Do Equity and Options Markets Agree about Volatility?*

Carsten H. Chong[†] Viktor Todorov[‡]

Abstract

We address this question by deriving tight pricing kernel restrictions from zero-date options, which are options that expire on the same day they are traded. These restrictions concern the volatility of small and frequent asset price moves that the equity and options markets must agree on in a frictionless economy where the two markets are integrated. We show that violations of such restrictions lead to local arbitrage opportunities that can be exploited using a static portfolio of zero-date options and a dynamic position in the underlying asset. These local arbitrage opportunities are characterized by arbitrarily high reward-to-risk ratios and cause local explosion of conditional moments of the aggregate pricing kernel. Empirically, we find no evidence of such local arbitrage opportunities. Thus, in spite of the nontrivial risk premium embedded in zero-date options, their prices correctly reflect the time-varying volatility of the underlying asset.

Keywords: arbitrage; market segmentation; options; pricing kernel; risk-neutral probability; stochastic discount factor; stochastic volatility.

1 Introduction

Option prices contain rich information about the pricing of risk, particularly for equilibrium models that feature volatility and jump risks in asset prices. Earlier work has documented various puzzling features of observed option data relative to standard asset pricing models. These include the U-shape pattern of the pricing kernel when projected onto the asset return space as well as the large returns, compared to investing in the stock market, from selling puts.[1] To rationalize the observed option data, the literature has primarily adopted two different approaches. The conventional approach involves adding more complicated features to the underlying asset dynamics and/or generalizing investors' risk preferences.[2] An alternative approach is to view the equity and options markets as partially segmented (Bates (2022)).

In the partially segmented story, the law of one price need not hold, and option prices might reflect arbitrage opportunities that cannot be fully eliminated due to factors such as trading restrictions or capital constraints. Evidence that a significant proportion of option trading is

*We would like to thank Nicola Fusari for the option pricing codes used in the Monte Carlo and Chun He for help with extracting the data used in the empirical analysis. We would also like to thank Caio Almeida, Torben Andersen and Gustavo Freire for helpful comments and suggestions.

†Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, e-mail: carstenchong@ust.hk.

‡Department of Finance, Northwestern University, e-mail: v-todorov@kellogg.northwestern.edu.

[1] See e.g., Jackwerth (2000), Aït-Sahalia and Lo (2000) and Rosenberg and Engle (2002) as well as Coval and Shumway (2001), Bondarenko (2014) and Orłowski et al. (2024).

[2] Examples of this line of work include Bollerslev et al. (2009), Drechsler and Yaron (2011), Du (2011), Seo and Wachter (2019), Eraker and Yang (2022) and Fournier et al. (2024).

due to retail investors, who prefer cheap speculative investments and lose money on average (Bryzgalova et al. (2023)), further reinforces the plausibility of such a market segmentation hypothesis.[3] If this turns out to be true, then any effort to rationalize the underlying and option markets jointly should take this segmentation into account.

The question whether there exists an aggregate pricing kernel that can rationalize observed prices in both the equity and options markets is a fundamental one in financial economics. It is well known that a joint pricing kernel, if it exists, implies the absence of arbitrage opportunities and reflects the marginal utility of wealth of investors in a frictionless market. Characterizing the properties of the pricing kernel is the cornerstone of modern asset pricing research, see e.g., the works of Hansen and Jagannathan (1991), Alvarez and Jermann (2005), Hansen and Scheinkman (2009) and Christensen (2017), as well as the many references therein.

In this paper, we address the question whether the equity and options markets are integrated or segmented by analyzing high-frequency data on equity prices and prices of options that expire on the same day they are traded. These options are called *zero-date options*, or *0DTEs* for short.[4] Our focus on options with the shortest times to maturity brings a distinct advantage. Mainly, due to potentially complicated underlying asset dynamics, classical no-arbitrage conditions[5] cannot provide sufficiently tight bounds for options with nontrivial time to expiration, e.g., monthly options. This has led Cochrane and Saa-Requejo (2000) to propose "no good deal bounds" instead, that is, to form restrictions for option prices based on ruling out trading strategies that can generate very high Sharpe ratios. These bounds have been extended by Bernardo and Ledoit (2000) by ruling out the existence of strategies with high gain-loss ratios.[6]

In this paper, we show that the concepts of no arbitrage and no good deals converge and become essentially equivalent for portfolios with an investment horizon that shrinks to zero. In other words, market segmentation corresponds to the presence of *local arbitrage* opportunities, which we define as good deals that can be realized via trading in the underlying asset and options with a shrinking time to maturity.

We start by recasting the existence of a joint pricing kernel for equity and option prices as a requirement that the statistical probability measure, $\mathbb{P}$, and the risk-neutral probability measure, $\mathbb{Q}$ (under which option prices are determined), agree about the small and frequent moves in the asset prices. These can come from diffusive moves but also from small and frequent jumps in the asset prices. As a result, market integration imposes restrictions on the diffusion coefficient in the asset price (spot volatility) and the small jumps under $\mathbb{P}$ and $\mathbb{Q}$. The law of the big jumps, on the other hand, remains essentially unrestricted and may differ under $\mathbb{P}$ and $\mathbb{Q}$.

In order to draw conclusions about market integration from equity and option prices, it is therefore crucial to consider option portfolios that can separate big jumps from small jumps. A

[3] Previous literature detailing occurrences of arbitrage opportunities in other markets include Pontiff (1997) for mutual funds, Shleifer and Vishny (1997) and Du et al. (2018) for currencies and Lamont and Thaler (2003) for stocks. Retail investors' tendency to engage in risky and losing investments has also been documented by Barber and Odean (2001) and Barberis and Huang (2008) for the stock market and by Xiong and Yu (2011) and Boyer and Vorkink (2014) for derivatives markets.

[4] 0DTEs have increased sharply in popularity recently. Currently, nearly half of the trading volume of all options written on the S&P 500 index that are traded on the CBOE options exchange is in the zero-date maturity domain, see the *Financial Times* article "Zero-day options now account for 50% of S&P options volumes" from September 25, 2023.

[5] See e.g., Harrison and Kreps (1979), Harrison and Pliska (1981), Delbaen and Schachermayer (1994) and the books by Duffie (2010) and Back (2010) for the relation between no-arbitrage (and extensions of this concept) and the existence of a pricing kernel.

[6] Almeida and Freire (2022) sharpen the "no good deal bounds" of Cochrane and Saa-Requejo (2000) by further imposing some economically-motivated (but nevertheless model-based) restrictions on the set of possible stochastic discount factors.

number of earlier studies have studied the proximity of option-implied volatility measures (e.g., the VIX, CBOE's volatility index) and historical volatility measures. Since the horizon of the options used for such measures is nontrivial (e.g., 30 days for the VIX), one typically finds a gap between the two, which is due to the so-called variance risk premium.[7] Without jumps in asset prices, if there is a pricing kernel relating $\mathbb{P}$ and $\mathbb{Q}$, this gap will disappear as the time-to-maturity of the options shrinks to zero. This, however, is no longer true if there are jumps in the asset price, and there is plenty of empirical evidence for the latter. In particular, the presence of jumps drives a wedge between $\mathbb{P}$ and $\mathbb{Q}$ measures of total return variance, even over short time horizons, and hence such a gap is no longer an indicator for market segmentation. Empirically, we also find a gap between a total measure of volatility in the spirit of the VIX, computed from 0DTEs, and historical volatility.[8]

Our proposal, therefore, is to consider short-horizon *truncated volatility* instead, which removes the contribution to price volatility due to big and infrequent asset moves. Risk-neutral truncated volatility can be spanned from a portfolio of zero-date options with different strikes in a manner similar to the construction of the VIX index but with significantly less weight assigned to deep out-of-the-money options. Truncated volatility under $\mathbb{P}$, on the other hand, can be recovered from high-frequency records of the underlying price in a short time window prior to observing the options. Market integration then boils down to an agreement between option-implied and return-based truncated volatilities in the limit as the tenor of the considered options shrinks to zero. This is something that can be easily assessed empirically.[9]

By contrast, if there is a gap between truncated volatilities under $\mathbb{P}$ and $\mathbb{Q}$, we show that this gives rise to local arbitrage opportunities involving the options and the underlying asset markets. These, being good deals over a shrinking time interval, in turn imply that a joint pricing kernel either does not exist or is highly variable and oscillatory, e.g., with an infinite conditional variance. In order to construct local arbitrage portfolios that can exploit gaps between $\mathbb{P}$- and $\mathbb{Q}$-truncated volatilities, we enter a long or a short position in the option portfolio spanning $\mathbb{Q}$-truncated volatility and a dynamic position in the underlying asset until the time to expiration of the options. This dynamic portfolio in the underlying asset has a zero cost and eliminates the directional risk that accumulates in the option portfolio until the expiration date.

Our empirical analysis shows that the answer to the question posed in the title is "Yes": Equity and options markets do agree about volatility due to small and frequent moves in the S&P 500 market index. This is in spite of the nontrivial risk premium embedded in zero-date options. Thus, the rich information contained in these options can be readily used for studying properties of the aggregate pricing kernel. From an econometric perspective, our empirical results justify using estimators of volatility based on short-dated options which are significantly more efficient than their return-based counterparts, see e.g., Andersen et al. (2017), Todorov (2019) and Bandi et al. (2023b), among others.

The rest of the paper is organized as follows. We establish the no-arbitrage conditions relating the $\mathbb{P}$ and $\mathbb{Q}$ dynamics for the asset price in Section 2. We introduce the notion of local arbitrage in Section 3 and of truncated volatility in Section 4. Section 5 shows how to construct local arbitrage portfolios of options and the underlying asset that exploit differences in perceptions about volatility in the equity and options markets due to small and frequent

[7] See e.g., Bakshi and Kapadia (2003), Bollerslev et al. (2009), Carr and Wu (2009), Bollerslev and Todorov (2011) and Drechsler and Yaron (2011).

[8] This is also in line with the empirical evidence reported in Bandi et al. (2023a), Vilkov (2023) and Almeida et al. (2024) concerning the variance risk premium in zero-date options.

[9] By contrast, estimates of the variance risk premium need a dynamic model for volatility in order to form the expectations of future volatility under $\mathbb{P}$, and this can have a nontrivial impact on the variance risk premium estimates as documented in Bekaert and Hoerova (2014).

asset moves. Section 6 makes the local arbitrage portfolios of Section 5 feasible in practice. We evaluate the various volatility-related trading strategies in Section 7 on simulated data. Section 8 contains our empirical application. Section 9 concludes. The Appendix contains the derivation of the theoretical results in the paper as well as additional technical details about the implementation of the volatility strategies.

2 Asset Price Dynamics and Market Integration

We consider an asset on a finite time interval $[0, \tau]$, where $\tau > 0$. Its price at time t is denoted by X_t and is defined on a standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a right-continuous filtration $(\mathcal{F}_t)_{t\in[0,\tau]}$. Under the physical probability measure $\mathbb{P}$, we assume that the asset price evolves according to a general semimartingale process given by

$$\frac{dX_t}{X_{t-}} = \alpha_t dt + \sigma_t^{\mathbb{P}} dW_t + \int_{\mathbb{R}} (e^z - 1)\widetilde{\mu}^{\mathbb{P}}(dt, dz), \tag{2.1}$$

where W_t is a standard $\mathbb{P}$-Brownian motion, μ is an integer-valued measure on $[0, \tau] \times \mathbb{R}$ that counts the jumps in the log-asset price, $\nu^{\mathbb{P}}(dt, dz) = \nu_t^{\mathbb{P}}(dz)dt$ is the jump compensator measuring the intensity of jumps of different sizes, and $\widetilde{\mu}^{\mathbb{P}} = \mu - \nu^{\mathbb{P}}$ is the associated martingale jump measure.

We further consider standard European-style options written on this asset. As the goal of this paper is to examine whether the equity and options markets are integrated, we are interested in the family $\mathcal{Q}$ of all risk-neutral probability measures $\mathbb{Q}$ that can jointly price the underlying asset and the associated options, which means that $\mathcal{Q}$ is the class of probability measures $\mathbb{Q}$ such that both asset and option prices are martingales under $\mathbb{Q}$ if discounted at the risk-free rate. For simplicity of exposition, and because we only work with short time horizons, we set the risk-free rate and the dividend yield of the underlying asset to zero throughout the rest of the paper. In this case, a risk-neutral measure $\mathbb{Q} \in \mathcal{Q}$ satisfies

$$X_t = \mathbb{E}_t^{\mathbb{Q}}(X_{t+T}), \quad P_{t,T}(K) = \mathbb{E}_t^{\mathbb{Q}}(K - X_{t+T})^+ \text{ and } C_{t,T}(K) = \mathbb{E}_t^{\mathbb{Q}}(X_{t+T} - K)^+, \tag{2.2}$$

for $K > 0$ and $t, T \geq 0$ with $t + T \leq \tau$, where $P_{t,T}(K)$ and $C_{t,T}(K)$ denote the prices at time t of a put and a call option, respectively, with strike K and time to maturity T.[10]

The risk-neutral measures $\mathbb{Q}$ may differ but they must all generate the set of observable stock prices and options, that is, satisfy (2.2) above. If we assume that we observe the options on a continuum of strikes covering $(0, \infty)$, then due to a well-known result by Breeden and Litzenberger (1978), all risk-neutral measures $\mathbb{Q}$ generate the same $\mathcal{F}_t$-conditional distribution of X_{t+T}.

Within the set $\mathcal{Q}$, of special interest are the ones for which the $\mathcal{F}_t$-conditional $\mathbb{Q}$-distribution of X_{t+T} is equivalent to its $\mathbb{P}$-counterpart. This is because such equivalence implies the existence of an aggregate pricing kernel that can rationalize the observed stock price and the options written on it at time t. In general, $\mathbb{P}$-$\mathbb{Q}$ equivalence of the $\mathcal{F}_t$-conditional distribution of X_{t+T} does not imply the equivalence of the $\mathbb{P}$ and $\mathbb{Q}$ probability measures. However, since we let $T \downarrow 0$ in our asymptotic setup later on, we analyze the slightly stronger condition of the equivalence of the $\mathbb{P}$ and $\mathbb{Q}$ probability measures. As we show later, the violation of $\mathbb{P}$-$\mathbb{Q}$ equivalence of the probability measures results in the existence of so-called local arbitrage opportunities.

On an intuitive level, the $\mathbb{P}$-$\mathbb{Q}$ equivalence means that the $\mathcal{F}_\tau$-measurable events with zero probability under $\mathbb{P}$ should be precisely those events that are also assigned zero probability under

[10] A put/call option with these parameters is the right to sell/buy the asset at time $t + T$ at price K.

$\mathbb{Q}$. We use $\mathcal{Q}_e$ to denote the subset of risk-neutral measures in $\mathcal{Q}$ that are equivalent to $\mathbb{P}$ on $\mathcal{F}_\tau$. If $\mathbb{Q} \in \mathcal{Q}_e$, the asset price X_t is also a semimartingale process under $\mathbb{Q}$, with dynamics given by

$$\frac{dX_t}{X_{t-}} = \sigma_t^{\mathbb{Q}} dW_t^{\mathbb{Q}} + \int_{\mathbb{R}} (e^z - 1)\widetilde{\mu}^{\mathbb{Q}}(dt, dz), \tag{2.3}$$

where $W_t^{\mathbb{Q}}$ is $\mathbb{Q}$-Brownian motion, the jump compensator of μ under $\mathbb{Q}$ is denoted by $\nu^{\mathbb{Q}}$ and $\widetilde{\mu}^{\mathbb{Q}} = \mu - \nu^{\mathbb{Q}}$ is the martingale jump measure under $\mathbb{Q}$. Note that the drift in the dynamics of X under $\mathbb{Q}$ is absent because X is a martingale under $\mathbb{Q}$.

By Girsanov's theorem, see e.g., Theorem III.3.24 in Jacod and Shiryaev (2003), the equivalence of $\mathbb{Q}$ and $\mathbb{P}$ on $\mathcal{F}_\tau$ further implies that

$$\sigma_t^{\mathbb{Q}} = \sigma_t^{\mathbb{P}} \quad \text{and} \quad \nu^{\mathbb{Q}}(dt, dz) = Y_t(z)\nu_t^{\mathbb{P}}(dz)dt \;\; \text{a.s.} \;\; \text{for } t \in [0, \tau], \tag{2.4}$$

where $Y_t(z)$ is a nonnegative predictable function such that

$$\int_{|z|<1} |z(Y_t(z) - 1)|\nu_t^{\mathbb{P}}(dz) < \infty \;\; \text{a.s.} \;\; \text{for } t \in [0, \tau]. \tag{2.5}$$

The first of the two conditions in (2.4) is the well-known requirement that the diffusion coefficients in the dynamics of X must be the same under $\mathbb{P}$ and $\mathbb{Q}$. The second condition, together with (2.5), imposes conditions on how the distribution of asset price jumps can change from $\mathbb{P}$ to $\mathbb{Q}$. This is captured by the nonnegative predictable function $Y_t(z)$ that reweighs the intensity of jumps of various sizes under the two probability measures. The requirement for Y only concerns the frequent and small jumps of X_t (of absolute value less than 1): If $\int_{|z|<1}|z|\nu_t^{\mathbb{P}}(dz)$ is finite, then the integral condition in (2.5) is automatically fulfilled as long as Y satisfies minimal boundedness conditions. If $\int_{|z|<1}|z|\nu_t^{\mathbb{P}}(dz)$ is infinite, however, then Y must not deviate too much from 1. Both conditions in (2.4) are intuitive as payoffs that depend on the small and frequent moves in X should earn no risk premium as the time horizon of these payoffs shrinks. This is because most of the risk in these payoffs can be hedged by a dynamic position in the underlying asset at no cost, as we show in the subsequent sections.

Definition 1. We introduce the following two related notions:

1. If for some $\mathbb{Q} \in \mathcal{Q}$, we have (2.4) together with (2.5), we say that *the equity and options markets agree about volatility*.

2. If $\mathcal{Q}_e \neq \emptyset$, we say that the equity and options markets are *integrated*. If $\mathcal{Q}_e = \emptyset$, we say that they are *segmented*.

The first definition explains the question posed in the title of this paper: we are interested in whether there is a risk-neutral measure that prices both asset and options and that agrees with $\mathbb{P}$ about both diffusive volatility and the small jumps in the asset prices. The motivation behind the second definition becomes clear once we notice that if $\mathbb{Q} \in \mathcal{Q}_e$, then there is an *aggregate pricing kernel* M_t for both asset and option prices such that (2.2) can be rewritten as[11]

$$X_t = \mathbb{E}_t^{\mathbb{P}}(D_{t,T}X_{t+T}),\; P_{t,T}(K) = \mathbb{E}_t^{\mathbb{P}}(D_{t,T}(K - X_{t+T})^+),\; C_{t,T}(K) = \mathbb{E}_t^{\mathbb{P}}(D_{t,T}(X_{t+T} - K)^+), \tag{2.6}$$

[11]In probability theory terms, M_t is the Radon–Nikodym derivative of $\mathbb{Q}|_{\mathcal{F}_t}$ with respect to $\mathbb{P}|_{\mathcal{F}_t}$. It is well-known that M_t is a martingale under $\mathbb{P}$. This fact as well as (2.6) can be found in Chapter III, Section 3, of Jacod and Shiryaev (2003).

where $D_{t,T} = M_{t+T}/M_t$ denotes the *stochastic discount factor*. Both M_t and $D_{t,T}$ depend on $\mathbb{Q}$ but we suppress the dependence in the notation. Therefore, market integration means that there is an aggregate pricing kernel that can be used to price the considered asset and options simultaneously. The two markets are segmented in the absence of such a joint pricing kernel.[12] Both notions introduced above are local in the sense that they are about what happens in a local neighborhood of time t.

From the preceding exposition, it is clear that market integration implies that the equity and options markets agree about volatility. Under reasonable assumptions, one can show that (2.4), together with a refinement of (2.5), is also sufficient for market integration; see Theorem A.1 in Appendix A.1. In the case of market integration, if $\mathbb{Q} \in \mathcal{Q}_e$, the stochastic discount factor has the following properties as the time horizon T shrinks.

Lemma 2.1. *If $\mathbb{Q} \in \mathcal{Q}_e$ and $t \in [0, \tau)$, the stochastic discount factor satisfies*

$$D_{t,T} \to 1 \quad \text{a.s. as } T \downarrow 0. \tag{2.7}$$

Moreover, if $\mathbb{E}_t^{\mathbb{P}}(|M_s|^q) < \infty$ for some $s \in (t, \tau]$ and $q \in (1, \infty)$, then

$$\mathbb{E}_t^{\mathbb{P}}(|D_{t,T} - 1|^q) \to 0 \quad \text{a.s. as } T \downarrow 0. \tag{2.8}$$

Intuitively, the result of the above lemma says that since the risks in the asset price are shrinking asymptotically as $T \downarrow 0$, then the same should happen to their prices. We note in this regard that in our specification for X, we have ruled out jumps that arrive at predictable times in $[0, \tau]$ such as those triggered by economic announcements. In the presence of such events, even over short intervals, the risk in the asset price might be nontrivial and the same applies to its corresponding price. It is easy to generalize the setup to allow for such event risk. We do not do this here to keep notation simple as in our application we look at small time intervals before market close when there are typically no pre-scheduled announcements.

3 Generalized Sharpe Ratios and Local Arbitrage

In this section, we devise a practical criterion that not only can be used to verify from observed asset and option prices whether (2.4) is satisfied or not, but also allows us to conclude whether a potential segmentation of the equity and options markets is *economically* significant. To this end, we start from a variant of the classical method of Hansen and Jagannathan (1991) to derive lower bounds on the stochastic discount factor. Instead of estimating the standard deviation of the stochastic discount factor using the Sharpe ratio of portfolio returns, we consider an extension due to Snow (1991) that uses reward-to-risk ratios of risky asset positions based on pth moments.

Definition 2. Given a portfolio with simple return $R_{t,T}$ from time t to time $t+T$ and some exponent $p \in [1, \infty)$, we define the *generalized Sharpe ratio (GSR)* of order p of this portfolio as

$$GSR_{t,T}(p) = \frac{\mathbb{E}_t^{\mathbb{P}}\left(R_{t,T}\right)}{\left(\mathbb{E}_t^{\mathbb{P}}\left(\left|R_{t,T} - \mathbb{E}_t^{\mathbb{P}}\left(R_{t,T}\right)\right|^p\right)\right)^{1/p}}. \tag{3.1}$$

[12]Even if equity and option markets are not segmented, there can be nevertheless partial segmentation where a segment of the option market, e.g., the one for deep out-of-the-money puts, is separated from the equity market and the rest of the option market, cf. Almeida et al. (2024). Such partial segmentation of deep out-of-the-money puts, however, cannot be separately identified using equity and option data alone as there exists a valid aggregate pricing kernel that can rationalize all observed equity and option prices.

The case $p = 2$ corresponds to the Sharpe ratio, $SR_{t,T}$, that is used in classical mean–variance portfolio analysis. The case $p = 1$ corresponds to the mean absolute deviation ratio of Konno and Yamazaki (1991). By Minkowski's inequality, we trivially have

$$GLR_{t,T} \geq GSR_{t,T}(1) \geq GSR_{t,T}(p_1) \geq SR_{t,T} \geq GSR_{t,T}(p_2), \tag{3.2}$$

for any $1 \leq p_1 \leq 2 \leq p_2 < \infty$, where $GLR_{t,T}$ denotes the conditional gain-loss ratio of Bernardo and Ledoit (2000) associated with the return $R_{t,T}$. The following result, due to Snow (1991), gives a lower bound on the variability of the pricing kernel in terms of generalized Sharpe ratios. We remind the reader that for simplicity, the risk-free rate in our analysis is set to zero.

Lemma 3.1. *Consider a portfolio with simple return $R_{t,T}$ from time t to time $t+T$ and assume that this portfolio can be priced by some risk-neutral measure $\mathbb{Q} \in \mathcal{Q}_e$, that is,*

$$\mathbb{E}_t^{\mathbb{Q}}(R_{t,T}) = 0. \tag{3.3}$$

For any $p \in (1, \infty)$, we have

$$\left(\mathbb{E}_t^{\mathbb{P}}\left(|D_{t,T} - 1|^q\right)\right)^{1/q} \geq |GSR_{t,T}(p)|, \tag{3.4}$$

where $D_{t,T}$ is the stochastic discount factor associated with $\mathbb{Q}$ and $q = \frac{p}{p-1}$ is the Hölder conjugate of p.

For the reader's convenience, we reproduce the short proof in Appendix A.1. The pricing equation (3.3) automatically holds as a consequence of (2.2) if the payoff of the portfolio is linear in the considered asset and option prices.

If $p = 2$, inequality (3.4) reduces to the classical bound of Hansen and Jagannathan (1991): the conditional standard deviation of the stochastic discount factor $D_{t,T}$ exceeds the Sharpe ratio attained by any portfolio that can be priced by $\mathbb{Q}$. For general $p \in (1, \infty)$, inequality (3.4) asserts that the qth-order central moment of the stochastic discount factor is bounded from below by the GSR of order p of any such portfolio, where p and q are related to each other through the equation $\frac{1}{p} + \frac{1}{q} = 1$. In particular, generalized Sharpe ratios with $p \in (1, 2)$ provide bounds on the moments of the stochastic discount factor of higher order than variance.

The generalized Hansen–Jagannathan bound of Lemma 3.1 means that portfolios with high GSR imply highly varying pricing kernels. This alone does not constitute an arbitrage opportunity, or a free lunch with vanishing risk, in the sense of Delbaen and Schachermayer (1994). Nonetheless, investment opportunities with high Sharpe ratios are "good deals" (Cochrane and Saa-Requejo (2000)) and, therefore, ought not to exist as investors would be eager to act upon them. The same clearly applies to portfolios with high generalized Sharpe ratios.

Contrary to the situation on a fixed time interval, as we show below, if a good deal can be realized over an asymptotically shrinking time interval, that is, if we can find portfolios whose (generalized) Sharpe ratios exceed a given positive threshold over shorter and shorter time horizons, then this essentially implies arbitrage.

Definition 3. Given $p \in (1, \infty)$, we say that *local arbitrage of order p* exists if for some $t \in [0, \tau)$ there is a sequence of portfolios with investment horizon T whose returns $R_{t,T}$ are priced according to (3.3) by a single risk-neutral measure $\mathbb{Q} \in \mathcal{Q}$ and whose generalized Sharpe ratios satisfy

$$\liminf_{T \downarrow 0} GSR_{t,T}(p) > 0 \quad \text{a.s.} \tag{3.5}$$

If $p = 2$, the notion of local arbitrage is related to the concept of quasi-arbitrage by Huberman and Stanzl (2004). The latter refers to a sequence of portfolios whose Sharpe ratios diverge to infinity. Clearly, in our context, if there is quasi-arbitrage of order p in the sense that $GSR_{t,T}(p) \xrightarrow{\mathbb{P}} \infty$ as $T \downarrow 0$, we also have local arbitrage of order p. But local arbitrage is strictly weaker because it already occurs if, for instance, $GSR_{t,T}(p)$ converges in probability to a strictly positive random variable. The next theorem is the main result of this section and explains why the situation in Definition 3 essentially constitutes arbitrage.

Theorem 3.2. *If for a given risk-neutral measure $\mathbb{Q} \in \mathcal{Q}$, there is local arbitrage of order $p \in (1, \infty)$ at time $t \in [0, \tau)$ with a sequence of portfolios whose payoffs can all be priced by $\mathbb{Q}$, then either $\mathbb{Q} \notin \mathcal{Q}_e$ or $\mathbb{E}_t^{\mathbb{P}}(|M_s|^q) = \infty$ for all $s \in (t, \tau]$, where $\frac{1}{p} + \frac{1}{q} = 1$ and M is the aggregate pricing kernel of $\mathbb{Q}$.*

Local arbitrage therefore implies that either there is no equivalent martingale measure that can price the considered asset and options simultaneously, so that the equity and options markets are segmented, or the qth moment of the aggregate pricing kernel must be infinite, implying highly unstable and oscillatory behavior.[13] To reach such a strong conclusion, it suffices by Definition 3 to exhibit a sequence of portfolios with GSR above a strictly positive threshold as $T \downarrow 0$. For this purpose, a sequence of good deals (uniformly in T) in the sense of Cochrane and Saa-Requejo (2000) is sufficient. Therefore, in our asymptotic setup where $T \downarrow 0$, we observe convergence of two different no-arbitrage related concepts considered in the previous literature: (strict) no-arbitrage in the sense that there exists an aggregate pricing kernel that can rationalize observed prices and the absence of good deals (or portfolios with large Sharpe ratios).

In a non-asymptotic setting, or in an asymptotic setting where the time horizon of the considered portfolios does not shrink, there is a substantial wedge between these two notions, as we know from previous work, see Cochrane and Saa-Requejo (2000) and Bernardo and Ledoit (2000) who show that pricing kernel bounds obtained from classical no-arbitrage concepts are much looser than the no-good-deal bounds. One would have to resort to the stronger notion of quasi-arbitrage (Huberman and Stanzl (2004)) in the case of a non-shrinking time interval in order to reach the conclusion of Theorem 3.2.

As Figure 1 shows, the notions of absence of local arbitrage, market integration and agreement about volatility in equity and options markets are essentially equivalent. The equivalence between market integration and agreement of asset and option prices about volatility follows from Girsanov's theorem and its converse in the form of Theorem A.1, which holds under reasonable regularity assumptions. The main theoretical contribution of the paper is to establish that these two notions, which are central to the financial economics of equity and options markets, are equivalent to the absence of local arbitrage opportunities, an economically interpretable criterion that can be feasibly implemented in practice. One direction is shown in Theorem 3.2: market integration rules out local arbitrage opportunities as long as the aggregate pricing kernel has a certain number of finite moments. The other direction is the content of Theorems 5.1 and 5.2 below: a disagreement between $\mathbb{P}$ and $\mathbb{Q}$ about volatility leads to local arbitrage opportunities, so an absence of local arbitrage implies that $\mathbb{P}$ and $\mathbb{Q}$ must agree about volatility.

[13] In finite dimensions (e.g., if we only consider the price of the underlying asset), it is well known that the absence of an equivalent local martingale measure implies a free lunch with vanishing risk (Delbaen and Schachermayer (1994)). In the presence of infinitely many assets (e.g., if we include options with a continuum of strikes and maturities), the absence of an equivalent local martingale measure almost implies a free lunch; see Cuchiero et al. (2016) for a precise statement.

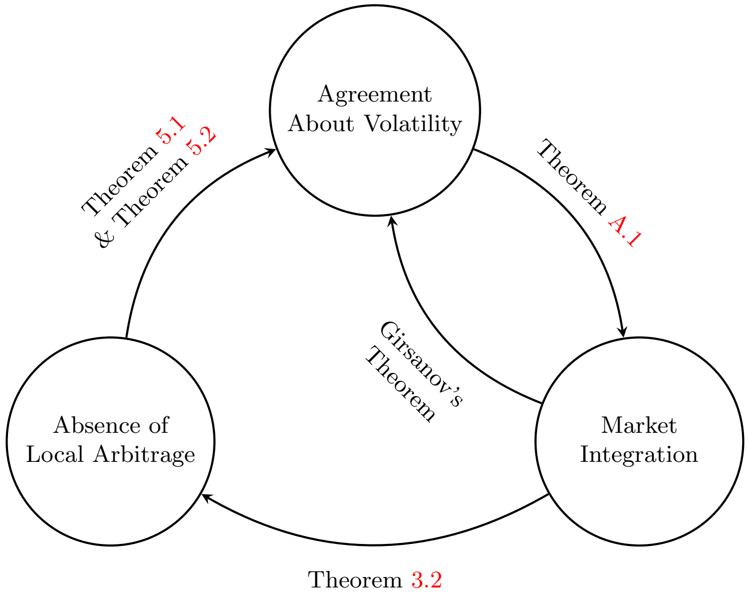

Figure 1: **Relationship between market integration, absence of local arbitrage and agreement about volatility in equity and options markets.**

4 Truncated Volatility and $\mathbb{P}$-$\mathbb{Q}$ Equivalence

Our goal now is to link the condition in (2.4) to a payoff that depends on the value of the underlying asset in a local neighborhood of time t. We will then use this connection in the following section to show the existence of local arbitrage if (2.4) is violated. As already discussed in Section 2, the condition in (2.4) concerns the small and frequent moves in the asset price, which is why we aim for payoffs that capture exactly those. A natural candidate would be the log-price payoff which is behind the construction of the popular VIX index. That is, to consider $\log(X_{t+T}) - \log(X_t)$, for small T. An application of Itô's formula leads to

$$-\frac{2}{T}\mathbb{E}_t^{\mathbb{Q}}(\log(X_{t+T}) - \log(X_t)) \approx (\sigma_t^{\mathbb{Q}})^2 + 2\int_{\mathbb{R}}(e^z - 1 - z)\nu_t^{\mathbb{Q}}(dz), \quad \text{as } T \downarrow 0. \tag{4.1}$$

The problem for our purposes is that such a payoff depends on all jumps and not only the small ones. Indeed, note that the integral above is over the entire real line. Hence, due to the risk premium for big jumps, $\mathbb{E}_t^{\mathbb{Q}}(\log(X_{t+T}) - \log(X_t))$ and $\mathbb{E}_t^{\mathbb{P}}(\log(X_{t+T}) - \log(X_t))$ can differ even in the absence of (local) arbitrage and if T is very small. We confirm that this is the case empirically later on. Therefore, we consider a different payoff which depends only on the small and frequent moves in the underlying asset price.

Definition 4. We call *truncated volatility* the payoff $f_T(X_{t+T}; X_t)$, where

$$f_T(x; x_0) = \phi\left(\eta_T \frac{\log(x) - \log(x_0)}{\sqrt{T}}\right)(\log(x) - \log(x_0))^2, \tag{4.2}$$

and ϕ is a truncation function that decays exponentially fast for values of its argument away from zero and the truncation parameter satisfies $\eta_T \downarrow 0$ as $T \downarrow 0$.

The conditional expectation of this payoff under $\mathbb{S} \in \{\mathbb{P}, \mathbb{Q}\}$ is given by

$$TV_{t,T}^{\mathbb{S}}(\eta) = \mathbb{E}_t^{\mathbb{S}}\left(\phi\left(\eta\frac{\log(X_{t+T}) - \log(X_t)}{\sqrt{T}}\right)(\log(X_{t+T}) - \log(X_t))^2\right), \quad \eta \geq 0. \tag{4.3}$$

If $\eta_T \downarrow 0$ slowly, e.g., if $\eta_T \sim 1/\sqrt{\log(1/T)}$, one can show the following expansion of truncated volatility as $T \downarrow 0$:

$$\frac{1}{T} TV_{t,T}^{\mathbb{S}}(\eta_T) \approx (\sigma_t^{\mathbb{S}})^2 + \int_{\mathbb{R}} \phi(\eta_T z/\sqrt{T}) z^2 \nu_t^{\mathbb{S}}(dz). \tag{4.4}$$

Unlike the case of the log-price payoff, which depends on the big jumps, the truncated volatility payoff eliminates the effect of the big jumps and they become negligible as $T \downarrow 0$. The condition in (2.4), therefore, implies

$$TV_{t,T}^{\mathbb{P}}(\eta_T) \approx TV_{t,T}^{\mathbb{Q}}(\eta_T), \quad \text{as } T \downarrow 0. \tag{4.5}$$

Conversely, as we show below, if $\sigma_t^{\mathbb{P}}$ and $\sigma_t^{\mathbb{Q}}$ are different and/or $\nu_t^{\mathbb{P}}$ and $\nu_t^{\mathbb{Q}}$ differ too much around zero, then there is η_T such that the above approximate equality does not hold. In this regard, (4.4) captures the volatility due to small and frequent moves in the asset price.

We can assess the quality of the approximation in (4.5) in a numerical example using the parametric model from our Monte Carlo experiment that we present in Section 7. In Figure 2, we plot the ratio $TV_{0,T}^{\mathbb{Q}}(\eta_T)/TV_{0,T}^{\mathbb{P}}(\eta_T)$ for different levels of the truncation parameter in the case of $\mathbb{P}$-$\mathbb{Q}$ equivalence. When we do not truncate ($\eta_T = 0$), then there is a significant gap between the two volatility measures, which is due to the compensation for the big negative jumps in the underlying asset price. As we increase the level of truncation, the gap between $TV_{0,T}^{\mathbb{P}}(\eta_T)$ and $TV_{0,T}^{\mathbb{Q}}(\eta_T)$ shrinks in relative terms. This is because truncated volatility depends less on the "big" jumps in the asset prices. Indeed, for the highest level of truncation considered here, $TV_{0,T}^{\mathbb{Q}}(\eta_T)$ is only around 3% higher than $TV_{0,T}^{\mathbb{P}}(\eta_T)$ for $T = 30$ minutes, which is a negligible difference. We also note that for every level of truncation $\eta_T > 0$, the upward bias in $TV_{0,T}^{\mathbb{Q}}(\eta_T)$ due to the compensation for big jumps shrinks in relative terms as T shrinks. For example, for the highest considered level of truncation, the ratio $TV_{0,T}^{\mathbb{Q}}(\eta_T)/TV_{0,T}^{\mathbb{P}}(\eta_T)$ shrinks from around 1.05 for T equal to 120 minutes to 1.03 for T equal to 30 minutes. We will consider the same levels of truncation as the ones used in Figure 2 in our applications.

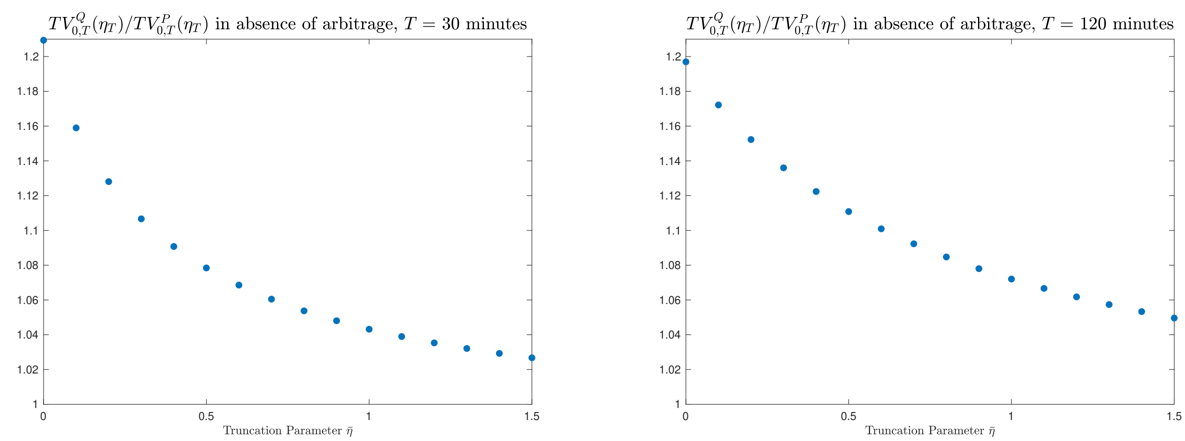

Figure 2: **Risk-Neutral versus True Conditional Expected Truncated Volatility in the Absence of Arbitrage**. The parameter settings for the no-arbitrage scenario are given in the Monte Carlo study. The conditional truncated volatilities are computed via simulation. The truncation function is $\phi(x) = e^{-x^2}$ and the truncation parameter is set to $\eta_T = \frac{\sqrt{\bar{\eta}}}{\sqrt{\log(1/T)}} \times \frac{1}{\sigma_0^{\mathbb{P}}}$.

5 Volatility Arbitrage Portfolios

We are now ready to show that if condition (2.4) does not hold, then this implies the existence of local arbitrage opportunities involving short-dated options with tenor shrinking to zero. As we discussed in the previous section, a violation of (2.4) implies that a truncated volatility payoff will be nontrivial even over shrinking time horizons. In this section, we construct portfolios that include short-dated options and the underlying asset that yield this truncated volatility payoff, and we further show that they constitute local arbitrage as the tenor of the options shrinks to zero if (2.4) is violated.

According to the option spanning results of Bakshi and Madan (2000) and Carr and Madan (2001), payoffs that are nonlinear in the terminal price (such as truncated volatility) can be obtained via portfolios of options over a continuum of strikes and with the same expiration date. More precisely, using Taylor expansion with remainder for the function $f_T(x; x_0)$ in (4.2), we obtain

$$\begin{aligned} f_T(X_{t+T}; X_t) &= \phi(\eta_T(\log(X_{t+T}) - \log(X_t))/\sqrt{T})(\log(X_{t+T}) - \log(X_t))^2 \\ &= \int_0^{X_t} f_T''(K; X_t)(K - X_{t+T})^+ dK + \int_{X_t}^{\infty} f_T''(K; X_t)(X_{t+T} - K)^+ dK. \end{aligned} \tag{5.1}$$

Note that $(K - X_{t+T})^+$ and $(X_{t+T} - K)^+$ are the payoffs at time $t + T$ of European-style put and call options, respectively, written on the asset at time t with strike K and time to maturity T. This means that the nonlinear payoff $f_T(X_{t+T}; X_t)$ can be replicated by a static portfolio of options created at time t. By (2.2), the cost of this portfolio at time t is given by

$$\mathcal{C}_{t,T}(\eta_T) \equiv TV_{t,T}^{\mathbb{Q}}(\eta_T) = \int_0^{\infty} f_T''(K; X_t) O_{t,T}(K) dK, \tag{5.2}$$

where $O_{t,T}(K)$ denotes the price of the out-of-the-money option for strike K. This is either the call or the put option with the same strike depending on whether the current asset price is below or above K. Its price is therefore

$$O_{t,T}(K) = \min\{P_{t,T}(K), C_{t,T}(K)\}.$$

We can compare the weights assigned to the options with different strikes in the above portfolio that generates the truncated volatility payoff for different levels of truncation. The case $\eta_T = 0$ (no truncation) corresponds to the second moment of the log-return that has been studied by Bakshi et al. (2003) (and which is very close to the VIX index discussed above). Higher truncation means shifting more weight to the smaller moves. As we can see from Figure 3, this translates into putting higher weights in (5.2) on options with strikes in the vicinity of the current stock price. Interestingly, the truncated volatility option portfolio assigns negative weight to options which are slightly away from the money. Intuitively, this removes an upward bias of at-the-money options due to the the larger asset price jumps. From a practical point of view, this also implies somewhat higher transaction costs for replicating the truncated volatility payoff ($\eta_T > 0$) versus replicating the total volatility payoff ($\eta_T = 0$).

As evident from the figure, the contribution of deep out-of-the-money puts and calls to the value of the option-based truncated volatility portfolio is negligible. In fact, one can show that due to the truncation in the payoff $f_T(X_{t+T}, X_t)$, one can consider only integration over a finite range of strikes covering the current stock price when computing the integral on the left-hand side of (5.2), without any change to the results that will follow. This is intuitive as

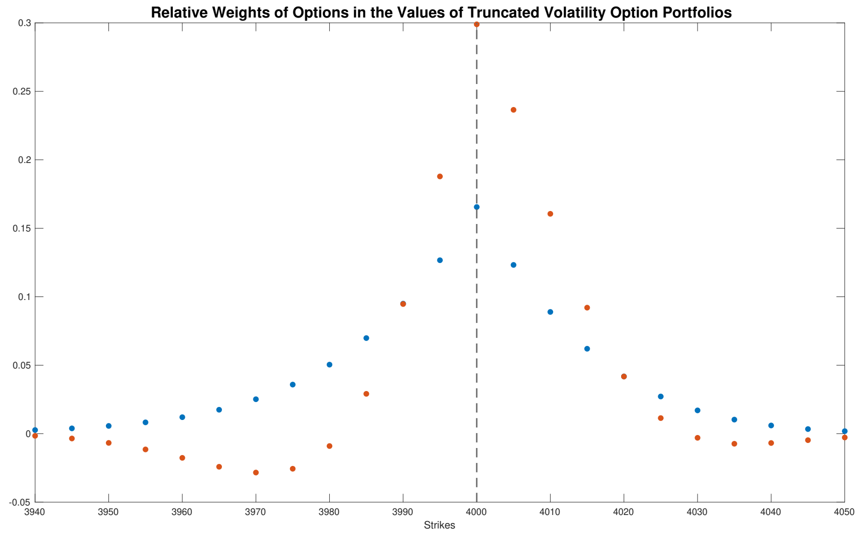

Figure 3: **Weights in Truncated Volatility Option Portfolios**. Each point on the figure represents the fraction of wealth in the truncated volatility option portfolio invested in the option of the given strike. Options are generated from the no-arbitrage model for the risk-neutral dynamics given in the Monte Carlo study. The starting value of the underlying price and the spot volatility are 4,000 and 0.02, respectively. The time to maturity is 2 hours. The truncation function is $\phi(x) = e^{-x^2}$ and the truncation parameter is set as in (7.5) with $\overline{\eta} = 0$ (blue dots) and $\overline{\eta} = 1.5$ (red dots).

the information contained in deep out-of-the-money puts and calls about the small and frequent moves in the asset price is negligible as T shrinks.

We can now analyze the properties of the truncated volatility option portfolio. The simple return of this portfolio is given by

$$R_{t,T} = \frac{f_T(X_{t+T}; X_t)}{\mathcal{C}_{t,T}(\eta_T)} - 1. \tag{5.3}$$

The conditional expected return of the portfolio is

$$\mathbb{E}_t^{\mathbb{P}}\left(R_{t,T}\right) = \frac{\mathbb{E}_t^{\mathbb{P}}\left(f_T(X_{t+T}; X_t)\right)}{\mathbb{E}_t^{\mathbb{Q}}\left(f_T(X_{t+T}; X_t)\right)} - 1. \tag{5.4}$$

Given the discussion in the previous section, condition (2.4) implies that

$$\mathbb{E}_t^{\mathbb{P}}\left(R_{t,T}\right) \approx 0, \quad \text{as} \ \ T \downarrow 0, \tag{5.5}$$

that is, there should be no reward for the risk of this portfolio asymptotically as $T \downarrow 0$. A violation of (2.4), on the other hand, means that either a long or a short position in the option portfolio should earn a non-negligible (i.e., non-vanishing) positive return. To check if this is a local arbitrage opportunity as $T \downarrow 0$, we need to analyze the generalized Sharpe ratio of the strategy, which we denote by

$$GSR_{t,T}(p) = \frac{\mathbb{E}_t^{\mathbb{P}}\left(R_{t,T}\right)}{\left(\mathbb{E}_t^{\mathbb{P}}\left(\left|R_{t,T} - \mathbb{E}_t^{\mathbb{P}}\left(R_{t,T}\right)\right|^p\right)\right)^{1/p}}. \tag{5.6}$$

One can show that even if there is no $\mathbb{P}$-$\mathbb{Q}$ equivalence, that is, if (2.4) does not hold, $GSR_{0,T}(p)$ converges to a finite number. This is sufficient for qualifying as local arbitrage, which by Theorem 3.2 implies that an aggregate pricing kernel either does not exist or it is not locally square-integrable (is highly volatile).

We can, however, strengthen the above result by modifying the above option portfolio strategy in a way that leads to even explosive generalized Sharpe ratios as T shrinks. To achieve this, we need to eliminate the directional risk in the strategy. This can be done by adding a dynamic portfolio position in the underlying asset, where, at each point in time $s \in [t, t+T]$, the exposure to the position in the underlying asset is $-f_T'(X_{s-}; X_t)$. Altogether, the cumulative realized gains of the combined positions in the options and the underlying asset from t to $t+T$ are given by

$$\mathcal{RG}_{t,T} = f_T(X_{t+T}; X_t) - \int_t^{t+T} f_T'(X_{s-}; X_t) dX_s, \tag{5.7}$$

while the initial cost of the position is still $\mathcal{C}_{t,T}(\eta_T)$. Thus, the cumulative rate of return over the interval $[t, t+T]$ becomes

$$\mathcal{R}_{t,T} = \frac{\mathcal{RG}_{t,T}}{\mathcal{C}_{t,T}(\eta_T)} - 1. \tag{5.8}$$

We have that

$$\mathbb{E}_t^{\mathbb{P}}(\mathcal{R}_{t,T}) \approx \mathbb{E}_t^{\mathbb{P}}(R_{t,T}), \tag{5.9}$$

with the reason for the approximation sign above being due to the equity risk premium in the underlying asset over the interval $[t, t+T]$, which shrinks as $T \downarrow 0$.[14] However, the wedge $\mathbb{E}_t^{\mathbb{P}}(\mathcal{R}_{t,T}) - \mathbb{E}_t^{\mathbb{P}}(R_{t,T})$ is much smaller than $\mathbb{E}_t^{\mathbb{P}}(R_{t,T})$ whenever $\mathbb{P}$-$\mathbb{Q}$ equivalence fails.

By contrast, adding the dynamic position in the underlying asset significantly reduces the risk in the option portfolio position and one can show that

$$\mathbb{E}_t^{\mathbb{P}}\left(\left|\mathcal{R}_{t,T} - \mathbb{E}_t^{\mathbb{P}}(\mathcal{R}_{t,T})\right|^p\right) \ll \mathbb{E}_t^{\mathbb{P}}\left(\left|R_{t,T} - \mathbb{E}_t^{\mathbb{P}}(R_{t,T})\right|^p\right), \tag{5.10}$$

for any $p \in [1, 2)$. For this reason, we refer to the combined option portfolio position with the dynamic position in the underlying asset described above as a *hedged long position in truncated volatility*. When the signs of the positions in the above strategy are flipped, then this will be referred to as a *hedged short position in truncated volatility*.

We denote the (conditional) generalized Sharpe ratio of the hedged long position in truncated volatility by

$$\mathcal{GSR}_{t,T}(p) = \frac{\mathbb{E}_t^{\mathbb{P}}(\mathcal{R}_{t,T})}{\left(\mathbb{E}_t^{\mathbb{P}}\left(\left|\mathcal{R}_{t,T} - \mathbb{E}_t^{\mathbb{P}}(\mathcal{R}_{t,T})\right|^p\right)\right)^{1/p}}. \tag{5.11}$$

We will show next that the above reduction in the risk of the strategy leads to explosive $\mathcal{GSR}_{t,T}(p)$ as T shrinks if the $\mathbb{P}$-$\mathbb{Q}$ equivalence condition in (2.4) is violated.

5.1 Local Arbitrage When Diffusive Volatility is Different

In this section, we formally show how local arbitrage arises if diffusive spot volatility is different under the physical and the risk-neutral probability measure. The behavior of the generalized Sharpe ratio of the hedged positions in truncated volatility when $\sigma_t^{\mathbb{P}} \neq \sigma_t^{\mathbb{Q}}$ is given in the next theorem.

[14] The conditional equity risk premium is given by $\mathbb{E}_t^{\mathbb{P}}\left(\int_t^{t+T} \alpha_s ds\right) = O_p(T)$.

Theorem 5.1. *Suppose that $\mathbb{Q} \in \mathcal{Q}$ is such that $\sigma_t^{\mathbb{P}} \neq \sigma_t^{\mathbb{Q}}$ for some $t \in [0, \tau)$. Further suppose that we have a continuous record of the asset price X over the interval $[t, t+T]$ as well as the prices $O_{t,T}(K)$ at time t of out-of-the-money options expiring at time $t+T$ over a continuum of strikes $K \in (0, \infty)$, so that we can form the truncated volatility portfolio whose return from time t to time $t+T$ is given by (5.8). Under Assumptions 1 and 2 detailed in Appendix A.3 and if $\eta_T \sim \eta/\sqrt{\log T^{-1}}$ for some $\eta > 0$, we have*

$$|\mathcal{GSR}_{t,T}(p)| \xrightarrow{\mathbb{P}} \infty \tag{5.12}$$

as $T \to 0$ for all $p \in [1, 2)$. For $p = 2$, we have

$$|\mathcal{GSR}_{t,T}(2)| = |\mathcal{SR}_{t,T}| \xrightarrow{\mathbb{P}} \frac{|(\sigma_t^{\mathbb{P}})^2 - (\sigma_t^{\mathbb{Q}})^2|}{\sigma_t^{\mathbb{P}}(\int_{\mathbb{R}} z^2 \nu_t^{\mathbb{P}}(dz))^{1/2}}. \tag{5.13}$$

In particular, we have local arbitrage of order p for any $p \in (1, 2]$.

We note that the Sharpe ratio of the strategy does not explode as T shrinks but the generalized Sharpe ratio for values of $p \in (1, 2)$ does. The reason for this is the jumps in the asset price. Their effect on the portfolio returns is dampened but not completely eliminated. This is due to their impact on the dynamic position in the underlying asset. When we consider higher moments of the portfolio returns, the jump contribution increases and this is the reason for the different asymptotic behavior of $|\mathcal{GSR}_{t,T}(p)|$ for $p < 2$ and $p = 2$. Nevertheless, $|\mathcal{GSR}_{t,T}(2)|$ does not shrink as T shrinks, which is enough to constitute a local arbitrage opportunity. This makes intuitive sense as other portfolios typically have a shrinking Sharpe ratio over such short time horizons as we consider here. A long (or short) static position in the underlying asset is one such example. But also the hedged volatility position without truncation, i.e., the position with truncation parameter η set to zero, has a shrinking Sharpe ratio. One can show that in the setting of Theorem 5.1, the Sharpe ratio $|\mathcal{SR}_{t,T}^{\eta=0}|$ of this strategy behaves like

$$\frac{|\mathcal{SR}_{t,T}^{\eta=0}|}{\sqrt{T}} \to \frac{|(\sigma_t^{\mathbb{P}})^2 - (\sigma_t^{\mathbb{Q}})^2 + \int_{\mathbb{R}} z^2(\nu_t^{\mathbb{P}}(dz) - \nu_t^{\mathbb{Q}}(dz))|}{\int_{\mathbb{R}} z^4 \nu_t^{\mathbb{P}}(dz)}, \quad \text{as } T \to 0. \tag{5.14}$$

In words, the Sharpe ratio of the hedged volatility strategy shrinks at rate $\sqrt{T}$. In Figure 4, we plot the asymptotic approximations for the Sharpe ratios of the hedged long or short positions in truncated volatility, with and without truncation, when $\sigma_t^{\mathbb{P}} \neq \sigma_t^{\mathbb{Q}}$. We use the parametric model in the Monte Carlo to generate the figure. This model matches key features of real data and hence the numbers in Figure 4 are representative of the possible Sharpe ratios that can be achieved if $\sigma_t^{\mathbb{P}} \neq \sigma_t^{\mathbb{Q}}$ (without taking into account trading costs). As seen from the figure, even relatively small differences in $\sigma_t^{\mathbb{P}}$ and $\sigma_t^{\mathbb{Q}}$ can generate very high Sharpe ratios using the hedged truncated volatility strategy with $\eta > 0$. The Sharpe ratio of the strategy without truncation ($\eta = 0$) is also large when T is relatively high but it quickly drops when T becomes small. The reason for this is that the payoff of this position is significantly exposed to big jumps in the asset price.

5.2 Local Arbitrage When Small Jumps are Different

In this section, we show how to construct local arbitrage portfolios if the physical and risk-neutral probability measures differ in the distribution of small asset price jumps. Note that this can occur only if assets have jumps of infinite variation. This is rarely the case for most models in finance considered in prior work. Nevertheless, we consider this case for completeness.

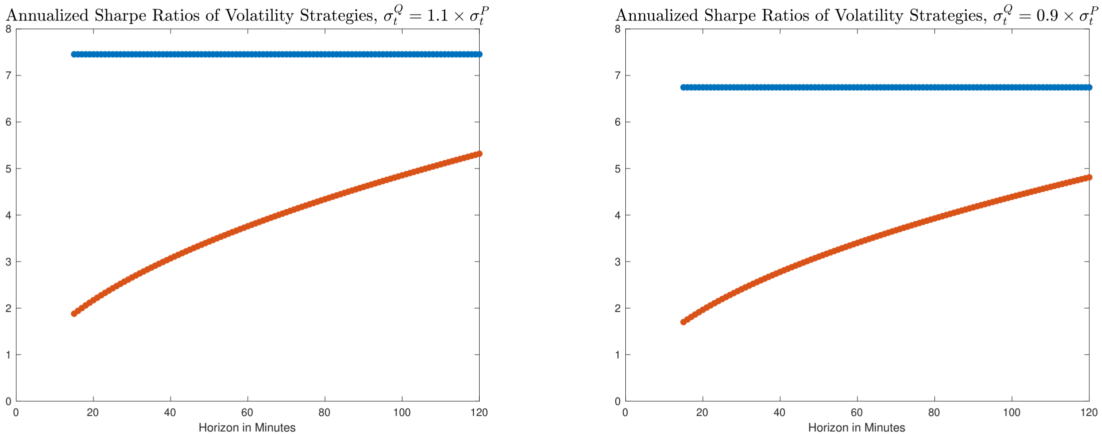

Figure 4: **Sharpe Ratios of Volatility Trading Strategies for Different Horizons**. The parameter settings for the different scenarios are given in the Monte Carlo study. The Sharpe ratios of the strategy with truncation ($\eta > 0$) and without truncation ($\eta = 0$) are computed using the asymptotic limits in (5.13) and (5.14). Blue line corresponds to $\eta > 0$ and red line to $\eta = 0$. The per-period Sharpe ratios are annualized by multiplying them by $\sqrt{252}$.

To this end, we set up a multi-period version of the volatility arbitrage portfolio from Section 5.1 and consider k_n time points $\tau_i = t + (i-1)\tau$, $i = 1, \ldots, k_n$, where $\tau \geq T$. For example, if T is the time to maturity of zero-date options, τ could be the length of one trading day. Now, at each time point τ_i, we invest in the volatility arbitrage portfolio from Section 5 (without compounding, to simplify the analysis), so that the return of this strategy after k_n periods is

$$\mathcal{R}_T^{(k_n)} = \sum_{i=1}^{k_n} \mathcal{R}_{\tau_i,T}, \qquad \mathcal{R}_{\tau_i,T} = \frac{\mathcal{RG}_{\tau_i,T}}{\mathcal{C}_{\tau_i,T}(\eta_T)} - 1, \tag{5.15}$$

where

$$\mathcal{RG}_{\tau_i,T} = f_T(X_{\tau_i+T}; X_{\tau_i}) - \int_{\tau_i}^{\tau_i+T} f_T'(X_{t-}; X_{\tau_i}) dX_t \tag{5.16}$$

and

$$\mathcal{C}_{\tau_i,T}(\eta_T) = \int_0^{\infty} f_T''(K; X_{\tau_i}) O_{\tau_i,T}(K) dK. \tag{5.17}$$

While the number of periods k_n increases to infinity asymptotically, we assume that $k_n \tau \to 0$, which means that we are still looking to detect local arbitrage violations with $\mathcal{R}_T^{(k_n)}$.

Theorem 5.2. *Suppose that $\sigma^{\mathbb{P}} = \sigma^{\mathbb{Q}}$ for some $\mathbb{Q} \in \mathcal{Q}$ but $\nu_t^{\mathbb{P}}$ and $\nu_t^{\mathbb{Q}}$ differ at time $t \in [0, \tau)$ in such a way that either $\beta^{\mathbb{P}} \neq \beta^{\mathbb{Q}}$ or $c_t^{\mathbb{P}} \neq c_t^{\mathbb{Q}}$ for the jump measure parameters introduced in Assumption 4. Further suppose that we have a continuous record of the asset price X over the time intervals $[\tau_i, \tau_i + T]$ and a record of the out-of-the-money option prices $O_{\tau_i,T}(K)$ at time τ_i over a continuum of strikes $K \in (0, \infty)$ for all $i = 1, \ldots, k_n$, so that we can form the multi-period truncated volatility portfolio whose return is given by (5.15). Under Assumptions 3 and 4 detailed in Appendix A.3 and if $\eta_T \sim \eta T^{1/6}$ for some $\eta > 0$, $k_n \sim \theta T^{-\kappa}$ for some $\theta > 0$ and $\kappa \in [\frac{2}{3}, 1)$ and $\tau = O(T)$, then the generalized Sharpe ratio of the multi-period portfolio satisfies*

$$|\mathcal{GSR}_{t,T}^{(k_n)}(p)| = \frac{|\mathbb{E}_t^{\mathbb{P}}(\mathcal{R}_T^{(k_n)})|}{(\mathbb{E}_t^{\mathbb{P}}(|\mathcal{R}_T^{(k_n)} - \mathbb{E}_t^{\mathbb{P}}(\mathcal{R}_T^{(k_n)})|^p))^{1/p}} \xrightarrow{\mathbb{P}} \infty, \tag{5.18}$$

for all $p \in [1, \beta^{\mathbb{P}}]$, implying the existence of local arbitrage of any order $p \in (1, \beta^{\mathbb{P}}]$.

Remark 1. In the situation of Theorem 5.2, one can show that if $\sigma_t^{\mathbb{P}} \neq \sigma_t^{\mathbb{Q}}$, then (5.18) holds irrespective of whether $(\beta^{\mathbb{P}}, c_t^{\mathbb{P}}) \neq (\beta^{\mathbb{Q}}, c_t^{\mathbb{Q}})$ or not. Therefore, the multi-period hedged volatility portfolio is able to exploit local arbitrage violations no matter whether they are due to differences in volatility or small jumps.

Remark 2. Our proof of Theorem 5.2 shows that it suffices to choose $\kappa \in (1 - \frac{\beta}{3}, 1)$. The proof can also be refined to show that if ϕ satisfies $|\phi(x) - 1| \leq C_1 \exp(-C_2/|x|^2)$ for $|x| \leq 1$ and some $C_1, C_2 > 0$, then the assertion of Theorem 5.2 continues to hold for $\eta_T \sim \eta/\sqrt{\log T^{-1}}$ (as in Theorem 5.1) and $\kappa \in (1 - \frac{\beta}{2}, 1)$.

6 Implementing Volatility Arbitrage Strategies

In the previous section, we showed how to take advantage if equity and option markets disagree about volatility due to frequent and small asset moves. In reality, the underlying asset is traded discretely (even though this can happen at a very high frequency), options are observed on a discrete strike grid, and there are nontrivial bid–ask spreads when trading options and the underlying asset. All this has the potential to reduce the profitability of local arbitrage portfolios (cf. Pontiff (1996)). In addition, the investor needs to use past information to decide if there are local arbitrage opportunities. In this section, we take care of all of these issues to propose volatility arbitrage trading strategies that are feasible in practice and which allow us, as econometricians, to decide if such local arbitrage opportunities exist empirically.

6.1 Transaction Costs

We start with the option portfolio that replicates the truncated volatility. We assume that at time t, we have access to options on the asset expiring at time $t + T$ on the discrete strike grid

$$K_{t,1} < \cdots < K_{t,N_t}, \tag{6.1}$$

for some positive integer N_t. Our estimate of $TV_{t,T}^{\mathbb{Q}}(\eta_T)$ is then given by

$$\widehat{TV}_{t,T}^{\mathbb{Q}}(\eta_T) = \sum_{j=2}^{N_t} f_T''(K_{t,j-1}; X_t) O_{t,T}(K_{t,j-1})(K_{t,j} - K_{t,j-1}). \tag{6.2}$$

In the above, we assume that we have access to the true option price. In reality, investors face nontrivial bid–ask spreads when trading options. We can then form two estimates of $TV_{t,T}^{\mathbb{Q}}(\eta_T)$: one that corresponds to a long position in truncated volatility, denoted by $\widehat{TV}_{t,T}^{\mathbb{Q}}(\eta_T)^a$, and another one that corresponds to a short position in truncated volatility, denoted by $\widehat{TV}_{t,T}^{\mathbb{Q}}(\eta_T)^b$. We have

$$\widehat{TV}_{t,T}^{\mathbb{Q}}(\eta_T)^b < \widehat{TV}_{t,T}^{\mathbb{Q}}(\eta_T)^a,$$

with the gap determined by the bid–ask spreads in options markets.

For our volatility strategy, we need to enter into a dynamic position in the underlying asset. Instead of trading continuously, investors trade at intervals of length $0 < \Delta_n < T$ in practice.

This leads to the following feasible counterpart of $\mathcal{RG}_{t,T}$ from (5.7):

$$\widehat{\mathcal{RG}}_{t,T} = \sum_{j=2}^{N_t} f''_T(K_{t,j-1}; X_t)\widehat{O}_{t+T}(K_{t,j-1})(K_{t,j} - K_{t,j-1}) - \sum_{j=1}^{k_n} f'_T(X_{t+(j-1)\Delta_n}; X_t)(X_{t+j\Delta_n} - X_{t+(j-1)\Delta_n}), \tag{6.3}$$

where $k_n \sim T/\Delta_n$. The first part of the realized gains is due to the payoffs of the static position in the portfolio of options. Here, $\widehat{O}_{t+T}(K)$ is the payoff of the option at expiration. The second part is due to the cumulative gains from trading the underlying asset. This part can be represented equivalently as

$$-f'_T(X_{t+k_n\Delta_n}; X_t)X_{t+k_n\Delta_n} + \sum_{j=1}^{k_n}(f'_T(X_{t+j\Delta_n}; X_t) - f'_T(X_{t+(j-1)\Delta_n}; X_t))X_{t+j\Delta_n}. \tag{6.4}$$

The first component in the above expression is the value of the position in the underlying asset at the terminal date $t + T$, while the second one contains the gains/losses from rebalancing the position in the underlying asset over the time window $[t, t + T]$. When trading the underlying asset, investors have to deal with bid–ask spreads. We account for that by using $X^b_{t+j\Delta_n}$ if the position at time $t + j\Delta_n$ is short and $X^a_{t+j\Delta_n}$ if the position at time $t + j\Delta_n$ is long, where $X^b_{t+j\Delta_n}$ and $X^a_{t+j\Delta_n}$ denote the best bid and best ask, respectively. For the amount held in the underlying asset at each point in time, we use the mid-quote $0.5 \times X^b_{t+j\Delta_n} + 0.5 \times X^a_{t+j\Delta_n}$.

We can derive an asymptotic approximation for the size of the trading costs due to the dynamic trading part of the strategy.

Lemma 6.1. *Let $\widehat{DTC}_{t,T}$ denote the trading costs due to the dynamic trading part of the above strategy, that is,*

$$\widehat{DTC}_{t,T} = |f'_T(X_{t+k_n\Delta_n}; X_t)|\epsilon_{t+k_n\Delta_n} + \sum_{j=1}^{k_n}|f'_T(X_{t+j\Delta_n}; X_t) - f'_T(X_{t+(j-1)\Delta_n}; X_t)|\epsilon_{t+j\Delta_n}, \tag{6.5}$$

where $\epsilon_{t+j\Delta_n} = X^a_{t+j\Delta_n} - X_t$ if the position at time $t + j\Delta_n$ is short and $\epsilon_{t+j\Delta_n} = X^b_{t+j\Delta_n} - X_t$ if the position at time $t + j\Delta_n$ is long. If the true underlying asset price is the mid-quote and positive and the relative bid–ask spread is a constant denoted with rba*, that is,*

$$X_s = \frac{1}{2}(X^a_s + X^b_s) \quad \text{and} \quad \mathrm{rba} = \frac{X^a_s - X^b_s}{X_s} \quad \text{for all } s \in [t, t + T],$$

then

$$\frac{\sqrt{\Delta_n}}{T}\widehat{DTC}_{t,T} \xrightarrow{\mathbb{P}} \mathrm{rba}\sqrt{\frac{2}{\pi}}\sigma^{\mathbb{P}}_t, \quad \text{as } \Delta_n,\, T \to 0 \text{ and } \Delta_n/T \to 0. \tag{6.6}$$

The lemma, which is proved in Appendix A.4, leads to the approximation

$$\widehat{DTC}_{t,T} \approx \frac{T}{\sqrt{\Delta_n}}\mathrm{rba}\sqrt{\frac{2}{\pi}}\sigma^{\mathbb{P}}_t. \tag{6.7}$$

We note that the size of the transaction costs are governed by the diffusive volatility and do not depend on the level of truncation in the volatility arbitrage strategy. This is because the frequent moves in the asset price, which trigger most of the rebalancing of the position in the underlying asset, are dominated by the diffusive part of the price. Naturally, the transaction costs are higher for higher rebalancing frequency.

The above approximation of the trading costs due to trading in the underlying asset can allow us to gauge their effect on the profitability of the volatility arbitrage strategies. For this, purpose, in Figure 5 we plot the relative transaction costs due to the dynamic part of the position. The latter are computed by dividing $\widehat{DTC}_{0,T}$ by $T \times (\sigma_0^{\mathbb{P}})^2$. These numbers should be compared with the relative expected gains of the trading strategy without trsaction costs which are given by $\max\{(\sigma_0^{\mathbb{Q}}/\sigma_0^{\mathbb{P}})^2 - 1, 1 - (\sigma_0^{\mathbb{P}}/\sigma_0^{\mathbb{Q}})^2\}$. For generating the numbers in the figure, we have set T to 30 minutes, Δ_n to 5 minutes, and rba to the median value we find in the real data in our empirical application. As seen from the figure, even though the bid–ask spreads for trading in the underlying asset are relatively small, they nevertheless generate nontrivial transaction costs for the volatility arbitrage strategy. These costs are bigger for smaller values of volatility, which suggests that potential arbitrage violations will be harder to exploit for lower levels of volatility.

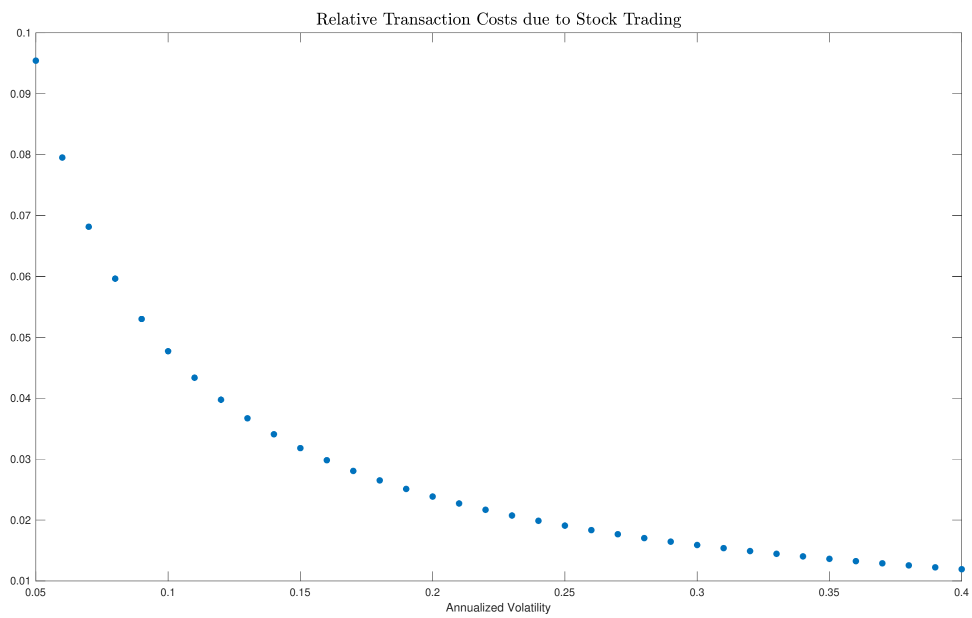

Figure 5: **Relative Transaction Costs due to Stock Trading**. The time window is $T = 30$ minutes and Δ_n corresponds to 5 minute frequency of rebalancing. The relative bid–ask spread (i.e., bid–ask spread over mid-quote) for trading in the underlying asset is set to $2 \times 2.1326e-05$. The transaction cost due to trading in the underlying asset is computed using the asymptotic result in (6.7). It is converted to a relative one by dividing by $T \times (\sigma_0^{\mathbb{P}})^2$.

The transaction costs due to trading in the options are easier to account for as the position in the options is static. Assuming constant proportional bid–ask spreads for trading options with different strikes, the relative transaction costs due to trading options based on numbers calibrated to the real data used in the application are around 4–6%. These numbers can be added to those reported in Figure 5 to get the total transaction costs of implementing the volatility arbitrage strategies. They are nontrivial and determine the size of a potential gap between $(\sigma_0^{\mathbb{P}})^2$ and $(\sigma_0^{\mathbb{Q}})^2$ that cannot be acted upon in practice.

6.2 Determining the Direction of the Volatility Strategy

Finally, the investor needs to know if the options market over- or under-values volatility. For this, we need $\mathbb{P}$- and $\mathbb{Q}$-estimates of volatility due to the small and frequent moves in the asset price at time t using information only up to that time. We denote these estimates with $\widehat{V}_{t,T}^{\mathbb{P}}$ and $\widehat{V}_{t,T}^{\mathbb{Q}}$. In general, conditional expectations under $\mathbb{P}$ are difficult to estimate as they require modeling the dynamics of the underlying asset. This is not the case, however, in our situation. The reason for this is that we need estimates of volatility that is due to the small and frequent asset price moves. In fact, if there was a continuous price record, then the investor should be able to infer exactly $\sigma_t^{\mathbb{P}}$ and $\nu_t^{\mathbb{P}}(dz)$, for small z, from the past trajectory of the asset price. In reality, the econometrician, and possibly the investor as well, has a discrete price record only.

There are various candidates for $\widehat{V}_{t,T}^{\mathbb{P}}$ and $\widehat{V}_{t,T}^{\mathbb{Q}}$. In Appendix A.2, we present the estimators that we use in our application. As we show in the Monte Carlo study, they appear to do a good job in identifying the potential existence of local arbitrage.

Given the volatility estimates $\widehat{V}_{t,T}^{\mathbb{P}}$ and $\widehat{V}_{t,T}^{\mathbb{Q}}$, the feasible strategy that exploits potential volatility arbitrage opportunities is given by:

- If $\widehat{V}_{t,T}^{\mathbb{P}} < 0.8 \times \widehat{V}_{t,T}^{\mathbb{Q}}$, engage in a hedged short position in truncated volatility.
- If $\widehat{V}_{t,T}^{\mathbb{P}} > 1.2 \times \widehat{V}_{t,T}^{\mathbb{Q}}$, engage in a hedged long position in truncated volatility.

We refer to the above strategy as the *volatility arbitrage strategy.* The reason for multiplying $\widehat{V}_{t,T}^{\mathbb{P}}$ by 1.2 or 0.8 when deciding whether to go short or long volatility is to account for the estimation uncertainty in $\widehat{V}_{t,T}^{\mathbb{P}}$ and $\widehat{V}_{t,T}^{\mathbb{Q}}$ as well as for accounting for the transaction costs when implementing the arbitrage strategies. A more formal approach for dealing with the estimation uncertainty would be to use the asymptotic distribution of these estimators. Similarly, one can use the asymptotic approximations for the transaction costs derived above to account for their effect and determine the rebalancing frequency. To keep things simple, we do not do this here and leave such analysis for future work.

We compare the above volatility arbitrage strategy with the popular one where one always sells total volatility, i.e., takes a short hedged position in truncated volatility with truncation parameter set to zero. We refer to this strategy as *selling volatility.* Note that this strategy does not exploit an arbitrage opportunity as it is exposed (loses money) to the big jumps in the asset price. If we do not hedge the above short position in volatility, as has been often done in earlier work, then we refer to this strategy as *naked selling of volatility.*

7 Monte Carlo Study

We now evaluate the performance of the volatility strategies discussed in the previous section in a simulation study that is designed to mimic the key properties of the real data that we are going to use later on.

7.1 Model

The dynamics of the underlying stock price under $\mathbb{P}$ in the Monte Carlo is given by (2.1) with

$$\alpha_t = 0.05, \quad \sigma_t^{\mathbb{P}} = \sigma_t, \quad d\sigma_t^2 = 8(0.02 - \sigma_t^2)dt + 0.2\sigma_t dB_t^{\mathbb{P}}, \tag{7.1}$$

where $B_t^{\mathbb{P}}$ is a $\mathbb{P}$-Brownian motion with $\text{corr}(W_t^{\mathbb{P}}, B_t^{\mathbb{P}}) = -0.9$. Our unit of time is one year, so the above parameters imply a half-life of a volatility shock of approximately one month. The jump compensator under $\mathbb{P}$ is given by

$$\nu^{\mathbb{P}}(dt, dz) = \nu^{ts}(dz)\, \sigma_t^2 dt, \tag{7.2}$$

where ν^{ts} is the compensator of a tempered stable process given by

$$\nu^{ts}(dz) = c_- \frac{e^{-\lambda_-|z|}}{|z|^{\alpha+1}} 1_{\{z<0\}} dz + c_+ \frac{e^{-\lambda_+|z|}}{|z|^{\alpha+1}} 1_{\{z>0\}} dz, \quad c_\pm \geq 0,\ \lambda_\pm > 0,\ \alpha < 2, \tag{7.3}$$

with parameters set to

$$\alpha = -1, \quad c_- = 0.1 \times \frac{(\lambda_-^{\mathbb{P}})^{2-\alpha}}{\Gamma(2-\alpha)}, \quad c_+ = 0.1 \times \frac{(\lambda_+^{\mathbb{P}})^{2-\alpha}}{\Gamma(2-\alpha)}, \quad \lambda_\pm^{\mathbb{P}} = 500. \tag{7.4}$$

This specification of the jumps corresponds to a time-changed double-exponential jump model that has been commonly used in existing parametric option pricing work.

Turning next to the risk-neutral probability, we assume that $\sigma_t^{\mathbb{Q}}$ has the same dynamics as σ_t above. Furthermore, the jump compensator under $\mathbb{Q}$ is the same as its $\mathbb{P}$-counterpart, with only the parameter $\lambda_-^{\mathbb{P}}$ changing to $\lambda_-^{\mathbb{Q}}$. For the specifications for which there is local arbitrage, we keep $\lambda_-^{\mathbb{Q}} = \lambda_-^{\mathbb{P}}$ while for the ones without such arbitrage opportunities, we determine $\lambda_-^{\mathbb{Q}}$ from the value of $\lambda_-^{\mathbb{P}}$ and

$$\left(\frac{\lambda_-^{\mathbb{P}}}{\lambda_-^{\mathbb{Q}}}\right)^{2-\alpha} = \frac{0.1 + 0.2 \times 1.2}{0.1}.$$

This choice of $\lambda_-^{\mathbb{Q}}$ implies a variance risk premium of 20%, which is similar to that observed in the real data. Thus, for all model specifications considered in our analysis, there is a variance risk premium (defined as difference between $\mathbb{Q}$- and $\mathbb{P}$-conditional expectations of future total return volatility) but its source differs. For the specifications with no arbitrage, the source of the variance risk premium is the pricing of negative jump risk in the asset price. For the specifications with local arbitrage opportunities, the variance risk premium is solely due to difference in diffusive spot volatility. The performance of the different volatility strategies designed in the previous sections should be able to identify the source of the variance risk premium.

Finally, in the free of arbitrage setting we set $\sigma_t^{\mathbb{Q}} = \sigma_t$, while we generate local volatility arbitrage opportunities by setting $\sigma_t^{\mathbb{Q}} \neq \sigma_t$. We consider situations in which the option market either overestimates or underestimates spot volatility.

7.2 Sampling Scheme and Choice of Truncation

The sampling scheme in the Monte Carlo mimics that of the real data we use. More specifically, we sample the underlying stock price at five second frequency. This corresponds to $\Delta_n = (1/252) \times (1/4680)$ in a 6.5 hour trading day when using business time convention. The starting values of the stock price and the volatility are $X_0 = 4000$ and $\sigma_0 = 0.02$, and we simulate price and volatility paths over 10,000 days. We generate zero-date option prices given the $\mathbb{Q}$ dynamics of X_t described above at every minute during the trading day. We set the bid/ask quotes to be 5% below/above the true option price. At each point in time, the available strike grid is equidistant with mesh of \$5. The highest and lowest strikes are determined as the lowest and highest multiples of \$5, respectively, for which the corresponding out-of-the-money option mid-quote falls below \$0.075. Similarly, for the underlying asset, we set the bid and ask to be 2.1326e-05% below and above, respectively, the true price. The size of the relative bid–ask spread in the underlying asset equals the average one for the data used in our empirical application.

We use the highest frequency of five seconds when forming our estimates of $\mathbb{P}$-volatility. We use a coarser frequency of five minutes for the dynamic position in the underlying asset. As discussed in the previous section, when deciding at what frequency to rebalance the position in the underlying asset we face a trade-off. A higher frequency means better hedging of the risks in the option portfolio but this comes at the expense of higher transaction costs. Our choice of frequency to update the position in the underlying asset aims to strike a balance between these two effects.

Finally, we use the truncation function $\phi(x) = e^{-x^2}$ and set the truncation parameter η_T in the following data-adaptive way:

$$\widehat{\eta}_{t,T} = \sqrt{\frac{\overline{\eta}}{\log(1/T)}} \times \frac{1}{\sqrt{\widehat{TV}^{\mathbb{Q}}_{t,T}(0)}}, \tag{7.5}$$

for some constant $\overline{\eta}$. We experiment with several different values of $\overline{\eta}$, with higher values corresponding to more aggressive truncation. The above choice of the truncation parameter means that the truncation is relative to the overall level of volatility.

7.3 Results

In Table 1, we report the percentage of times our estimates of $\mathbb{P}$- and $\mathbb{Q}$-volatility trigger activation of the volatility arbitrage strategy. In the case of no arbitrage and when the time horizon T is relatively long (2 hours), the strategy is erroneously triggered nearly half of the times. The reason for this is that for such relatively long horizon T, it is difficult to infer from the data whether the variance risk premium embedded in the options is due to compensation for big jumps in the asset price or due to the existence of local arbitrage opportunities. When the horizon T shrinks, this separation is easier. As a result for T equal to 30 minutes, the activation ratio drops to 10.3% only when markets are free of arbitrage. On the other hand, when there are large local arbitrage opportunities ($\sigma_t^{\mathbb{Q}}$ deviates from $\sigma_t^{\mathbb{P}}$ by more than 20%), the arbitrage volatility strategy is triggered almost every day. The activation ratio naturally drops when the size of the arbitrage violation gets smaller. This is particularly true for the case $\sigma_t^{\mathbb{Q}} = 0.9 \times \sigma_t^{\mathbb{P}}$ and illustrates the difficulties in identifying the existence of small local arbitrage opportunities from the data.

In Table 2, we report the performance of the various volatility strategies in terms of annualized Sharpe ratios in the Monte Carlo. The performance of the volatility arbitrage strategy is computed only on the days when the strategy is triggered. For simplicity, we report only unconditional Sharpe ratios. For computing conditional Sharpe ratios, one would need to estimate the conditional jump distribution which is significantly more difficult. Moreover, in presence of local arbitrage, the unconditional Sharpe ratio, just like the conditional one, should also remain high as T shrinks. For brevity, we also do not report estimates of $GSR(p)$, for different values of $p \neq 2$. To compare the performance of the different volatility strategies, with and without local arbitrage, the unconditional Sharpe ratio suffices.

We can draw several conclusions from the reported simulation results. First, when the market is free of local arbitrage, then the feasible volatility arbitrage strategy always looses money on average. This is due to the transaction costs. In contrast, the strategy which sells volatility generates positive and nontrivial Sharpe ratios in this case. Consistent with our theoretical analysis, the Sharpe ratios of selling volatility shrink as T shrinks. Second, when $\sigma_t^{\mathbb{P}}$ and $\sigma_t^{\mathbb{Q}}$ differ by around 10%, then the infeasible volatility arbitrage strategy generates positive Sharpe ratios as long as the truncation is not too high (i.e., $\overline{\eta}$ is below 1.5). The feasible volatility

Table 1: **Monte Carlo Results: Activation Ratios for the Volatility Arbitrage Strategy**

Scenario	TTM	Activation Ratio	Scenario	TTM	Activation Ratio
no arbitrage	30 minutes	10.3%			
	60 minutes	24.4%			
	90 minutes	37.8%			
	120 minutes	48.4%			
$\sigma_t^{\mathbb{Q}} = 1.1 \times \sigma_t^{\mathbb{P}}$	30 minutes	51.1%	$\sigma_t^{\mathbb{Q}} = 1.2 \times \sigma_t^{\mathbb{P}}$	30 minutes	83.1%
	60 minutes	66.0%		60 minutes	85.2%
	90 minutes	71.3%		90 minutes	86.8%
	120 minutes	74.0%		120 minutes	87.6%
$\sigma_t^{\mathbb{Q}} = 0.9 \times \sigma_t^{\mathbb{P}}$	30 minutes	44.3%	$\sigma_t^{\mathbb{Q}} = 0.8 \times \sigma_t^{\mathbb{P}}$	30 minutes	99.2%
	60 minutes	33.0%		60 minutes	99.2%
	90 minutes	28.1%		90 minutes	98.2%
	120 minutes	26.2%		120 minutes	96.4%

Note: TTM stands for time to maturity. The activation ratio for each feasible volatility arbitrage strategy corresponds to the percentage of days that the strategy is implemented from 10,000 days.

strategies also generate positive Sharpe ratios for any considered T when $\overline{\eta} = 0.5$ and for $T = 30$ minutes when $\overline{\eta} = 1.0$. The comparison of the feasible and infeasible volatility strategy reveals the cost of learning about the existence of a gap between $\sigma_t^{\mathbb{P}}$ and $\sigma_t^{\mathbb{Q}}$. Focusing on the shortest considered T of 30 minutes, for which our asymptotic analysis should apply best, we can see that for mild truncation, i.e., $\overline{\eta} = 0.5$, the feasible and infeasible volatility arbitrage strategies outperform a strategy of hedged selling of volatility if $\sigma_t^{\mathbb{P}}$ and $\sigma_t^{\mathbb{Q}}$ differ by around 20%. The gap is not as big as the one reported in Figure 4 and this shows the impact of discrete hedging, transaction costs and not applying conditioning when computing the Sharpe Ratio.

Overall, the Monte Carlo analysis shows two distinguishing features of the volatility arbitrage strategies in the presence versus absence of local arbitrage opportunities. One is the fact that, in the presence of arbitrage, the volatility arbitrage strategies deliver nontrivial Sharpe ratios even for moderate levels of truncation ($\overline{\eta} < 1.5$). The other is the fact that when there is a local arbitrage opportunity, the performance of the volatility arbitrage strategies does not deteriorate as the horizon T shrinks. Finally, and not surprisingly, the gap between the feasible and infeasible volatility strategies is higher for the smaller local arbitrage opportunities, as they are harder to detect.

8 Do Equity and Options Markets Agree about Volatility?

8.1 Data

We use three sets of data in the empirical analysis. The first one consists of intraday best bid and best ask quotes for the SPY ETF tracking the S&P 500 index. This dataset is extracted from the TAQ database. We sample the SPY every 5 seconds during the trading hours range of 9:30–16:00 EST. The second dataset consists of prices of zero-date S&P 500 index (SPX) options. The source of this dataset is the CBOE DataShop. We sample the option prices at one minute frequency and keep only out-of-the-money options with positive bid quotes. We perform the analysis at a point in time during the day only if there are at least five strikes with non-zero

Table 2: **Monte Carlo Results: Sharpe Ratios**

Scenario	TTM	Strategy							
		infeasible vol. arb.			feasible vol. arb.			selling	naked short
		$\overline{\eta} = 0.5$	$\overline{\eta} = 1.0$	$\overline{\eta} = 1.5$	$\overline{\eta} = 0.5$	$\overline{\eta} = 1.0$	$\overline{\eta} = 1.5$	volatility	volatility
no arbitrage	30 minutes				−2.34	−2.56	−2.75	1.15	1.04
	60 minutes				−1.15	−1.46	−1.64	1.57	1.01
	90 minutes				−0.54	−1.08	−1.40	2.18	1.20
	120 minutes				−0.05	−0.84	−1.31	2.63	1.32
$\sigma_t^{\mathbb{Q}} = 1.1 \times \sigma_t^{\mathbb{P}}$	30 minutes	0.90	0.51	0.01	0.68	0.28	−0.19	0.98	0.92
	60 minutes	0.92	0.42	−0.16	0.41	−0.04	−0.50	1.22	0.83
	90 minutes	1.41	0.68	−0.01	0.54	−0.06	−0.61	1.70	0.98
	120 minutes	1.57	0.72	−0.01	0.74	0.08	−0.48	2.03	1.08
$\sigma_t^{\mathbb{Q}} = 0.9 \times \sigma_t^{\mathbb{P}}$	30 minutes	0.98	0.60	0.13	1.06	0.78	0.36	−3.33	−1.95
	60 minutes	1.25	0.63	−0.01	1.09	0.57	−0.06	−4.52	−2.16
	90 minutes	1.16	0.31	−0.32	0.80	−0.12	−0.80	−5.23	−2.14
	120 minutes	1.15	0.16	−0.52	1.26	0.16	−0.67	−5.87	−2.18
$\sigma_t^{\mathbb{Q}} = 1.2 \times \sigma_t^{\mathbb{P}}$	30 minutes	3.71	3.28	2.46	3.62	3.15	2.32	3.33	2.49
	60 minutes	4.20	3.47	2.40	3.82	3.17	2.19	4.38	2.49
	90 minutes	5.37	4.05	2.59	4.51	3.35	2.06	5.57	2.74
	120 minutes	5.92	4.26	2.61	4.97	3.56	2.13	6.45	2.93
$\sigma_t^{\mathbb{Q}} = 0.8 \times \sigma_t^{\mathbb{P}}$	30 minutes	3.62	2.97	1.98	3.63	2.97	1.99	−5.27	−3.23
	60 minutes	4.17	3.03	1.74	4.18	3.03	1.74	−7.10	−3.48
	90 minutes	4.53	2.72	1.35	4.56	2.74	1.36	−8.27	−3.51
	120 minutes	4.64	2.50	1.06	4.64	2.49	1.04	−9.36	−3.60

Note: TTM stands for time to maturity. The reported numbers for each strategy correspond to annualized Sharpe ratios estimated from 10,000 days of implementing the strategies. The per-period Sharpe ratios are annualized by multiplying them by $\sqrt{252}$.

bid quotes for out-of-the-money options. The moneyness of the options is determined by implied forward, which in turn is recovered via put–call parity from pairs of call and put mid-quotes with the same strike for which the difference in the put–call premium is the smallest in absolute value. We use three such pairs with smallest put–call premium gap and take the median as our implied forward estimate. Finally, we obtain closing prices for the S&P 500 index from the CRSP database.

The sample period for the study is 2020–2023. Prior to Spring 2022, the available SPX option expiration dates were Monday, Wednesday and Friday. After Spring 2022, there are expiration dates on every trading day of the week. We remove from the analysis days with partial trading around holidays. We further remove days with FOMC announcements as the zero-date options on these days can be partially affected (depending on the time of the day) by the event risk contained in these announcements. To account for the well-known and pronounced intraday volatility pattern, we adjust the forecast of volatility $\widehat{V}_{t,T}^{\mathbb{P}}$ by the ratio of the sample average variance over the period $[t, t+k_n\Delta_n]$ by that over the period $[t, t-k_n\Delta_n]$ from which the forecast is constructed.

8.2 Empirical Evidence

We implement the feasible volatility arbitrage strategies in exactly the same way as in the Monte Carlo. Our estimates for $\mathbb{P}$- and $\mathbb{Q}$-volatility which determine the activation of these strategies are plotted in Figure 6. As seen from the figure, for all considered values of T, the estimates $\widehat{V}_{t,T}^{\mathbb{P}}$ and $\widehat{V}_{t,T}^{\mathbb{Q}}$ are close to each other suggesting no significant local volatility arbitrage opportunities.

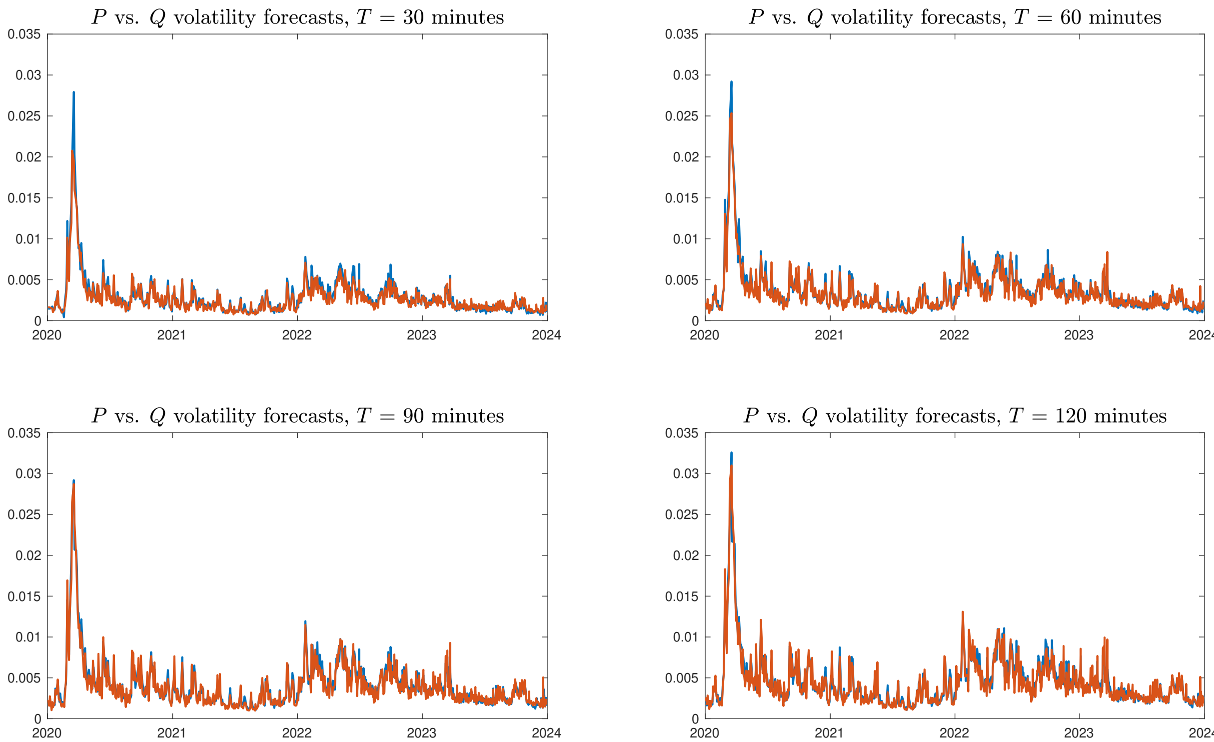

Figure 6: **Risk-Neutral versus True Conditional Expected Truncated Volatility**. The estimates of the conditional expected volatility under $\mathbb{P}$ and $\mathbb{Q}$ are $\sqrt{\widehat{V}_{t,T}^{\mathbb{P}}}$ (blue line) and $\sqrt{\widehat{V}_{t,T}^{\mathbb{Q}}}$ (red line), respectively.

The performance of the various volatility strategies on the real data are reported in Table 3. They indicate a lack of local arbitrage opportunities due to different perceptions of volatility by equity and option markets. Indeed, the volatility arbitrage strategies yield negative Sharpe ratios for all considered levels of $\overline{\eta}$ and all horizons T. This performance is in sharp contrast to that of the hedged selling of volatility strategy which generates high Sharpe ratios. Comparing the performance of the different strategies on the real data and in the Monte Carlo, we can conclude that the real data performance is in line with that in the Monte Carlo for the case $\sigma_t^{\mathbb{P}} = \sigma_t^{\mathbb{Q}}$. One notable difference between the performance of the volatility arbitrage strategy in the real data and in the Monte Carlo for the case $\sigma_t^{\mathbb{P}} = \sigma_t^{\mathbb{Q}}$ is the slightly higher activation ratios of the volatility arbitrage strategies in the real data. This reveals the slightly higher level of uncertainty about short-term future volatility in the real data than in the Monte Carlo, either due to volatility jumps and/or to time-varying time-of-day volatility effects.

In Figure 7, we display the returns from the different volatility strategies for a horizon of one hour. We can see from the figure that the volatility arbitrage strategies have nearly symmetric returns. Also, there is no apparent clustering of the periods when the strategies are inactive.

Table 3: **Empirical Results**

Strategy	TTM	SR	% active	Skewness	Kurtosis
vol. arb., $\overline{\eta} = 0.5$	30 minutes	-0.10 (1.22)	71%	0.57	7.62
	60 minutes	-0.91 (1.90)	68%	0.24	7.87
	90 minutes	-0.55 (1.48)	62%	-0.06	4.84
	120 minutes	-2.15 (1.95)	61%	0.13	4.12
vol. arb., $\overline{\eta} = 1.0$	30 minutes	-0.52 (1.40)	71%	0.61	9.81
	60 minutes	-1.13 (1.69)	68%	0.23	7.76
	90 minutes	-0.99 (1.83)	62%	-0.89	7.67
	120 minutes	-2.37 (2.56)	61%	-1.48	11.44
vol. arb., $\overline{\eta} = 1.5$	30 minutes	-0.79 (1.45)	71%	0.44	11.49
	60 minutes	-1.14 (1.42)	68%	0.27	6.91
	90 minutes	-1.08 (1.91)	62%	-1.40	11.15
	120 minutes	-2.40 (2.29)	61%	-1.80	13.22
selling volatility	30 minutes	2.56 (2.27)	100%	-3.24	23.56
	60 minutes	4.73 (3.47)	100%	-4.62	42.94
	90 minutes	4.15 (4.82)	100%	-7.30	106.05
	120 minutes	3.61 (4.60)	100%	-7.77	108.80
naked short volatility	30 minutes	0.67 (2.27)	100%	-8.96	142.82
	60 minutes	0.84 (2.05)	100%	-7.88	100.79
	90 minutes	0.39 (2.25)	100%	-11.50	205.48
	120 minutes	-0.10 (1.99)	100%	-11.71	196.30

Note: TTM stands for time to maturity and SR for annualized Sharpe Ratio. The per-period Sharpe ratios are annualized by multiplying them by $\sqrt{252}$. The numbers in brackets are Newey–West HAC standard errors.

Finally, the returns from these strategies do not appear significantly fat-tailed. That said, jump risk is apparently still affecting these strategies, which is in line with our theoretical analysis. On the other hand, the strategies involving selling of total volatility (the bottom two panels of the figure) are extremely heavy tailed and highly negatively skewed. The reason for this is the fact that these strategies are exposed significantly more to the big jumps in the underlying asset price than the truncated volatility strategies. This is further confirmed by comparing the skewness and kurtosis of the various strategies reported in Table 3. Thus, in particular, the high Sharpe ratio that the hedged short volatility strategy generates comes at the cost of very fat-tailed and skewed distribution.

9 Conclusion

In this paper we consider the question whether equity and option markets are integrated, i.e., whether there is an aggregate pricing kernel that can jointly rationalize observed equity and option prices. We show that such market integration is equivalent to an agreement between the two markets about the volatility due to small and frequent moves in the asset price. Absence of such an agreement leads to local arbitrage opportunities, which can be exploited via a portfolio of short-dated options together with a dynamic position in the underlying asset that aims to hedge some of the risk in the options. Empirically, we find that, in spite of the large premium embedded in short-dated S&P 500 index options, their prices correctly reflect the volatility due

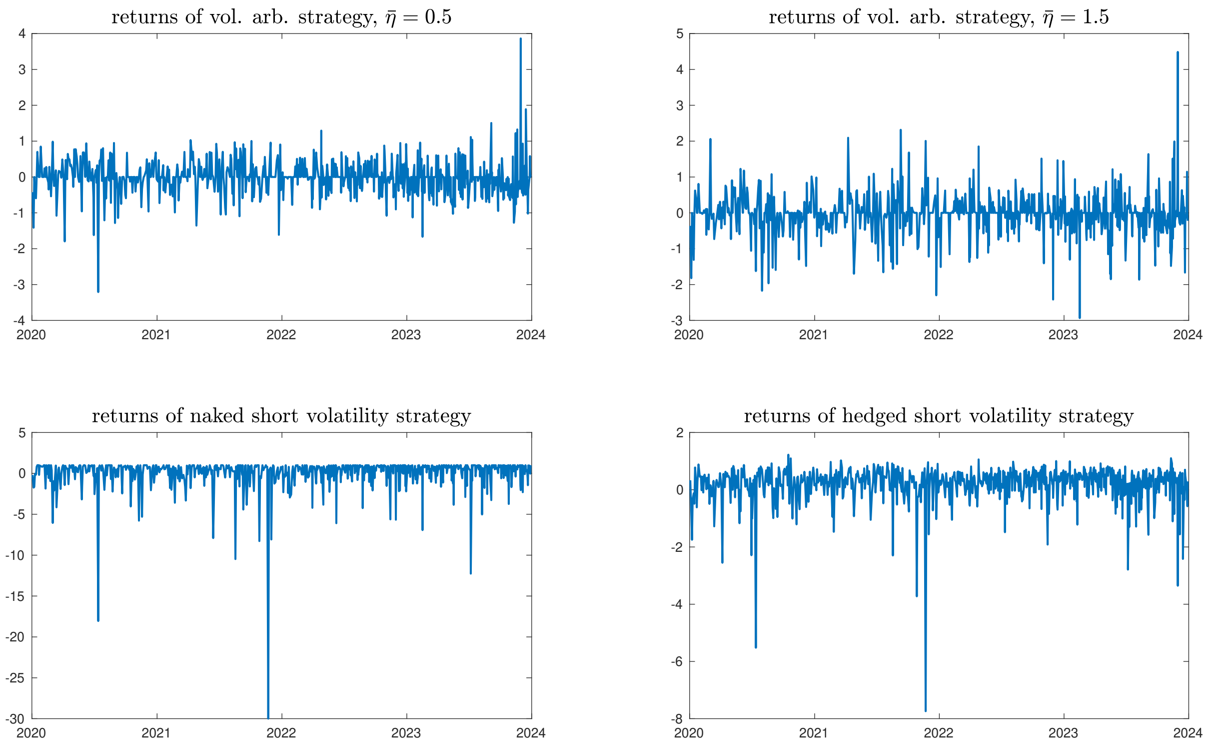

Figure 7: **Returns of Volatility Arbitrage Strategies**. If a volatility strategy is not activated on a given day, then we put a return of zero on that day. The time horizon is $T = 60$ minutes.

to small and frequent asset price moves. Thus, zero-date options are expensive but do not appear mis-priced.

A Appendix

A.1 Auxiliary Results and Proofs for Sections 2 and 3

We first explain how a refinement of (2.4) and (2.5) can be turned into a necessary and sufficient condition for market integration. As we are only interested in equivalence of $\mathbb{P}$ and $\mathbb{Q}$ in restriction to $\mathcal{F}_\tau$, there is no loss of generality to assume that $\mathcal{F} = \mathcal{F}_\tau$ and $\mathbb{P} = \mathbb{P}|_{\mathcal{F}_\tau}$.

Theorem A.1. *Suppose that $\mathcal{G}_0$ is a sub-σ-field of $\mathcal{F}_0$, equipped with a probability measure P_0 (initial condition). Further suppose that $\mathbb{P}$ is the locally unique probability measure (see Definition III.2.37 in Jacod and Shiryaev (2003)) that satisfies $\mathbb{P}|_{\mathcal{G}_0} = P_0$ and that renders X a special semimartingale*[15] *with local characteristics $(B^{X,\mathbb{P}}, C^{X,\mathbb{P}}, \nu^{X,\mathbb{P}})$, where for all $t \in [0, \tau]$,*

$$B_t^{X,\mathbb{P}} = \int_0^t X_s \alpha_s ds, \quad C_t^{X,\mathbb{P}} = \int_0^t (X_s \sigma_s^{\mathbb{P}})^2 ds, \quad \nu^{X,\mathbb{P}}(dt, dz) = \left[\nu_t^{\mathbb{P}} \circ (X_{t-}(e^z - 1))^{-1}\right](dz)dt. \tag{A.1}$$

[15] A semimartingale X is special if it is locally integrable. In this case, it can be decomposed into the sum of a local martingale and a predictable process of finite variation (called the drift of X). Furthermore, in writing the semimartingale characteristics of X, one can then take the identity as truncation function (i.e., no truncation), which is implicitly assumed in (A.1) and (A.2).

Then $\mathcal{Q}_e$ is equal to the family of risk-neutral measures $\mathbb{Q} \in \mathcal{Q}$ for which $\mathbb{Q}|_{\mathcal{G}_0}$ is equivalent to P_0 and X is a martingale under $\mathbb{Q}$ with local characteristics $(B^{X,\mathbb{Q}}, C^{X,\mathbb{Q}}, \nu^{X,\mathbb{Q}})$ given by

$$B_t^{X,\mathbb{Q}} = 0, \quad C_t^{X,\mathbb{Q}} = \int_0^t (X_s \sigma_s^{\mathbb{Q}})^2 ds, \quad \nu^{X,\mathbb{Q}}(dt, dz) = \left[\nu_t^{\mathbb{Q}} \circ (X_{t-}(e^z - 1))^{-1}\right](dz)dt \tag{A.2}$$

for $t \in [0, \tau]$, where

$$\sigma_t^{\mathbb{Q}} = \sigma_t^{\mathbb{P}} \quad \text{and} \quad \nu_t^{\mathbb{Q}}(dz) = Y_t(z)\nu_t^{\mathbb{P}}(dz), \tag{A.3}$$

and $Y_t(z)$ is a strictly positive predictable function such that $\mathbb{P}$- and $\mathbb{Q}$-almost surely,

$$\begin{aligned} &\int_0^\tau \int_{\mathbb{R}} \left(\frac{\alpha_t + \int_{\mathbb{R}} (e^z - 1)(Y_t(z) - 1)\nu_t^{\mathbb{P}}(dz)}{\sigma_t^{\mathbb{P}}} \right)^2 dt < \infty, \\ &\int_0^\tau \int_{\mathbb{R}} (|Y_t(z) - 1|^2 \wedge |Y_t(z) - 1|)\nu_t^{\mathbb{P}}(dz)dt < \infty. \end{aligned} \tag{A.4}$$

Proof. Suppose that $\mathbb{Q} \in \mathcal{Q}_e$. By definition, we have $\mathbb{P} \sim \mathbb{Q}$ ("$\mathbb{P}$ is equivalent to $\mathbb{Q}$"), which implies $\mathbb{Q}|_{\mathcal{G}_0} \sim P_0$. Moreover, by Girsanov's theorem (see Theorem III.3.24 in Jacod and Shiryaev (2003)), X is a $\mathbb{Q}$-semimartingale with characteristics $(B^{X,\mathbb{Q}}, C^{X,\mathbb{Q}}, \nu^{X,\mathbb{Q}})$ where $C^{X,\mathbb{Q}}$ and $\nu^{X,\mathbb{Q}}$ are given by (A.2) and satisfy (A.3) for some nonnegative measurable $Y_t(z)$, and

$$B_t^{X,\mathbb{Q}} = \int_0^t X_s \left(\alpha_s + \beta_s X_s (\sigma_s^{\mathbb{P}})^2 + \int_{\mathbb{R}} (e^z - 1)(Y_s(z) - 1)\nu_s^{\mathbb{P}}(dz) \right) ds \tag{A.5}$$

for some predictable process β_t. Because $\mathbb{Q} \in \mathcal{Q}_e \subseteq \mathcal{Q}$, we know that X is a $\mathbb{Q}$-martingale, hence $B^{X,\mathbb{Q}} \equiv 0$, which implies that β_t can be chosen such that

$$\beta_t X_t = -\frac{\alpha_t + \int_{\mathbb{R}} (e^z - 1)(Y_t(z) - 1)\nu_t^{\mathbb{P}}(dz)}{(\sigma_t^{\mathbb{P}})^2}, \quad t \in [0, \tau]. \tag{A.6}$$

Together with the assumption $\mathbb{P} \sim \mathbb{Q}$, we can use Théorème 4.1 in Jacod and Mémin (1976) to deduce that $Y_t(z)$ can be chosen as strictly positive and must satisfy (A.4) $\mathbb{P}$- and $\mathbb{Q}$-almost surely.

Conversely, suppose now that $\mathbb{Q} \in \mathcal{Q}$ satisfies the properties listed in the lemma. By assumption, $\mathbb{P}$ is the locally unique probability measure that equals P_0 if restricted on $\mathcal{G}_0$ and makes X a semimartingale with local characteristics given in (A.1). Therefore, by Théorème 4.3 in Jacod and Mémin (1976), we have $\mathbb{P} \sim \mathbb{Q}$, which means $\mathbb{Q} \in \mathcal{Q}_e$. □

Proof of Lemma 2.1. Because $\mathbb{Q}$ is equivalent to $\mathbb{P}$, we have $M_t > 0$ almost surely for all $t \in [0, \tau]$. As M is a $\mathbb{P}$-martingale, by right-continuity, we obtain (2.7). To upgrade this to (2.8), by dominated convergence, it suffices to note that $\mathbb{E}_t^{\mathbb{P}}(\sup_{r \in [t,s]} |M_r|^q) < \infty$ by Doob's martingale inequality and our assumption that $\mathbb{E}_t^{\mathbb{P}}(|M_s|^q) < \infty$. □

Proof of Lemma 3.1. By the definition of the stochastic discount factor, we have

$$\mathbb{E}_t^{\mathbb{Q}}(R_{t,T}) = \mathbb{E}_t^{\mathbb{P}}(D_{t,T} R_{t,T}),$$

and therefore,

$$\mathbb{E}_t^{\mathbb{Q}}(R_{t,T}) - \mathbb{E}_t^{\mathbb{P}}(R_{t,T}) = \mathbb{E}_t^{\mathbb{P}}\left[(D_{t,T} - 1)R_{t,T}\right] = \mathbb{E}_t^{\mathbb{P}}\left[(D_{t,T} - 1)\left(R_{t,T} - \mathbb{E}_t^{\mathbb{P}}(R_{t,T})\right)\right],$$

because $\mathbb{E}_t^{\mathbb{P}}(D_{t,T}-1)=0$. Applying Hölder's inequality, we derive the estimate

$$\begin{aligned}&\left|\mathbb{E}_t^{\mathbb{P}}\left[(D_{t,T}-1)\left(R_{t,T}-\mathbb{E}_t^{\mathbb{P}}(R_{t,T})\right)\right]\right|\\&\quad\leq\left(\mathbb{E}_t^{\mathbb{P}}\left(\left|R_{t,T}-\mathbb{E}_t^{\mathbb{P}}(R_{t,T})\right|^p\right)\right)^{\frac{1}{p}}\left(\mathbb{E}_t^{\mathbb{P}}\left(|D_{t,T}-1|^{\frac{p}{p-1}}\right)\right)^{1-\frac{1}{p}}.\end{aligned}\tag{A.7}$$

As $\mathbb{E}_t^{\mathbb{Q}}(R_{t,T})=0$ by (3.3), we obtain

$$\left(\mathbb{E}_t^{\mathbb{P}}\left(|D_{t,T}-1|^{\frac{p}{p-1}}\right)\right)^{1-\frac{1}{p}}\geq\frac{\left|\mathbb{E}_t^{\mathbb{Q}}(R_{t,T})-\mathbb{E}_t^{\mathbb{P}}(R_{t,T})\right|}{\left(\mathbb{E}_t^{\mathbb{P}}\left(\left|R_{t,T}-\mathbb{E}_t^{\mathbb{P}}(R_{t,T})\right|^p\right)\right)^{\frac{1}{p}}}=|GSR_{t,T}(p)|,\tag{A.8}$$

proving the lemma. □

Proof of Theorem 3.2. Suppose that $\mathbb{Q}\in\mathcal{Q}_e$. By Lemma 3.1 and Definition 3, local arbitrage implies that for a subsequence of time horizons, which we still denote by T,

$$\left(\mathbb{E}_t^{\mathbb{P}}\left(|D_{t,T}-1|^q\right)\right)^{1/q}\geq|GSR_{t,T}(p)|\to G_t(p)\quad\text{a.s. as } T\downarrow 0,$$

where $G_t(p)$ denotes the left-hand side of (3.5) and is therefore a strictly positive random variable. As this contradicts (2.8), the assumptions behind (2.8) must be violated, which means that we must have $\mathbb{E}_t^{\mathbb{P}}(|M_s|^q)=\infty$ for all $s\in(t,\tau]$. □

A.2 Estimating Volatility Under $\mathbb{P}$ and $\mathbb{Q}$

In this section, we provide the two volatility estimators, $\widehat{V}_{t,T}^{\mathbb{P}}$ and $\widehat{V}_{t,T}^{\mathbb{Q}}$, that we use to determine if there is an arbitrage opportunity. These estimators are based on the characteristic function of the price increments, which, similarly to the truncated volatility, can separate small and frequent moves from the big and infrequent ones, see Jacod and Todorov (2014) and Todorov (2019). We start with $\widehat{V}_{t,T}^{\mathbb{Q}}$. It is given by

$$\widehat{V}_{t,T}^{\mathbb{Q}}=-\frac{2T}{\widehat{u}_t^2}\log\left|\widehat{\mathcal{L}}_{t,T}^{\mathbb{Q}}(\widehat{u}_t)\right|,\tag{A.9}$$

where the risk-neutral characteristic function of the price increments is inferred from the options via

$$\widehat{\mathcal{L}}_{t,T}^{\mathbb{Q}}(u)=1-(u^2+iu)\sum_{j=2}^{N_t}\frac{e^{iu(\log(K_{t,j-1})-\log(X_t))}}{K_{t,j-1}^2}O_{t,T}(K_{t,j-1})(K_{t,j}-K_{t,j-1}),\tag{A.10}$$

and the characteristic exponent is set to

$$\widehat{u}_t=\sqrt{\frac{-2T\log(0.3)}{\widehat{TV}_{t,T}(0)}}.\tag{A.11}$$

This estimator was proposed by Todorov (2019) and the above construction is behind the spot volatility index reported by the CBOE options exchange.

We continue next with $\widehat{V}_{t,T}^{\mathbb{P}}$. It is constructed as the return counterpart of the option-based estimator above. More specifically, we set

$$\widehat{V}_{t,T}^{\mathbb{P}}=-\frac{2Tk_n}{\widehat{u}_t^2}\log\left|\frac{1}{m_n}\sum_{j=1}^{m_n}e^{i\widehat{u}_t(x_{t-(j-1)\Delta_n}-x_{t-j\Delta_n})/\sqrt{T}}\right|,\tag{A.12}$$

where $k_n=\lfloor T/\Delta_n\rfloor$ and $m_n=12\times180$, which corresponds to a window for volatility estimation of 3 hours. The theoretical properties of $\widehat{V}_{t,T}^{\mathbb{P}}$ are investigated in Jacod and Todorov (2014).

A.3 Assumptions and Proofs for Theorem 5.1 and Theorem 5.2

For the proof of Theorems 5.1 and 5.2, we can assume without loss of generality that $t = 0$ and that $X_0 = 1$. We also introduce

$$\theta(x) = \phi(x)x^2 \qquad \text{and} \qquad \Phi(\eta, x)_T = \phi(\eta x/\sqrt{T})x^2 = \frac{T}{\eta^2}\theta(\eta x/\sqrt{T}). \tag{A.13}$$

Then

$$\begin{aligned} \theta'(x) &= \phi'(x)x^2 + 2x\phi(x), & \theta''(x) &= \phi''(x)x^2 + 4\phi'(x)x + 2\phi(x), \\ \theta'''(x) &= \phi'''(x)x^2 + 6\phi''(x)x + 6\phi'(x), & \theta^{(4)}(x) &= \phi^{(4)}(x)x^2 + 8\phi'''(x)x + 12\phi''(x) \end{aligned} \tag{A.14}$$

and

$$\frac{\partial}{\partial x}\Phi(\eta, x)_T = \frac{\sqrt{T}}{\eta}\theta'(\eta x/\sqrt{T}), \qquad \frac{\partial^2}{\partial x^2}\Phi(\eta, x)_T = \theta''(\eta x/\sqrt{T}). \tag{A.15}$$

By Itô's formula, the log-price $x_t = \log X_t$ satisfies

$$dx_t = a_t^{\mathbb{S}} dt + \sigma_t^{\mathbb{S}} dW_t^{\mathbb{S}} + \int_{\mathbb{R}} z\widetilde{\mu}^{\mathbb{S}}(dt, dz), \quad x_0 = \log X_0,$$

where $a_t^{\mathbb{S}} = \alpha_t^{\mathbb{S}} - \frac{1}{2}(\sigma_t^{\mathbb{S}})^2 - \int_{\mathbb{R}}(e^z - 1 - z)\nu_t^{\mathbb{S}}(dz)$ with $\alpha_t^{\mathbb{S}} = \alpha_t$ if $\mathbb{S} = \mathbb{P}$ and $\alpha_t^{\mathbb{S}} = 0$ if $\mathbb{S} = \mathbb{Q}$. For Theorem 5.1, we assume the following conditions on ϕ and the coefficients of x:

Assumption 1. *The function $\phi(x)$ is nonnegative, symmetric, equal to unity at zero and twice continuously differentiable with all its derivatives (including ϕ itself) decreasing exponentially fast as $|x| \to \infty$.*

Assumption 2. *For both $\mathbb{S} = \mathbb{P}$ and $\mathbb{S} = \mathbb{Q}$, we have the following:*

1. *There is $r^{\mathbb{S}} \in (0, 2)$ such that for all $0 \leq s \leq t$,*

$$\mathbb{E}_s^{\mathbb{P}}\left(\int_{\mathbb{R}} |z|^{r^{\mathbb{S}}} \nu_t^{\mathbb{S}}(dz)\right) < \infty. \tag{A.16}$$

2. *There are positive $\mathcal{F}_t$-measurable random variables C_t, only depending on t, such that*

$$\sup_{0 \leq s \leq t \leq s+1} \mathbb{E}_s^{\mathbb{S}}(z_t^4) < C_s \quad \textit{for } z \in \{a^{\mathbb{S}}, (\sigma^{\mathbb{S}})^{-4}\} \tag{A.17}$$

and

$$\sup_{0 \leq s \leq t \leq s+1} \mathbb{E}_s^{\mathbb{S}}((z_t - z_s)^2) \leq C_s(t - s) \quad \textit{for } z \in \{\sigma^{\mathbb{S}}, (\sigma^{\mathbb{S}})^2\} \tag{A.18}$$

and

$$\sup_{0 \leq s \leq t \leq s+1} \sup_{\psi} \mathbb{E}_s^{\mathbb{S}}\left(\left(\int_{\mathbb{R}} \psi_t(z)(|z|^{r^{\mathbb{S}}} \vee |z|^2)(\nu_t^{\mathbb{S}} - \nu_s^{\mathbb{S}})(dz)\right)^2\right) \leq C_s(t - s), \tag{A.19}$$

where in the last line the second supremum is taken over all adapted processes $\psi = \psi_t(z)$ that are uniformly bounded by 1.

There are many choices of the truncation function ϕ that satisfy Assumption 1. For example, the one that we use in our application is $\phi(x) = e^{-x^2}$. Assumption 2 is a weak restriction on the asset price dynamics. This assumption is satisfied by most models considered in finance. In particular, we do not restrict the activity of the jumps (e.g., jumps of infinite variation are allowed).

Proof of Theorem 5.1. We compute the expected return of the hedged truncated volatility portfolio: By (4.3), (5.2), (5.7) and (5.8), we have

$$
\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}(\mathcal{R}_{0,T}) &= \frac{\mathbb{E}_0^{\mathbb{P}}(\mathcal{RG}_{0,T}) - \mathcal{C}_{0,T}(\eta_T)}{\mathcal{C}_{0,T}(\eta_T)} \\
&= \frac{TV_{0,T}^{\mathbb{P}}(\eta_T) - TV_{0,T}^{\mathbb{Q}}(\eta_T)}{TV_{0,T}^{\mathbb{Q}}(\eta_T)} - \frac{\int_0^T \mathbb{E}_0^{\mathbb{P}}(f_T'(X_t;X_0)\alpha_t X_t)dt}{TV_{0,T}^{\mathbb{Q}}(\eta_T)}.
\end{aligned}
\tag{A.20}
$$

By Itô's formula, for both $\mathbb{S}=\mathbb{P}$ and $\mathbb{S}=\mathbb{Q}$,

$$
\begin{aligned}
TV_{0,T}^{\mathbb{S}}(\eta_T) &= \mathbb{E}_0^{\mathbb{S}}(\Phi(\eta_T, x_T)_T) \\
&= \mathbb{E}_0^{\mathbb{S}}\Bigg(\frac{1}{2}\int_0^T \theta''(\tfrac{\eta_T}{\sqrt{T}}x_t)(\sigma_t^{\mathbb{S}})^2 dt + \frac{\sqrt{T}}{\eta_T}\int_0^T \theta'(\tfrac{\eta_T}{\sqrt{T}}x_t)a_t^{\mathbb{S}} dt \\
&\quad + \int_0^T \int_{\mathbb{R}} \left[\Phi(\eta_T, x_t+z) - \Phi(\eta_T, x_t) - \frac{\partial}{\partial x}\Phi(\eta_T, x_t)_T z\right]\nu_t^{\mathbb{S}}(dz)dt\Bigg) \\
&= \mathbb{E}_0^{\mathbb{S}}\Bigg(\frac{1}{2}\int_0^T \theta''(\tfrac{\eta_T}{\sqrt{T}}x_t)(\sigma_t^{\mathbb{S}})^2 dt + \frac{\sqrt{T}}{\eta_T}\int_0^T \theta'(\tfrac{\eta_T}{\sqrt{T}}x_t)a_t^{\mathbb{S}} dt \\
&\quad + \int_0^T \int_{\mathbb{R}} \int_0^1 \theta''(\tfrac{\eta_T}{\sqrt{T}}(x_t+vz))(1-v)dv\, z^2 \nu_t^{\mathbb{S}}(dz)dt\Bigg).
\end{aligned}
\tag{A.21}
$$

Note that θ and its derivatives are bounded functions and $|\theta'(x)| \le C|x|$. Therefore, by (A.18), (A.19) and the fact that $\mathbb{E}_0^{\mathbb{S}}(|x_t/\sqrt{T}|) \le (\mathbb{E}_0^{\mathbb{S}}(|x_t/\sqrt{T}|^2))^{1/2} \le C$ for all $t \in [0,T]$, we have

$$
\begin{aligned}
\frac{1}{T}TV_{0,T}^{\mathbb{S}}(\eta_T) &= \frac{(\sigma_0^{\mathbb{S}})^2}{2T}\int_0^T \mathbb{E}_0^{\mathbb{S}}\left(\theta''(\tfrac{\eta_T}{\sqrt{T}}x_t)\right)dt \\
&\quad + \frac{1}{T}\int_0^T \int_{\mathbb{R}} \int_0^1 \mathbb{E}_0^{\mathbb{S}}\left(\theta''(\tfrac{\eta_T}{\sqrt{T}}(x_t+vz))\right)(1-v)dv\, z^2\nu_0^{\mathbb{S}}(dz)dt + O_p(\sqrt{T}).
\end{aligned}
\tag{A.22}
$$

As $\theta''(0)=2$, using the mean-value theorem, we further obtain

$$
\frac{1}{T}TV_{0,T}^{\mathbb{S}}(\eta_T) = (\sigma_0^{\mathbb{S}})^2 + \int_{\mathbb{R}}\int_0^1 \theta''(\tfrac{\eta_T}{\sqrt{T}}vz)(1-v)dv\, z^2\nu_0^{\mathbb{S}}(dz) + O_p(\eta_T).
$$

We have $\int_0^1 \theta''(\frac{\eta_T}{\sqrt{T}}vz)(1-v)dv = \phi(\frac{\eta_T}{\sqrt{T}}z)$ and $\phi(x) \to 0$ as $|x| \to \infty$, so the dominated convergence theorem implies that

$$
\int_{\mathbb{R}}\int_0^1 \theta''(\tfrac{\eta_T}{\sqrt{T}}vz)(1-v)dv\, z^2\nu_0^{\mathbb{S}}(dz) = \int_{\mathbb{R}} \phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2\nu_0^{\mathbb{S}}(dz) = o_p(1).
$$

In summary, we have shown that

$$
\frac{1}{T}TV_{0,T}^{\mathbb{S}}(\eta_T) = (\sigma_0^{\mathbb{S}})^2 + o_p(1). \tag{A.23}
$$

Since $\sigma_0^{\mathbb{P}} \neq \sigma_0^{\mathbb{Q}}$ by assumption, it follows that

$$
\mathbb{E}_0^{\mathbb{P}}(\mathcal{R}_{0,T}) = \frac{(\sigma_0^{\mathbb{P}})^2 - (\sigma_0^{\mathbb{Q}})^2}{(\sigma_0^{\mathbb{Q}})^2} + o_p(1). \tag{A.24}
$$

Next, we turn to the L_p-risk of the hedged volatility portfolio, which is given by

$$\left(\mathbb{E}_0^{\mathbb{P}}\left(|\mathcal{R}_{0,T}-\mathbb{E}_0^{\mathbb{P}}(\mathcal{R}_{0,T})|^p\right)\right)^{1/p}=\frac{\left(\mathbb{E}_0^{\mathbb{P}}\left(|\mathcal{RG}_{0,T}-\mathbb{E}_0^{\mathbb{P}}(\mathcal{RG}_{0,T})|^p\right)\right)^{1/p}}{TV_{0,T}^{\mathbb{Q}}(\eta_T)}. \tag{A.25}$$

By (A.14), we have $\Phi(\eta_T,x_{t-}+z)-\Phi(\eta_T,x_{t-})-\frac{\partial}{\partial x}\Phi(\eta_T,x_{t-})z=\int_0^1\theta''(\frac{\eta_T}{\sqrt{T}}(x_t+vz))(1-v)dv$. Thus, Itô's formula implies that

$$\begin{aligned}
&\mathcal{RG}_{0,T}-\mathbb{E}_0^{\mathbb{P}}(\mathcal{RG}_{0,T})\\
&\quad=\frac{1}{2}\int_0^T\left[\theta''(\tfrac{\eta_T}{\sqrt{T}}x_t)(\sigma_t^{\mathbb{P}})^2-\mathbb{E}_0^{\mathbb{P}}(\theta''(\tfrac{\eta_T}{\sqrt{T}}x_t)(\sigma_t^{\mathbb{P}})^2)\right]dt\\
&\qquad+\int_0^T\int_{\mathbb{R}}\left[\Phi(\eta_T,x_{t-}+z)-\Phi(\eta_T,x_{t-})-\frac{\partial}{\partial x}\Phi(\eta_T,x_{t-})z\right]\widetilde{\mu}^{\mathbb{P}}(dt,dz)\\
&\qquad+\int_0^T\int_{\mathbb{R}}\int_0^1\left[\theta''(\tfrac{\eta_T}{\sqrt{T}}(x_t+vz))-\mathbb{E}_0^{\mathbb{P}}(\theta''(\tfrac{\eta_T}{\sqrt{T}}(x_t+vz)))\right](1-v)dv\,z^2\nu_t^{\mathbb{P}}(dz)dt.
\end{aligned} \tag{A.26}$$

We denote the three terms on the right-hand side by $I_{0,T}^{(1)}(\eta_T)$, $I_{0,T}^{(2)}(\eta_T)$ and $I_{0,T}^{(3)}(\eta_T)$, respectively. By the mean-value theorem, $|\theta''(\frac{\eta_T}{\sqrt{T}}x_t)-\theta''(0)|\leq C\frac{\eta_T}{\sqrt{T}}|x_t|$, so the first and the third term satisfy

$$\left(\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(1)}(\eta_T)|^p\right)\right)^{1/p}+\left(\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(3)}(\eta_T)|^p\right)\right)^{1/p}=O_p(T\eta_T). \tag{A.27}$$

To analyze the second term, we further split it into the sum of

$$\begin{aligned}
I_{0,T}^{(2,1)}(\eta_T)&=\int_0^T\int_{\mathbb{R}}\left[\Phi(\eta_T,x_{t-}+z)-\Phi(\eta_T,x_{t-})\right]\widetilde{\mu}^{\mathbb{P}}(dt,dz)\\
&=\frac{\sqrt{T}}{\eta_T}\int_0^T\int_{\mathbb{R}}\int_0^1\theta'(\tfrac{\eta_T}{\sqrt{T}}(x_{t-}+vz)dv\,z\widetilde{\mu}^{\mathbb{P}}(dt,dz),\\
I_{0,T}^{(2,2)}(\eta_T)&=-\int_0^T\int_{\mathbb{R}}\frac{\partial}{\partial x}\Phi(\eta_T,x_{t-})z\widetilde{\mu}^{\mathbb{P}}(dt,dz)=-\frac{\sqrt{T}}{\eta_T}\int_0^T\int_{\mathbb{R}}\theta'(\tfrac{\eta_T}{\sqrt{T}}x_{t-})z\widetilde{\mu}^{\mathbb{P}}(dt,dz).
\end{aligned} \tag{A.28}$$

By the Burkholder–Davis–Gundy inequality and (A.19), we have

$$\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(2,1)}(\eta_T)|^p\right)&\leq\frac{CT^{p/2}}{\eta_T^p}\mathbb{E}_0^{\mathbb{P}}\left(\int_0^T\int_{\mathbb{R}}\left|\int_0^1\theta'(\tfrac{\eta_T}{\sqrt{T}}(x_t+vz))dv\right|^p|z|^p\nu_t^{\mathbb{P}}(dz)dt\right)\\
&\leq\frac{CT^{p/2}}{\eta_T^p}\mathbb{E}_0^{\mathbb{P}}\left(\int_0^T\int_{\mathbb{R}}\left(\int_0^{\frac{\eta_T|z|}{\sqrt{T}}}|\theta'(\tfrac{\eta_T}{\sqrt{T}}x_t+\operatorname{sgn}(z)u)|du\frac{\sqrt{T}}{\eta_T|z|}\right)^p|z|^p\nu_t^{\mathbb{P}}(dz)dt\right)\\
&\leq\frac{CT^{p/2}}{\eta_T^p}\mathbb{E}_0^{\mathbb{P}}\left(\int_0^T\int_{\mathbb{R}}\left(\int_{\mathbb{R}}|\theta'(u)|du\frac{\sqrt{T}}{\eta_T|z|}\right)^{p-r^{\mathbb{P}}}\left(\max_{u\in\mathbb{R}}|\theta'(u)|^{r^{\mathbb{P}}}\right)|z|^p\nu_t^{\mathbb{P}}(dz)dt\right)\\
&\leq\frac{CT^{p-r^{\mathbb{P}}/2}}{\eta_T^{2p-r^{\mathbb{P}}}}\mathbb{E}_0^{\mathbb{P}}\left(\int_0^T\int_{\mathbb{R}}|z|^{r^{\mathbb{P}}}\nu_t^{\mathbb{P}}(dz)dt\right)=O_p(T^{1+p-r^{\mathbb{P}}/2}/\eta_T^{2p-r^{\mathbb{P}}})
\end{aligned}$$

for any $p\in(r^{\mathbb{P}},2]$. In particular, Jensen's inequality shows that for all $p\in[1,2]$,

$$\left(\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(2,1)}(\eta_T)|^p\right)\right)^{1/p}\leq\left(\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(2,1)}(\eta_T)|^2\right)\right)^{1/2}=O_p(T^{3/2-r^{\mathbb{P}}/4}/\eta_T^{2-r^{\mathbb{P}}/2})=o_p(T). \tag{A.29}$$

Next, using the Burkholder–Davis–Gundy inequality, (A.19) and the bounds $|\theta'(x)| \leq C|x|$ and $\mathbb{E}_0^{\mathbb{P}}(|x_t/\sqrt{T}|^p) \leq C$, we derive the estimate

$$\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(2,2)}(\eta_T)|^p\right) &\leq \frac{CT^{p/2}}{\eta_T^p}\mathbb{E}_0^{\mathbb{P}}\left(\int_0^T\int_{\mathbb{R}}|\theta'(\tfrac{\eta_T}{\sqrt{T}}x_t)|^p|z|^p\nu_t^{\mathbb{P}}(dz)dt\right)\\
&\leq \frac{CT^{p/2}}{\eta_T^p}\mathbb{E}_0^{\mathbb{P}}\left(\int_0^T\int_{\mathbb{R}}|\theta'(\tfrac{\eta_T}{\sqrt{T}}x_t)|^p|z|^p\nu_0^{\mathbb{P}}(dz)dt\right) + O_p(T^{(p+3)/2}/\eta_T^p)\\
&\leq CT^{1+p/2}\int_{\mathbb{R}}|z|^p\nu_0^{\mathbb{P}}(dz) + O_p(T^{(p+3)/2}/\eta_T^p) = O_p(T^{1+p/2})
\end{aligned}$$

for all $p \in (r^{\mathbb{P}}, 2]$. Therefore, for any $p \in (r^{\mathbb{P}}, 2)$,

$$\left(\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(2,2)}(\eta_T)|^p\right)\right)^{1/p} = O_p(T^{1/p+1/2}) = o_p(T). \tag{A.30}$$

By Jensen's inequality, the final bound of the last line extends to all $p \in [1, 2)$. For $p = 2$, we can use (A.14) to explicitly compute the limit as

$$\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}\left(|I_{0,T}^{(2,2)}(\eta_T)|^2\right) &= \frac{T}{\eta_T^2}\mathbb{E}_0^{\mathbb{P}}\left(\int_0^T\int_{\mathbb{R}}|\theta'(\tfrac{\eta_T}{\sqrt{T}}x_t)|^2z^2\nu_0^{\mathbb{P}}(dz)dt\right) + O_p(T^{5/2}/\eta_T^2)\\
&= 2\int_0^T\int_{\mathbb{R}}\mathbb{E}_0^{\mathbb{P}}((x_t^c)^2)z^2\nu_0^{\mathbb{P}}(dz)dt + o_p(T^2)\\
&= T^2(\sigma_0^{\mathbb{P}})^2\int_{\mathbb{R}}z^2\nu_0^{\mathbb{P}}(dz) + o_p(T^2).
\end{aligned} \tag{A.31}$$

Combining (A.23)–(A.25), (A.27) and (A.29)–(A.31) concludes the proof of the theorem. □

For Theorem 5.2, we make the following assumptions.

Assumption 3. *The function $\phi(x)$ is nonnegative, symmetric, equal to unity at zero, decreasing in $|x|$ and six times continuously differentiable with all its derivatives (including ϕ itself) decreasing exponentially fast as $|x| \to \infty$.*

Assumption 4. *In addition to Assumption 2, assume that under both $\mathbb{S} = \mathbb{P}$ and $\mathbb{S} = \mathbb{Q}$,*

$$|\nu_t^{\mathbb{S}}(z : |z| > u) - c_t^{\mathbb{S}}u^{-\beta^{\mathbb{S}}}\mathbf{1}_{(0,1)}(u)| \leq c_t'^{\mathbb{S}}(u^{-r^{\mathbb{S}}} \wedge u^{-r'^{\mathbb{S}}}), \quad t \in [0, \tau], \quad u > 0, \tag{A.32}$$

where $\beta^{\mathbb{S}} \in (1, 2)$, $r^{\mathbb{S}} \in (0, \beta^{\mathbb{S}})$, $r'^{\mathbb{S}} \in (4, \infty)$ and $c^{\mathbb{S}}$ and $c'^{\mathbb{S}}$ are predictable processes such that (A.17) and (A.18) also hold for $z \in \{c^{\mathbb{S}}, c'^{\mathbb{S}}\}$.

Proof of Theorem 5.2. Clearly, we have

$$\mathbb{E}_0^{\mathbb{P}}(\mathcal{R}_T^{(k_n)}) = \sum_{i=1}^{k_n}\mathbb{E}_0^{\mathbb{P}}\left(\mathbb{E}_{\tau_i}^{\mathbb{P}}(\mathcal{R}_{\tau_i,T})\right) \tag{A.33}$$

and, for $p \geq 1$,

$$\begin{aligned}
\mathbb{E}_0^{\mathbb{P}}\left(\left|\mathcal{R}_T^{(k_n)} - \mathbb{E}_0^{\mathbb{P}}(\mathcal{R}_T^{(k_n)})\right|^p\right)^{1/p} &\leq \mathbb{E}_0^{\mathbb{P}}\left(\left|\sum_{i=1}^{k_n}\left[\mathcal{R}_{\tau_i,T} - \mathbb{E}_{\tau_i}^{\mathbb{P}}(\mathcal{R}_{\tau_i,T})\right]\right|^p\right)^{1/p}\\
&\quad + \mathbb{E}_0^{\mathbb{P}}\left(\left|\sum_{i=1}^{k_n}\left[\mathbb{E}_{\tau_i}^{\mathbb{P}}(\mathcal{R}_{\tau_i,T}) - \mathbb{E}_0^{\mathbb{P}}(\mathcal{R}_{\tau_i,T})\right]\right|^p\right)^{1/p}.
\end{aligned} \tag{A.34}$$

We first derive an asymptotic expansion of $\mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{R}_{\tau_i,T})$ and note that (A.20)–(A.22) remain valid (with obvious modifications to accommodate a general τ_i). Writing $x^{(i)}_t = x_{\tau_i+t} - x_{\tau_i}$ and using Itô's formula as well as (A.18) and (A.19), we further find that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta''(\tfrac{\eta_T}{\sqrt{T}}(x^{(i)}_t + vz))\right) \\
&\quad = \theta''(\tfrac{\eta_T}{\sqrt{T}}vz) + \frac{\eta_T^2(\sigma^{\mathbb{S}}_{\tau_i})^2}{2T}\int_{\tau_i}^{\tau_i+t}\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i)}_s + vz))\right)ds \\
&\qquad + \frac{\eta_T^2}{T}\iint_{\tau_i}^{\tau_i+t}\int_0^1 \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i)}_s + vz + wy))\right)(1-w)dw\, y^2\nu^{\mathbb{S}}_{\tau_i}(dy)ds + O_p(\sqrt{T})
\end{aligned} \tag{A.35}$$

for every $v \in [0,1]$. In the last line, and for the remainder of this proof, we use $O_p(g(T))$ to denote a random variable whose $\mathcal{F}_0$-conditional L_1-norm is $O(g(T))$. Clearly, we have

$$\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i)}_s + vz + wy))\right) = \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),m}_s + vz + wy))\right) + O_p(\eta_T\sqrt{T}), \tag{A.36}$$

where $x^{(i),m}_t = x^{(i)}_t - \int_{\tau_i}^{\tau_i+t} a^{\mathbb{S}}_s ds$ is the martingale component of $x^{(i)}$. Next, writing $x^{(i),c}_t = \int_{\tau_i}^{\tau_i+t}\sigma^{\mathbb{S}}_s dW^{\mathbb{S}}_s$ for the continuous martingale part and $x^{(i),d}_t = \int_{\tau_i}^{\tau_i+t}\int_{\mathbb{R}} z\widetilde{\mu}^{\mathbb{S}}(ds,dz)$ for the discontinuous martingale part of x, we represent $\mathbb{E}^{\mathbb{S}}_{\tau_i}(\theta^{(4)}(\frac{\eta_T}{\sqrt{T}}(x^{(i),m}_s + vz + wy))$ as a sum of three terms, which we analyze separately:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),m}_s + vz + wy))\right) \\
&\quad = \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),c}_s + vz + wy))\right) + \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(5)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),c}_s + vz + wy))\frac{\eta_T x^{(i),d}_s}{\sqrt{T}}\right) \\
&\qquad + \mathbb{E}^{\mathbb{S}}_{\tau_i}\Bigg(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),m}_s + vz + wy)) - \theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),c}_s + vz + wy)) \\
&\qquad\qquad - \theta^{(5)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),c}_s + vz + wy))\frac{\eta_T x^{(i),d}_s}{\sqrt{T}}\Bigg).
\end{aligned} \tag{A.37}$$

By (A.18), the first term on the right-hand side of (A.37) satisfies

$$\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),c}_s + vz + wy))\right) = \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s + vz + wy))\right) + O_p(\eta_T\sqrt{T}),$$

where $W^{(i),\mathbb{S}}_t = W^{\mathbb{S}}_{\tau_i+t} - W^{\mathbb{S}}_{\tau_i}$. The second term on the right-hand side of (A.37) satisfies

$$\begin{aligned}
\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(5)}(\tfrac{\eta_T}{\sqrt{T}}(x^{(i),c}_s + vz + wy))\frac{\eta_T x^{(i),d}_s}{\sqrt{T}}\right) &= \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(5)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s + vz + wy))\frac{\eta_T x^{(i),d}_s}{\sqrt{T}}\right) \\
&\quad + O_p(\eta_T\sqrt{T}).
\end{aligned}$$

By the Grigelionis representation theorem for Itô semimartingales (see e.g., Theorem 2.1.2 in Jacod and Protter (2012)), one can assume that $x^{(i),d}$ is a compensated Poisson integral where the Poisson measure is independent of $W^{(i),\mathbb{S}}$. So, if we condition on $W^{(i),\mathbb{S}}$, we realize that the conditional expectation on the right-hand side of the last display is identically zero. We turn to the third term on the right-hand side of (A.37). By Assumption 3, both $\theta^{(5)}$ and $\theta^{(6)}$ are bounded functions, so it is easily established using Taylor's theorem that the absolute value of this term is bounded by

$$C\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\left|\frac{\eta_T x_s^{(i),d}}{\sqrt{T}}\right| \wedge \left|\frac{\eta_T x_s^{(i),d}}{\sqrt{T}}\right|^2\right) \leq C\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\left|\frac{\eta_T x_s^{(i),d}}{\sqrt{T}}\right|^p\right)$$

for any choice of $p \in [1,2]$. Applying the Burkholder–Davis–Gundy inequality in the first step and Jensen's inequality and the elementary inequality $(a+b)^p \leq a^p + b^p$ for $p \in [0,1]$ and $a, b \geq 0$ in the second step, we derive the estimates

$$\begin{aligned}
\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\left|\frac{\eta_T x_s^{(i),d}}{\sqrt{T}}\right|^p\right) &\leq \frac{C\eta_T^p}{T^{p/2}}\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\left(\int_{\tau_i}^{\tau_i+s}\int_{\mathbb{R}} z^2\mu(dr,dz)\right)^{p/2}\right) \\
&\leq \frac{C\eta_T^p}{T^{p/2}}\Bigg[\left(\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\int_{\tau_i}^{\tau_i+s}\int_{\mathbb{R}}|z|^2\mathbf{1}_{\{|z|\leq T^{1/\beta^{\mathbb{S}}}\}}\nu_r^{\mathbb{S}}(dz)dr\right)\right)^{p/2} \\
&\qquad + \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\int_{\tau_i}^{\tau_i+s}\int_{\mathbb{R}}|z|^p\mathbf{1}_{\{|z|>T^{1/\beta^{\mathbb{S}}}\}}\nu_r^{\mathbb{S}}(dz)dr\right)\Bigg].
\end{aligned}$$

Using (A.32), we have for $p = \beta^{\mathbb{S}}$ that

$$\begin{aligned}
\int_{\mathbb{R}}|z|^{\beta^{\mathbb{S}}}\mathbf{1}_{\{|z|>T^{1/\beta^{\mathbb{S}}}\}}\nu_r^{\mathbb{S}}(dz) &= \beta^{\mathbb{S}}\int_{\mathbb{R}}\int_0^{|z|} u^{\beta^{\mathbb{S}}-1}du\,\mathbf{1}_{\{|z|>T^{1/\beta^{\mathbb{S}}}\}}\nu_r^{\mathbb{S}}(dz) \\
&= \beta^{\mathbb{S}}\int_0^\infty u^{\beta^{\mathbb{S}}-1}\nu_r^{\mathbb{S}}(z:|z|>u\vee T^{1/\beta^{\mathbb{S}}})du \\
&= T\nu_r^{\mathbb{S}}(z:|z|>T^{1/\beta^{\mathbb{S}}}) + \beta^{\mathbb{S}}\int_{T^{1/\beta^{\mathbb{S}}}}^\infty u^{\beta^{\mathbb{S}}-1}\nu_r^{\mathbb{S}}(z:|z|>u)du \\
&= c_r^{\mathbb{S}}(1+\log T^{-1}) + O_p(1).
\end{aligned} \tag{A.38}$$

Similarly, we have

$$\begin{aligned}
\int_{\mathbb{R}}|z|^2\mathbf{1}_{\{|z|\leq T^{1/\beta^{\mathbb{S}}}\}}\nu_r^{\mathbb{S}}(dz) &= 2\int_0^{T^{1/\beta^{\mathbb{S}}}} u\nu_r^{\mathbb{S}}(z:u\leq|z|\leq T^{1/\beta})du \\
&= 2c_r^{\mathbb{S}}\int_0^{T^{1/\beta^{\mathbb{S}}}} u(u^{-\beta^{\mathbb{S}}}-T^{-1})du + O_p(T^{2/r^{\mathbb{S}}-1}) \\
&= \frac{2\beta^{\mathbb{S}}c_r^{\mathbb{S}}}{4-2\beta^{\mathbb{S}}}T^{2/\beta^{\mathbb{S}}-1} + O_p(T^{2/r^{\mathbb{S}}-1}).
\end{aligned} \tag{A.39}$$

We have thus proved the bound

$$\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\left|\frac{\eta_T x_s^{(i),d}}{\sqrt{T}}\right|^{\beta^{\mathbb{S}}}\right) = O_p(T^{1-\beta^{\mathbb{S}}/2}\eta_T^{\beta^{\mathbb{S}}}\log T^{-1}),$$

which in turn shows that

$$\begin{aligned}
\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(x_s^{(i)}+vz+wy))\right) &= \mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W_s^{(i),\mathbb{S}}+vz+wy))\right) \\
&\qquad + O_p(T^{1-\beta^{\mathbb{S}}/2}\eta_T^{\beta^{\mathbb{S}}}\log T^{-1}).
\end{aligned} \tag{A.40}$$

Inserting this result in (A.35), we obtain

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta''(\tfrac{\eta_T}{\sqrt{T}}(x^{(i)}_t+vz))\right)\\
&\quad=\theta''(\tfrac{\eta_T}{\sqrt{T}}vz)+\frac{\eta_T^2(\sigma^{\mathbb{S}}_{\tau_i})^2}{2T}\int_{\tau_i}^{\tau_i+t}\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s+vz))\right)ds\\
&\qquad+\frac{\eta_T^2}{T}\int_{\tau_i}^{\tau_i+t}\int_{\mathbb{R}}\int_0^1\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s+vz+wy))\right)(1-w)dw\,y^2\nu^{\mathbb{S}}_{\tau_i}(dy)ds\\
&\qquad+O_p(T^{1-\beta^{\mathbb{S}}/2}\eta_T^{\beta^{\mathbb{S}}+2}\log T^{-1}).
\end{aligned}
$$

By (A.22), it follows that

$$
\begin{aligned}
\frac{1}{T}TV^{\mathbb{S}}_{\tau_i,T}(\eta_T)&=\frac{(\sigma^{\mathbb{S}}_{\tau_i})^2}{2T}\int_{\tau_i}^{\tau_i+T}\left(2+\frac{\eta_T^2(\sigma^{\mathbb{S}}_{\tau_i})^2}{2T}\int_{\tau_i}^{\tau_i+t}\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s))\right)ds\right)dt\\
&+\frac{\eta_T^2(\sigma^{\mathbb{S}}_{\tau_i})^2}{2T^2}\int_{\tau_i}^{\tau_i+T}\int_{\tau_i}^{\tau_i+t}\int_{\mathbb{R}}\int_0^1\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s+wy))\right)(1-w)dw\,y^2\nu^{\mathbb{S}}_{\tau_i}(dy)dsdt\\
&+\frac{1}{T}\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\int_0^1\theta''(\tfrac{\eta_T}{\sqrt{T}}vz)(1-v)dv\,z^2\nu^{\mathbb{S}}_{\tau_i}(dz)dt\\
&+\frac{\eta_T^2(\sigma^{\mathbb{S}}_{\tau_i})^2}{2T^2}\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\int_0^1\int_{\tau_i}^{\tau_i+t}\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s+vz))\right)ds\,(1-v)dv\,z^2\nu^{\mathbb{S}}_{\tau_i}(dz)dt\\
&+\frac{\eta_T^2}{T^2}\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\int_0^1\int_{\tau_i}^{\tau_i+t}\int_{\mathbb{R}}\int_0^1\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s+vz+wy))\right)\\
&\qquad\qquad\times(1-w)dw\,y^2\nu^{\mathbb{S}}_{\tau_i}(dy)ds\,(1-v)dv\,z^2\nu^{\mathbb{S}}_{\tau_i}(dz)dt\\
&+O_p(T^{1-\beta^{\mathbb{S}}/2}\eta_T^{\beta^{\mathbb{S}}+2}\log T^{-1}).
\end{aligned}
$$

We can now determine the leading order terms of $\frac{1}{T}(TV^{\mathbb{P}}_{\tau_i,T}(\eta_T)-\frac{1}{T}TV^{\mathbb{Q}}_{\tau_i,T}(\eta_T)$. Because $\sigma^{\mathbb{P}}_{\tau_i}=\sigma^{\mathbb{Q}}_{\tau_i}$ by assumption, the first term on the right-hand side of the previous display cancels out. The second and the fourth term are identical and satisfy

$$
\begin{aligned}
&\left|\frac{\eta_T^2(\sigma^{\mathbb{S}}_{\tau_i})^2}{2T^2}\int_{\tau_i}^{\tau_i+T}\int_{\tau_i}^{\tau_i+t}\int_{\mathbb{R}}\int_0^1\mathbb{E}^{\mathbb{S}}_{\tau_i}\left(\theta^{(4)}(\tfrac{\eta_T}{\sqrt{T}}(\sigma^{\mathbb{S}}_{\tau_i}W^{(i),\mathbb{S}}_s+wy))\right)(1-w)dw\,y^2\nu^{\mathbb{S}}_{\tau_i}(dy)dsdt\right|\\
&\quad\le\frac{\eta_T^2(\sigma^{\mathbb{S}}_{\tau_i})^2}{4}\max_{x\in\mathbb{R}}|\theta^{(4)}(x)|\int_{\mathbb{R}}y^2\mathbf{1}_{\{|y|\le\sqrt{T}/\eta_T\}}\nu^{\mathbb{S}}_{\tau_i}(dy)\\
&\qquad+\frac{\eta_T(\sigma^{\mathbb{S}}_{\tau_i})^2T^{1/2}}{4}\int_{\mathbb{R}}|\theta^{(4)}(x)|dx\int_{\mathbb{R}}|y|\mathbf{1}_{\{|y|>\sqrt{T}/\eta_T\}}\nu^{\mathbb{S}}_{\tau_i}(dy)\\
&\quad=O_p(T^{1-\beta^{\mathbb{S}}/2}\eta_T^{\beta^{\mathbb{S}}}).
\end{aligned}
$$

Using similar arguments, one can derive the same bound for the fifth term on the right-hand side of the penultimate display. The leading term is the third one. Because $\int_0^1\theta''(\frac{\eta_T}{\sqrt{T}}vz)(1-v)dv=$

$\phi(\frac{\eta_T}{\sqrt{T}}z)$ and ϕ is symmetric and exponentially decreasing by assumption, it can be written as

$$\begin{aligned}\int_{\mathbb{R}} \phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2\nu^{\mathbb{S}}_{\tau_i}(dz) &= \int_{\mathbb{R}}\int_0^{|z|} \tfrac{d}{dz}[\phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2]\Big|_{z=y} dy\nu^{\mathbb{S}}_{\tau_i}(dz) \\ &= \int_0^\infty \nu^{\mathbb{S}}_{\tau_i}(z:|z|\geq y)\tfrac{d}{dz}[\phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2]\Big|_{z=y} dy \\ &= c^{\mathbb{S}}_{\tau_i}\int_0^1 y^{-\beta^{\mathbb{S}}}\tfrac{d}{dz}[\phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2]\Big|_{z=y} dy \\ &\quad + \int_1^\infty \nu^{\mathbb{S}}_{\tau_i}(z:|z|\geq y)\tfrac{d}{dz}[\phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2]\Big|_{z=y} dy \\ &\quad + \int_0^1 \left(\nu^{\mathbb{S}}_{\tau_i}(z:|z|\geq y) - c^{\mathbb{S}}_{\tau_i}y^{-\beta^{\mathbb{S}}}\right)\tfrac{d}{dz}[\phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2]\Big|_{z=y} dy \\ &= c^{\mathbb{S}}_{\tau_i}\phi(\tfrac{\eta_T}{\sqrt{T}}) + \beta^{\mathbb{S}}c^{\mathbb{S}}_{\tau_i}\int_0^1 \phi(\tfrac{\eta_T}{\sqrt{T}}z)z^{1-\beta^{\mathbb{S}}}dz \\ &\quad + O_p\left(\int_0^\infty (y^{-r^{\mathbb{S}}}\wedge y^{-r'^{\mathbb{S}}})\left|\phi'(\tfrac{\eta_T}{\sqrt{T}}y)\tfrac{\eta_T}{\sqrt{T}}y^2 + 2\phi(\tfrac{\eta_T}{\sqrt{T}}y)y\right| dy\right) \\ &= \frac{\beta^{\mathbb{S}}c^{\mathbb{S}}_{\tau_i}T^{1-\beta^{\mathbb{S}}/2}}{\eta_T^{2-\beta^{\mathbb{S}}}}\int_0^\infty \phi(w)w^{1-\beta^{\mathbb{S}}}dw + O_p(T^{1-r^{\mathbb{S}}/2}/\eta_T^{2-r^{\mathbb{S}}}).\end{aligned} \tag{A.41}$$

Therefore, using the notation $\beta = \beta^{\mathbb{P}}\vee\beta^{\mathbb{Q}}$ and $c_{\tau_i} = c^{\mathbb{P}}_{\tau_i}\mathbf{1}_{\{\beta^{\mathbb{P}}=\beta\}} - c^{\mathbb{Q}}_{\tau_i}\mathbf{1}_{\{\beta^{\mathbb{Q}}=\beta\}}$ and recalling (A.23), we obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{R}_{\tau_i,T}) &= \frac{1}{(\sigma^{\mathbb{Q}}_{\tau_i})^2}\int_{\mathbb{R}} \phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2(\nu^{\mathbb{P}}_{\tau_i} - \nu^{\mathbb{Q}}_{\tau_i})(dz) + o_p(T^{1-\beta/2}/\eta_T^{2-\beta}) \\ &= \frac{T^{1-\beta/2}}{\eta_T^{2-\beta}}\frac{\beta c_{\tau_i}}{(\sigma^{\mathbb{Q}}_{\tau_i})^2}\int_0^\infty \phi(w)w^{1-\beta}dw + o_p(T^{1-\beta/2}/\eta_T^{2-\beta}).\end{aligned} \tag{A.42}$$

By (A.33) and (A.19), it follows that

$$\frac{1}{k_n}\mathbb{E}^{\mathbb{P}}_0(\mathcal{R}^{(k_n)}_T) = \frac{T^{1-\beta/2}}{\eta_T^{2-\beta}}\frac{\beta c_0}{(\sigma^{\mathbb{Q}}_0)^2}\int_0^\infty \phi(w)w^{1-\beta}dw + o_p(T^{1-\beta/2}/\eta_T^{2-\beta}). \tag{A.43}$$

Next, we estimate the L_p-risk of the multi-period hedged volatility portfolio. By (A.18) and (A.42), the second term on the right-hand side of (A.34) satisfies

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}_0\left(\left|\sum_{i=1}^{k_n}\left[\mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{R}_{\tau_i,T}) - \mathbb{E}^{\mathbb{P}}_0(\mathcal{R}_{\tau_i,T})\right]\right|^p\right)^{1/p} &= O_p(k_n(k_nT)^{1/2}T^{1-\beta/2}/\eta_T^{2-\beta}) + o_p(k_nT^{1-\beta/2}/\eta_T^{2-\beta}) \\ &= o_p(k_nT^{1-\beta/2}/\eta_T^{2-\beta}).\end{aligned} \tag{A.44}$$

The first term on the right-hand side of (A.34) is a martingale sum, so we can apply the Burkholder–Davis–Gundy inequality to obtain

$$\mathbb{E}^{\mathbb{P}}_0\left(\left|\sum_{i=1}^{k_n}\left[\mathcal{R}_{\tau_i,T} - \mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{R}_{\tau_i,T})\right]\right|^p\right)^{1/p} \leq C\left(\sum_{i=1}^{k_n}\mathbb{E}^{\mathbb{P}}_0\left(\left|\mathcal{R}_{\tau_i,T} - \mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{R}_{\tau_i,T})\right|^p\right)\right)^{1/p}. \tag{A.45}$$

Clearly, we have

$$\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(\left|\mathcal{R}_{\tau_i,T}-\mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{R}_{\tau_i,T})\right|^p\right)\right)^{1/p}=\frac{\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(|\mathcal{RG}_{\tau_i,T}-\mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{RG}_{\tau_i,T})|^p\right)\right)^{1/p}}{TV^{\mathbb{Q}}_{\tau_i,T}(\eta_T)}. \tag{A.46}$$

By Itô's formula and (A.14),

$$\begin{aligned}
&\mathcal{RG}_{\tau_i,T}-\mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{RG}_{\tau_i,T})\\
&=\frac{1}{2}\int_{\tau_i}^{\tau_i+T}\left[\theta''(\tfrac{\eta_T}{\sqrt{T}}x_t^{(i)})(\sigma_t^{\mathbb{P}})^2-\mathbb{E}^{\mathbb{P}}_{\tau_i}(\theta''(\tfrac{\eta_T}{\sqrt{T}}x_t^{(i)})(\sigma_t^{\mathbb{P}})^2)\right]dt\\
&\quad+\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\left[\Phi(\eta_T,x_t^{(i)}+z)_T-\Phi(\eta_T,x_t^{(i)})_T\right]\widetilde{\mu}^{\mathbb{P}}(dt,dz)\\
&\quad+\int_{\tau_j}^{\tau_i+T}\int_{\mathbb{R}}\left[-\frac{\partial}{\partial x}\Phi(\eta_T,x_t^{(i)})_T z\right]\widetilde{\mu}^{\mathbb{P}}(dt,dz)\\
&\quad+\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\int_0^1\left[\theta''(\tfrac{\eta_T}{\sqrt{T}}(x_t^{(i)}+vz))-\mathbb{E}^{\mathbb{P}}_{\tau_i}(\theta''(\tfrac{\eta_T}{\sqrt{T}}(x_t^{(i)}+vz)))\right](1-v)dv\,z^2\nu_t^{\mathbb{P}}(dz)dt.
\end{aligned}$$

We denote the four terms on the right-hand side by $I^{(1)}_{\tau_i,T}(\eta_T)$, $I^{(2)}_{\tau_i,T}(\eta_T)$, $I^{(3)}_{\tau_i,T}(\eta_T)$ and $I^{(4)}_{\tau_i,T}(\eta_T)$.

Because

$$\begin{aligned}
\Phi(\eta_T,x_t^{(i)}+z)_T-\Phi(\eta_T,x_t^{(i)})_T&=\frac{\sqrt{T}}{\eta_T}z\int_0^1\theta'(\tfrac{\eta_T}{\sqrt{T}}(x_t^{(i)}+vz)dv\\
&=\frac{\sqrt{T}}{\eta_T}z\int_0^1\theta'(\tfrac{\eta_T}{\sqrt{T}}vz)dv+x_t^{(i)}z\int_0^1\int_0^1\theta''(\tfrac{\eta_T}{\sqrt{T}}(vz+ux_t^{(i)})dudv\\
&=\phi(\tfrac{\eta_T}{\sqrt{T}}z)z^2+x_t^{(i)}z\int_0^1\int_0^1\theta''(\tfrac{\eta_T}{\sqrt{T}}(vz+ux_t^{(i)})dudv,
\end{aligned}$$

the Burkholder–Davis–Gundy inequality implies that for $p\in[1,2]$,

$$\begin{aligned}
&\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(|I^{(2)}_{\tau_i,T}(\eta_T)|^p\right)\right)^{1/p}\\
&\le C\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\phi(\tfrac{\eta_T}{\sqrt{T}}z)^p|z|^{2p}\nu_t^{\mathbb{P}}(dz)dt\right)\right)^{1/p}\\
&\quad+C\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\left|\int_0^1\int_0^1\theta''(\tfrac{\eta_T}{\sqrt{T}}(vz+ux_t^{(i)}))dudv\,zx_t^{(i)}\right|^p\mathbf{1}_{\{|z|>T^{1/\beta^{\mathbb{P}}}\}}\nu_t^{\mathbb{P}}(dz)dt\right)\right)^{1/p}\\
&\quad+C\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\left|\int_0^1\int_0^1\theta''(\tfrac{\eta_T}{\sqrt{T}}(vz+ux_t^{(i)}))dudv\,zx_t^{(i)}\right|^2\mathbf{1}_{\{|z|\le T^{1/\beta^{\mathbb{P}}}\}}\nu_t^{\mathbb{P}}(dz)dt\right)\right)^{1/2}.
\end{aligned}$$

Analogously to (A.42), one can show that

$$\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\phi(\tfrac{\eta_T}{\sqrt{T}}z)^p|z|^{2p}\nu_t^{\mathbb{P}}(dz)dt\right)\right)^{1/p}=O_p(T^{1+(1-\beta^{\mathbb{P}}/2)/p}/\eta_T^{2-\beta^{\mathbb{P}}/p})$$

for all $p\in[1,2]$. Regarding the other two terms term, we first note that a refinement of (A.38) and (A.39) leads to the bounds

$$\int_{\mathbb{R}}|z|^{\beta^{\mathbb{P}}}\mathbf{1}_{\{|z|>T^{1/\beta^{\mathbb{P}}}\}}\nu_t^{\mathbb{P}}(dz)\le c_t^{\mathbb{P}}(1+\log T^{-1})+c_t'^{\mathbb{P}}\left(\frac{\beta^{\mathbb{P}}}{\beta^{\mathbb{P}}-r^{\mathbb{P}}}+\frac{\beta^{\mathbb{P}}}{r'^{\mathbb{P}}-\beta^{\mathbb{P}}}+T^{1-r^{\mathbb{P}}/\beta^{\mathbb{P}}}\right)$$

$$\leq 2c_t^{\mathbb{P}} \log T^{-1} + 2\beta^{\mathbb{P}} \left(\frac{1}{\beta^{\mathbb{P}} - r^{\mathbb{P}}} + \frac{1}{r'^{\mathbb{P}} - \beta^{\mathbb{P}}} \right) c_t'^{\mathbb{P}},$$

$$\int_{\mathbb{R}} |z|^2 \mathbf{1}_{\{|z| \leq T^{1/\beta^{\mathbb{P}}}\}} \nu_t^{\mathbb{P}}(dz) = \frac{2\beta^{\mathbb{P}} c_t^{\mathbb{P}}}{4 - 2\beta^{\mathbb{P}}} T^{2/\beta^{\mathbb{P}} - 1} + \frac{2c_t'^{\mathbb{P}}}{2 - r^{\mathbb{P}}} T^{(2 - r^{\mathbb{P}})/\beta^{\mathbb{P}}}.$$

Therefore, for $p = \beta^{\mathbb{P}}$,

$$\begin{aligned}
&\left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(\int_{\tau_i}^{\tau_i+T} \int_{\mathbb{R}} \left| \int_0^1 \int_0^1 \theta''(\tfrac{\eta_T}{\sqrt{T}}(vz + ux_t^{(i)})) du dv \, z x_t^{(i)} \right|^{\beta^{\mathbb{P}}} \mathbf{1}_{\{|z| > T^{1/\beta^{\mathbb{P}}}\}} \nu_t^{\mathbb{P}}(dz) dt \right) \right)^{1/\beta^{\mathbb{P}}} \\
&\leq C \left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(\int_{\tau_i}^{\tau_i+T} |x_t^{(i)}|^{\beta^{\mathbb{P}}} \int_{\mathbb{R}} |z|^{\beta^{\mathbb{P}}} \mathbf{1}_{\{|z| > T^{1/\beta^{\mathbb{P}}}\}} \nu_t^{\mathbb{P}}(dz) dt \right) \right)^{1/\beta^{\mathbb{P}}} \\
&\leq C \left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(\int_{\tau_i}^{\tau_i+T} |x_t^{(i)}|^{\beta^{\mathbb{P}}} (c_t^{\mathbb{P}} \log T^{-1} + c_t'^{\mathbb{P}}) dt \right) \right)^{1/\beta^{\mathbb{P}}} \\
&\leq C \Bigg[\left((c_{\tau_i}^{\mathbb{P}} \log T^{-1} + c_{\tau_i}'^{\mathbb{P}}) \int_{\tau_i}^{\tau_i+T} \mathbb{E}_{\tau_i}^{\mathbb{P}} (|x_t^{(i)}|^{\beta^{\mathbb{P}}}) dt \right)^{1/\beta^{\mathbb{P}}} \\
&\quad + \left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(\int_{\tau_i}^{\tau_i+T} |x_t^{(i)}|^{\beta^{\mathbb{P}}} \left[(c_t^{\mathbb{P}} - c_{\tau_i}^{\mathbb{P}}) \log T^{-1} + (c_t'^{\mathbb{P}} - c_{\tau_i}'^{\mathbb{P}}) \right] dt \right) \right)^{1/\beta^{\mathbb{P}}} \Bigg].
\end{aligned}$$

As $\mathbb{E}_{\tau_i}^{\mathbb{P}}(|x_t^{(i)}|^{\beta^{\mathbb{P}}}) \leq C_{\tau_i} t^{\beta^{\mathbb{P}}/2}$ and $\mathbb{E}_{\tau_i}^{\mathbb{P}}(|x_t^{(i)}|^{2\beta^{\mathbb{P}}}) \leq C_{\tau_i} t$, applying the Cauchy–Schwarz inequality to the second term and using (A.18) for $z = c^{\mathbb{P}}$ and $z = c'^{\mathbb{P}}$, we obtain

$$\begin{aligned}
&\left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(\int_{\tau_i}^{\tau_i+T} \int_{\mathbb{R}} \left| \int_0^1 \int_0^1 \theta''(\tfrac{\eta_T}{\sqrt{T}}(vz + ux_t^{(i)})) du dv \, z x_t^{(i)} \right|^{\beta^{\mathbb{P}}} \mathbf{1}_{\{|z| > T^{1/\beta^{\mathbb{P}}}\}} \nu_t^{\mathbb{P}}(dz) dt \right) \right)^{1/\beta^{\mathbb{P}}} \\
&= O_p(T^{1/2 + 1/\beta^{\mathbb{P}}} \log T^{-1}).
\end{aligned}$$

A similar argument shows that

$$\begin{aligned}
&\left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(\int_{\tau_i}^{\tau_i+T} \int_{\mathbb{R}} \left| \int_0^1 \int_0^1 \theta''(\tfrac{\eta_T}{\sqrt{T}}(vz + ux_t^{(i)})) du dv \, z x_t^{(i)} \right|^2 \mathbf{1}_{\{|z| \leq T^{1/\beta^{\mathbb{P}}}\}} \nu_t^{\mathbb{P}}(dz) dt \right) \right)^{1/2} \\
&= O_p(T^{1/2 + 1/\beta^{\mathbb{P}}}).
\end{aligned}$$

Altogether, we have shown that

$$\left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(|I_{\tau_i,T}^{(2)}(\eta_T)|^{\beta^{\mathbb{P}}} \right) \right)^{1/\beta^{\mathbb{P}}} = O_p(T^{1/2 + 1/\beta^{\mathbb{P}}} / \eta_T + T^{1/2 + 1/\beta^{\mathbb{P}}} \log T^{-1}). \tag{A.47}$$

Because $\frac{\partial}{\partial x} \Phi(\eta_T, x_t^{(i)})_T = \frac{\sqrt{T}}{\eta_T} \theta'(\frac{\eta_T}{\sqrt{T}} x_t^{(i)})$, a similar calculation shows that

$$\left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(|I_{\tau_i,T}^{(3)}(\eta_T)|^{\beta^{\mathbb{P}}} \right) \right)^{1/\beta^{\mathbb{P}}} = O_p(T^{1/2 + 1/\beta^{\mathbb{P}}} \log T^{-1}). \tag{A.48}$$

Next, we turn to $I_{\tau_i,T}^{(1)}(\eta_T)$ and $I_{\tau_i,T}^{(4)}(\eta_T)$. Because θ'' is bounded, (A.18) for $(\sigma^{\mathbb{P}})^2$ implies that for $p \in [1, 2]$,

$$\left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(|I_{\tau_i,T}^{(1)}(\eta_T)|^p \right) \right)^{1/p} = \frac{(\sigma_{\tau_i}^{\mathbb{P}})^2}{2} \left(\mathbb{E}_{\tau_i}^{\mathbb{P}} \left(\left| \frac{1}{2} \int_{\tau_i}^{\tau_i+T} \left[\theta''(\tfrac{\eta_T}{\sqrt{T}} x_t^{(i)}) - \mathbb{E}_{\tau_i}^{\mathbb{P}}(\theta''(\tfrac{\eta_T}{\sqrt{T}} x_t^{(i)})) \right] dt \right|^p \right) \right)^{1/p}$$

$$+ O_p(T^{3/2}).$$

Since $\theta'''(0) = 0$ and $\theta^{(4)}$ is bounded, Taylor's theorem implies that $|\theta''(\frac{\eta_T}{\sqrt{T}} x_t^{(i)}) - 2| \leq C \frac{\eta_T^2}{T}(x_t^{(i)})^2$. As $\mathbb{E}^{\mathbb{P}}_{\tau_i}(|x_t^{(i)}|^{2p}) \leq C_{\tau_i} t$, it follows that

$$\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(|I^{(1)}_{\tau_i,T}(\eta_T)|^p\right)\right)^{1/p} = O_p(T^{1/p}\eta_T^2). \tag{A.49}$$

With the same arguments, one can show that

$$\begin{aligned}
&\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(|I^{(4)}_{\tau_i,T}(\eta_T)|^p\right)\right)^{1/p} \\
&\leq C\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(\left|\int_{\tau_i}^{\tau_i+T}\int_{\mathbb{R}}\int_0^1 \theta'''(\tfrac{\eta_T}{\sqrt{T}} vz)\frac{\eta_T}{\sqrt{T}}\left[x_t^{(i)} - \mathbb{E}^{\mathbb{P}}_{\tau_i}(x_t^{(i)})\right](1-v)dv\, z^2 \nu^{\mathbb{P}}_{\tau_i}(dz)dt\right|^p\right)\right)^{1/p} \\
&\quad + O_p(T^{1/p}\eta_T^2).
\end{aligned}$$

Since $\int_0^1 \theta'''(\frac{\eta_T}{\sqrt{T}} vz)(1-v)dv = \phi'(\frac{\eta_T}{\sqrt{T}} z) + 2(\phi(\frac{\eta_T}{\sqrt{T}} z) - 1)\frac{\sqrt{T}}{\eta_T z}$, a straightforward computation along the lines of (A.41) yields

$$\int_{\mathbb{R}}\int_0^1 \theta'''(\tfrac{\eta_T}{\sqrt{T}} vz)(1-v)dv\, z^2 \nu^{\mathbb{P}}_{\tau_i}(dz) = O_p(T^{1-\beta^{\mathbb{P}}/2}/\eta_T^{2-\beta^{\mathbb{P}}}).$$

As a result,

$$\left(\mathbb{E}^{\mathbb{P}}_{\tau_i}\left(|I^{(4)}_{\tau_i,T}(\eta_T)|^p\right)\right)^{1/p} = O_p(T^{2-\beta^{\mathbb{P}}/2}\eta_T^{\beta^{\mathbb{P}}-1} + T^{1/p}\eta_T^2). \tag{A.50}$$

Combining (A.47)–(A.50) and recalling (A.23), (A.45) and (A.46), we conclude that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}_0\left(\left|\sum_{i=1}^{k_n}\left[\mathcal{R}_{\tau_i,T} - \mathbb{E}^{\mathbb{P}}_{\tau_i}(\mathcal{R}_{\tau_i,T})\right]\right|^{\beta^{\mathbb{P}}}\right)^{1/\beta^{\mathbb{P}}} \\
&\quad = O_p(k_n^{1/\beta^{\mathbb{P}}}[T^{1/\beta^{\mathbb{P}}-1/2}(1/\eta_T + \log T^{-1}) + T^{-(1-1/\beta^{\mathbb{P}})}\eta_T^2 + T^{1-\beta^{\mathbb{P}}/2}\eta_T^{\beta^{\mathbb{P}}-1}]).
\end{aligned} \tag{A.51}$$

In conjunction with (A.44), we have shown that for $p \in [1, \beta^{\mathbb{P}}]$,

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}_0\left(\left|\mathcal{R}^{(k_n)}_T - \mathbb{E}^{\mathbb{P}}_0(\mathcal{R}^{(k_n)}_T)\right|^p\right)^{1/p} &\leq \mathbb{E}^{\mathbb{P}}_0\left(\left|\mathcal{R}^{(k_n)}_T - \mathbb{E}^{\mathbb{P}}_0(\mathcal{R}^{(k_n)}_T)\right|^{\beta^{\mathbb{P}}}\right)^{1/\beta^{\mathbb{P}}} \\
&= O_p(k_n^{1/\beta^{\mathbb{P}}}T^{1/\beta^{\mathbb{P}}-1/2}(1/\eta_T + \log T^{-1}) + k_n^{1/\beta^{\mathbb{P}}}T^{-(1-1/\beta^{\mathbb{P}})}\eta_T^2 \\
&\quad + k_n^{1/\beta^{\mathbb{P}}}T^{1-\beta^{\mathbb{P}}/2}\eta_T^{\beta^{\mathbb{P}}-1}) + o_p(k_n T^{1-\beta/2}/\eta_T^{2-\beta}).
\end{aligned}$$

Up to now, all the derivations are valid for general k_n and η_T (as long as $\sqrt{T}/\eta_T \to 0$). By specializing the last display to the case where $\eta_T \sim \eta T^{1/6}$ for some $\eta > 0$, we obtain

$$\mathbb{E}^{\mathbb{P}}_0\left(\left|\mathcal{R}^{(k_n)}_T - \mathbb{E}^{\mathbb{P}}_0(\mathcal{R}^{(k_n)}_T)\right|^p\right)^{1/p} = O_p(k_n^{1/\beta^{\mathbb{P}}}T^{1/\beta^{\mathbb{P}}-1/2}/\eta_T) + o_p(k_n T^{1-\beta/2}/\eta_T^{2-\beta}).$$

Further taking $k_n \sim \theta T^{-\kappa}$ for some $\theta > 0$ and $\kappa \in [\frac{2}{3}, 1)$, we have

$$\mathbb{E}^{\mathbb{P}}_0\left(\left|\mathcal{R}^{(k_n)}_T - \mathbb{E}^{\mathbb{P}}_0(\mathcal{R}^{(k_n)}_T)\right|^p\right)^{1/p} = o_p(k_n T^{1-\beta/2}/\eta_T^{2-\beta}),$$

which by (A.43) shows that the L_p-risk of the multi-period hedged volatility portfolio is much smaller asymptotically than the absolute value of the expected gain. □

A.4 Proof of Lemma 6.1

Proof of Lemma 6.1. We suppress the dependence of f_T on the second argument (which is always X_t) in this proof. By assumption, we have $\epsilon_{t+j\Delta_n} = \frac{1}{2}\text{rba}X_{t+j\Delta_n}$, so

$$\widehat{DTC}_{t,T} = \frac{1}{2}\text{rba}\left(|f'_T(X_{t+k_n\Delta_n})|X_{t+k_n\Delta_n} + \sum_{j=1}^{k_n}|f'_T(X_{t+j\Delta_n}) - f'_T(X_{t+(j-1)\Delta_n})|X_{t+j\Delta_n}\right)$$
$$= \frac{1}{2}\text{rba}X_t\left(|f'_T(X_{t+k_n\Delta_n})|X_{t+k_n\Delta_n} + \sum_{j=1}^{k_n}|f''_T(X_{t+(j-1)\Delta_n})||X_{t+j\Delta_n} - X_{t+(j-1)\Delta_n}|\right)$$
$$+ o_p(k_n\sqrt{\Delta_n}).$$

Because $f'_T(X_{t+k_n\Delta_n}) = o_p(1)$ and $f''_T(X_{t+(j-1)\Delta_n}) = 2X_t^{-2} + o_p(1)$, we can use Theorem 7.2.2 in Jacod and Protter (2012) to show that

$$\widehat{DTC}_{t,T} = \frac{\text{rba}}{X_t}\sum_{j=1}^{k_n}\mathbb{E}^{\mathbb{P}}_{t+(j-1)\Delta_n}\left[|X_{t+j\Delta_n} - X_{t+(j-1)\Delta_n}|\right] + o_p(k_n\sqrt{\Delta_n})$$
$$= \sqrt{\frac{2}{\pi}}\sigma_t^{\mathbb{P}}k_n\sqrt{\Delta_n}\text{rba} + o_p(k_n\sqrt{\Delta_n}).$$

As $k_n \sim T/\Delta_n$, the lemma is proved. □

References

Aït-Sahalia, Y. and A. W. Lo (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics 94*(1-2), 9–51.

Almeida, C. and G. Freire (2022). Pricing of index options in incomplete markets. *Journal of Financial Economics 144*(1), 174–205.

Almeida, C., G. Freire, and R. Hizmeri (2024). 0DTE asset pricing. *Available at SSRN*.

Alvarez, F. and U. J. Jermann (2005). Using asset prices to measure the persistence of the marginal utility of wealth. *Econometrica 73*(6), 1977–2016.

Andersen, T. G., N. Fusari, and V. Todorov (2017). Short-term market risks implied by weekly options. *The Journal of Finance 72*(3), 1335–1386.

Back, K. (2010). *Asset Pricing and Portfolio Choice Theory.* Oxford University Press.

Bakshi, G. and N. Kapadia (2003). Delta-hedged gains and the negative market volatility risk premium. *The Review of Financial Studies 16*(2), 527–566.

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

Bakshi, G. and D. Madan (2000). Spanning and derivative-security valuation. *Journal of Financial Economics 55*(2), 205–238.

Bandi, F. M., N. Fusari, and R. Renò (2023a). 0DTE option pricing. *Available at SSRN*.

Bandi, F. M., N. Fusari, and R. Renò (2023b). Structural stochastic volatility. *Available at SSRN*.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

Barberis, N. and M. Huang (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review 98*(5), 2066–2100.

Bates, D. S. (2022). Empirical option pricing models. *Annual Review of Financial Economics 14*, 369–389.

Bekaert, G. and M. Hoerova (2014). The VIX, the variance premium and stock market volatility. *Journal of Econometrics 183*(2), 181–192.

Bernardo, A. E. and O. Ledoit (2000). Gain, loss, and asset pricing. *Journal of Political Economy 108*(1), 144–172.

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

Bollerslev, T. and V. Todorov (2011). Tails, fears, and risk premia. *The Journal of Finance 66*(6), 2165–2211.

Bondarenko, O. (2014). Why are put options so expensive? *The Quarterly Journal of Finance 4*(03), 1450015.

Boyer, B. H. and K. Vorkink (2014). Stock options as lotteries. *The Journal of Finance 69*(4), 1485–1527.

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

Bryzgalova, S., A. Pavlova, and T. Sikorskaya (2023). Retail trading in options and the rise of the big three wholesalers. *The Journal of Finance 78*(6), 3465–3514.

Carr, P. and D. Madan (2001). Optimal positioning in derivative securities. *Quantitative Finance 1*, 19–37.

Carr, P. and L. Wu (2009). Variance risk premiums. *The Review of Financial Studies 22*(3), 1311–1341.

Christensen, T. M. (2017). Nonparametric stochastic discount factor decomposition. *Econometrica 85*(5), 1501–1536.

Cochrane, J. H. and J. Saa-Requejo (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy 108*(1), 79–119.

Coval, J. D. and T. Shumway (2001). Expected option returns. *The Journal of Finance 56*(3), 983–1009.

Cuchiero, C., I. Klein, and J. Teichmann (2016). A new perspective on the fundamental theorem of asset pricing for large financial markets. *Theory of Probability & Its Applications 60*(4), 561–579.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen 300*(1), 463–520.

Drechsler, I. and A. Yaron (2011). What's vol got to do with it. *The Review of Financial Studies 24*(1), 1–45.

Du, D. (2011). General equilibrium pricing of options with habit formation and event risks. *Journal of Financial Economics 99*(2), 400–426.

Du, W., A. Tepper, and A. Verdelhan (2018). Deviations from covered interest rate parity. *The Journal of Finance 73*(3), 915–957.

Duffie, D. (2010). *Dynamic Asset Pricing Theory*. Princeton University Press.

Eraker, B. and A. Yang (2022). The price of higher order catastrophe insurance: The case of vix options. *The Journal of Finance 77*(6), 3289–3337.

Fournier, M., K. Jacobs, and P. Orłowski (2024). Modeling conditional factor risk premia implied by index option returns. *The Journal of Finance 79*(3), 2289–2338.

Hansen, L. P. and R. Jagannathan (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy 99*(2), 225–262.

Hansen, L. P. and J. A. Scheinkman (2009). Long-term risk: An operator approach. *Econometrica 77*(1), 177–234.

Harrison, J. M. and D. M. Kreps (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory 20*(3), 381–408.

Harrison, J. M. and S. R. Pliska (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications 11*(3), 215–260.

Huberman, G. and W. Stanzl (2004). Price manipulation and quasi-arbitrage. *Econometrica 72*(4), 1247–1275.

Jackwerth, J. C. (2000). Recovering risk aversion from option prices and realized returns. *The Review of Financial Studies 13*(2), 433–451.

Jacod, J. and J. Mémin (1976). Caractéristiques locales et conditions de continuité absolue pour les semi-martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 35*(1), 1–37.

Jacod, J. and P. Protter (2012). *Discretization of Processes*, Volume 67 of *Stochastic Modelling and Applied Probability*. Springer, Heidelberg.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (Second ed.), Volume 288 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin.

Jacod, J. and V. Todorov (2014). Efficient estimation of integrated volatility in presence of infinite variation jumps. *The Annals of Statistics 42*(3), 1029–1069.

Konno, H. and H. Yamazaki (1991). Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market. *Management Science 37*(5), 519–531.

Lamont, O. and R. Thaler (2003). Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy 111*(2), 227–268.

Orłowski, P., P. Schneider, and F. Trojani (2024). On the nature of (jump) skewness risk premia. *Management Science 70*(2), 1154–1174.

Pontiff, J. (1996). Costly arbitrage: Evidence from closed-end funds. *The Quarterly Journal of Economics 111*(4), 1135–1151.

Pontiff, J. (1997). Excess volatility and closed-end funds. *American Economic Review 87*(1), 155–169.

Rosenberg, J. V. and R. F. Engle (2002). Empirical pricing kernels. *Journal of Financial Economics 64*(3), 341–372.

Seo, S. B. and J. A. Wachter (2019). Option prices in a model with stochastic disaster risk. *Management Science 65*(8), 3449–3469.

Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. *The Journal of Finance 52*(1), 35–55.

Snow, K. N. (1991). Diagnosing asset pricing models using the distribution of asset returns. *The Journal of Finance 46*(3), 955–983.

Todorov, V. (2019). Nonparametric spot volatility from options. *The Annals of Applied Probability 29*(6), 3590–3636.

Vilkov, G. (2023). 0DTE trading rules. *Available at SSRN*.

Xiong, W. and J. Yu (2011). The Chinese warrants bubble. *American Economic Review 101*(6), 2723–2753.