sigma} \). This TDI characteristic describes the covariance of the value increments in the futures and square contract. The fair value swap rate can be derived using the replication theorem of Carr and Madan [2001] as \( E\left[(F_T - F_0)(\Sigma_T - \Sigma_0)\right] = Q_0 - F_0\Sigma_0 \). The P&L on this swap is obtained from setting \( \alpha = 0 \) in Theorem 3, with \( \Omega_{12} + \Omega_{21} = 1 \), the other elements of \( \Omega \) being zero, yielding \( \hat{V}_{Pt} := \hat{Q}_t - F_{t-1}\hat{\Sigma}_t - \Sigma_{t-1}\hat{F}_t \). Hence, this covariance swap can be hedged by selling a cube contract and holding \( F_{t-1} \) squared contracts plus \( \Sigma_{t-1} \) futures from time \( t-1 \) to \( t \).

Example 3: Third Moment Swap. Again set \( z = (F, \Sigma)' \). We construct a third moment swap by combining a long position in the covariance swap with \( 2F_0 \) short positions in the variance swap. The characteristic becomes \( f(\hat{z}) = \hat{F}\hat{\Sigma} - 2F_0\hat{F}^2 \) and since

\[
E[f(z_T - z_0)] = E\left[(F_T - F_0)(\Sigma_T - \Sigma_0) - 2F_0(F_T - F_0)^2\right]
= E\left[F_T^3 - F_0F_T^2 - 2F_0(F_T^2 - F_0^2)\right]
= E\left[F_T^3 - 3F_T^2F_0 + 3F_TF_0^2 - F_0^3\right] = E\left[(F_T - F_0)^3\right].
\]

the fair-value swap rate is a third moment. The unrealised P&L for this swap may be written:

\[
\hat{V}_{Qt} := \hat{V}_{Pt} - 2F_0\hat{V}_{\Sigma t} = \hat{Q}_t - (F_{t-1} + 2F_0)\hat{\Sigma}_t - (\Sigma_{t-1} - 4F_0F_{t-1})\hat{F}_t.
\]
Hence, the swap can be hedged exactly by selling a cube contract and dynamically holding $F_{t-1} + 2F_0$ square contracts as well as $\Sigma_{t-1} - 4F_0F_{t-1}$ futures. The nice association of this swap with the third moment comes at the price that the hedge ratios depend on the initial price level $F_0$.

**Example 4: Fourth Moment Swap.** Let $z = (F, \Sigma, Q)$ and consider the characteristic $f(\hat{z}) = \hat{F}\hat{Q} - 3F_0\hat{F}\hat{\Sigma} + 3F_0^2\hat{F}^2 - 4F_0^3\hat{F}$. The corresponding fair-value swap rate is

$$\mathbb{E} \left[ (F_T - F_0)(Q_T - Q_0) - 3F_0(F_T - F_0)(\Sigma_T - \Sigma_0) + 3F_0^2(F_T - F_0)^2 - 4F_0^3(F_T - F_0) \right]$$

$$= \mathbb{E} \left[ F_T^4 - 4F_0F_T^3 + 6F_0^2F_T^2 - 4F_0^3F_T + F_0^4 \right] = \mathbb{E} \left[ (F_T - F_0)^4 \right]$$

Following Theorem 3 we have

$$\hat{V}_{Ht} := \hat{H}_t + (F_{t-1} - 3F_0)\hat{Q}_t + 3F_0(F_0 - F_{t-1})\hat{\Sigma}_t + (Q_{t-1} - 3F_0\Sigma_{t-1} + 6F_0^2F_{t-1} - 4F_0^3)\hat{F}_t.$$

Hence, this swap can be hedged exactly by selling a quartic contract and dynamically holding cubed, squared and forward contracts according to the above hedging ratios.

**Example 5: Generalised Variance Swaps.** Generalised moment swaps may be defined that capture essentially the same risk premiums as the aforementioned moment swaps. For instance, examples of generalised variance swaps include Neuberger’s log variance swap or the entropy swap. For the latter $z = (F, \Lambda)'$ and $f(\hat{z}) = \hat{F}\hat{\Lambda}$ and the fair-value swap rate is $\mathbb{E} \left[ (F_T - F_0)(\Lambda_T - \Lambda_0) \right] = \Psi_0 - F_0\Lambda_0$. From Theorem 3, the unrealised P&L is

$$\hat{V}_{\Psi t} := \hat{\Psi}_t - F_{t-1}\hat{\Lambda}_t - \Lambda_{t-1}\hat{F}_t,$$

so the swap can be hedged exactly by selling an entropy contract and dynamically holding $F_{t-1}$ log contracts as well as $\Lambda_{t-1}$ futures contracts. For another example set $z = \Lambda$ and $f(\hat{z}) = \hat{\Lambda}^2$. Then we have a *squared-log swap*, so-called because the fair-value swap rate is $\mathbb{E} \left[ (\Lambda_T - \Lambda_0)^2 \right]$, which may be expressed in the form $\Upsilon_0 - \Lambda_0^2$. This swap rate corresponds to the variance of the log price since $\Upsilon_0 - \Lambda_0^2 = \mathbb{E} \left[ x_T^2 \right] - \mathbb{E} \left[ x_T \right]^2$. 

16
Following Theorem 3 we have $\tilde{V}_t := \tilde{Y}_t - 2\Lambda_{t-1}\tilde{\Lambda}_t$. Hence, this swap can be hedged exactly by selling a squared log contract and dynamically holding $\Lambda_{t-1}$ log contracts.

The pricing and hedging of all swaps presented in the above examples relies on the replication of certain fundamental contracts. Table 1 summarises the pricing formulas and Table 2 the hedge portfolios for these contracts, all obtained by applying the theorem of Carr and Madan [2001] to the implied characteristic. The difference between the two replication methods is that the pricing formula is based on OTM options which are more liquidly traded – but not investable due to the stochastic separation strike – and the hedge portfolio involves options that are OTM only at inception but remain investable. In particular, the hedge portfolios describe buy-and-hold strategies that require no dynamic rebalancing. However, for an ex-post analysis the two representations are exchangeable and we choose to implement the pricing formula for liquidity sake.

| Contract | Variable | Fair Value | Pricing Formula |
|----------|----------|------------|-----------------|
| Log      | $\Lambda := \{\Lambda_t\}_{t \in \Pi}$ | $\Lambda_t := E_t [x_T]$ | $x_t - \int_{\mathbb{R}^+} k^{-2} q_t(k) dk$ |
| Squared-Log | $\Upsilon := \{\Upsilon_t\}_{t \in \Pi}$ | $\Upsilon_t = E_t [x_T^2]$ | $x_t^2 + 2 \int_{\mathbb{R}^+} (1 - \ln k) k^{-2} q_t(k) dk$ |
| Entropy  | $\Psi := \{\Psi_t\}_{t \in \Pi}$ | $\Psi_t := E_t [F_T x_T]$ | $F_T x_t + \int_{\mathbb{R}^+} k^{-1} q_t(k) dk$ |
| Square   | $\Sigma := \{\Sigma_t\}_{t \in \Pi}$ | $\Sigma_t := E_t [F_T^2]$ | $F_T^2 + 2 \int_{\mathbb{R}^+} q_t(k) dk$ |
| Cube     | $Q := \{Q_t\}_{t \in \Pi}$ | $Q_t := E_t [F_T^3]$ | $F_T^3 + 6 \int_{\mathbb{R}^+} k q_t(k) dk$ |
| Quartic  | $H := \{H_t\}_{t \in \Pi}$ | $H_t := E_t [F_T^4]$ | $F_T^4 + 12 \int_{\mathbb{R}^+} k^2 q_t(k) dk$ |

Table 1: Fundamental contract specifications and the corresponding pricing formulas based on european OTM options, derived from the replication theorem of Carr and Madan [2001].

One of the challenges faced by issuers of standard variance swaps is to hedge the realised variance through dynamic rebalancing of an options portfolio which is tilted towards the low
Hedge Portfolio

|                | Cash          | Futures       | Puts                                      | Calls                                    |
|----------------|---------------|---------------|-------------------------------------------|------------------------------------------|
| $\Lambda_t$   | $x_0 - 1$     | $+F_0^{-1}F_t$| $- \int_0^{F_0} k^{-2} P_t(k)dk$          | $- \int_{F_0}^{\infty} k^{-2} C_t(k)dk$ |
| $\Upsilon_t$  | $x_0^2 - 2x_0$| $+2x_0F_0^{-1}F_t$| $+2 \int_0^{F_0} (1 - \ln k) k^{-2} P_t(k)dk$ | $+2 \int_{F_0}^{\infty} (1 - \ln k) k^{-2} C_t(k)dk$ |
| $\Psi_t$      | $-F_0$        | $+x_0F_t + F_t$| $+ \int_0^{F_0} k^{-1} P_t(k)dk$          | $+ \int_{F_0}^{\infty} k^{-1} C_t(k)dk$ |
| $\Sigma_t$    | $-F_0^2$      | $+2F_0F_t$    | $+2 \int_0^{F_0} P_t(k)dk$               | $+2 \int_{F_0}^{\infty} C_t(k)dk$ |
| $Q_t$          | $-2F_0^3$     | $+3F_0^2F_t$  | $+6 \int_0^{F_0} kP_t(k)dk$              | $+6 \int_{F_0}^{\infty} kC_t(k)dk$ |
| $H_t$          | $-3F_0^4$     | $+4F_0^3F_t$  | $+12 \int_0^{F_0} k^2P_t(k)dk$           | $+12 \int_{F_0}^{\infty} k^2C_t(k)dk$ |

Table 2: Hedge portfolios for fundamental contracts in terms of cash, futures and European options.

strike options via the weight $k^{-2}$ in the replication formula for the log contract. This factor is also present, though moderated by the factor $(1 - \ln k)$, in the squared-log contract. In the entropy contract, which is another generalized variance swap contract, the weight is only $k^{-1}$, but still this causes problems because it places most weight on very low strike options. These are the illiquid and expensive deep-OTM put options for which demand much exceeds supply during market crashes, because they provide insurance for risk-averse investors. The illiquid market in such options on single-name equities during the financial crisis of 2008-9 is the main reason why equity variance swaps are now focussed mainly on indices, rather than individual stocks. An additional advantage of the new cube and quartic fundamental contracts is that they are tilted more toward high-strike options, the OTM calls where transaction costs are lower and the market is more liquid.

In all cases integration over a continuum of option strikes is necessary, but in practice vanilla options are traded on a relatively small number of discrete strikes. We now introduce a class of strike-discretisation invariant swaps that can be priced and replicated exactly based only on the available vanilla option prices, while preserving the time-discretisation invariance
property. We refer to these swaps as discretisation-invariant (DI) swaps because, like TDI
swaps they have the same fair value, independent of the partition \( \Pi_N \), which is free from
both discrete monitoring and model-specific (e.g. jump) error. But also, their fair-value can
be computed exactly without requiring an integral approximation based on a continuum of
traded option strikes. By the same token, these swaps can be hedged exactly, without having
to replicate (imperfectly) the log, or entropy, or other fundamental contract.

Let \( P := \{ P_t \}_{t \in \Pi} \) and \( C := \{ C_t \}_{t \in \Pi} \) describe the forward price processes of \( d \) vanilla put
options and \( d \) vanilla call options, with identical, traded strikes \( k \), on an underlying futures
\( F \) with maturity \( T \). That is,

\[
P_t := E_t[(k - F_T 1)^+] \quad \text{and} \quad C_t := E_t[(F_T 1 - k)^+]
\]

where \( 1 := (1, \ldots, 1)' \in \mathbb{R}^d \). Assume w.l.o.g. that the traded strikes \( k := (k_1, \ldots, k_d)' \in \mathbb{R}^n \)
are ordered such that \( k_1 < k_2 < \ldots < k_d \), and denote by \( \hat{P} \) and \( \hat{C} \) the increments in \( P \) and
\( C \), respectively, along some partition of \([0, T] \). By Theorem 2, the vector space

\[
\{ f : \mathbb{R}^n \to \mathbb{R} \mid f(\hat{z}) = \alpha' \hat{P} + \hat{P}' \Omega \hat{C} + \gamma' \hat{C}, \ \alpha \in \mathbb{R}^d, \ \Omega \in \mathbb{R}^{d \times d}, \ \gamma \in \mathbb{R}^d \}
\]

contains all TDI characteristics \( f \) for \( z = (P, C) \). The fixed leg of the corresponding swap
can be derived as

\[
E \left[ (P_T - P_0)' \Omega (C_T - C_0) \right] = E \left[ P_T' \Omega C_T \right] - P_0' \Omega C_0.
\]

However, in the case that \( \Omega \) is a lower triangular matrix, we have

\[
E \left[ P_T' \Omega C_T \right] = E \left[ (k' - F_T 1)^+ \Omega (F_T 1 - k)^+ \right] = 0
\]

since the strikes are in ascending order and hence either the put pay-off or the call pay-off
is zero, for each component. Therefore an exact swap rate can be derived only based on the
prices $P_0$ and $C_0$ of traded vanilla options with strikes $k$, without using the replication theorem of Carr and Madan [2001]. Now, by Theorem 3, the value increments of this swap are

$$
\hat{V}_t = [\alpha' - C_{t-1}' \Omega] \hat{P}_t + [\gamma' - P_{t-1}' \Omega] \hat{C}_t.
$$

Hence, the swap can be hedged exactly by dynamically holding $[\alpha' - C_{t-1}' \Omega]_j$ puts and $[\gamma' - P_{t-1}' \Omega]_j$ calls with strike $k_j$ for $j = 1, 2, \ldots, d$.

**Example 6: Straddle and Strangle Swaps.** Let $P := \{P_t\}_{t \in \mathbb{H}}$ and $C := \{C_t\}_{t \in \mathbb{H}}$ describe the forward price processes of a vanilla put option with strike $k_p$ and a vanilla call option with strike $k_c$, i.e. $P_t := \mathbb{E}_t [(k_p - F_t)^+]$ and $C_t := \mathbb{E}_t [(F_t - k_c)^+]$. Since $z = (P, C)'$ follows a $Q$-martingale, $f(\hat{z}) = \hat{P}\hat{C}$ is a TDI characteristic and the corresponding fair-value swap rate is $\mathbb{E}[(P_T - P_0)(C_T - C_0)] = \mathbb{E}[P_TC_T] - P_0C_0$. Under the additional restriction $k_p \leq k_c$, $\mathbb{E}[P_TC_T] = 0$ so the fair value swap rate is simply the negative of the product of the call price and put price. Following Theorem 3, $\hat{V}_{st} = -P_{t-1}\hat{C}_t - C_{t-1}\hat{P}_t$ and the swap can be hedged exactly by dynamically holding $P_{t-1}$ calls and $C_{t-1}$ puts from time $t - 1$ to $t$.

## 4 Risk Premiums

By definition, the expected realised characteristic for a TDI swap is independent of the monitoring frequency under $Q$. However, its expectation under the physical measure $\mathbb{P}$, and therefore the risk premium that it captures, is not. The relationship between model assumption and risk premiums for TDI swaps is an interesting area for future research which goes beyond the scope of this paper. In this section we are content to derive an expression relating the risk premium to the monitoring frequency for a simple TDI variance swap, based on a geometric Brownian motion for the underlying price. To this end, consider the variance swap of Example 1, under the assumption

$$
\text{d}F_t = F_t \sigma \text{d}W_t^Q = F_t (\mu \text{d}t + \sigma \text{d}W_t^P), \quad F_0 > 0,
$$

20
with $\mu \neq 0$ so that the variance risk premium may also be non-zero. Assume further that the swap is monitored under the regular partition $\Pi_N = \{i\tau\}_{i=0,\ldots,N}$ for some $N \in \mathbb{N}_+$, with $\tau = T/N$ so that the monitoring frequency increases as $\tau$ decreases. We compute the expectation of the realised characteristic under $\mathbb{P}$ and examine how the deviation of this expectation from the fair-value swap rate (i.e. the risk premium) depends on $\tau$.

In Example 1 we have $z = F$ and $f(\hat{z}) = \hat{z}^2$ so the fair-value swap rate is $\mathbb{E}^\mathbb{Q} \left[ (F_T - F_0)^2 \right] = F_0^2 \left( e^{\sigma^2 T} - 1 \right) > 0$. In the Appendix we prove that the expected realised characteristic under $\mathbb{P}$ may be expressed in the form:

$$
\mathbb{E}^\mathbb{P} \left[ \sum_{\Pi_N} \hat{F}_i^2 \right] = F_0^2 \times \begin{cases} 
2T (1 - e^{\mu \tau}) \tau^{-1} & \text{if } \kappa = 0, \\
(e^{\kappa T} - 1) \left( 1 - 2 \frac{e^{\mu \tau} - 1}{e^{\kappa \tau} - 1} \right) & \text{otherwise},
\end{cases}
$$

(10)

where $\kappa := 2\mu + \sigma^2$, so that $\kappa = 0$ corresponds to the case where $\ln F$ follows a martingale under the physical measure. Now using Taylor expansion of the exponential terms above and letting $\tau \to 0$ we obtain the following expression for the expected realised characteristic of a continuously monitored variance swap:

$$
\mathbb{E}^\mathbb{P} \left[ \langle F \rangle_T \right] = F_0^2 \times \begin{cases} 
\sigma^2 T & \text{if } \kappa = 0, \\
\frac{\sigma^2}{\kappa} (e^{\kappa T} - 1) & \text{otherwise}.
\end{cases}
$$

The risk premium for a continuously monitored swap may thus be written

$$
\mathbb{E}^\mathbb{P} \left[ \langle F \rangle_T \right] - \mathbb{E}^\mathbb{Q} \left[ (F_T - F_0)^2 \right] = F_0^2 \times \begin{cases} 
\sigma^2 T - \left( e^{\sigma^2 T} - 1 \right) & \text{if } \kappa = 0, \\
\frac{\sigma^2}{\kappa} (e^{\kappa T} - 1) - \left( e^{\sigma^2 T} - 1 \right) & \text{otherwise}.
\end{cases}
$$

(11)

We show in the Appendix that the sign of the risk premium coincides with the sign of $\kappa - \sigma^2$. 

21
Now, to determine the relationship between $\tau$ and the risk premium, consider

\[
\frac{d}{d\tau} \mathbb{E}^p \left[ \sum_n \hat{F}_{i,n}^2 \right] = \begin{cases} 
\frac{2}{\sigma^2} \left( e^{\mu\tau} - \mu \tau e^{\mu\tau} - 1 \right) / \tau^2 & \text{if } \kappa = 0, \\
\frac{2}{\sigma^2} \kappa \frac{(e^{\mu\tau} - 1) ke^{\kappa \tau} - (e^{\kappa \tau} - 1) \mu e^{\mu\tau}}{(e^{\kappa \tau} - 1)^2} & \text{otherwise.}
\end{cases}
\]

(12)

The sign of the derivative coincides with the sign of $|\kappa| - \sigma^2$, as follows from some tedious algebra. That is, for small negative drifts ($-\sigma^2 < \mu < 0$) the risk premium on the discretely monitored variance swap increases with the monitoring frequency, converging to a negative limit. For both $\mu = 0$ and $\mu = -\sigma^2$ the risk premium does not depend on the monitoring frequency, being equal to zero in the first and negative in the second case. For positive or strongly negative drifts ($\mu < -\sigma^2$) the risk premium decreases as the monitoring frequency increases.

5 Empirical Analysis

Our empirical study is based on the daily closing prices $P_t$ and $C_t$ of all traded European put and call options on the S&P 500 between January 1997 and December 2013. We eliminate quotes that fulfil any of the following criteria: less than seven calendar days to maturity, more than 365 calendar days to maturity, zero trading volume, mid-price $\leq 0.5$ or an implied Black Scholes volatility $\leq 1\%$ or $\geq 1$. For each trading day, we further delete all quotes that refer to the same maturity if less than three different strikes are traded. The forward price is backed out via put-call-parity for each maturity from the pair of quotes whose strike minimises $|P_t - C_t|$. This forward price is also used as the separation strike between OTM put and call options, i.e. we use the put price for $k < F_t$ and the call price for $k \geq F_t$.

We then apply a cubic spline smoothing algorithm according to Fengler [2009] in order to preclude static arbitrage in both the strike and maturity dimension (calendar arbitrage). This yields OTM option prices $q_t(k)$ that cover an equally distributed grid of 2000 strikes $k_j$ on a six-$\sigma$-range around the forward price, $\sigma$ being the average implied volatility of the
sample. Outside the domain of the spline we assume the implied volatility to be constant and equal to the implied volatility at the closest strike. These data are integrated numerically w.r.t. $k$ to derive time series of daily prices for the contracts with fixed maturity dates in Table 1. We then calculate trading day, weekly (five trading days) and monthly (20 trading days) value increments (i.e. profit and loss, P&L) for these contracts and the corresponding increments for the log variance, the variance, covariance, third moment, fourth moment, entropy, squared-log and straddle swaps based on Theorem 3 and the example applications.

We aim to provide a large-sample time-series analysis of the properties of TDI and DI swaps, examining risk premiums over a period of 16 years. To this end we convert the contracts with fixed expiry dates into synthetic, constant-maturity contracts that are still investable, i.e. replicable by holding a portfolio of futures and options. For this purpose we follow Galai (1979), applying linear interpolation of the option-price value increments, not the prices themselves, between the two adjacent maturities. We apply a similar constant-maturity transformation to the swap value increments, i.e. the value increment between time $t - 1$ and time $t$ of a contract $\Phi$ with constant time-to-maturity $t_c$ is

$$\hat{\Phi}_t := (T_u - T_t)^{-1} \left[ (T_u - t - t_c) \hat{\Phi}_{lt} - (T_l - t - t_c) \hat{\Phi}_{ut} \right]$$

where $\hat{\Phi}_{lt}$ and $\hat{\Phi}_{ut}$ denote the increments in the prices of the contracts with fixed maturity dates $T_l$ and $T_u$. Now, by definition,

$$\mathbb{E}^Q \left[ \hat{F}_t \right] = \mathbb{E}^Q \left[ \hat{\Lambda}_t \right] = \mathbb{E}^Q \left[ \hat{\Upsilon}_t \right] = \mathbb{E}^Q \left[ \hat{\Psi}_t \right] = \mathbb{E}^Q \left[ \hat{\Sigma}_t \right] = \mathbb{E}^Q \left[ \hat{Q}_t \right] = \mathbb{E}^Q \left[ \hat{H}_t \right] = 0, \quad \forall t \in \Pi,$$

and the same holds for our TDI swaps, since they are portfolios of these contracts. However, under the physical probability measure the expected increments in these contracts and swaps need not be zero, in the presence of a risk premium.

Table 3 shows the average risk premiums for 30-day swaps that are monitored daily.

---

9For example, the entropy contract is approximated by $\Psi_t \approx F_t x_t - \sum_{j=2}^{2000} k_j^{-1} q_t (k_j - k_j - 1)$ and similar approximations apply for $\Lambda$, $\Upsilon$, $\Sigma$, $Q$ and $H$.

10Note that it is only possible to determine a value increment if at least three options with different strikes and the same fixed maturity are quoted on both trading days.

11From now on there is a slight of notation here in that the increments $\hat{F}_t$, $\hat{\Lambda}_t$, etc. refer now to increments in the constant-maturity time series, rather than the fixed maturity series that we have used for developing the theory.
Table 3: Average risk premiums over 16 years, for 30-day constant-maturity contracts based on daily, weekly and monthly monitoring, for: futures (\(F\)), covariance swap (\(V_P\)), third moment swap (\(V_Q\)), fourth moment swap (\(V_H\)), log variance swap (\(V_\Lambda\)), variance swap (\(V_\Sigma\)), squared-log swap (\(V_\Upsilon\)), entropy swap (\(V_\Psi\)) and straddle swaps with strikes 1000, 1100 and 1200.

| \(\Pi\) | \(F\) | \(V_P\) | \(V_Q\) | \(V_H\) | \(V_\Lambda\) | \(V_\Sigma\) | \(V_\Upsilon\) | \(V_\Psi\) | \(1000\) | \(1100\) | \(1200\) |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| \(\Pi_D\) | 0.23  | -0.78 | 0.39  | -0.04 | -0.58 | -0.72 | -0.56 | -0.64 | 0.32  | 0.42  | 0.69  |
| \(\Pi_W\) | 0.22  | -1.07 | 0.51  | -0.04 | -0.95 | -1.04 | -0.94 | -0.99 | 0.60  | 0.58  | 0.97  |
| \(\Pi_M\) | 0.22  | -0.95 | 0.39  | -0.04 | -0.80 | -0.88 | -0.77 | -0.82 | 0.45  | 0.43  | 0.64  |

weekly and monthly over the entire 16-year sample period. Besides the daily partition \(\Pi_D\), we include results for the weekly and monthly partitions, denoted \(\Pi_W\) and \(\Pi_M\). Each premium is computed as the average value increment divided by its standard deviation, and annualised to enable comparison between daily, weekly and monthly increments. The premium for the variance contracts are negative, those for the third moment swaps and straddle swaps are positive, and the fourth-moment risk premium is small and negative, on average.

Figure \[\text{I}\] depicts the cumulative P&L for our 30-day, synthetic but investable, constant-maturity contracts over the sample period. These time series give more information on the dependence of risk premiums on both the characteristic \(f\) that defines the swap and the monitoring frequency of the realised leg. The top-left diagram shows the P&L of our synthetic but investable constant-maturity futures, showing a black line for daily, blue for weekly and red for monthly increments. Clearly, the monitoring frequency has little impact on the performance of this portfolio. By contrast, the performance of the constant-maturity variance swap becomes more positive as the monitoring frequency increases. The risk premium is usually small and negative, the exception being the periods of equity market turmoil surrounding the onset of a crisis, such as the collapse of Lehman Brothers in September 2008 and the breaking news in August 2011 of a European sovereign debt crisis.

By contrast, the P&L on the third-moment swap decreases only during these crisis periods, when the negative skew in realised returns on equity indices becomes pronounced. During crisis periods it is also more negative for daily-monitored third-moment swaps than it is for
weekly or monthly monitored swaps. Otherwise, during stable trending markets it is small and generally positive, particularly since the crash in September 2008, and does not depend much on the monitoring frequency. During the last 5 years of the sample, except during August 2011, the skew in realised equity index has been more positive than the skew in the implied characteristic, which is inferred from option price distributions.

The risk premium on the fourth-moment swap is more variable over time. The returns on this swap are sensitive to downward jumps in the index and the P&L is strongly positive during the crisis periods. It also tends to exhibit a highly variable but generally negative P&L immediately prior to a market crash. For instance, this is evident during the period leading up to the technology crash in year 2000 and during 2007, the period preceding the banking crisis. At these times realised returns on S&P500 futures have significantly less kurtosis than the kurtosis in the distribution implied by option prices, with the implied distribution reflecting a high degree of uncertainty about forward prices. Notice that the
fourth-moment swap has a particularly variable, negative P&L during 2013, the last year of the sample, a period when S&P500 futures prices exhibited a very strong upward trend amid much uncertainty in the global economy.

Figure 2: Time series for the cumulative MtM P&L of the constant-maturity straddle swaps with strikes 1000 ($V_{1000}$), 1100 ($V_{1100}$) and 1200 ($V_{1200}$). Black, blue and red lines refer to swaps with realised characteristics that are monitored on a daily, weekly and monthly basis, respectively.

The risk premium on these swaps can be large and negative during a crisis, e.g. in September 2008 and August 2011. Otherwise, the risk premium is small and positive and greater for straddle swaps that are monitored weekly or monthly than for straddle swaps that are monitored daily.

Table 4 presents correlations between the daily (and weekly and monthly) unrealised P&Ls on the moment and straddle swaps that we have previously examined. In each case the choice of strike here allows us to investigate the behaviour of the swaps over the 16-year sample period because call and put options at these strikes were traded most of the time. We exclude strangle swaps from this analysis since they are more expensive to trade, due to the concentration of liquidity at the money, but results are available from the author on request.
| $\Pi_D$ | moments | variances | straddles |
|---|---|---|---|
| $F$ | 1 | -0.54 | 0.82 | -0.88 | -0.62 | -0.60 | -0.61 | -0.62 | 0.29 | 0.35 | 0.43 |
| $V_F$ | -0.54 | 1 | -0.70 | 0.56 | 0.78 | 0.98 | 0.78 | 0.90 | -0.68 | -0.69 | -0.67 |
| $V_Q$ | 0.82 | -0.70 | 1 | -0.67 | -0.86 | -0.79 | -0.85 | -0.84 | 0.44 | 0.45 | 0.48 |
| $V_H$ | -0.88 | 0.56 | -0.67 | 1 | 0.43 | 0.54 | 0.42 | 0.49 | -0.27 | -0.37 | -0.51 |
| $V_A$ | -0.62 | 0.78 | -0.86 | 0.43 | 1 | 0.89 | 1.00 | 0.97 | -0.55 | -0.47 | -0.40 |
| $V_S$ | -0.60 | 0.98 | -0.79 | 0.54 | 0.89 | 1 | 0.89 | 0.97 | -0.70 | -0.67 | -0.62 |
| $V_T$ | -0.61 | 0.78 | -0.85 | 0.42 | 1.00 | 0.89 | 1 | 0.97 | -0.55 | -0.46 | -0.39 |
| $V_\Psi$ | -0.62 | 0.90 | -0.84 | 0.49 | 0.97 | 0.97 | 0.97 | 1 | -0.66 | -0.59 | -0.51 |
| $V_{1000}$ | 0.29 | -0.68 | 0.44 | -0.27 | -0.55 | -0.70 | -0.55 | -0.66 | 1 | 0.81 | 0.42 |
| $V_{1100}$ | 0.35 | -0.69 | 0.45 | -0.37 | -0.47 | -0.67 | -0.46 | -0.59 | 0.81 | 1 | 0.71 |
| $V_{1200}$ | 0.43 | -0.67 | 0.48 | -0.51 | -0.40 | -0.62 | -0.39 | -0.51 | 0.42 | 0.71 | 1 |

| $\Pi_W$ | moments | variances | straddles |
|---|---|---|---|
| $F$ | 1 | -0.53 | 0.78 | -0.88 | -0.57 | -0.56 | -0.56 | -0.57 | 0.43 | 0.45 | 0.46 |
| $V_F$ | -0.53 | 1 | -0.80 | 0.52 | 0.85 | 0.98 | 0.85 | 0.93 | -0.77 | -0.80 | -0.81 |
| $V_Q$ | 0.78 | -0.80 | 1 | -0.67 | -0.83 | -0.85 | -0.84 | -0.86 | 0.79 | 0.79 | 0.72 |
| $V_H$ | -0.88 | 0.52 | -0.67 | 1 | 0.42 | 0.51 | 0.42 | 0.47 | -0.33 | -0.39 | -0.48 |
| $V_A$ | -0.57 | 0.85 | -0.83 | 0.42 | 1 | 0.93 | 1.00 | 0.98 | -0.77 | -0.72 | -0.62 |
| $V_S$ | -0.56 | 0.98 | -0.85 | 0.51 | 0.93 | 1 | 0.93 | 0.98 | -0.82 | -0.82 | -0.78 |
| $V_T$ | -0.56 | 0.85 | -0.84 | 0.42 | 1.00 | 0.93 | 1 | 0.98 | -0.78 | -0.73 | -0.62 |
| $V_\Psi$ | -0.57 | 0.93 | -0.86 | 0.47 | 0.98 | 0.98 | 0.98 | 1 | -0.83 | -0.80 | -0.72 |
| $V_{1000}$ | 0.43 | -0.77 | 0.79 | -0.33 | -0.77 | -0.82 | -0.78 | -0.83 | 1 | 0.93 | 0.67 |
| $V_{1100}$ | 0.45 | -0.80 | 0.79 | -0.39 | -0.72 | -0.82 | -0.73 | -0.80 | 0.93 | 1 | 0.83 |
| $V_{1200}$ | 0.46 | -0.81 | 0.72 | -0.48 | -0.62 | -0.78 | -0.62 | -0.72 | 0.67 | 0.83 | 1 |

| $\Pi_M$ | moments | variances | straddles |
|---|---|---|---|
| $F$ | 1 | -0.52 | 0.70 | -0.89 | -0.53 | -0.48 | -0.51 | -0.52 | 0.49 | 0.48 | 0.44 |
| $V_F$ | -0.52 | 1 | -0.88 | 0.43 | 0.95 | 0.99 | 0.95 | 0.99 | -0.89 | -0.86 | -0.81 |
| $V_Q$ | 0.70 | -0.88 | 1 | -0.60 | -0.86 | -0.84 | -0.86 | -0.88 | 0.90 | 0.87 | 0.77 |
| $V_H$ | -0.89 | 0.43 | -0.60 | 1 | 0.41 | 0.41 | 0.40 | 0.42 | -0.38 | -0.43 | -0.44 |
| $V_A$ | -0.53 | 0.95 | -0.86 | 0.41 | 1 | 0.89 | 1.00 | 0.99 | -0.91 | -0.78 | -0.64 |
| $V_S$ | -0.48 | 0.99 | -0.84 | 0.41 | 0.89 | 1 | 0.88 | 0.94 | -0.83 | -0.86 | -0.85 |
| $V_T$ | -0.51 | 0.95 | -0.86 | 0.40 | 1.00 | 0.88 | 1 | 0.99 | -0.91 | -0.78 | -0.64 |
| $V_\Psi$ | -0.52 | 0.99 | -0.88 | 0.42 | 0.99 | 0.94 | 0.99 | 1 | -0.92 | -0.84 | -0.73 |
| $V_{1000}$ | 0.49 | -0.89 | 0.90 | -0.38 | -0.91 | -0.83 | -0.91 | -0.92 | 1 | 0.90 | 0.67 |
| $V_{1100}$ | 0.48 | -0.86 | 0.87 | -0.43 | -0.78 | -0.86 | -0.78 | -0.84 | 0.90 | 1 | 0.87 |
| $V_{1200}$ | 0.44 | -0.81 | 0.77 | -0.44 | -0.64 | -0.85 | -0.64 | -0.73 | 0.67 | 0.87 | 1 |

Table 4: Correlations of the MtM P&L on the constant-maturity contracts for daily, weekly and monthly monitoring.
the top-left sub-matrix presents the correlations for the futures, covariance, third-moment
and fourth moment swaps described in Examples 2, 3 and 4 of Section 3; the middle sub-
matrix presents correlations for the log variance swap in Neuberger [2012], the variance swap
of Example 1 and the two generalised variance swaps of Example 5; and the bottom-right
sub-matrix represents correlations for the straddle swaps of Example 6 for the three different
strikes \( k = 1000, 1100 \) and 1200. The other elements of the matrix are cross-correlations
between the different TDI and DI swaps and the S&P500 futures P&L.

The P&L on the covariance swap is negatively correlated with the P&Ls on futures and
the third-moment swap and positively correlated with the P&Ls on the variance and fourth-
moment swaps. As expected, given its fair value decomposition into third-moment and
variance components in Example 2, the strongest negative correlation is with the P&L on
the third-moment swap and the strongest positive correlations are with the variance swaps.
For instance, at the daily frequency \( V_P \) has a correlation of 0.98 with \( V_\Sigma \), 0.78 with \( V_\Lambda \) and
0.90 with \( V_\Psi \). Hence, the risk premium on the covariance swap is already accessible through
trading variance swaps.

As is expected from several previous empirical studies the correlation between the daily
P&L on the futures and the variance swaps is around ~0.6; this decreases (in magnitude)
marginally when measured (and monitored) at the weekly and monthly frequencies. The
P&L on all variance swaps are very highly correlated at every frequency. Thus, as is also
evident from Figure 1 these swaps compensate the investor for downward shocks in the
futures by a strongly positive realised variance.

The third- and fourth-moment swaps are more interesting. As expected there is a strong
positive correlation between the futures P&L and the P&L on the third-moment swap, but
this decreases with frequency so that the monthly P&Ls have a correlation of 0.7. The
empirical features of this third-moment swaps and the one considered by Kozhan et al. [2013]
are very similar. In particular, the third-moment swap may be attractive to variance-swap
issuers as a hedge given its strong positive performance during crisis periods, when large losses
are made on short variance swaps positions. However, given the very high negative correlation between the P&Ls on third-moment and variance swaps no attractive new diversification characteristics have been identified here. By contrast, the fourth-moment swap could be an excellent diversifier of both equity and variance. It has an unusually significant negative correlation with the futures which also is largely independent at frequency. The correlation between the futures P&L and $V_H$ is $-0.88$ at both daily and weekly frequency and $-0.89$ at the monthly frequency. Clearly, fourth-moment swaps have the potential to offer equity diversification, but their risk premium is clearly distinct from the variance premium. For instance, the correlation between $V_H$ and $V_Σ$ is only $0.54$ at the daily frequency, falling to $0.41$ at the monthly frequency. Similar figures can be observed in the correlations with other variance swaps.

A further source of diversification is provided by the straddle swaps. Distinct from the swaps considered so far, they have a relatively low, positive correlations with the futures, a strong negative correlation with variance and a relatively small negative correlation with the fourth-moment swap. Thus, the fourth-moment and straddle swaps appear to access different drivers of return, different from both equity and variance. The optimal diversification weights and the empirical properties of equity portfolios that are diversified with variance, fourth-moment and straddle swaps are and interesting subject for further research.

Until now we have considered swaps with a synthetic, constant maturity of 30 days. At each monitoring period (daily, weekly or monthly) we have re-balanced our position using the investable methodology of Galai [1979] to roll over into another swap with 30-days to maturity. To investigate how the characteristics of these synthetic contracts compare with longer maturities, Table 5 presents the standardized risk premiums on the S&P500 futures contract and the TDI moment swaps of Examples 1, 3 and 4, with realised legs monitored under the daily, weekly and monthly partitions.

The risk premium decreases in magnitude with the maturity of each swap. For instance, the risk premium on the 30-day variance swap $Σ$ is $-0.72$ compared with $-0.46$ for the 180-
Table 5: Average risk premiums for 30-day, 90-day and 180-day to maturity contracts.

| $t_c$ | $F$ | $V_\Sigma$ | $V_Q$ | $V_H$ |
|-------|-----|-------------|-------|-------|
|       | 30  | 90  | 180  | 30  | 90  | 180  | 30  | 90  | 180  |
| $\Pi_D$ | 0.23 | 0.22 | 0.21 | -0.72 | -0.61 | -0.46 | 0.39 | 0.15 | 0.12 | -0.04 | -0.03 | -0.02 |
| $\Pi_W$ | 0.22 | 0.22 | 0.21 | -1.04 | -0.90 | -0.79 | 0.51 | 0.23 | 0.12 | -0.05 | -0.04 | -0.03 |
| $\Pi_M$ | 0.22 | 0.21 | 0.20 | -0.88 | -0.79 | -0.67 | 0.39 | 0.23 | 0.04 | -0.04 | -0.04 | -0.03 |

Figure 3: Time series for the unrealised P&L of monthly-daily frequency swaps on the variance, third-moment, fourth-moment and 1000-strike straddle characteristics. Black, blue and red lines refer to swaps with maturities 30-day, 90-day and 180-day, respectively.

day swap, when monitored at the daily frequency. This finding of different risk premiums even when measured on average over the 16-year period motivates one to consider two types of TDI floating-floating swaps:

(a) **Frequency Swaps**: These exchange two realised legs of the same maturity that are monitored at different frequencies. Conveniently, the AP implies that the fair-value rate on this type of swap is zero, by definition, but the risk premium may be positive or negative depending on the sample period; and
(b) **Calendar Swaps**: These exchange two realised legs of different maturities that are monitored at the same frequency. Since all payments up to the lower maturity cancel out, these swaps give access to forward variance, skewness and kurtosis and therefore the corresponding risk premium term structure. From Table 5, the risk premium is positive for variance calendar swaps, negative for third moment calendar swaps and close to zero (no more than 0.02) for fourth moment calendar swaps when receiving the longer and paying the shorter maturity. The largest absolute premium (-0.39) is paid on a weekly monitored 180-for-30-day third moment swap, indicating strong backwardation of the skewness term structure.

Figure 3 depicts the monthly unrealised P&L on monthly-daily frequency swaps for the variance, third-moment, fourth-moment and 1000-strike straddle swaps of example 1, 3, 4 and 6. The three lines on each graph indicate the maturity of the swaps: black for 30 days, blue for 90 days and red for 180 days. Whether it refers to the 30-day, 90-day or 180-day maturity, the P&L on each frequency swap is very close to zero during the credit boom years 2003–2007. This is indicative of a calm period when realised moments were very close to implied moments all along the term structure, at least up to 180 days. At other times the P&L fluctuates between positive and negative values. It is interesting that the frequency swaps have non-zero P&L during the upward trending equity market towards the end of the sample. By contrast with the credit boom years, greater uncertainty is evident in the options market during this period as implied variance, skewness and kurtosis term structures for the S&P500 vary between contango and backwardation.

Another interesting feature about the fourth-moment frequency swap is that the P&Ls from these swaps are very highly correlated as we move along the term structure, as is evident from the fact that the black, blue and red lines in this chart from Figure 3 almost coincide. A possible explanation for this is that the risk premium on the fourth moment swap is dominated by expectations about jumps in the index, and that such expectations are only held over the short-term. In other words, all the jump risk premium is contained already in the 30-day fourth moment swap.
6 Conclusion

The advantages to a prospective issuer of TDI swaps (and DI swaps in particular) are evident: the residual hedging risks are much smaller than they are for standard variance swaps; the theoretical fair-value swap rate is independent of the monitoring frequency (it is the same for swaps that are monitored daily as for those that are monitored weekly or monthly based on the same realised characteristic); and not only is the fair-value swap rate free of a discrete monitoring error, it is model independent, indeed we can derive it assuming only that the futures price follows a martingale. For DI swaps exact fair values can be computed from traded option prices even without requiring the approximation of numerical integration.

Theorem 1 allows us to find all characteristics which have this property, by solving a second order system of partial differential equations, for any set of deterministic functions of a multivariate martingale process. Theorem 2 focusses on a particular sub-class of these swaps, i.e. those for which the characteristic depends only on a multivariate martingale itself, and its logarithm. In this case, the non-trivial TDI characteristics are quadratic forms of the martingale processes plus exponentials for the log-martingale processes. The rich variety of TDI characteristics that capture third, fourth and higher-moment risk premiums is accessed by letting the underlying multivariate martingale processes contain these conditional moments. Within this variety of time-discretisation invariant (TDI) swaps we have focussed on some special swaps corresponding to second, third and fourth-order moments of a single futures price or log return distribution. These swaps have fair-values that can be computed from those of so-called ‘fundamental contracts’, each of which can be derived from vanilla option prices using the standard replication theorem.

The fair-value computation still has an error, though relatively small, because we must use numerical integration over the option prices at traded strikes to approximate the option-price integral formula for the fair-value of the fundamental contracts. However, a second sub-class of TDI swaps can be defined for which this strike-discretisation error is zero. These
‘discretisation invariant’ swaps have characteristics defined by bilinear forms of traded call and put prices. Again, an infinite variety of such DI swaps exists and we have only examined ‘straddle’ swaps in depth, which are based on a single put and call at the same strike.

Our empirical analysis, spanning a 16-year sample period, demonstrates that a diverse variety of risk premiums are available to trade via these swaps. By contrast with the variance and third-moment swaps introduced by Neuberger [2012], and later analysed empirically by Kozhan et al. [2013], the fourth-moment and straddle-swap risk premiums that we introduce are not highly correlated with the variance risk premium. Indeed, some novel sources of risk become tradable via the creative use of these new swaps and they should be attractive to investors seeking new sources of diversification as equities, commodities and bonds become more highly correlated. Furthermore, the lack of error in the pricing formulas for discretisation-invariant swaps, plus the exact dynamic hedging portfolios that can be used to replicate them, considerably reduce the uncertainties faced by their issuers.

References

Y. Ait-Sahalia, M. Karaman, and L. Mancini. The term structure of variance swaps, risk premia and the expectation hypothesis. Working Paper, 2012.

C. Alexander and S. Leontsinis. Model risk in variance swaps. Working Paper, 2011.

G. Bakshi, N. Kapadia, and D. Madan. Stock return characteristics, skew laws and the differential pricing of individual equity options. Review of Financial Studies, 16(1):101–143, 2003.

C. Bernard and Z. Cui. Prices and asymptotics for discrete variance swaps. Applied Mathematical Finance, 21:140–173, 2014.

D.T. Breeden and R.H. Litzenberger. Prices of state-contingent claims implicit in option prices. Journal of Business, 51(4):621–651, 1978.

M. Broadie and A. Jain. The effect of jumps and discrete sampling on volatility and variance swaps. International Journal of Theoretical and Applied Finance, 11(8):761–979, 2008.

P. Carr and R. Lee. Volatility derivatives. The Annual Review of Financial Economics, 1:1–21, 2009.

P. Carr and D. Madan. Optimal positioning in derivative securities. Quantitative Finance, 1(1):19–37, 2001.
P. Carr and L. Wu. Variance risk premiums. *Review of Financial Studies*, 22(3):1311–1341, 2009.

M. Davis, J. Obloj, and V. Ravel. Arbitrage bounds for weighted variance swap prices. *Mathematical Finance*, forthcoming.

K. Demeterfi, E. Derman, M. Kamal, and J. Zou. A guide to volatility and variance swaps. *Journal of Derivatives*, 6(4):9–32, 1999.

M.R. Fengler. Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9(4):417–428, 2009.

D. Galai. A proposal for indexes for traded call options. *The Journal of Finance*, 34(5):1157–1172, 1979.

R. Hafner and M. Wallmeier. Financial markets and portfolio management. *Journal of Economic Theory*, 22(2):147–167, 2008.

J.M. Harrison and D.M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

D. Hobson and M. Klimmek. Model independent hedging strategies for variance swaps. *Finance and Stochastics*, 16:611–649, 2012.

R. Jarrow, Y. Kchia, M. Larsson, and P. Protter. Discretely sampled variance and volatility swaps versus their continuous approximations. *Finance and Stochastics*, 17:305–324, 2013.

G. Jiang and Y. Tian. The model-free implied volatility and its information content. *Review of Financial Studies*, 18(4):1305–1342, 2005.

R. Kozhan, A. Neuberger, and P. Schneider. The skew risk premium in the equity index market. *Review of Financial Studies*, 2013.

R. Lee. Gamma swaps. *Encyclopedia of Quantitative Finance*, 2010.

I. Martin. Simple variance swaps. *Working Paper*, 2013.

A. Neuberger. The log contract. *The Journal of Portfolio Management*, 20(2):74–80, 1994.

A. Neuberger. Realized skewness. *Review of Financial Studies*, 25(11):3423–3455, 2012.

L.S. Rompolis and E. Tzavalis. Retrieving risk neutral moments and expected quadratic variation from option prices. *Working Paper*, 2013.
A Appendix

A.1 Proof of Theorem 1

Let the forward price process $\mathbf{F}$ follow the $\mathbb{Q}$-dynamics $d\mathbf{F}_t = \sigma_t d\mathbf{W}_t$ where $\sigma = \{\sigma_t\}_{t \in \Pi} \in \mathbb{R}^{d \times d}$ denotes the local volatility matrix and $\mathbf{W} = \{\mathbf{W}_t\}_{t \in \Pi} \in \mathbb{R}^d$ is a multivariate Wiener process with $T^{-1}\langle \mathbf{W} \rangle_t = I$, the identity matrix. Then $d\langle \mathbf{F} \rangle_t = \sigma_t \sigma'_t dt$ is the quadratic covariation process of $\mathbf{F}$.

Let $\Delta := \nabla'_p \mathbf{z} \in \mathbb{R}^{n \times d}$ and $\Gamma := \nabla''_p \Delta \in \mathbb{R}^{n \times d \times d}$ denote the first and second partial derivatives of $\mathbf{z}$ w.r.t. $\mathbf{F}$ where $\nabla_p \hat{=} \left( \frac{\partial}{\partial F_1}, \ldots, \frac{\partial}{\partial F_d} \right)'$. Then, applying Itô’s Lemma and the cyclic property of the trace operator, we have

$$dz_t = \Delta_t d\mathbf{F}_t + \frac{1}{2} \text{tr} (\Gamma_t d\langle \mathbf{F} \rangle_t)$$

and the quadratic covariation process of $\mathbf{z}$ follows the dynamics

$$d\langle \mathbf{z} \rangle_t = \Delta_t \sigma_t \sigma'_t dt.$$ [13]

Since we want the discrete monitoring error to be zero for all possible forward price processes, it must hold in particular for any specific martingale. We can therefore derive a necessary condition for the functions spanning $\mathcal{V}$ by starting from the assumptions that (7) holds w.r.t. $(f, \mathbf{z})$ and that $\mathbf{z}$ follows the dynamics specified in (13).

Let us denote the Jacobian vector of first partial derivatives of $f$ by $\mathbf{J}(\hat{\mathbf{z}}) := \nabla'_{\hat{\mathbf{z}}} f (\hat{\mathbf{z}}) \in \mathbb{R}^n$ and the Hessian matrix of second partial derivatives of $f$ by $\mathbf{H}(\hat{\mathbf{z}}) := \nabla''_{\hat{\mathbf{z}}} \mathbf{J}(\hat{\mathbf{z}}) \in \mathbb{R}^{n \times n}$ where $\nabla_{\hat{\mathbf{z}}} \hat{=} \left( \frac{\partial}{\partial z_1}, \ldots, \frac{\partial}{\partial z_n} \right)'$. Then Itô’s Lemma yields

$$f (\mathbf{z}_T - \mathbf{z}_0) = \int_{\Pi} \mathbf{J}' (\mathbf{z}_t - \mathbf{z}_0) d\mathbf{z}_t + \frac{1}{2} \text{tr} \int_{\Pi} \mathbf{H} (\mathbf{z}_t - \mathbf{z}_0) d\langle \mathbf{z} \rangle_t.$$ [15]

[13]The quadratic covariation is a straightforward generalisation of the quadratic variation for multivariate processes and is defined as $\langle \mathbf{z} \rangle_t := \lim_{\Pi N \rightarrow \Pi} \sum_{\Pi N} \hat{\mathbf{z}}_i \hat{\mathbf{z}}'_i = \int_{\Pi} d\mathbf{z}_t d\mathbf{z}'_t$. Note that the quadratic covariation $\langle \mathbf{z} \rangle$ is a matrix while the $f$-variation $\langle \mathbf{z} \rangle'$ is a scalar.
Applying (15) to each summand in (4) yields

$$\sum_{i=1}^{N} \left\{ \int_{t_{i-1}}^{t_{i}} J' (z_t - z_{t_{i-1}}) \, dz_t + \frac{1}{2} \text{tr} \int_{t_{i-1}}^{t_{i}} H (z_t - z_{t_{i-1}}) \, d\langle z \rangle_t \right\}$$

$$\int_{\Pi} J' (z_t - z_{m(t)}) \, dz_t + \frac{1}{2} \text{tr} \int_{\Pi} H (z_t - z_{m(t)}) \, d\langle z \rangle_t$$

(16)

where $m(t) := \max\{t_i \in \Pi_N | t_i \leq t\}$. Taking the limit as $\Pi_N \to \Pi$ yields the $f$-variation

$$\langle z \rangle_f = \int_{\Pi} J' dz_t + \frac{1}{2} \text{tr} \int_{\Pi} H d\langle z \rangle_t$$

(17)

where $J := J(0)$ and $H := H(0)$. With (15) and (17), the condition (7) is equivalent to

$$\text{E}_Q \left[ \int_{\Pi} \left[ J (z_t - z_0) - J' \Gamma_t + \Delta'_t [H (z_t - z_0) - H] \Delta_t \right] \sigma_t \sigma'_t dt \right] = 0.$$  

(18)

Substituting (13) and (14) in the above, and considering that $\text{E}_Q [dF_t] = 0$, shows that (7) is equivalent to

$$\text{tr} \text{E}_Q \left[ \int_{\Pi} \left\{ [J (z_t - z_0) - J'] \Gamma_t + \Delta'_t [H (z_t - z_0) - H] \Delta_t \right\} \sigma_t \sigma'_t dt \right] = 0.$$  

(19)

Applying a spectral decomposition to the symmetric matrix in curly brackets above yields

$$[J (z_t - z_0) - J'] \Gamma_t + \Delta'_t [H (z_t - z_0) - H] \Delta_t =: E_t A_t E'_t$$

(19)

where $A_t = \text{diag} \{ \lambda^1_t, \ldots, \lambda^d_t \}$ is a diagonal matrix of eigenvalues and $E_t$ is an orthogonal matrix of eigenvectors. In order to derive a necessary condition for (7), following the same logic of refinement as before, we can select a specific local volatility process. Let us now assume that

$$\sigma_t := \exp \left\{ \frac{\xi}{2} E_t A_t E'_t \right\} = E_t \exp \left\{ \frac{\xi}{2} A_t \right\} E'_t$$

where $\xi \in \mathbb{R}$ is an arbitrary constant. Because $\exp \{ YXY^{-1} \} = Y \exp \{ X \} Y^{-1}$ for $X, Y \in \mathbb{R}^d$.
\( \mathbb{R}^{d \times d} \) as long as \( Y \) is invertible, we have

\[
\sigma_t \sigma'_t = E_t \exp \{ \xi A_t \} E'_t. \tag{20}
\]

Inserting (19) and (20) into (18) and differentiating w.r.t. \( T \), then using the cyclic property of the trace yields

\[
E^Q [\text{tr} (A_t \exp \{ \xi A_t \})] = 0.
\]

Differentiating once w.r.t. \( \xi \) and evaluating the equation at \( \xi = 0 \) yields the condition

\[
E^Q [\text{tr} (A_t^2)] = \sum_{i=1}^{d} E^Q [(\lambda_i^t)^2] = 0,
\]

which implies that all eigenvalues in \( A_t \) must be equal to zero. Hence we know that both sides in (19) are zero and, given that this must hold for all \( F_t \) and \( z_0 \), we can write

\[
[J (\hat{z}) - J] ' \Gamma + \Delta ' [H (\hat{z}) - H] \Delta = 0 \tag{21}
\]

where \( F \) and \( \hat{z} \) are independent variables. We have derived this \( d \times d \) system of partial differential equation based on the assumption that \( F \) follows a particular local volatility process, so it represents a necessary condition for the more general case where \( F \) can be any martingale diffusion. However, since (21) is also sufficient for (18) to hold, the two conditions are equivalent.

A.2 Proof of Corollary 1

The proof of Theorem 1 can be performed analogously for the AP by substituting (15) and (16) into condition (8) and yields the same solution (21). This version does not require the existence of the \( f \)-variation.
A.3 Proof of Theorem 2

When \( z = (F, x)' \) we have \( \Delta(F) = (I, \text{diag}(F)^{-1})' \in \mathbb{R}^{2d \times d} \) and \( \Gamma(F) = (0, -\text{diag}_3(F)^{-2})' \in \mathbb{R}^{2d \times d \times d} \) where \( \text{diag}_3(F) \) denotes a three dimensional tensor with the elements of \( F \) on the diagonal and zeros everywhere else. We shall further use the following decompositions:

\[
\begin{bmatrix}
J(\hat{z}) - J(0)
\end{bmatrix} =
\begin{pmatrix}
J_F(\hat{z}) & J_x(\hat{z}) \\
J_F(\hat{z}) & J_x(\hat{z})
\end{pmatrix} \in \mathbb{R}^{2d}
\]

and

\[
\begin{bmatrix}
H(\hat{z}) - H(0)
\end{bmatrix} =
\begin{pmatrix}
H_F(\hat{z}) & G(\hat{z}) \\
G(\hat{z}) & H_x(\hat{z})
\end{pmatrix} \in \mathbb{R}^{2d \times 2d}.
\]

Then (21) may be written:

\[
- J_x(\hat{z})' \text{diag}(F)^{-2} + H_F(\hat{z}) + G(\hat{z}) \text{diag}(F)^{-1}
+ \text{diag}(F)^{-1} G(\hat{z})' + \text{diag}(F)^{-1} H_x(\hat{z}) \text{diag}(F)^{-1} = 0
\]

and multiplying from left and right with \( \text{diag}(F) \) yields

\[
- \text{diag}(J_x(\hat{z})) + \text{diag}(F) H_F(\hat{z}) \text{diag}(F)
+ \text{diag}(F) G(\hat{z}) + G(\hat{z})' + \text{diag}(F)^{-1} H_x(\hat{z}) = 0.
\]

Since this condition must be fulfilled for all martingale Itô processes \( F \) (and for \( F = 1c \) in particular) this implies \( H_F(\hat{z}) = G(\hat{z}) = 0 \) as well as \( H_x(\hat{z}) = \text{diag}(J_x(\hat{z})) \). Therefore the solution must be of the form

\[
f(\hat{z}) = \alpha' \hat{z} + \hat{F}' \Omega \hat{F} + \beta'(e^x - 1), \quad \alpha \in \mathbb{R}^{2d}, \quad \Omega \in \mathbb{R}^{d \times d}, \quad \beta \in \mathbb{R}^d.
\]

Swaps associated with \( \alpha \) are TDI since \( \lim_{\Pi_N \to \Pi} \sum_{\Pi_N} \alpha' \hat{z}_i = \alpha' (z_t - z_0) \) even without
expectation for any process. For the swaps associated with $\Omega$ we can apply

$$
\mathbb{E}^Q \left[ \lim_{\Pi_N \to \Pi} \sum_{\Pi_N} \hat{\mathbf{F}}_i' \Omega \hat{\mathbf{F}}_i \right] = \mathbb{E}^Q \left[ \lim_{\Pi_N \to \Pi} \sum_{\Pi_N} \text{tr} \left( \Omega \hat{\mathbf{F}}_i' \hat{\mathbf{F}}_i' \right) \right]
$$

$$
= \text{tr} \mathbb{E}^Q \left[ \Omega \lim_{\Pi_N \to \Pi} \sum_{\Pi_N} (\mathbf{F}_{t_i} - \mathbf{F}_{t_{i-1}}) (\mathbf{F}_{t_i} - \mathbf{F}_{t_{i-1}})' \right]
$$

$$
= \text{tr} \mathbb{E}^Q \left[ \Omega \lim_{\Pi_N \to \Pi} \sum_{\Pi_N} (\mathbf{F}_{t_i} \mathbf{F}_{t_i}' - \mathbf{F}_{t_{i-1}} \mathbf{F}_{t_{i-1}}') \right]
$$

$$
= \text{tr} \mathbb{E}^Q \left[ \Omega (\mathbf{F}_T \mathbf{F}_T' - \mathbf{F}_0 \mathbf{F}_0') \right]
$$

$$
= \text{tr} \mathbb{E}^Q \left[ \Omega (\mathbf{F}_T - \mathbf{F}_0) (\mathbf{F}_T - \mathbf{F}_0)' \right]
$$

$$
= \mathbb{E}^Q \left[ (\mathbf{F}_T - \mathbf{F}_0)' \Omega (\mathbf{F}_T - \mathbf{F}_0) \right],
$$

where the only requirement is that $\mathbf{F}$ follows a martingale (not necessarily an Itô process).

Finally, for all swaps associated with $\beta$ we have

$$
\mathbb{E}^Q \left[ \lim_{\Pi_N \to \Pi} \sum_{\Pi_N} \beta' (e^{\hat{x}} - 1) \right] = \mathbb{E}^Q \left[ \beta' (e^{\hat{x}_T - \hat{x}_0} - 1) \right] = 0.
$$

Therefore, if $\mathbf{z} = (\mathbf{F}, \mathbf{x})'$, the necessary condition (21) is sufficient for all martingales. \qed
A.4 Proof of Theorem 3

With the value process of a TDI swap being defined as

\[ V_t := E_t^Q \left[ \sum_{i=1}^N f \left( \hat{z}_i \right) \right] - E_{t-1}^Q \left[ \sum_{i=1}^N f \left( \hat{z}_i \right) \right], \]

the increments of the swap along the partition \( \Pi_N \) are given by

\[
\hat{V}_i = V_i - V_{i-1} = E_{t_i}^Q \left[ \sum_{i=1}^N f \left( \hat{z}_i \right) \right] - E_{t_{i-1}}^Q \left[ \sum_{i=1}^N f \left( \hat{z}_i \right) \right]
\]

\[
= \sum_{i=1}^i f \left( \hat{z}_i \right) + E_{t_i}^Q \left[ \sum_{i=i+1}^N f \left( \hat{z}_i \right) \right] - E_{t_{i-1}}^Q \left[ \sum_{i=i}^N f \left( \hat{z}_i \right) \right]
\]

\[
= f \left( \hat{z}_i \right) + E_{t_i}^Q \left[ f \left( z_T - z_{t_i} \right) \right] - E_{t_{i-1}}^Q \left[ f \left( z_T - z_{t_{i-1}} \right) \right]
\]

where \( \hat{u}_i = u_{t_i} - u_{t_{i-1}} \) and \( u_t = E_t^Q \left[ f \left( z_T - z_t \right) \right] \). Now

\[
\hat{u}_i = E_{t_i}^Q \left[ \alpha' \left( z_T - z_{t_i} \right) + \left( F_T - F_{t_i} \right)' \Omega \left( F_T - F_{t_i} \right) + \beta' \left( e^{x_T - x_{t_i}} - 1 \right) \right]
\]

\[-E_{t_{i-1}}^Q \left[ \alpha' \left( z_T - z_{t_{i-1}} \right) + \left( F_T - F_{t_{i-1}} \right)' \Omega \left( F_T - F_{t_{i-1}} \right) + \beta' \left( e^{x_T - x_{t_{i-1}}} - 1 \right) \right]
\]

\[
= E_{t_i}^Q \left[ \alpha' z_T + F_T' \Omega F_T \right] - \alpha' z_{t_i} - F_{t_i}' \Omega F_{t_i}
\]

\[-E_{t_{i-1}}^Q \left[ \alpha' z_T + F_T' \Omega F_T \right] + \alpha' z_{t_{i-1}} + F_{t_{i-1}}' \Omega F_{t_{i-1}}
\]

\[
= \hat{U}_i - \alpha' \hat{z}_i - F_{t_i}' \Omega F_{t_i} + F_{t_{i-1}}' \Omega F_{t_{i-1}}
\]

where \( \hat{U}_i = U_{t_i} - U_{t_{i-1}} \) and \( U_t = E_t^Q \left[ \alpha' z_T + F_T' \Omega F_T \right] \). Therefore

\[
\hat{V}_i = \alpha' \hat{z}_i + \left( F_{t_i} - F_{t_{i-1}} \right)' \Omega \left( F_{t_i} - F_{t_{i-1}} \right) + \beta' \left( e^{\hat{x}_i} - 1 \right) + \hat{u}_i
\]

\[
= \hat{U}_i - F_{t_{i-1}}' \left( \Omega' + \Omega \right) \hat{F}_{t_i} + \beta' \left( e^{\hat{x}_i} - 1 \right). \quad \Box
\]
A.5 Risk Premiums

Let the forward price follow the dynamics

\[ dF_t = F_t \sigma dW_t^Q = F_t \left( \mu dt + \sigma dW_t^P \right), \quad F_0 > 0, \]

for some \( \sigma > 0 \) and consider the regular partition \( \Pi_N = \{ i\tau \}_{i=1}^{N} \) with \( \tau = T/N \) for some \( N \in \mathbb{N}_+ \). Recall that \( \mathbb{E}^P[F_t] = F_t e^{\mu(t-r)} \) for \( r \leq t \) and \( \mathbb{E}^P[F^2_t] = F^2_0 e^{\kappa t} \) with \( \kappa = 2\mu + \sigma^2 \), where setting \( \mu = 0 \) yields the corresponding expectations under the risk-neutral measure.

The fair-value swap rate of the TDI variance swap of Example 1 is

\[ \mathbb{E}^Q \left( \sum_{i\in \Pi_N} \hat{F}^2_i \right) = \mathbb{E}^Q \left[ (F_T - F_0)^2 \right] = \mathbb{E}^Q \left[ F^2_T - F^2_0 \right] = F^2_0 \left( e^{\sigma^2 T} - 1 \right). \]

The expectation under \( \mathbb{P} \) of the discretely monitored realised characteristic is

\[
\mathbb{E}^P \left[ \sum_{i\in \Pi_N} \hat{F}^2_i \right] = \sum_{i=1}^{N} \mathbb{E}^P \left[ (F_{i\tau} - F_{(i-1)\tau})^2 \right] = \sum_{i=1}^{N} \mathbb{E}^P \left[ F^2_{i\tau} - 2F_{i\tau}F_{(i-1)\tau} + F^2_{(i-1)\tau} \right]
\]

\[
= \sum_{i=1}^{N} \mathbb{E}^P \left[ F^2_{i\tau} - (2e^{\mu\tau} - 1) F^2_{(i-1)\tau} \right] = F^2_0 \sum_{i=1}^{N} \left( e^{\kappa\tau} - (2e^{\mu\tau} - 1) e^{\kappa(i-1)\tau} \right)
\]

\[
= F^2_0 \left( e^{\kappa\tau} - 2e^{\mu\tau} + 1 \right) \sum_{i=1}^{N} e^{\kappa(i-1)\tau}
\]

By multiplying each element of the original sum with \( (e^{\kappa\tau} - 1)^{-1} (e^{\kappa\tau} - 1) \) in the case \( \kappa \neq 0 \), this can be simplified to

\[
\mathbb{E}^P \left[ \sum_{i\in \Pi_N} \hat{F}^2_i \right] = F^2_0 \times \begin{cases} 
2T \left( 1 - e^{\mu\tau} \right) \tau^{-1} & \text{if } \kappa = 0, \\
(e^{\kappa T} - 1) \left( 1 - 2e^{\mu\tau} - 1 \right) & \text{otherwise.}
\end{cases}
\]
The risk premium on a continuously monitored swap

\[
\mathbb{E}^P \left[ (F)_{T} \right] - \mathbb{E}^Q \left[ (F_T - F_0)^2 \right] = F_0^2 \times \begin{cases} 
\sigma^2 T - \left( e^{\sigma^2 T} - 1 \right) & \text{if } \kappa = 0, \\
\frac{\sigma^2}{\kappa} (e^{\kappa T} - 1) - \left( e^{\sigma^2 T} - 1 \right) & \text{otherwise},
\end{cases}
\]  

(22)

has the same sign as \( \kappa - \sigma^2 \). For \( \kappa = 0 \) we can apply the inequality \( \ln (1 + \sigma^2 T) < \sigma^2 T \) to show that the risk premium is negative. If \( \kappa \neq 0 \) we observe that the risk premium is zero for \( \kappa = \sigma^2 \) (i.e. \( \mu = 0 \)). Due to the dominance of exponential growth (decay) the risk premium is positive (negative) for \( \kappa < \sigma^2 \) (\( \kappa > \sigma^2 \)).

In order to determine the asymptotic of the risk premium as \( \tau \rightarrow 0 \) we look at

\[
\frac{\mathbb{E}^P \left[ \sum_{i} \hat{F}_i^2 \right]}{\mathbb{E}^P \left[ (F)_{T} \right]} = \begin{cases} 
\frac{2}{\sigma^2} (1 - \mu \tau) \tau^{-1} & \text{if } \kappa = 0, \\
\frac{\kappa}{\sigma^2} (1 - \frac{e^{\mu \tau} - 1}{e^{\mu \tau} - 1}) & \text{otherwise}.
\end{cases}
\]

Taking the derivative w.r.t. \( \tau \) yields

\[
\frac{d}{d\tau} \frac{\mathbb{E}^P \left[ \sum_{i} \hat{F}_i^2 \right]}{\mathbb{E}^P \left[ (F)_{T} \right]} = \begin{cases} 
\frac{2}{\sigma^2} \frac{(e^{\mu \tau} - \mu \tau e^{\mu \tau} - 1) / \tau^2}{(e^{\mu \tau} - 1) \mu e^{\mu \tau}} & \text{if } \kappa = 0, \\
\frac{2}{\sigma^2} \frac{\kappa (e^{\mu \tau} - 1) \kappa e^{\mu \tau} - (e^{\mu \tau} - 1) \mu e^{\mu \tau}}{(e^{\mu \tau} - 1)^2} & \text{otherwise},
\end{cases}
\]

which has the same sign as \( |\kappa| - \sigma^2 \). It is easy to verify that the derivative is zero for the case \( \kappa = \sigma^2 \), which corresponds to \( \mu = 0 \). For \( \kappa = 0 \) we apply the inequality \( \ln (1 - \mu \tau) < -\mu \tau \) to show that the derivative is negative. For all other cases we consider the monotonic function \( (1 - e^{\mu \tau}) / x \), with a removable discontinuity at the origin, to argue that \( (1 - e^{\mu \tau}) \kappa = (1 - e^{\mu \tau}) \mu \) if and only if \( \kappa = \mu \), which corresponds to the case \( \kappa = -\sigma^2 \) where the derivative again equals zero. The conclusion about the sign of the derivative follows from continuity.