Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

Routledge
Taylor & Francis Group

The Fundamental Theorem of Derivative Trading - exposition, extensions and experiments

Simon Ellersgaard, Martin Jönsson & Rolf Poulsen

To cite this article: Simon Ellersgaard, Martin Jönsson & Rolf Poulsen (2017) The Fundamental Theorem of Derivative Trading - exposition, extensions and experiments, Quantitative Finance, 17:4, 515-529, DOI: 10.1080/14697688.2016.1222078

To link to this article: https://doi.org/10.1080/14697688.2016.1222078

Published online: 14 Sep 2016.

Quantitative Finance, 2017
Vol. 17, No. 4, 515–529, http://dx.doi.org/10.1080/14697688.2016.1222078

Routledge
Taylor & Francis Group

The Fundamental Theorem of Derivative Trading - exposition, extensions and experiments

SIMON ELLERSGAARD, MARTIN JÖNSSON and ROLF POULSEN*

Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark

(*Received 14 December 2015; accepted 28 July 2016; published online 14 September 2016*)

When estimated volatilities are not in perfect agreement with reality, delta-hedged option portfolios will incur a non-zero profit-and-loss over time. However, there is a surprisingly simple formula for the resulting hedge error, which has been known since the late 1990s. We call this The Fundamental Theorem of Derivative Trading. This paper is a survey with twists on that result. We prove a more general version of it and discuss various extensions and applications, from incorporating a multi-dimensional jump framework to deriving the Dupire–Gyöngy–Derman–Kani formula. We also consider its practical consequences, both in simulation experiments and on empirical data, thus demonstrating the benefits of hedging with implied volatility.

Keywords: Delta hedging; Model uncertainty; Volatility arbitrage

JEL Classification: G13, G17

1. A meditation on the art of derivative hedging

1.1. Introduction

Of all possible concepts within the field mathematical finance, that of *continuous time derivative hedging* indubitably emerges as the central pillar. First used in the seminal work by Black and Scholes (1973),† it has become the cornerstone in the determination of no-arbitrage prices for new financial products. Yet a disconnect between this body of abstract mathematical theory and real-world practice prevails. Specifically, successful hedging relies crucially on us having near-perfect information about the model that drives the underlying asset. Even if we boldly adopt the standard stochastic differential equation paradigm of asset pricing, it remains to make exact specifications for the degree to which the price process reacts to market fluctuations (i.e. to specify the diffusion term, the volatility). Alas, volatility blatantly transcends direct human observation, being, as it were, a Kantian *Ding an sich*‡ of which we only have approximate knowledge.

One such source supervenes upon historical realizations of the underlying price process: for example, assuming that the governing model can at least *locally* be approximated as a geometric Brownian motion, one can proceed to measure the standard deviation of past log returns over time. Yet this procedure raises uncomfortable questions pertaining to statistical measurement: under ordinary circumstances, increasing the sample space should narrow the confidence interval around our sample parameter. Only here, there is no a priori way of telling if and when a model undergoes a structural change.§ Inevitably, this implies that extending the time series of log returns too far into the past might lead to a less accurate estimator as we might end up sampling from a governing dynamics that is no longer valid. Of course, we may take *some* measures against this issue by trying our luck with ever more intricate time series analyses, until we stumble upon a model, the parameters of which satisfy our arbitrary tolerance for statistical significance. Nevertheless, in practice this procedure invariably boils down to checking some finite basket of models and selecting the best one from the lot. Furthermore, unknown structural breaks continue to pose a problem no matter what.

Alternatively, we might try to extract an implied volatility from the market by fitting our model to observed option prices.

*Corresponding author. Email: rolf@math.ku.dk

†That Black and Scholes along with Merton were the first is the general consensus, although the paper by Haug and Taleb (2011) shows that the view is not universal.

‡Literally, *thing in itself* or the *noumenon*. Kant held that there is a distinction between the way things appear to observers (phenomena) and the way reality actually is construed (noumena).

§This scenario is not at all implausible. Unlike the physical sciences where the fundamental laws are assumed to have no *sufficient reason* to change (in the Leibnizian or Occamian sense), this philosophical principle would hardly withstand scrutiny in a social science context. Asset price processes are fundamentally governed by market agents and their reactions to various events (be they self-induced or exogenous). There is really no reason to assume that these market players will not drastically change their opinions at some point (for one reason or the other).

Nevertheless, the inadequacy of this methodology quickly becomes apparent: first, implied volatility might be ill-defined as it is the case for certain exotic products such as barrier options. Secondly, it is quite clear that the market hysteria which drives the prices of traded options need not capture the market hysteria which drives the corresponding market for the underlying asset. Fair pricing ultimately boils down to understanding the true nature of the underlying product: not to mimic the collective madness of option traders.

Whilst volatility at its core remains elusive to us, the situation is perhaps not as dire as one might think. Specifically, we can develop a formal understanding of the profit-&-loss we incur upon hedging a portfolio with an erroneous volatility—at least insofar as we make some moderate assumptions of the dynamical form of the underlying assets. To give a concrete example of this, consider the simple interest rate free framework presented in Andreasen (2003) where the price process of a single non-dividend paying asset is assumed to follow the real dynamics

$$\mathrm{d}X_t = X_t[\mu_{t,r}\mathrm{d}t + \sigma_{t,r}\mathrm{d}W_t].$$

Let V_t^i be the value of an option that trades in the market at a certain implied volatility σ_i—possibly quite different from the epistemically inaccessible $\sigma_{t,r}$. Now if we were to set up a hedge of a long position on such an option, using σ_i as our hedge volatility, an application of Itō's formula, coupled with the Black–Scholes equation, shows that the infinitesimal value change in the hedge portfolio $\Pi_t = V_t^i - \partial_x V_t^i \cdot X_t$, is of the form

$$\mathrm{d}\Pi_t = \tfrac{1}{2}(\sigma_{t,r}^2 - \sigma_i^2)X_t^2\partial_{xx}^2 V_t \mathrm{d}t, \qquad (1)$$

which generally is non-zero unless $\sigma_i = \sigma_{t,r}$. For reasons that will become clearer below, the importance of this result is of such magnitude that Andreasen dubs it the Fundamental Theorem of Derivative Trading. Indeed, a more abstract variation of it will be the central object of study in this paper.

To the best of our knowledge, quantitative studies into the effect of hedging with an erroneous volatility first appeared in a paper on the robustness of the Black–Scholes formula by El Karoui *et al.* (1998). They viewed the result as a largely negative one: unless volatility is bounded (which it is not in any stochastic volatility model) then there is no simple super-replication strategy. Easier derivations of the Fundamental Theorem are encountered in the papers by Gibson *et al.* (1999), Mahayni *et al.* (2001), and Rasmussen (2001). Indeed, there seems to be an added awareness of the result being a positive one: a decent forecast of volatility gives rise to a small hedge error. Jointly, what characterizes these papers is their binary partition of volatility into a 'wrong' and 'right' category. Somewhat subtler treatments can be found in the unpublished works by Carr (2002) and Henrard (2001) in which the partition becomes tripartite: concretely, these papers explicitly differentiate between a true governing volatility of the underlying asset, an implied volatility characterizing market consensus and a hedge volatility which captures the personal belief of the option hedger. The benefits of this three-part structure are clearly enunciated in Ahmad and Wilmott (2005) and Wilmott (2007) in which the associated P&L paths under the various choices of volatility are exhibited, alongside mathematical (non-analytic) expressions for the mean implied P&L and its variance.

The Fundamental Theorem has in other words received extensive treatment in the academic literature, yet it has never quite reached the 'textbook' status we reckon it deserves (in particular, it rarely finds its way into the curriculum of aspiring financial engineers). To end this regrettable situation, we here provide a thorough exposition. Taking our vantage point in the tripartite philosophy of Henrard and Carr, we endeavour to provide a simple proof of the Fundamental Theorem, whilst simultaneously spicing up the result by considering a more general dynamics in a full-fledged multi-asset framework. Specifically, the structure of this paper is as follows: in section 2, we state and prove a generalized version of the Fundamental Theorem of Derivative Trading for basket options and discuss its various implications for hedging strategies and applications. Key insights here include the benefit of hedging with the implied volatility with regards to attaining 'smooth' P&L paths, and a somewhat surprising connection to Dupire's local volatility formula. In section 3, we expose the implications of adding a multi-dimensional jump process to the dynamics, thus emphasizing the relative ease with which the original proof can be adapted. In particular, we argue that long option positions with convex pay-out profiles profit from a discontinuous movement in one of the underlying stocks. Finally, section 4 presents an empirical investigation into what actually happens to our portfolio when we hedge using various volatilities. In particular, using actual quotes for stocks and call options, we demonstrate that self-financing portfolios hedged at the implied volatility indeed give rise to a smoother P&L over time, whilst allowing for some amount of volatility arbitrage to be picked up in the process.

2. The Fundamental Theorem of Derivative Trading

2.1. *Model set-up*

Consider a financial market comprised of a risk-free money account as well as n risky assets, each of which pays out a continuous dividend yield. We assume all assets to be infinitely divisible as to the amount which may be held, that trading takes place continuously in time, and that no trade is subject to financial friction. Formally, we imagine the information flow of this world to be captured by the stochastic basis $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$, where Ω represents all possible states of the economy, $\mathbb{P}$ is the physical probability measure and $\mathbb{F} = \{\mathscr{F}_t\}_{t\geq 0}$ is a filtration which satisfies the *usual conditions*.† The price processes of the risky assets, $\boldsymbol{X}_t = (X_{1t}, X_{2t}, \dots, X_{nt})^\intercal$, are assumed to follow the real dynamics‡

$$\mathrm{d}\boldsymbol{X}_t = \mathbf{D}_{\boldsymbol{X}_t}[\boldsymbol{\mu}_r(t, \widetilde{\boldsymbol{X}}_t)\mathrm{d}t + \boldsymbol{\sigma}_r(t, \widetilde{\boldsymbol{X}}_t)\mathrm{d}\boldsymbol{W}_t], \qquad (2)$$

where $\mathbf{D}_{\boldsymbol{X}}$ is the $n\times n$ diagonal matrix $\mathrm{diag}(X_{1t}, X_{2t}, \dots, X_{nt})$, and $\boldsymbol{W}_t = (W_{1t}, W_{2t}, \dots, W_{nt})^\intercal$ is an n-dimensional standard Brownian motion adapted to $\mathbb{F}$. Furthermore, $\boldsymbol{\mu}_r : [0, \infty) \times \mathbb{R}^{n+m} \mapsto \mathbb{R}^n$ and $\boldsymbol{\sigma} : [0, \infty) \times \mathbb{R}^{n+m} \mapsto \mathbb{R}^{n\times n}$ are deterministic functions, sufficiently well-behaved for the SDE to have a

†Specifically, it satisfies right-continuity, $\cap_{s\geq t}\mathscr{F}_s = \mathscr{F}_t \ \forall t \geq 0$ (if we move incrementally forward in time there will be no jump in information), and completeness, i.e. $\mathscr{F}_0$ contains all $\mathbb{P}$ null sets.

‡The nomenclature 'real dynamics' is ripe with unfortunate connotations of Platonic realism (ontological significance of mathematical objects). Strictly speaking, this is not what we require, but rather the *expressive adequacy* of a model: i.e. it's ability to adequately capture the financial events unfolding.

unique strong solution (in particular, we assume the regularity conditions

$$\int_t^s |\mathbf{D}_{X_u}\boldsymbol{\mu}_r(u,\widetilde{X}_u)|\mathrm{d}u < \infty, \quad \int_t^s |\mathbf{D}_{X_u}\boldsymbol{\sigma}_r(u,\widetilde{X}_u)|^2\mathrm{d}u < \infty, \tag{3}$$

hold a.s. $\forall t \le s$, where the first norm is to be understood in the Euclidian sense, whilst the latter should be construed in the matrical sense).† Finally, we define $\widetilde{X}_t$ as the $n+m$ dimensional vector $(X_t; \chi_t)$ where $\chi_t = (\chi_{1t}, \chi_{2t}, \dots, \chi_{mt})^\intercal$ has the interpretation of an m-dimensional state variable, the exact dynamical nature of which is not integral to what follows.‡

In what , we consider the scenario of what happens when we hedge an option on X_t, ignorant of the existence of the state variable χ_t, as well as the form of $\boldsymbol{\mu}_r(\cdot,\cdot)$ and $\boldsymbol{\sigma}_r(\cdot,\cdot)$. Specifically, we shall imagine that we are misguided to the extent that we would model the dynamics of X_t as a local volatility model with diffusion matrix $\boldsymbol{\sigma}_h(t, X_t)$. Similar assumptions pertain to the market, although here we label the 'implied' diffusion matrix $\boldsymbol{\sigma}_i(t, X_t)$ to distinguish it from our personal belief. Irrespective of which dynamical specification is being made, we maintain that regularity conditions analogous to (3) remain satisfied. Finally, a cautionary remark: throughout these pages we use r and i to emphasize that the volatility is *real* and *implied*, respectively, whilst h refers to an arbitrary *hedge* volatility. For a comprehensible reading, it is incumbent that the reader keeps these definitions in mind.

2.2. Theorem and derivation

THEOREM 1 The Fundamental Theorem of Derivative Trading *Let $V_t = V(t, X_t) \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^n)$ be the price process of a European option with terminal pay-off $V_T = g(X_T)$, the underlying of which follows the real dynamics (2). We assume the option trades in the market at the (not necessarily uniquely determined) implied volatility $\boldsymbol{\sigma}_i = \boldsymbol{\sigma}_i(t, X_t)$. Now suppose we at time $t = 0$ acquire such an option for the implied price V_0^i and set out to Δ-hedge our position. Said hedge is performed under the notion that the volatility function ought, in fact, to be of the form $\boldsymbol{\sigma}_h = \boldsymbol{\sigma}_h(t, X_t)$, leading to the fair price process V_t^h. Then the present value of the* ***profit-&-loss*** *we incur from holding such a portfolio over the interval $\mathbb{T} = [0, T]$ is*

$$\mathrm{P\&L}_{\mathbb{T}}^h = V_0^h - V_0^i + \frac{1}{2}\int_0^T \times e^{-\int_0^t r_u \mathrm{d}u} tr[\mathbf{D}_{X_t}\boldsymbol{\Sigma}_{rh}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla_{xx}^2 V_t^h]\mathrm{d}t, \tag{4}$$

where $r_u = r(u, X_u)$ is the locally risk free rate, ∇_{xx}^2 is the Hessian operator, and

$$\boldsymbol{\Sigma}_{rh}(t,\widetilde{X}_t) \equiv \boldsymbol{\sigma}_r(t,\widetilde{X}_t)\boldsymbol{\sigma}_r^\intercal(t,\widetilde{X}_t) - \boldsymbol{\sigma}_h(t,X_t)\boldsymbol{\sigma}_h^\intercal(t,X_t), \tag{5}$$

is a matrix which takes values in $\mathbb{R}^{n\times n}$.

Proof Let $\{\Pi_t^h\}_{t\in[0,T]}$ be the value process of the hedge portfolio long one option valued according to the implied market conception, $\{V_t^i\}_{t\in[0,T]}$, and short $\{\Delta_t^h = \nabla_x V_t^h\}_{t\in[0,T]}$ units of the underlying with value process $\{X_t\}_{t\in[0,T]}$, where ∇_x is the gradient operator. We suppose the money account B is chosen such that the net value of the position is zero:

$$\Pi_t^h = V_t^i + B_t - \nabla_x V_t^h \bullet X_t = 0,$$

where $\bullet$ is the dot product. Now consider the infinitesimal change to the value of this portfolio over the interval $[t, t+\mathrm{d}t]$, where $t \in [0, T)$. From the *self-financing condition* we have that

$$\mathrm{d}\Pi_t^h = \mathrm{d}V_t^i + r_t B_t \mathrm{d}t - \nabla_x V_t^h \bullet (\mathrm{d}X_t + q_t \circ X_t \mathrm{d}t),$$

where $q_t = (q_1(t, X_{1t}), q_2(t, X_{2t}), \dots, q_n(t, X_{nt}))^\intercal$ codifies the continuous dividend yields and $\circ$ is the Hadamard product.§ Jointly, the two previous equations entail that

$$\mathrm{d}\Pi_t^h = \mathrm{d}V_t^i - \nabla_x V_t^h \bullet (\mathrm{d}X_t - (r_t\boldsymbol{\iota} - q_t)\circ X_t\mathrm{d}t) - r_t V_t^i \mathrm{d}t, \tag{6}$$

where $\boldsymbol{\iota} = (1, 1, \dots, 1)^\intercal \in \mathbb{R}^n$.

Now consider the option valued under $\boldsymbol{\sigma}_h(t, X_t)$; from the multi-dimensional Itō formula¶ we have that

$$\mathrm{d}V_t^h = \{\partial_t V_t^h + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}_r^\intercal(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla_{xx}^2 V_t^h \mathbf{D}_{X_t}\boldsymbol{\sigma}_r(t,\widetilde{X}_t)]\}\mathrm{d}t + \nabla_x V_t^h \bullet \mathrm{d}X_t, \tag{7}$$

where we have used the fact that X_t is governed by (2). Meanwhile, V_t^h satisfies the multi-dimensional Black–Scholes equation for dividend paying underlyings,‖

$$r_t V_t^h = \partial_t V_t^h + \nabla_x V_t^h \bullet ((r_t\boldsymbol{\iota} - q_t)\circ X_t) + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}_h^\intercal(t,X_t)\mathbf{D}_{X_t}\nabla_{xx}^2 V_t^h \mathbf{D}_{X_t}\boldsymbol{\sigma}_h(t,X_t)]. \tag{8}$$

Combining this expression with the Itō expansion we obtain,

$$0 = -\mathrm{d}V_t^h + r_t V_t^h \mathrm{d}t + \nabla_x V_t^h \bullet (\mathrm{d}X_t - (r_t\boldsymbol{\iota} - q_t)\circ X_t \mathrm{d}t) + \tfrac{1}{2}\mathrm{tr}[\mathbf{D}_{X_t}\big(\boldsymbol{\sigma}_r(t,\widetilde{X}_t)\boldsymbol{\sigma}_r^\intercal(t,\widetilde{X}_t) - \boldsymbol{\sigma}_h(t,X_t)\boldsymbol{\sigma}_h^\intercal(t,X_t)\big)\mathbf{D}_{X_t}\nabla_{xx}^2 V_t^h]\mathrm{d}t, \tag{9}$$

where we have used the fact that the trace is invariant under cyclic permutations of its constituent matrices. Finally, defining $\boldsymbol{\Sigma}_{rh}(t,\widetilde{X}_t)$ as in (5), and adding (9) to (6) we obtain

$$\begin{aligned}\mathrm{d}\Pi_t^h &= \mathrm{d}V_t^i - \mathrm{d}V_t^h - r_t(V_t^i - V_t^h)\mathrm{d}t + \tfrac{1}{2}\mathrm{tr}[\mathbf{D}_{X_t}\boldsymbol{\Sigma}_{rh}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla_{xx}^2 V_t^h]\mathrm{d}t \\ &= e^{\int_0^t r_u \mathrm{d}u} d(e^{-\int_0^t r_u \mathrm{d}u}(V_t^i - V_t^h)) + \tfrac{1}{2}\mathrm{tr}[\mathbf{D}_{X_t}\boldsymbol{\Sigma}_{rh}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla_{xx}^2 V_t^h]\mathrm{d}t.\end{aligned} \tag{10}$$

Whilst a perfect hedge would render this infinitesimal value-change in the portfolio *zero*, this is clearly not the case here. In fact, upon discounting (10) back to the present ($t = 0$) and integrating up the infinitesimal components, we find that net profit-&-loss incurred over the lifetime of the portfolio is

$$P\&L_{\mathbb{T}}^h = \int_0^T d(e^{-\int_0^t r_u \mathrm{d}u}(V_t^i - V_t^h))$$

†Specifically, if $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times d}$, the Euclidian norm is defined as $|x| \equiv (\sum_{i=1}^n x_i^2)^{1/2}$, whilst the matrical norm is $|A| \equiv (\sum_{i=1}^n \sum_{j=1}^d A_{ij}^2)^{1/2}$.

‡Nonetheless, a common assumption in the stochastic volatility literature is obviously to let χ be driven by a stochastic differential equation of the form $\mathrm{d}\chi_t = m(\chi_t)\mathrm{d}t + v(\chi_t)\mathrm{d}W_t + \bar{v}(\chi_t)\mathrm{d}\bar{W}_t$, where $\bar{W}_t$ is second standard Brownian motion (independent of the first), and m, v and $\bar{v}$ are dimensionally consistent, regularity conforming vectors and matrices.

§Otherwise known as the entry-wise product. As per definition, if A and B are matrices of equal dimensions, then $(A\circ B)_{ij} = A_{ij}B_{ij}$.

¶See for instance Björk (2009), p. 65.

‖See for instance Björk, Theorem 13.1 and Proposition 16.7.

$$+\int_0^T e^{-\int_0^t r_u \mathrm{d}u}\tfrac{1}{2}\mathrm{tr}[\mathbf{D}_{X_t}\mathbf{\Sigma}_{rh}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla^2_{xx}V_t^h]\mathrm{d}t$$
$$= V_0^h - V_0^i + \tfrac{1}{2}\int_0^T e^{-\int_0^t r_u \mathrm{d}u}$$
$$\times\ \mathrm{tr}[\mathbf{D}_{X_t}\mathbf{\Sigma}_{rh}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla^2_{xx}V_t^h]\mathrm{d}t.$$

where $P\&L_{\mathbb{T}}^h \equiv \int_0^T e^{-\int_0^t r_u \mathrm{d}u}\mathrm{d}\Pi_t^h$, and the last line makes use of the fact that $V_T^i = V_T^h = g(X_T)$. This is the desired result. □

REMARK 1 A few observations on this proof are in order: first, the relative simplicity of (4) clearly boils down to the assumption that the market is perceived to be driven by a local volatility model. If this assumption is dropped, equation (8) no longer holds. Secondly, it should be clear that the value of the *P&L* changes sign if we are short on the derivative and long on the underlying. Thirdly, the market price of the derivative enters only though the initial price V_0. That is because we look at the profit-&-loss accrued over the entire lifetime of the portfolio. The case of marking-to-market requires further analysis and/or assumption. We will elaborate on this in the following subsection.

REMARK 2 From a generalist's perspective, theorem 1 suffers from a number of glaring limitations: for instance, the governing asset price dynamics only considers Brownian stochasticity, the hedge is assumed to be a workaday Δ-hedge, and the option type is vanilla European in the sense that the terminal pay-off is determined by the instantaneous price of the underlying assets. Fortunately, the Fundamental Theorem can readily be extended in various directions: e.g. it can be shown that if $V_t = V(t, X_t, A_t)$ is an Asian option written on the continuous average A_t of the underlying process X_t, then the Fundamental Theorem remains form invariant. In section 3, we consider one particularly topical dynamical modification viz. the incorporation of possible market crashes through jump diffusion.

2.3. *The implications for Δ-hedging*

From a first inspection, the Fundamental Theorem quite clearly demonstrates that reasonably successful hedging is possible *even* under significant model uncertainty. Indeed, as Davis (2010) puts it 'without some robustness property of this kind, it is hard to imagine that the derivatives industry could exist at all'. In this section, we dive further into the implications of what happens to our portfolio, by considering the case where we hedge with (a) the real volatility, and (b) the implied volatility. It is a standing assumption in this subsection that whatever is used as hedge volatility is of the form that allows the use of theorem 1 (see the discussion in remark 1 above). The reader can think of the cases of constant volatility.

2.3.1. Hedging with the real volatility. Suppose we happen to be bang on our estimate of the real volatility matrix in our Δ-hedge, i.e. let $\boldsymbol{\sigma}_h(t, X_t) = \boldsymbol{\sigma}_r(t, \widetilde{X}_t)$ a.s. $\forall t \in [0, T]$, then $\mathbf{\Sigma}_{rr}(t, \widetilde{X}_t) = \mathbf{0}$ and the present valued profit-&-loss amounts to

$$P\&L_{\mathbb{T}}^r = V_0^r - V_0^i,$$

which is manifestly deterministic.† However, we observe that this relies crucially on us holding the portfolio until expiry of the option. Day-to-day fluctuations of the profit-&-loss still vary stochastically (erratically) as it is vividly demonstrated by combining equation (9) (where $h = i$) with equation (6) (where $h = r$):

$$\mathrm{d}\Pi_t^r = \tfrac{1}{2}\mathrm{tr}[\mathbf{D}_{X_t}\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla^2_{xx}V_t^i]\mathrm{d}t + \nabla_x\left(V_t^i - V_t^r\right)\bullet\Big\{(\boldsymbol{\mu}_t^r - r_t\boldsymbol{\iota} + \boldsymbol{q}_t)\circ X_t\mathrm{d}t + \mathbf{D}_{X_t}\boldsymbol{\sigma}_r(t,\widetilde{X}_t)\mathrm{d}\boldsymbol{W}_t\Big\},$$

cf. the explicit dependence of the Brownian increment. As for the *profitability* of the Δ-hedging strategy, this is a complex issue which ultimately must be studied on a case-by-case basis. However, for options with positive *vega*,‡ it suffices to require that the real volatility everywhere exceeds the implied volatility.

2.3.2. Hedging with the implied volatility. Suppose instead we hedge the portfolio using the implied volatility matrix $\boldsymbol{\sigma}_i(t, X_t)\ \forall t \in [0, T]$, then the associated present-valued profit-&-loss is of the form

$$P\&L_{\mathbb{T}}^i = \tfrac{1}{2}\int_0^T e^{-\int_0^t r_u \mathrm{d}u}\mathrm{tr}[\mathbf{D}_{X_t}\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla^2_{xx}V_t^i]\mathrm{d}t.$$

As we find ourselves integrating over the stochastic process X_t, this profit-&-loss is manifestly stochastic. Notice though that $\mathrm{d}\Pi_t^i$ here does *not* depend explicitly on the Brownian increment (the daily profit-and-loss is $\mathcal{O}(\mathrm{d}t)$) which gives rise to point that 'bad models cause bleeding—not blow-ups'. As for the *profitability* of the strategy, again this is a complex issue: however, insofar as $\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)\circ\nabla^2_{xx}V_t^i$ is *positive definite* a.s. for all $t \in [0, T]$, then we are making a profit with probability one. To see this, recall that the trace can be written as§

$$\mathrm{tr}[\mathbf{D}_{X_t}\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)\mathbf{D}_{X_t}\nabla^2_{xx}V_t^i] = X_t^{\mathsf{T}}(\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)\circ\nabla^2_{xx}V_t^i)X_t,$$

In particular, if $\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)\circ\nabla^2_{xx}V_t^i$ is positive definite at all times, i.e.

$$\forall t \in [0, T]\ \forall X_t \in \mathbb{R}^n : X_t^{\mathsf{T}}(\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)\circ\nabla^2_{xx}V_t^i)X_t > 0,$$

then $P\&L_{\mathbb{T}}^i > 0$. A sufficient condition for this to be the case is that $\mathbf{\Sigma}_{ri}(t,\widetilde{X}_t)$ and $\nabla^2_{xx}V_t^i$ individually are positive definite $\forall t$, as demonstrated by the *Schur Product Theorem*.

2.3.3. Wilmott's hedge experiment. The points imbued in the previous two paragraphs are forcefully demonstrated in the event that there is only one risky asset in existence, the

†Obviously, this can only be the case if there is no underlying state variable.

‡A clear example of vega being manifestly positive would be European calls and puts, which satisfy the assumptions needed to derive the Black–Scholes formula. Explicitly, $\nu \equiv \frac{\partial V}{\partial \sigma} = S_t e^{-\delta(T-t)}\phi(d_1)\sqrt{T-t} > 0$ where ϕ is the standard normal pdf and d_1 has the usual definition.

§This follows from the general identity for matrices $\mathbf{A}$ and $\mathbf{B}$ of corresponding dimensions: $\mathbf{x}^{\mathsf{T}}(\mathbf{A}\circ\mathbf{B})\mathbf{y} = \mathrm{tr}[\mathbf{D_x}\mathbf{A}\mathbf{D_y}\mathbf{B}^{\mathsf{T}}]$ where $\mathbf{x}$ and $\mathbf{y}$ are vectors.

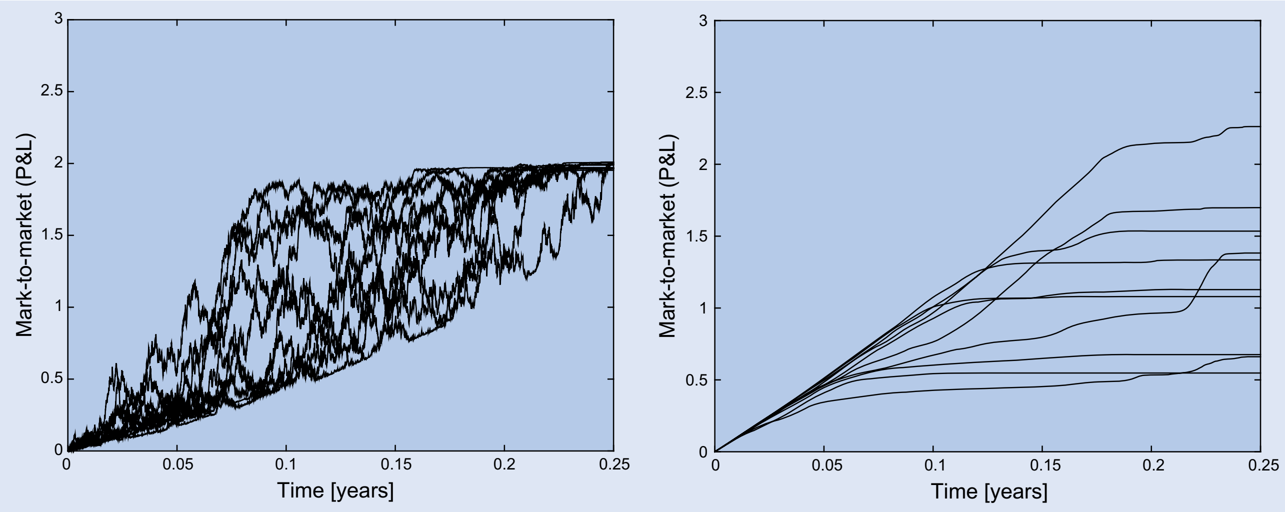

Figure 1. Left: Delta hedging a portfolio assuming that $\sigma_h = \sigma_r$. The parameter specifications are: $r = 0.05$, $\mu = 0.1$, $\sigma_i = 0.2$, $\sigma_r = 0.3$, $S_0 = 100$, $K = 100$, $q = 0$ and $T = 0.25$. The portfolio is rebalanced 5000 times during the lifetime of the option. Observe that whilst the P&L fluctuates randomly along the path of S_t due to the presence of $\mathrm{d}W_t$, the accumulated P&L at the maturity of the option is the deterministic quantity $\Pi_T = e^{rT}(V_0^r - V_0^i)$. From the Black–Scholes formula, it follows that $V_0^r = 6.583$ and $V_0^i = 4.615$ so $\Pi_{T=1} = 1.993$. The fact that our 10 paths only approximately hit this terminal value is attributable to the discretization of the hedging which should be done in continuous time. Right: Delta hedging a portfolio assuming that $\sigma_h = \sigma_i$. The parameter specifications are as before. Evidently, the accumulated P&L stays highly path dependent for the *entire* duration of the option. However, the curves per se are smooth, which highlights that $\mathrm{d}\Pi_t^i$ does not depend explicitly on the Brownian increment.

derivative is a European call option and all volatilities are assumed constant. Based on Wilmott and Ahmad, figure 1 clearly illustrates the behaviour of the profit-&-loss paths insofar as we hedge with (a) the real volatility, and (b) the implied volatility. Again, the main insights are as follows: hedging V_t^i with the real volatility causes the $P\&L$ of the portfolio to fluctuate erratically over time, only to land at a deterministic value at maturity. On the other hand, hedging V_t^i with the implied volatility yields smoother (albeit still stochastic) P&L curves. Nonetheless, here, there is no way of telling what the P&L actually amounts to at maturity.

Rather perturbingly, both strategies blatantly suggest the relative ease with which we can make *volatility arbitrage*. Specifically, assuming that the historical volatility is a reasonable proxy for the real volatility, $\sigma_{\text{hist}} \approx \sigma_r$, and that $\sigma_{\text{hist}} > \sigma_i$ ($\sigma_{\text{hist}} < \sigma_i$), it would suffice to go long (short) on the hedge portfolio for $\mathbb{P}(P\&L_{\mathbb{T}} \geq 0) = 1$ and $\mathbb{P}(P\&L_{\mathbb{T}} > 0) > 0$.

The Wilmott experiment reflects the concerting approach (Hull and White 2016) of traders: Hedge with market implied parameters/quantities. For the experiment's strong smoothness property of implied volatility hedging P&L to hold it is important that implied volatility is assumed constant, or at least that it does not diffuse randomly as it would in, e.g. the Heston model. Were the latter the case, then anything might happen. However, since implied volatility is intricately linked to (conditional expectation of) time-integrated real volatility (squared),† it is likely to be quantitatively smoother than real volatility, so there is hope that the conclusion of the idealized experiment carries over to empirical analysis. In section 4, we will see that this is indeed the case.

Remark 3 It is tempting to think of an option trade as a zero-sum game (up to a risk-premium) between the buyer and the seller; if one wins the other loses. That, however, is not true. Imagine a stock dynamics which is described by geometric Brownian motion with 15% volatility. A buyer is strongly bullish and makes a directional bet by buying an at-the-money call-option, which he is willing to pay 20% implied volatility on. The buyer does not hedge his position, the market rallies, and he makes a nice profit. Meanwhile, the seller is right in her 15%-volatility forecast, Δ-hedges and makes a nice (arbitrage) profit too. Needless to say, option writing can also imply losses for both the buyer and the seller as we saw it during the financial crisis.‡

2.4. Applications

Due to the presence of the real volatility, the exact nature of which transcends our epistemic domain, one might reasonably ponder whether the Fundamental Theorem conveys any practical points besides those of the preceding subsection. Using two poignant (even if somewhat eccentric) examples, we will argue that the gravity of the Fundamental Theorem propagates well into risk management and volatility surface calibration. Zero rates and dividends will be assumed throughout.

Example 1 Let $V_t(T, K)$ be the price process of a European strike K maturity T call or put option, written on an underlying which obeys geometric Brownian motion, $\mathrm{d}X_t = X_t[\mu_r \mathrm{d}t + \sigma_r \mathrm{d}W_t]$, where μ_r, σ_r are constants. Suppose we Δ-hedge a long position on V_t at the *implied* volatility, $\sigma_h = \sigma_i$, then the Fundamental Theorem implies that

$$P\&L_{\mathbb{T}}^i = \frac{1}{2}\int_0^T (\sigma_r^2 - \sigma_i^2) X_t^2 \Gamma_t^i \mathrm{d}t,$$

†Think of the Black–Scholes model with time-dependent volatility or see Romano and Touzi's Proposition 4.1 (Romano and Touzi 1997).

‡This observation is inspired by Antoine Savine.

where

$$\Gamma_t^i \equiv \frac{\phi(d_1^i)}{X_t \sigma_i \sqrt{T-t}},$$

is the option's gamma, $\phi : \mathbb{R} \mapsto \mathbb{R}_+$ is the standard normal pdf and

$$d_1^i \equiv \frac{1}{\sigma_i\sqrt{T-t}} \left\{ \ln(X_t/K) + \tfrac{1}{2}\sigma_i^2(T-t) \right\}.$$

Since $\forall t\ \Gamma_t^i > 0$ the strategy is profitable if and only if $\sigma_r^2 > \sigma_i^2$. Furthermore, by maximizing the integrand with respect to X_t, we find that the $P\&L_{\mathbb{T}}^i$ is maximal when

$$X_t^* = K e^{\frac{1}{2}\sigma_i^2(T-t)},$$

Specifically, upon evaluating the integral explicitly it can be shown that†

$$\max_{X_t} P\&L_{\mathbb{T}}^i = \sqrt{\frac{T}{2\pi}}\frac{K}{\sigma_i}(\sigma_r^2 - \sigma_i^2).$$

From a risk management point of view, the important point is that we can compute a confidence interval for the real volatility based on historical observations. Hence, we can compute a confidence interval for the maximal profit-&-loss we might face upon holding the hedge portfolio till expiry.

EXAMPLE 2 Let $V_t = C_t(T, K)$ be the price process of a European strike K maturity T call option written on an underlying price process X. As in (2), we assume the fundamental dynamics to be of the form $\mathrm{d}X_t = X_t[\mu_r(t, \widetilde{X}_t)\mathrm{d}t + \sigma_r(t, \widetilde{X}_t)\mathrm{d}W_t]$, where $\widetilde{X}_t$ is defined as the $(1+m)$-dimensional vector $(X_t; \boldsymbol{\chi}_t)$ and $\boldsymbol{\chi}$ is a state variable. Also, we suppose the integrability condition $\mathbb{E}[\int_0^T \sigma_r^2(t, \widetilde{X}_t)X_t^2 \mathrm{d}t] < \infty$, and that there exists an equivalent martingale measure, $\mathbb{Q}$, which renders X_t a martingale (recall the risk-free rate is assumed zero)‡:

$$\mathrm{d}X_t = \sigma_r(t, \widetilde{X}_t)X_t \mathrm{d}W_t^{\mathbb{Q}}.$$

Now consider the admittedly somewhat contrived scenario of a Δ-hedged portfolio, long one unit of the call, for which σ_h and σ_i are both zero.§ The associated value process is

$$\begin{aligned}\Pi_t^i &= C_t^i(T, K) + B_t - \partial_x C_t^h(T, K) \cdot X_t \\ &= (X_t - K)^+ + B_t - \mathbf{1}_{\{X_t > K\}} X_t, \end{aligned} \tag{11}$$

where $\mathbf{1}_{\{X_t>K\}}$ is the indicator function. The important point here is that $(X_t - K)^+$ may be reinterpreted as the terminal pay-off of a strike K maturity t call option (obviously, the specification $\sigma_h = \sigma_i = 0$ is paramount here). Substituting (11) into the infinitesimal form of the Fundamental Theorem,

$$\mathrm{d}\Pi_t^i = \tfrac{1}{2}(\sigma_r^2(t, \widetilde{X}_t) - \sigma_i^2)X_t^2 \partial_{xx}^2 C_t^i(T, K)\mathrm{d}t,$$

we find that

$$d((X_t - K)^+ + B_t - \mathbf{1}_{\{X_t>K\}}X_t) = \tfrac{1}{2}\sigma_r^2(t, \widetilde{X}_t)X_t^2\delta(X_t - K)\mathrm{d}t, \tag{12}$$

where we once again have made use of $\sigma_i = 0$, alongside the fact that $\partial_x \mathbf{1}_{\{X_t>K\}}$ is the Dirac delta-function $\delta(X_t - K)$. Taking the risk neutral expectation of (12), conditional on $\mathscr{F}_0$, the left-hand side reduces to

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[LHS] &= \mathbb{E}^{\mathbb{Q}}[d(X_t - K)^+] + \mathbb{E}^{\mathbb{Q}}[\mathrm{d}B_t - \mathbf{1}_{\{X_t>K\}}\mathrm{d}X_t] \\ &= \mathrm{d}\mathbb{E}^{\mathbb{Q}}[(X_t - K)^+] - \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{X_t>K\}}\mathrm{d}X_t] \\ &= \mathrm{d}C_0^r(t, K) - \mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{X_t>K\}}\mathrm{d}X_t|\mathscr{F}_t]] \\ &= \mathrm{d}C_0^r(t, K) - \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{X_t>K\}}\mathbb{E}^{\mathbb{Q}}[\mathrm{d}X_t|\mathscr{F}_t]] \\ &= \mathrm{d}C_0^r(t, K), \end{aligned} \tag{13}$$

where the second line uses $\mathrm{d}B_t = 0$, whilst the third line uses the law of iterated expectations and the fact that $\mathbb{E}^{\mathbb{Q}}[(X_t - K)^+]$ is the time zero price of a strike K maturity t call option. Finally, the fourth line follows from the $\mathscr{F}_t$-measurability of $\mathbf{1}_{\{X_t>K\}}$, whilst the fifth line exploits the martingale property $\mathbb{E}^{\mathbb{Q}}[dX_t] = 0$.

As for the right-hand side, define the joint density

$$f_{\sigma_r^2, X_t}^{\mathbb{Q}}(\sigma^2, x)\mathrm{d}\sigma^2\mathrm{d}x \equiv \mathbb{Q}(\{\sigma^2 \le \sigma_r^2 \le \sigma^2 + \mathrm{d}\sigma^2\} \cap \{x \le X_t \le x + \mathrm{d}x\}),$$

then

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[RHS] &= \tfrac{1}{2}\iint_{\mathbb{R}_+^2} \sigma^2 x^2 \delta(x - K) f_{\sigma_r^2, X_t}^{\mathbb{Q}}(\sigma^2, x)\mathrm{d}\sigma^2\,\mathrm{d}x\,\mathrm{d}t \\ &= \tfrac{1}{2}\iint_{\mathbb{R}_+^2} \sigma^2 x^2 \delta(x - K) f_{\sigma_r^2}^{\mathbb{Q}} \\ &\quad \times (\sigma^2|X_t = x) f_{X_t}^{\mathbb{Q}}(x)\mathrm{d}\sigma^2\,\mathrm{d}x\,\mathrm{d}t \\ &= \tfrac{1}{2}\int_{\mathbb{R}_+} x^2\delta(x - K) f_{X_t}^{\mathbb{Q}}(x) \\ &\quad \times \left\{ \int_{\mathbb{R}_+} \sigma^2 f_{\sigma_r^2}^{\mathbb{Q}}(\sigma^2|X_t = x)\mathrm{d}\sigma^2 \right\}\mathrm{d}x\,\mathrm{d}t \\ &\equiv \tfrac{1}{2}\int_{\mathbb{R}_+} x^2\delta(x - K) f_{X_t}^{\mathbb{Q}}(x) \\ &\quad \times \mathbb{E}^{\mathbb{Q}}[\sigma_r^2(t, \widetilde{X}_t)|X_t = x]\mathrm{d}x\,\mathrm{d}t \\ &= \tfrac{1}{2}K^2 f_{X_t}^{\mathbb{Q}}(K)\mathbb{E}^{\mathbb{Q}}[\sigma_r^2(t, \widetilde{X}_t)|X_t = K]\mathrm{d}t. \end{aligned} \tag{14}$$

Recalling that $\partial_K \mathbb{E}^{\mathbb{Q}}[(X_t - K)\mathbf{1}_{\{X_t>K\}}] = -\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{X_t>K\}}]$, and $-\partial_K \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{X_t>K\}}] = \mathbb{E}^{\mathbb{Q}}[\delta(X_t - K)]$ we arrive at the Breeden–Litzenberger formula

$$f_{X_t}^{\mathbb{Q}}(K) = \partial_{KK}^2 C_0^r(t, K). \tag{15}$$

Combining equations (13), (14) and (15), we thus have that

$$\frac{\mathrm{d}C_0^r}{\mathrm{d}t}(t, K) = \tfrac{1}{2}\partial_{KK}^2 C_0^r(t, K)K^2\mathbb{E}^{\mathbb{Q}}[\sigma_r^2(t, \widetilde{X}_t)|X_t = K],$$

†We observe that a similar result can be found in Derman (2008).

‡Obviously, such an existence claim is not altogether innocuous. Indeed, the measure change is here further complicated by the fact that we have not made formal specifications for the dynamical form of the state variable $\boldsymbol{\chi}_t$. However, insofar as we adopt the standard dynamical assumption $\mathrm{d}\boldsymbol{\chi}_t = \boldsymbol{m}(\boldsymbol{\chi}_t)\mathrm{d}t + \boldsymbol{v}(\boldsymbol{\chi}_t)dW_t + \bar{\boldsymbol{v}}(\boldsymbol{\chi}_t)\mathrm{d}\bar{\boldsymbol{W}}_t$, our existence claim is tantamount to positing the existence of a market price of risk vector $\boldsymbol{\theta} \in \mathbb{R}^m$ which renders $L(T) = L_X(T)L_{\boldsymbol{\chi}}(T)$ a true martingale, where

$$L_X(T) \equiv \exp\left\{ -\int_0^T \frac{\mu_r(t, \widetilde{X}_t)}{\sigma_r(t, \widetilde{X}_t)}\mathrm{d}W_t - \frac{1}{2}\int_0^T \frac{\mu_r^2(t, \widetilde{X}_t)}{\sigma_r^2(t, \widetilde{X}_t)}\mathrm{d}t \right\},$$

and

$$L_{\boldsymbol{\chi}}(T) \equiv \exp\left\{ -\int_0^T \boldsymbol{\theta}_t^{\intercal}\mathrm{d}\bar{\boldsymbol{W}}_t - \frac{1}{2}\int_0^T |\boldsymbol{\theta}_t|^2\mathrm{d}t \right\}.$$

§To be precise, the contrived part is the assumption that the call trades at zero volatility; less so that we hedge it at zero volatility. The latter corresponds to a so-called *stop-loss strategy*, see Carr and Jarrow (1990).

which using the change of notation† $t = T$ amounts to the celebrated Dupire–Gyöngy–Derman–Kani formula

$$\mathbb{E}^{\mathbb{Q}}[\sigma_r^2(T, \widetilde{X}_T)|X_T = K] = \frac{\partial_T C_0^r(T, K)}{\frac{1}{2}K^2\partial_{KK}^2 C_0^r(T, K)}, \quad (16)$$

– see Derman and Kani (1998), Dupire (1994), Gyöngy (1986). Conceptually, the important point here is that the right-hand side, modulo some amount of interpolation,‡ is empirically measurable, whence (16) provides a way of calibrating the volatility surface to observed call option prices in the market. Specifically, the formula enables us to square a local diffusion model with the infamous skew/smile effect of implied volatilities across different values for the strike and time to maturity, whilst delicately sidestepping the deeper issue as to why this phenomenon prevails.

REMARK 4 In Wittgensteinian terms, we must 'throw away the ladder' to arrive at this final conclusion, Wittgenstein (1922), prop. 6.54. Hitherto, we have assumed that the real parameters (r) are fundamentally unobservable, whilst the implied parameters (i) are those we are exposed to in the market. Yet, no such distinction exists in the works of Dupire et al., whence the r superscript in (16) really ought to be dropped.

REMARK 5 The above derivation is arguably unconventional and neither rigorous nor the quickest way to demonstrate (16). In fact, the entire point of setting $\sigma_i = 0$ is essentially to extract the Itō-(Tanaka) formula applied to $(X_t - K)^+$, from which Derman et al.'s derivation takes its starting point. We keep the derivation here, as it provides a curious glimpse into how two philosophically quite distinct theorems can be interconnected.§

3. The Gospel of the Jump

Following remark 2, it is worthwhile exploring how the Fundamental Theorem can be adapted to new terrain. For instance, it is well known that Brownian motion in itself does not adequately capture the sporadic discontinuities that emerge in stock price processes. Hence, it is opportune to scrutinize the effect of a jump diffusion process, which in turn will give rise to another valuable lesson on the profitability of imperfect hedging.

Already, it is a well-known fact that exact hedges generally do not exist in a jump economy where the true dynamics of the underlying is perfectly disseminated (see Merton (1976) or the expositions by Privault (2013) and Shreve (2008)). It is thus of some theoretical interest to see how this preexisting hedge error is further complicated under the model error framework of the Fundamental Theorem. Stepping stones towards answering this question are found in Andreasen (2003) and Davis (2010) both of whom consider a mono-dimensional implied-hedge scenario with perfect information about the jump diffusion component. Our contribution is to generalize their results to a multi-dimensional framework with arbitrary (mis)-specifications for the volatility and jump distribution. For an overview of multi-dimensional jump-diffusion theory, we refer the reader to the appendix.

Suppose the real dynamics of the underlying price process obeys

$$\mathrm{d}\boldsymbol{X}_t = \mathbf{D}_{\boldsymbol{X}_t}[\boldsymbol{\mu}_r(t, \widetilde{\boldsymbol{X}}_t)\mathrm{d}t + \boldsymbol{\sigma}_r(t, \widetilde{\boldsymbol{X}}_t)\mathrm{d}\boldsymbol{W}_t] + \mathbf{D}_{\boldsymbol{X}_{t-}}\mathrm{d}\boldsymbol{Y}_t, \quad (17)$$

where $\{\boldsymbol{Y}_t\}_{t\geq 0}$ is an n-dimensional vector of independent compound Poisson processes. Specifically, the jth component is given by

$$Y_t^j = \sum_{k=1}^{N_t^j} Z_k^j,$$

where $\{N_t^j\}_{t\geq 0}$ is an intensity-λ_j Poisson process, and $\{Z_k^j\}_{k\geq 1}$ is a sequence of relative jump-sizes, assumed to be i.i.d. square-integrable random variables with cumulative distribution function (cdf) $\nu_j : \mathbb{R} \mapsto [0, 1]$. For shorthand, we shall refer to the vectors $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)^\intercal$ and $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_n)^\intercal$ as the intensity and cdf of $\boldsymbol{Y}_t$.

Oblivious to the true nature of (17), we imagine that pricing and hedging should be performed (with obvious notation) under the tuple $\langle \boldsymbol{\phi}, \boldsymbol{\lambda}_h^{\mathbb{Q}}, \boldsymbol{\nu}_h^{\mathbb{Q}}, \boldsymbol{\sigma}_h(t, \boldsymbol{X}_t), \mathbb{Q}\rangle$, where $\mathbb{Q}$ is the risk neutral measure

$$\begin{aligned}\mathrm{d}\mathbb{Q}_{\boldsymbol{\phi},\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}} = \exp\Bigg\{ &\int_0^T \boldsymbol{\phi}_s \bullet \mathrm{d}\boldsymbol{W}_s \\ &- \frac{1}{2}\int_0^T |\boldsymbol{\phi}_s|^2 \mathrm{d}s - \sum_{j=1}^n (\lambda_{h,j}^{\mathbb{Q}} - \lambda_{h,j})T \Bigg\} \\ &\times \prod_{j=1}^n \prod_{k=1}^{N_t^j} \frac{\lambda_{h,j}^{\mathbb{Q}} \mathrm{d}\nu_{h,j}^{\mathbb{Q}}(Z_k^j)}{\lambda_{h,j}\mathrm{d}\nu_{h,j}(Z_k^j)} \mathrm{d}\mathbb{P}^h, \end{aligned} \quad (18)$$

such that $\{\boldsymbol{\phi}_t\}_{t\geq 0}$ is a bounded adapted n-dimensional process (the so-called Girsanov kernel), and $\boldsymbol{\lambda}_h^{\mathbb{Q}}$, $\boldsymbol{\nu}_h^{\mathbb{Q}}$, respectively, represent the jump intensity and jump-size distribution under $\mathbb{Q}$. Specifically, the price of an option with terminal pay-off $g(\boldsymbol{X}_T)$ is determined as

$$V_t^h = \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^T r_u \mathrm{d}u} g(\boldsymbol{X}_T)|\mathscr{F}_t^X],$$

with the underlying supposedly driven by

$$\begin{aligned}\mathrm{d}\boldsymbol{X}_t = \mathbf{D}_{\boldsymbol{X}_t}[r_t\boldsymbol{\iota}\mathrm{d}t + \boldsymbol{\sigma}_h(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t^{\mathbb{Q}}] \\ + \mathbf{D}_{\boldsymbol{X}_{t-}}[\mathrm{d}\boldsymbol{Y}_t - \boldsymbol{\lambda}_h^{\mathbb{Q}} \circ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1]],\end{aligned}$$

with $\boldsymbol{Z}_1 = (Z_1^1, Z_1^2, \ldots, Z_1^n)^\intercal$, and $\mathbb{Q}$ has been specified such that

$$\boldsymbol{\mu}_h(t, \boldsymbol{X}_t) + \boldsymbol{\lambda}_h^{\mathbb{Q}} \circ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1] + \boldsymbol{\sigma}_h(t, \boldsymbol{X}_t)\boldsymbol{\phi}_t = r_t\boldsymbol{\iota}, \quad (19)$$

is satisfied almost everywhere.¶

†We do this to emphasize that t is the *maturity* of the option (not its value at time t).

‡Exactly how to do this extrapolation has turned out to be sufficiently non-trivial to spurn numerous papers and successive quant-of-the-year awards a-decade-and-a-half later, see Andreasen and Huge (2011) (pure local volatility), Guyon and Henry-Labordère (2012) (decorated stochastic volatility models).

§Our motivation for establishing this curious connection stems from Andreasen (2003) who declares the Dupire–Gyöngy–Derman–Kani formula to be a corollary of the Fundamental Theorem (without proof). Whether this take on the situation ultimately stems from Dupire himself is perhaps dubious, but it is interesting to note that a variant of the Fundamental Theorem has appeared in his work at least since early 2003 (Dupire 2003).

¶It should be clear the $\mathbb{Q}$ is not uniquely determined. In fact, for (19) to admit only one solution, we would require that either (i) $\boldsymbol{\lambda}_h = \boldsymbol{\lambda}_h^{\mathbb{Q}} = 0$ (there are *no* jumps), in which case we recover the standard Girsanov

REMARK 6 We emphasize that (18) is a risk neutral measure transformation of the *hedge* dynamics with the associated measure $\mathbb{P}^h$. This is to be contrasted with example 2 in section 2.4 in which $\mathbb{Q}$ is the risk neutral measure of the *real* dynamics.

THEOREM 2 The Fundamental Theorem of Derivative Trading with Jumps *Let $V_t = V(t, \boldsymbol{X}_t) \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^n)$ be the price process of a European option with terminal payoff $V_T = g(\boldsymbol{X}_T)$, the underlying of which follows the real dynamics (17). We assume the option trades in the market at the (not necessarily uniquely determined) implied volatility $\boldsymbol{\sigma}_i = \boldsymbol{\sigma}_i(t, \boldsymbol{X}_t)$. Now suppose we at time $t = 0$ acquire such an option for the implied price V_0^i and set out to Δ-hedge our position. Said hedge is performed under the notion $\langle \boldsymbol{\phi}, \boldsymbol{\lambda}_h^{\mathbb{Q}}, \boldsymbol{\nu}_h^{\mathbb{Q}}, \boldsymbol{\sigma}_h(t, \boldsymbol{X}_t), \mathbb{Q} \rangle$, leading to the fair price process V_t^h. Then the present value of the **profit-&-loss** we incur from holding such a portfolio over the interval $\mathbb{T} = [0,T]$ is*

$$
\begin{aligned}
\mathrm{P\&L}_{\mathbb{T}}^h &= V_0^h - V_0^i + \tfrac{1}{2} \int_0^T e^{-\int_0^t r_u \mathrm{d}u} \\
&\quad \times tr[\mathbf{D}_{X_t} \boldsymbol{\Sigma}_{rh}(t, \widetilde{\boldsymbol{X}}_t) \mathbf{D}_{X_t} \nabla^2_{\boldsymbol{xx}} V_t^h] \mathrm{d}t, \\
&\quad + \int_0^T \sum_{j=1}^n e^{-\int_0^t r_u \mathrm{d}u} \bigg\{ \Big(\Delta_j V_t^h(t, \boldsymbol{X}_{t-}) \\
&\qquad - X_{j,t-} Z_{N_t^j} \partial_{x_j} V_t^h \Big) \mathrm{d}N_t^j \\
&\quad - \lambda_{h,j}^{\mathbb{Q}} \Big(\mathbb{E}^{\mathbb{Q}}[\Delta_j V_t^h(t, \boldsymbol{x})]|_{\boldsymbol{x}=\boldsymbol{X}_{t-}} \\
&\qquad - X_{j,t-} \mathbb{E}^{\mathbb{Q}}[Z_1^i] \partial_{x_j} V_t^h \Big) \mathrm{d}t \bigg\},
\end{aligned} \tag{20}
$$

where

$$\Delta_j V_t^h(t, \boldsymbol{X}_{t-}) \equiv V^h(t, \boldsymbol{X}_{t-} \circ (\boldsymbol{\iota} + \hat{\mathbf{e}}_j Z_{N_t^j}^j)) - V^h(t, \boldsymbol{X}_{t-}),$$

represents the change in value of the option when the underlying jumps in the jth component, and $\hat{\mathbf{e}}_j$ is a unit vector in $\mathbb{R}^n$ s.t. $[\hat{\mathbf{e}}_j]_k = \delta_{j,k}$.

Sketch Proof The proof runs in parallel with that of theorem 1. Specifically, the analogue of expression (6) is

$$
\begin{aligned}
\mathrm{d}\Pi_t^h &= \mathrm{d}V_t^i - \nabla_{\boldsymbol{x}} V_t^h \bullet (\mathrm{d}\boldsymbol{X}_t^{\text{cont.}} - (r_t \boldsymbol{\iota} - \boldsymbol{q}_t) \circ \boldsymbol{X}_t \mathrm{d}t) \\
&\quad - \nabla_{\boldsymbol{x}} V_t^h \bullet \mathrm{d}\boldsymbol{Y}_t - r_t V_t^i \mathrm{d}t,
\end{aligned}
$$

where $\mathrm{d}\boldsymbol{X}_t^{\text{cont.}}$ is the continuous part of (17) i.e.

$$\mathrm{d}\boldsymbol{X}_t^{\text{cont.}} = \mathbf{D}_{X_t}[\boldsymbol{\mu}_r(t, \widetilde{\boldsymbol{X}}_t)\mathrm{d}t + \boldsymbol{\sigma}_r(t, \widetilde{\boldsymbol{X}}_t)\mathrm{d}\boldsymbol{W}_t].$$

Furthermore, in analogy with (7) and (8) we have the Itō formula

$$
\begin{aligned}
\mathrm{d}V_t^h &= \{\partial_t V_t^h + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}_r^{\intercal}(t, \widetilde{\boldsymbol{X}}_t) \mathbf{D}_{X_t} \nabla^2_{\boldsymbol{xx}} V_t^h \mathbf{D}_{X_t} \boldsymbol{\sigma}_r(t, \widetilde{\boldsymbol{X}}_t)]\}\mathrm{d}t \\
&\quad + \nabla_{\boldsymbol{x}} V_t^h \bullet \mathrm{d}\boldsymbol{X}_t^{\text{cont.}}, \\
&\quad + \sum_{j=1}^n [V^h(t, \boldsymbol{X}_{t-} \circ (\boldsymbol{\iota} + \hat{\mathbf{e}}_j Z_{N_t^j}^j)) - V^h(t, \boldsymbol{X}_{t-})] \mathrm{d}N_t^j,
\end{aligned}
$$

and the partial integro-differential equation for pricing purposes

$$
\begin{aligned}
r_t V_t^h &= \partial_t V_t^h + \nabla_{\boldsymbol{x}} V_t^h \bullet ((r_t \boldsymbol{\iota} - \boldsymbol{q}_t) \circ \boldsymbol{X}_t) \\
&\quad + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}_h^{\intercal}(t, \boldsymbol{x}) \mathbf{D}_{X_t} \nabla^2_{\boldsymbol{xx}} V_t \mathbf{D}_{X_t} \boldsymbol{\sigma}_h(t, \boldsymbol{X})] \\
&\quad + \sum_{j=1}^n \lambda_{h,j}^{\mathbb{Q}} \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[V(t, \boldsymbol{x} \circ (\boldsymbol{\iota} + \hat{\mathbf{e}}_j Z_1^j)) \\
&\quad - V(t, \boldsymbol{x}) - x_j Z_1^j \partial_{x_j} V(t, \boldsymbol{x})]_{\boldsymbol{x}=\boldsymbol{X}_{t-}}.
\end{aligned}
$$

Combining these three expressions as above yields the desired result. □

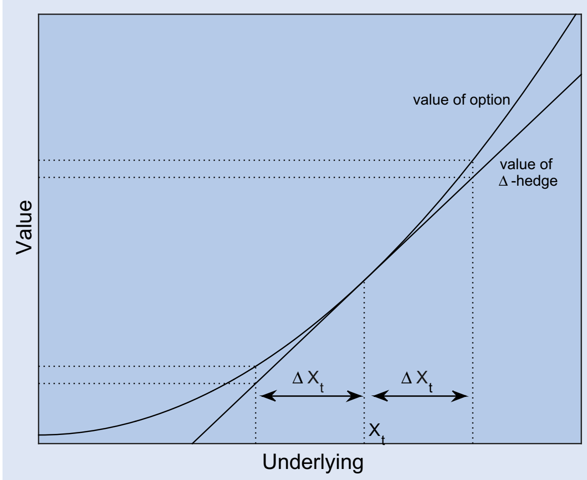

Figure 2. Suppose we Δ-hedge a long position in an option with a convex pricing function. Insofar as a jump in the underlying occurs, $X_t \mapsto X_t \pm \Delta X_t$, it follows that the value of the option will exceed the value of the Δ-position. Hence, our net *P&L* benefits from such an occurrence. Obviously, the converse will be true if we hold a short position in the option.

REMARK 7 The last four lines in (20), which we denote by $P\&L_J$, represent the present-valued profit-&-loss brought about by our inability to hedge the jump risk completely. A more compact way of writing this result is attained by considering the associated Poisson random measure $J_j(\mathrm{d}t \times \mathrm{d}z^j)$ with intensity measure $\mathbb{E}[J_j(\mathrm{d}t \times \mathrm{d}z^j)] = \lambda_j \mathrm{d}t \mathrm{d}\nu_j(z^j)$ for $j = 1, 2, \ldots, n$. Specifically, upon defining the pseudo-compensated random measure

$$\tilde{J}_{h,j}(\mathrm{d}t \times \mathrm{d}z^j) \equiv J_j(\mathrm{d}t \times \mathrm{d}z^j) - \lambda_{h,j}^{\mathbb{Q}} \mathrm{d}t\, \mathrm{d}\nu_{h,j}^{\mathbb{Q}}(z^j), \tag{21}$$

for $j = 1, 2, \ldots, n$, we see that the jump contribution to the profit-and-loss may be written as

$$
\begin{aligned}
P\&L_J = \int_0^T \int_{\mathbb{R}} \sum_{j=1}^n e^{-\int_0^t r_u \mathrm{d}u} \{ \Delta_j V_t^h(t, \boldsymbol{X}_{t-}) \\
- z^j X_{j,t-} \partial_{x_j} V_t^h \} \tilde{J}_{h,j}(\mathrm{d}t \times \mathrm{d}z_j),
\end{aligned} \tag{22}
$$

where 'pseudo' is used to emphasize that (21) is a real-world Poisson random measure compensated by a mis-specified intensity measure term: it is neither a martingale measure under $\mathbb{P}$ nor under the listed $\mathbb{Q}$. Only if we remove parameter uncertainty on the jump parameters in the sense $(\boldsymbol{\lambda}_h^{\mathbb{Q}}, \boldsymbol{\nu}_h^{\mathbb{Q}}) = (\boldsymbol{\lambda}^{\mathbb{Q}}, \boldsymbol{\nu}^{\mathbb{Q}})$ do we recover the $\mathbb{Q}$ martingale measure property. In particular, for the Merton specification $(\boldsymbol{\lambda}^{\mathbb{Q}}, \boldsymbol{\nu}^{\mathbb{Q}}) = (\boldsymbol{\lambda}, \boldsymbol{\nu})$ the compensated random measures become $\mathbb{P}$ martingale mea-

theorem with $\boldsymbol{\phi}_t = \boldsymbol{\sigma}_h^{-1}(r\boldsymbol{\iota} - \boldsymbol{\mu}_h)$, or (ii) when $\boldsymbol{\sigma}_h = 0$ (there are *only* jumps) and $\boldsymbol{\nu}^{\mathbb{Q}} = \boldsymbol{\nu} = \boldsymbol{\delta}_1$ (the simple Poisson process case) in which case $\boldsymbol{\lambda}_h^{\mathbb{Q}} = r_t \boldsymbol{\iota} - \boldsymbol{\mu}_h$.

sures, which in turn implies $\mathbb{E}[P\&L_J] = 0$ insofar as the integrand is square integrable.

Conceptually, the takeaway message from the formulation (22) is that if V is convex in all of its components (a property it will inherit from the pay-off function under mild conditions) then $\forall j : \Delta V > \partial_{x_j} V \Delta X_j$ whence the integrand in $P\&L_J$ is positive. Thus, our hedge portfolio actually benefits from jumps in either direction of any of the underlying price process. Conversely, if we had shorted the option, the hedge profit would obviously take a hit in the event of a jump (in Talebian terms, holding a hedge portfolio with a short option position corresponds to 'picking pennies in front of a steam roller').† A vivid illustration of this point is provided in figure 2 for an option written on a single underlying.

4. Insights from empirics: on arbitrage and erraticism

Inspired by Wilmott's theoretical hedge experiment, we now look into the empirical performance of Δ-hedging strategies based on (I) forecasted implied volatilities and (II) forecasted actual (i.e. historical) volatilities. Specifically, we are interested in the properties of the accumulated P&L, insofar as we Δ-hedge, till expiry, a three-month call-option on the S&P500 index, initially purchased at-the-money. We investigate a totality of 36 such portfolios over disjoint intervals between July 2004 and July 2013. This involves market data on both the underlying index and on options. Daily data on the S&P500 index are readily and freely available. For option data, we combine a 2004–2009 data-set from a major commercial bank‡ with more recent prices from OptionMetrics obtained via the Wharton Financial Database.

Whilst ATM call option prices straightforwardly are obtained from the data-set, the (forecasted) implied and actual volatilities require a bit of manipulation. In case of the former, we define the daily implied volatility, over the lifetime of the portfolio, as the ATM implied volatility of corresponding tenor obtained at the portfolio purchasing date (the resulting volatility process is illustrated by the black curve in figure 3). In case of the latter, we require a suitable volatility model fitted to historical data in order to predict the 'actual' volatility process. Specifically, we define the daily actual volatility, over the lifetime of the portfolio, as the conditional expectation of a volatility model which has been fitted to market data from the previous portfolio period. In this context, we observe that models with lognormal volatility dynamics generally have more empirical support than, say, Heston's model.§ The Exponential General Autoregressive Conditional Heteroskedasticity model (EGARCH(1,1)) has proven particularly felicitous in the context of S&P 500 forecasting¶—a result we assume applies universally for each of the 36 portfolios investigated. Thus, we hold it to be the case that daily log returns, r_t, can be modelled as $r_t = \mu + \varepsilon_t$, where μ is the mean return, and ε_t has the interpretation of a hetereoskedastic error. In particular, ε_t is construed to be the product between a white noise process, $z_t \sim N(0, 1)$ and a daily standard deviation, σ_t, which obeys the relation

$$\log \sigma_t^2 = \alpha_0 + \alpha_1 \log \sigma_{t-1}^2 + \alpha_2 \left[\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}} \right] + \alpha_3 \frac{\varepsilon_{t-1}}{\sigma_{t-1}}, \tag{23}$$

where α_0, α_1, α_2 and α_3 are constants. The resulting volatility process is illustrated by the light grey curve in figure 3.

A few remarks on the estimated volatility processes are in order. First, we clearly see that volatility can change dramatically during the lifetime of a portfolio. We also see that implied volatility typically is higher than actual volatility. This oft-reported result can be explained theoretically by the stochastic volatility having a market price of risk attached, see for instance Henderson *et al*. (2005). Finally, there is a clear negative correlation between stock returns and volatility during the financial turmoil which followed the Lehman default in September 2008. All in all, reality unsurprisingly turns out to be a bit more complicated than the set-up in Wilmott's experiment. Still and all, do its main messages carry over? To test this, we perform a hedge experiment with the following design:

- For any given portfolio, we compute the daily implied volatilities $\{\sigma_t^{\text{imp}}\}_{t=1}^{63}$ and the daily actual volatilities $\{\sigma_t^{\text{act}}\}_{t=1}^{63}$ as outlined above. We assume there are 63 trading days over a three months period (labelled by $t = 1, 2, \ldots, 63$) and let S_t, r_t and q_t denote the time t value of the index, interest rate and dividend yield.
- For each of the two hedging strategies $x \in \{\sigma^{\text{imp}}, \sigma^{\text{act}}\}$ we do the following: If $\sigma_1^{\text{act}} < \sigma_1^{\text{imp}}$ we short the call ($\gamma = -1$); otherwise, we go along the call ($\gamma = +1$) in accordance with the remark made in the section on Wilmott's hedge experiment. Then, we set up the delta neutral portfolio $\Pi_1 = B_1 - \gamma \Delta_1^{\text{BS}}(x_1) S_1 + \gamma C_1^{\text{BS}}(\sigma_1^{\text{imp}})$ s.t. $\Pi_1 = 0$, where $\Delta_1^{\text{BS}}(x_1)$ is the well-known Black–Scholes delta.
- For $t = 2, 3, \ldots, 63$ we do the following: compute the time t value of the portfolio set up the previous day: $\tilde{\Pi}_t = B_{t-1} e^{r_{t-1}\Delta t} - \gamma \Delta_t^{\text{BS}}(x_t) S_t e^{q_{t-1}\Delta t} + \gamma C_t^{\text{BS}}(\sigma_t^{\text{imp}})$. The quantity $dP\&L_t = \tilde{\Pi}_t - \Pi_{t-1}$ defines the profit-&-loss accrued over the interval $[t-1, t]$. Next, we rebalance the portfolio such that it, once again, is delta-neutral, $\Pi_t = B_t - \gamma \Delta_t^{\text{BS}}(x_t) S_t + \gamma C_t^{\text{BS}}(\sigma_t^{\text{imp}})$, where B_t is chosen in accordance with the self-financing condition: $\tilde{\Pi}_t = \Pi_t$.
- Finally, at the option expiry, we compute the terminal P&L, as well as its lifetime quadratic variation, $\sum_{t=1}^{63} |dP\&L_t|^2/63$.

The 36 hedge error (or P&L) paths and the distributions of the quadratic variation of the two methods are shown in figure 4. Table 1 reports descriptive statistics and statistical tests of various hypotheses.

First, we note (top panels figure 4) that even though implied volatility typically is above actual volatility, this far from creates arbitrage. Hedge errors for the two methods readily become negative. A primary explanation for this is the randomness of volatility. Our Δ-hedged strategy only makes us

†Taleb (2010), p.19, 'Most traders were just "picking pennies in front of a steam roller," exposing themselves to the high impact rare event yet sleeping like babies, unaware of it.'
‡The bank shall remain nameless, but the data can be downloaded from http://www.math.ku.dk/~rolf/Svend/http://www.math.ku.dk/~rolf/Svend/.
§See Gatheral *et al*. (2014) and their references.
¶See Awartani and Corradi (2005).

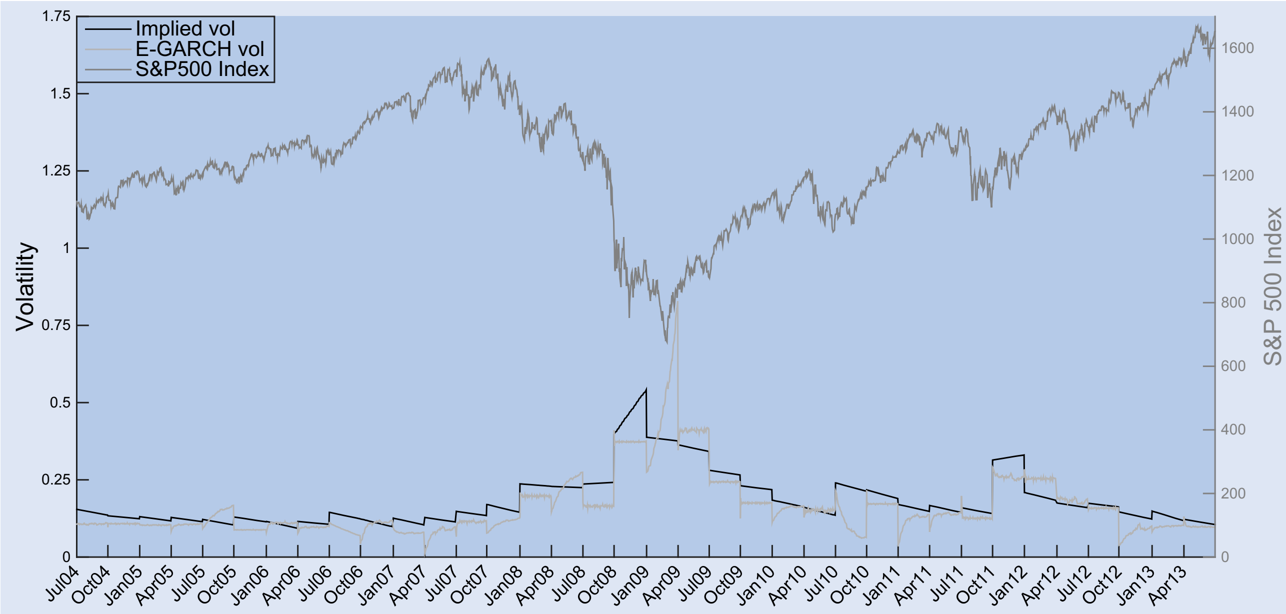

Figure 3. The top **grey curve** is the S&P500 Index plotted from July 2004 to July 2013 [units on right hand axis]. The tic-dates on the time axis have deliberately been chosen to match the purchasing dates $\{t_i\}_{i=1}^{36}$ of the 36 delta-hedged portfolios under investigation (each of which is of three months' duration). The **light grey curve** is the actual (stochastic) volatility estimated from a lognormal volatility model. Specifically, every time segment between purchasing dates $[t_i, t_{i+1})$ reflects a mean Monte Carlo simulated forecast based upon an EGARCH(1,1) fitted to market data from the previous time segment $[t_{i-1}, t_i)$. Finally, the **black curve** is the three-month ATM implied volatility. Specifically, every time segment between purchasing dates $[t_i, t_{i+1})$ is a static forecast based upon ATM implied volatility data from the purchasing date t_i. Both volatility curves have their units on the left hand axis.

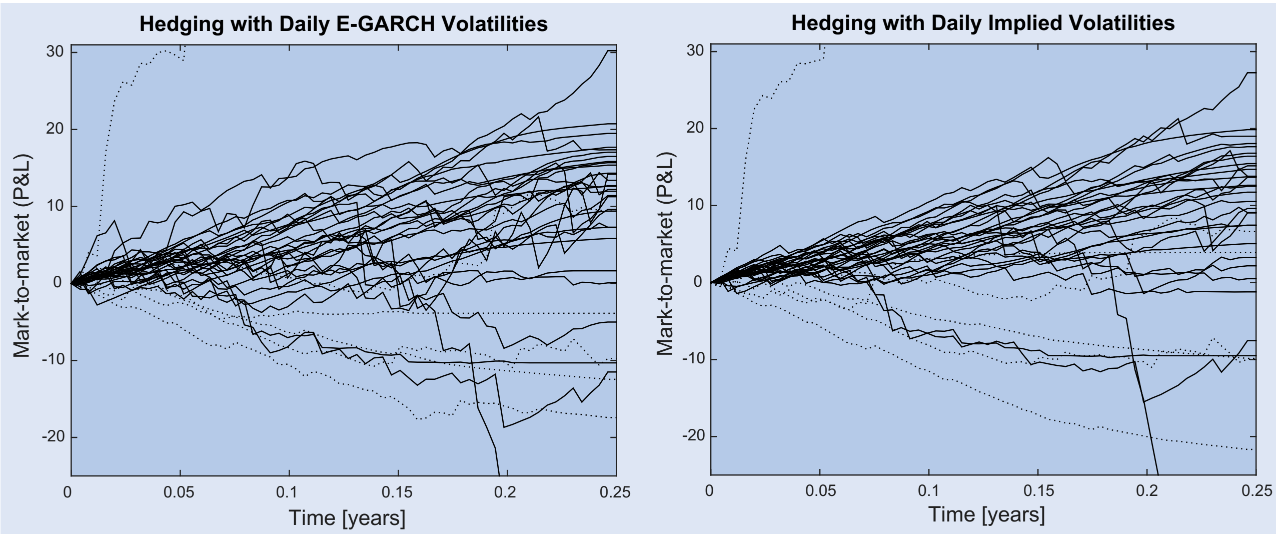

Figure 4. Panels (a) (actual) and (b) (implied) show the path-for-path hedge error behaviour for the 36 non-overlapping three-month hedges. Dotted paths correspond to cases where we initially take a long position in the option.

a profit if realized volatility ends up 'on the right side' of initial implied volatility. And that we don't know for sure until after the hedging period is over; we have to base our decisions on forecasts; initial forecasts even, for the fundamental theorem to apply. Notice though that the averages for both hedge errors are significantly positive. This shows that there is a risk premium that can be picked up, most often by selling options and Δ-hedging them. Because the hedge is not perfect, this compensation is anticipated. The question is, is it financially significant? In theory, the hedged portfolio has an initial cost of zero, so it is not obvious how to define a rate of return, but the initial option price would seem a reasonable (possibly conservative) benchmark for the collateral that would need to be posted on a hedged short call option position. From column three in table C1 the average option price is \$49.2. Comparing this to the means ($\sim$7.7; remember this is over a three-month horizon) and standard deviations ($\sim$15.5; ditto) of the hedge errors in table 1 shows that the gains are also significant in economic terms. Put differently, the crude calculation $(4 \cdot \frac{7.7}{49.2} - 0.02)/(\sqrt{4} \cdot \frac{15.5}{49.2})$ gives annualized Sharpe-ratios around 1.

If we look just at the terminal hedge errors, then the difference in riskiness (as measured by standard deviation) between hedging with actual and hedging with implied volatility is in no way statistically significant (the p-value for equality of

Table 1. Summary statistics and hypothesis tests for different hedge strategies.

Quantity	Mean (m)	Std. Dev. (sd)	Notes (hypotheses tests)
Hedge error, actual volatility	7.7	17.3	The mean hedge error is statistically greater than zero (p-value = 1%) when we hedge with the actual vol. forecast.
Hedge error, implied volatility	7.7	15.6	The mean hedge error is statistically greater than zero (p-value = 1%) when we hedge with the implied vol. forecast. We cannot reject the hypothesis that the standard deviations $sd_{act} = sd_{imp}$ are equal (p-value = 55%).
Quadratic var., actual volatility	1.2	2.1	
Quadratic var., implied volatility	0.81	2.0	The mean quadratic variation of the hedge error when we hedge with the actual vol. forecast is statistically less than the quadratic variation when we hedge with the implied vol. forecast (p-value = 1.4%).

variances is 55%). Also, the correlation between the terminal hedge error from the two approaches is 0.97. However, if we consider the quadratic variations as the measure of riskiness, then the picture changes. The average quadratic variation of the implied hedge error (0.81) is only two-thirds of the average quadratic variation of the actual hedge error (1.2) (a paired t-test for equality yields a p-value of 1.4%).

All in all this shows that volatility arbitrage is difficult, but the following insight from Wilmott's experiment stands: if you are in the business of hedging, then the use of implied volatility should make you sleep better at night.

5. Conclusion

In the world of finance, no issue is more pressing than that of hedging our risks, yet remarkably little attention has been paid to the risk brought about by the possibility that our models might be wrong. To remedy this deplorable situation, we have in this paper derived a meta-theorem that quantifies the P&L of a Δ-hedged portfolio with an erroneous volatility specification. *Meta-* to the extent that one of the constituent parameters (the real volatility) is transcendental; yet, also a theorem with some very concrete 'real world' corollaries. For instance, a specific case was investigated in which the implied volatility gives rise to smooth (i.e. $\mathcal{O}(\mathrm{d}t)$) P&L-paths, whilst any other hedge volatility yields erratic (i.e. $\mathcal{O}(\mathrm{d}W_t)$) P&L paths. In a somewhat quirkier context, the Dupire–Gyöngy–Derman–Kani formula for volatility surface calibration was shown to be a corollary.

Whilst the theorem proved in section 2 is more general than the versions typically found in the literature, it does not go *far enough*. Extensive empirical support has been added to the case of discontinuities in the stock price process: thus, in the Gospel of the Jump we extended the Fundamental Theorem to include compound Poisson processes, which came with the revelation that jumps unambiguously hurt you when you try to hedge short put and call option positions.

One of the most conspicuous implications of the Fundamental Theorem is undoubtedly the apparent ease with which arbitrage can be made: e.g. in the constant parameter framework of Wilmott's experiment, a free lunch is guaranteed insofar as we can establish $\max\{\sigma_r, \sigma_i\}$ (in case of the former, we go long on the option—otherwise, we short it). Studying this strategy empirically, we find that the mean P&L indeed is in the positive; nonetheless, qua a significant dispersion the profit readily turns negative: the statistical arbitrage accordingly relies on us being willing to take some significant hits along the way. Indeed, this is without even factoring in the non-negligible role of transaction costs. On the other hand, there is strong evidence that hedging at the implied volatility does yield smoother P&L paths.

One final remark: this paper can be seen as an exhaustive exposition of which volatility to use when delta hedging. To keep the length manageable and the presentation self-contained we have ignored an aspect that is of both theoretical and practical importance. It can be posed thus: Which delta should I use? In models where the underlying and the volatility are correlated a strong case can be made for using the so-called risk-minimizing delta, which is in broad terms is the usual delta plus the (underlying, volatility)-correlation times the volatility of volatility times the Vega of the target option, see for instance Poulsen *et al*. (2009), Andreasen (2013), or Hull and White (2016). We leave the connection of this theory to the Fundamental Theorem of Derivative Trading, theoretically as well empirically, to future research.

Acknowledgements

All authors gratefully acknowledge support from the Danish Strategic Research Council, Program Committee for Strategic Growth Technologies, via the research center HIPERFIT: Functional High Performance Computing for Financial Information Technology (hiperfit.dk) under contract number [10-092299].

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Danish Strategic Research Council, Program Committee for Strategic Growth Technologies, via the research center HIPERFIT: Functional High Performance Computing for Financial Information Technology (hiperfit.dk) under contract number [10-092299].

References

Ahmad, R. and Wilmott, P., Which free lunch would you like today, Sir?: Delta hedging, volatility arbitrage and optimal portfolios. *Wilmott Mag*., November, 2005, 64–79.
Andreasen, J., Derivatives – The view from the trenches. Presentation at the University of Copenhagen, 2003. Available online at: http://www.math.ku.dk/~rolf/jandreasen.pdf.
Andreasen, J., Option anatomy. Inagural lectures at Department of Mathematical Sciences, University of Copenhagen, 2013. Available online at: http://www.math.ku.dk/~rolf/Andreasen_Inagural.pdf.
Andreasen, J. and Huge, B., Volatility interpolation. *Risk Mag*., March, 2011, 76–79.
Awartani, B. and Corradi, V., Predicting the volatility of the S&P-500 stock index via GARCH models: The role of asymmetries. *Int. J. Forecasting*, 2005, **21**, 167–183.
Björk, T., *Arbitrage Theory in Continuous Time*, 3rd ed., 2009 (Oxford University Press).
Black, F. and Scholes, M., The pricing of options and corporate liabilities. *J. Polit. Econ*., 1973, **81**, 637–654.
Carr, P., FAQ's in option pricing theory. Working Paper, 2002. Available online at: http://www.math.nyu.edu/research/carrp/papers/pdf/faq2.pdf.
Carr, P. and Jarrow, R., The stop-loss start-gain paradox and option valuation: A new decomposition into intrinsic and time value. *Rev. Financial Stud*., 1990, **3**, 469–492.
Cont, R. and Tankov, P., *Financial Modelling with Jump Processes*, Financial Mathematics Series, 2004 (Chapman & Hall/CRC).
Davis, M.H.A., Black Scholes Formula. *Encycl. Quant. Finance*, 2010, **1**, 201–207.
Derman, E., Lectures on the smile. Columbia University, 2008. Available online at: http://www.emanuelderman.com/writing/entry/lectures-on-the-smile-2008.
Derman, E. and Kani, I., Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility. *Int. J. Theor. Appl. Finance*, 1998, 7–22.
Dupire, B., Stochastic volatility modelling. Presentation, Nice, 2003. Available online at: http://www-math.unice.fr/~diener/nicemathfi03/Bruno_Dupire_vol_presentation_1302c.ppt.
Dupire, B., Pricing with a smile. *Risk Mag*., January, 1994, 18–20.
El Karoui, N., Jeanblanc-Picque, M. and Shreve, S., Robustness of the Black and Scholes Formula. *Math. Finance*, 1998, **8**, 93–126.
Gatheral, J., Jaisson, T. and Rosenbaum, M., Volatility is rough. Working Paper, 2014. Available online at: http://arxiv.org/abs/1410.3394.
Gibson, R., Lhabitant, F., Pistre, N. and Talay, D., Interest rate model risk: What are we talking about? *J. Risk*, 1999, **1**, 37–62.
Guyon, J. and Henry-Labordère, P., Being particular about calibration. *Risk Mag*., January, 2012, 92–97.
Gyöngy, I., Mimicking the one-dimensional marginal distributions of processes having an Ito differential. *Probab. Theory Relat. Fields*, 1986, **71**(4), 501–516.
Haug, E. and Taleb, N., Option traders use (very) sophisticated heuristics, never the Black–Scholes–Merton Formula. *J. Econ. Behav. Organiz.* 2011, **77**, 97–106.
Henderson, V., Hobson, D., Howison, S. and Kluge, T., A comparison of option prices under different pricing measures in a stochastic volatility model with correlation. *Rev. Derivatives Research*, 2005, **8**, 5–25.
Henrard, M., Parameter risk in the Black and Scholes model. Working Paper, 2001. Available online at: http://128.118.178.162/eps/ri/papers/0310/0310002.pdf.
Hull, J. and White, A., Optimal delta hedging for options. SSRN Working Paper, 2016. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2658343.
Merton, R., Option pricing when underlying stock returns are discontinuous. *J. Financial Econ*., 1976, **3**, 125–144.
Mahayni, A., Schlögl, E. and Schlögl, L., Robustness of Gaussian hedges and the hedging of fixed income derivatives. Working Paper, 2001. Available online at: http://www.qfrc.uts.edu.au/research/research_papers/rp19.pdf.
Rasmussen, N., Hedging with a misspecified model. Working Paper at SSRN, 2001. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=260754.
Romano, M. and Touzi, N., Contingent claims and market completeness in a stochastic volatility model. *Math. Finance*, 1997, **7**, 399–412.
Runggaldier, W., Jump diffusion models. In *Handbook of Heavy Tailed Distributions in Finance*, 2003 (Elsevier Science B.V.).
Poulsen, R., Schenk-Hoppé, K.R. and Ewald, C.-O., Risk minimization in stochastic volatility models: Model risk and empirical performance. *Quant. Finance*, 2009, **9**, 693–704.
Privault, N., Notes on stochastic finance. Lecture Notes FE6516, 2013. Available online at: http://www.ntu.edu.sg/home/nprivault/indext.html.
Shreve, S., *Stochastic Calculus for Finance II*, 2008 (Springer).
Taleb, N., *The Black Swan – The Impact of the Highly Improbable*, Revised Edition, 2010 (Penguin Books).
Wilmott, P., *Paul Wilmott Introduces Quantitative Finance*, 2nd ed., 2007 (John Wiley & Sons Ltd).
Wittgenstein, L., *Tractatus Logico-Philosophicus*, 1922.

Appendix 1. Multi-dimensional jumps

In this section, we establish Girsanov's Theorem and a pricing PDE for multi-dimensional jump-diffusion models. The equivalent results for 1-dimensional models are ubiquitous—see for instance Cont (2004), Privault (2013) or Runggaldier (2003).

A.1. The Radon–Nikodym derivative

THEOREM 3 A generalized Girsanov Theorem for jump-diffusion processes *Let* $\{\boldsymbol{W}_t\}_{t\in[0,T]}$ *be a* d_w*-dimensional vector of independent Wiener processes on the filtered probability space* $(\Omega, \mathscr{F}, \mathbb{P}, \{\mathscr{F}_t\}_{t\in[0,T]})$. *One the same space, let* $\{\boldsymbol{Y}_t\}_{t\in[0,T]}$ *be a* d_y*-dimensional vector of independent compound Poisson processes, the* i*th component of which is*

$$Y_t^i = \sum_{k=1}^{N_t^i} Z_k^i,$$

where $N_t^i \sim Pois(\lambda_i t)$, $\lambda_i > 0$ *is an intensity parameter, and* $\{Z_k^i\}_{k\in\mathbb{N}}$ *is a sequence of i.i.d. random variables with jump distribution* $d\nu_i(z)$. *Finally, let* $\{\boldsymbol{\phi}_t\}_{t\in[0,T]}$ *be a* d_w*-dimensional Girsanov kernel (some bounded, adapted process), then the processes*

$$\left\{\boldsymbol{W}_t^{\mathbb{Q}} := \boldsymbol{W}_t - \int_0^t \boldsymbol{\phi}_s \mathrm{d}s\right\}_{t\in[0,T]}, \quad \text{and}$$
$$\left\{\tilde{\boldsymbol{Y}}_t^{\mathbb{Q}} := \boldsymbol{Y}_t - \boldsymbol{\lambda}^{\mathbb{Q}} \circ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1]t\right\}_{t\in[0,T]}$$

are martingales under the probability measure $\mathbb{Q}$ *defined as*

$$\mathrm{d}\mathbb{Q}_{\boldsymbol{\phi},\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}} = \mathscr{E}(\boldsymbol{\phi} \star \boldsymbol{W})(T) \times e^{-\sum_{i=1}^{d_y}(\lambda_i^{\mathbb{Q}}-\lambda_i)T} \prod_{i=1}^{d_y}\prod_{k=1}^{N_t^i} \frac{\lambda_i^{\mathbb{Q}} \mathrm{d}\nu_i^{\mathbb{Q}}(Z_k^i)}{\lambda_i \mathrm{d}\nu_i(Z_k^i)} \mathrm{d}\mathbb{P}$$

where $\mathscr{E}(\boldsymbol{\phi} \star \boldsymbol{W})(T) = e^{\int_0^T \boldsymbol{\phi}_s \bullet \mathrm{d}\boldsymbol{W}_s - \frac{1}{2}\int_0^T |\boldsymbol{\phi}_s|^2 \mathrm{d}s}$ *is the Doleans exponential with respect to* $\boldsymbol{W}_t$, *and we have defined* $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{d_y})^\intercal$, *and* $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_{d_y})^\intercal$.

Proof The diffusion part is well known from Girsanov's theorem and will not be treated here. Instead we will show that for any bounded measurable function $f : \mathbb{R}^{d_y} \mapsto \mathbb{R}$, the following equivalence obtains

$$\mathbb{E}_{\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}[f(\boldsymbol{Y}_T)] = \mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}\left[f(\boldsymbol{Y}_T)\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right],$$

for the defined measure $\mathbb{Q}$. To this end, define $\boldsymbol{Y}_t^h$ as the vector $\boldsymbol{Y}_t$ with the upper limit of the summation, N_t^i, replaced by some fixed number $h^i \in \mathbb{N}_0$ for all i. Then the RHS can be written as

$$
\begin{aligned}
&e^{-\sum_{i=1}^{d_y}(\lambda_i^{\mathbb{Q}}-\lambda_i)T}\mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}\left[f(\boldsymbol{Y}_t)\prod_{i=1}^{d_y}\prod_{k=1}^{N_t^i}\frac{\lambda_i^{\mathbb{Q}}\mathrm{d}\nu_i^{\mathbb{Q}}(Z_k^i)}{\lambda_i\mathrm{d}\nu_i(Z_k^i)}\right]\\
&= e^{-\sum_{i=1}^{d_y}(\lambda_i^{\mathbb{Q}}-\lambda_i)T}\sum_{h^1=0}^{\infty}\cdots\sum_{h^{d_y}=0}^{\infty}\mathbb{P}\left(\bigcap_{i=1}^{d_y}\{N_T^i=h^i\}\right)\\
&\quad\times\mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}\left[f(\boldsymbol{Y}_t)\prod_{i=1}^{d_y}\prod_{k=1}^{N_t^i}\frac{\lambda_i^{\mathbb{Q}}\mathrm{d}\nu_i^{\mathbb{Q}}(Z_k^i)}{\lambda_i\mathrm{d}\nu_i(Z_k^i)}\middle|\bigcap_{i=1}^{d_y}\{N_T^i=h^i\}\right]\\
&= e^{-\sum_{i=1}^{d_y}(\lambda_i^{\mathbb{Q}}-\lambda_i)T}\sum_{h^1=0}^{\infty}\cdots\sum_{h^{d_y}=0}^{\infty}\prod_{i=1}^{d_y}\mathbb{P}\left(N_T^i=h^i\right)\\
&\quad\times\mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}\left[f(\boldsymbol{Y}_t^h)\prod_{i=1}^{d_y}\prod_{k=1}^{h^i}\frac{\lambda_i^{\mathbb{Q}}\mathrm{d}\nu_i^{\mathbb{Q}}(Z_k^i)}{\lambda_i\mathrm{d}\nu_i(Z_k^i)}\right]\\
&= e^{-\sum_{i=1}^{d_y}(\lambda_i^{\mathbb{Q}}-\lambda_i)T}\sum_{h^1=0}^{\infty}\cdots\sum_{h^{d_y}=0}^{\infty}\prod_{i=1}^{d_y}\frac{e^{-\lambda_iT}(\lambda_iT)^{h_i}}{k^i!}\\
&\quad\times\mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}\left[f(\boldsymbol{Y}_t^h)\prod_{i=1}^{d_y}\prod_{k=1}^{h^i}\frac{\lambda_i^{\mathbb{Q}}\mathrm{d}\nu_i^{\mathbb{Q}}(Z_k^i)}{\lambda_i\mathrm{d}\nu_i(Z_k^i)}\right]\\
&= \sum_{h^1=0}^{\infty}\cdots\sum_{h^n=0}^{\infty}\prod_{i=1}^{d_y}\frac{e^{-\lambda_i^{\mathbb{Q}}T}(\lambda_i^{\mathbb{Q}}T)^{h_i}}{k^i!}\\
&\quad\times\mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}\left[f(\boldsymbol{Y}_t^h)\prod_{i=1}^{d_y}\prod_{k=1}^{h^i}\frac{\mathrm{d}\nu_i^{\mathbb{Q}}(Z_k^i)}{\mathrm{d}\nu_i(Z_k^i)}\right]\\
&= \sum_{h^1=0}^{\infty}\cdots\sum_{h^{d_y}=0}^{\infty}\prod_{i=1}^{d_y}\frac{e^{-\lambda_i^{\mathbb{Q}}T}(\lambda_i^{\mathbb{Q}}T)^{h_i}}{k^i!}\int_{\mathbb{R}^{h^1}}\cdots\int_{\mathbb{R}^{h^{d_y}}}\\
&\quad\times f(\boldsymbol{y}_t^h)\prod_{i=1}^{d_y}\prod_{k=1}^{h^i}\frac{\mathrm{d}\nu_i^{\mathbb{Q}}(z_k^i)}{\mathrm{d}\nu_i(z_k^i)}\mathrm{d}\nu_i(z_k^i)\\
&= \sum_{h^1=0}^{\infty}\cdots\sum_{h^{d_y}=0}^{\infty}\prod_{i=1}^{d_y}\frac{e^{-\lambda_i^{\mathbb{Q}}T}(\lambda_i^{\mathbb{Q}}T)^{h_i}}{k^i!}\int_{\mathbb{R}^{h^1}}\cdots\int_{\mathbb{R}^{h^{d_y}}}\\
&\quad\times f(\boldsymbol{y}_t^h)\prod_{i=1}^{d_y}\prod_{k=1}^{h^i}\mathrm{d}\nu_i^{\mathbb{Q}}(z_k^i)\\
&= \sum_{h^1=0}^{\infty}\cdots\sum_{h^n=0}^{\infty}\prod_{i=1}^{d_y}\mathbb{Q}\left(N_T^i=h^i\right)\mathbb{E}_{\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}\left[f(\boldsymbol{Y}_t^h)\right]\\
&= \sum_{h^1=0}^{\infty}\cdots\sum_{h^{d_y}=0}^{\infty}\mathbb{Q}\left(\bigcap_{i=1}^{d_y}\{N_T^i=h^i\}\right)\\
&\quad\times\mathbb{E}_{\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}\left[f(\boldsymbol{Y}_t)\middle|\bigcap_{i=1}^{d_y}\{N_T^i=h^i\}\right]
\end{aligned}
$$

which by the law of total probability corresponds to the LHS, $\mathbb{E}_{\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}[f(\boldsymbol{Y}_T)]$ as desired. To show independence of increments under $\mathbb{Q}$, let $\xi_s = \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}(s)$, and let f and g be two bounded measurable functions. Suppose $s < t \leq T$ then

$$
\begin{aligned}
\mathbb{E}_{\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}[f(\boldsymbol{Y}_s)g(\boldsymbol{Y}_t-\boldsymbol{Y}_s)] &= \mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}[f(\boldsymbol{Y}_s)g(\boldsymbol{Y}_t-\boldsymbol{Y}_s)\xi_t]\\
&= \mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}[f(\boldsymbol{Y}_s)\xi_s]\mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}[g(\boldsymbol{Y}_t-\boldsymbol{Y}_s)\xi_t/\xi_s]\\
&= \mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}[f(\boldsymbol{Y}_s)\xi_s]\mathbb{E}_{\boldsymbol{\lambda},\boldsymbol{\nu}}[g(\boldsymbol{Y}_t-\boldsymbol{Y}_s)\xi_t]\\
&= \mathbb{E}_{\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}[f(\boldsymbol{Y}_s)]\cdot\mathbb{E}_{\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}[g(\boldsymbol{Y}_t-\boldsymbol{Y}_s)].
\end{aligned}
$$

□

A.2. Measure changes in jump-diffusion dynamics

To appreciate the gravity of this result, consider the jump-diffusion dynamics of the n-dimensional stock price process

$$\mathrm{d}\boldsymbol{X}_t = \mathbf{D}_{\boldsymbol{X}_{t-}}[\boldsymbol{\mu}(t,\boldsymbol{X}_t)\mathrm{d}t + \boldsymbol{\sigma}(t,\boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t] + \mathbf{D}_{\boldsymbol{X}_t}\boldsymbol{\theta}(t,\boldsymbol{X}_{t-})\mathrm{d}\boldsymbol{Y}_t \tag{A1}$$

where $\mathbf{D}_{\boldsymbol{X}_{t-}} = \mathrm{diag}(X_{1,t-}, X_{2,t-}, \dots, X_{n,t-})$, $\boldsymbol{\mu} : [0,\infty)\times\mathbb{R}^n \mapsto \mathbb{R}^n$, $\boldsymbol{\sigma} : [0,\infty)\times\mathbb{R}^n \mapsto \mathbb{R}^{n\times d_w}$ and $\boldsymbol{\theta} : [0,\infty)\times\mathbb{R}^n \mapsto \mathbb{R}^{n\times d_y}$. For the purposes of no-arbitrage pricing we want to determine the measure $\mathbb{Q}$ such that the discounted process

$$\boldsymbol{X}_t^* := e^{-\int_0^t r_u \mathrm{d}u}\boldsymbol{X}_t,$$

is a martingale. From Ito's lemma, it can readily be deduced that

$$
\begin{aligned}
\mathrm{d}\boldsymbol{X}_t^* = \mathbf{D}_{\boldsymbol{X}_t}^*&[(\boldsymbol{\mu}(t,\boldsymbol{X}_t) - r_t\boldsymbol{\iota} + \boldsymbol{\theta}(t,\boldsymbol{X}_t)\boldsymbol{\lambda}\circ\mathbb{E}_{\boldsymbol{\nu}}[\boldsymbol{Z}_1])\mathrm{d}t\\
&+ \boldsymbol{\sigma}(t,\boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t]\\
&+ \mathbf{D}_{\boldsymbol{X}_{t-}}^*\boldsymbol{\theta}(t,\boldsymbol{X}_{t-})(\mathrm{d}\boldsymbol{Y}_t - \boldsymbol{\lambda}\circ\mathbb{E}_{\boldsymbol{\nu}}[\boldsymbol{Z}_1]\mathrm{d}t),
\end{aligned}
$$

where we have added and subtracted $\boldsymbol{\theta}(t,\boldsymbol{X}_t)\boldsymbol{\lambda}\circ\mathbb{E}_{\boldsymbol{\nu}}[\boldsymbol{Z}_1]\mathrm{d}t$, where $\boldsymbol{Z}_1 := (Z_1^1, Z_1^2, \dots, Z_1^{d_y})^{\intercal}$. Using the measure transformation in the theorem above this transforms to

$$
\begin{aligned}
\mathrm{d}\boldsymbol{X}_t^* = \mathbf{D}_{\boldsymbol{X}_t}^*&[(\boldsymbol{\mu}(t,\boldsymbol{X}_t) - r_t\boldsymbol{\iota} + \boldsymbol{\theta}(t,\boldsymbol{X}_t)\boldsymbol{\lambda}^{\mathbb{Q}}\circ\mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1]\\
&+ \boldsymbol{\sigma}(t,\boldsymbol{X}_t)\boldsymbol{\phi}_t)\mathrm{d}t + \boldsymbol{\sigma}(t,\boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t^{\mathbb{Q}}]\\
&+ \mathbf{D}_{\boldsymbol{X}_{t-}}^*\boldsymbol{\theta}(t,\boldsymbol{X}_{t-})\mathrm{d}\tilde{\boldsymbol{Y}}_t^{\mathbb{Q}}.
\end{aligned}
$$

Hence, $\boldsymbol{X}_t^*$ is a $\mathbb{Q}$-martingale iff the tuple $\langle\boldsymbol{\phi}_t, \boldsymbol{\lambda}^{\mathbb{Q}}, \boldsymbol{\nu}^{\mathbb{Q}}, \mathbb{Q}\rangle$ is chosen such that

$$\boldsymbol{\mu}(t,\boldsymbol{X}_t) + \boldsymbol{\theta}(t,\boldsymbol{X}_t)\boldsymbol{\lambda}^{\mathbb{Q}}\circ\mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1] + \boldsymbol{\sigma}(t,\boldsymbol{X}_t)\boldsymbol{\phi}_t = r_t\boldsymbol{\iota}, \tag{A2}$$

almost everywhere. Needless to say, the infinite number of tuples which satisfies the no-arbitrage condition (A2) ruins our chances of establishing *unique* prices for financial derivatives depending on the underlying jump diffusion dynamics. This is obvious by recalling that the discounted price process of V also should be a martingale, whence:

$$V(t,\boldsymbol{X}_t) = \mathbb{E}_{t,\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\nu}^{\mathbb{Q}}}\left[e^{-\int_t^T r_u\mathrm{d}u}g(\boldsymbol{X}_T)\right], \tag{A3}$$

where $g(\boldsymbol{X}_T)$ is the terminal pay-off.

Appendix 2. PIDE methods

THEOREM 4 The Partial Integro-Differential Pricing Equation (PIDE) *Consider a jump diffusion dynamics of the form* (A1)*, and let* $V_t = V(t,\boldsymbol{X}_t)$ *be a derivative the value of which is contingent upon it. Let* $\langle\boldsymbol{\phi}_t, \boldsymbol{\lambda}^{\mathbb{Q}}, \boldsymbol{\nu}^{\mathbb{Q}}, \mathbb{Q}\rangle$ *be a tuple such that the no-arbitrage condition* (A2) *is satisfied. Then*

$$
\begin{aligned}
\mathrm{d}\boldsymbol{X}_t = \mathbf{D}_{\boldsymbol{X}_t}&[r_t\boldsymbol{\iota}\mathrm{d}t + \boldsymbol{\sigma}(t,\boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t^{\mathbb{Q}}]\\
&+ \mathbf{D}_{\boldsymbol{X}_{t-}}\boldsymbol{\theta}(t,\boldsymbol{X}_{t-})[\mathrm{d}\boldsymbol{Y}_t - \boldsymbol{\lambda}^{\mathbb{Q}}\circ\mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1]\mathrm{d}t]],
\end{aligned}
$$

and $V_t = V(t,\boldsymbol{x})$ *satisfies the PIDE*

$$
\begin{aligned}
r_tV_t &= \partial_tV_t + r_t\boldsymbol{x}\bullet\nabla_{\boldsymbol{x}}V_t + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}^{\intercal}(t,\boldsymbol{x})\mathbf{D}_{\boldsymbol{x}}\nabla_{\boldsymbol{x}\boldsymbol{x}}^2V_t\mathbf{D}_{\boldsymbol{x}}\boldsymbol{\sigma}(t,\boldsymbol{x})]\\
&\quad+ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}\left[\sum_{i=1}^{d_y}\lambda_i^{\mathbb{Q}}\{V(t,\boldsymbol{x}\circ(\boldsymbol{\iota}+\boldsymbol{\theta}_{:,i}(t,\boldsymbol{x})Z_1^i)) - V(t,\boldsymbol{x})\}\right.\\
&\quad\left.-(\mathbf{D}_{\boldsymbol{x}}\boldsymbol{\theta}(t,\boldsymbol{x})\boldsymbol{\lambda}^{\mathbb{Q}}\circ\boldsymbol{Z}_1)\bullet\nabla_{\boldsymbol{x}}V(t,\boldsymbol{x})\right],
\end{aligned}
$$

where $\nabla_{\boldsymbol{x}}$ is the gradient operator, $\nabla^2_{\boldsymbol{xx}}$ is the Hessian operator, $\boldsymbol{\iota} := (1, 1, \ldots, 1)^\top \in \mathbb{R}^n$, and $\boldsymbol{\theta}_{:,i}(t, \boldsymbol{x})$ denotes the ith column of the matrix $\boldsymbol{\theta}(t, \boldsymbol{x})$. Particularly, when $n = d_y$ and $\boldsymbol{\theta} = \mathbb{I}$ is the identity matrix then

$$\begin{aligned} r_t V_t = {} & \partial_t V_t + r_t \boldsymbol{x} \bullet \nabla_{\boldsymbol{x}} V_t + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}^\top(t, \boldsymbol{x})\mathbf{D}_{\boldsymbol{x}} \nabla^2_{\boldsymbol{xx}} V_t \mathbf{D}_{\boldsymbol{x}} \boldsymbol{\sigma}(t, \boldsymbol{x})] \\ & + \sum_{i=1}^{n} \lambda_i^{\mathbb{Q}} \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[V(t, \boldsymbol{x} \circ (\boldsymbol{\iota} + \hat{\mathbf{e}}_i Z_1^i)) \\ & - V(t, \boldsymbol{x}) - x_i Z_1^i \partial_{x_i} V(t, \boldsymbol{x})], \end{aligned}$$

where $\hat{\mathbf{e}}_i$ is a unit vector in the ith direction.

Proof Suppose a jump occurs in the ith component of the compound Poisson process $\boldsymbol{Y}_t$: $Y_t^i = Y_{t-}^i + Z_t^i$. From the governing dynamics (A1), this means that the stock price process jumps by $\Delta \boldsymbol{X}_t = \boldsymbol{X}_{t-} \circ (\boldsymbol{\iota} + \boldsymbol{\theta}_{:,i}(t, \boldsymbol{X}_{t-}) Z_t^i)$. Defining the continuous SDE

$$\begin{aligned} \mathrm{d}\boldsymbol{X}_t^{\mathrm{cont.}} := {} & \mathbf{D}_{\boldsymbol{X}_t}[r_t \boldsymbol{\iota} \mathrm{d}t + \boldsymbol{\sigma}(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t^{\mathbb{Q}}] \\ & - \mathbf{D}_{\boldsymbol{X}_{t-}} \boldsymbol{\theta}(t, \boldsymbol{X}_{t-}) \boldsymbol{\lambda}^{\mathbb{Q}} \circ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1]\mathrm{d}t], \end{aligned}$$

we find by Ito's lemma,

$$\begin{aligned} & \mathrm{d}V(t, \boldsymbol{X}_t) \\ & = \partial_t V(t, \boldsymbol{X}_t)\mathrm{d}t + \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet \mathrm{d}\boldsymbol{X}_t^{\mathrm{cont.}} \\ & \quad + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}^\top(t, \boldsymbol{X}_t)\mathbf{D}_{\boldsymbol{X}_t} \nabla^2_{\boldsymbol{xx}} V(t, \boldsymbol{X}_t) \mathbf{D}_{\boldsymbol{X}_t} \boldsymbol{\sigma}(t, \boldsymbol{X}_t)]\mathrm{d}t \\ & \quad + \sum_{i=1}^{d_y} (V(t, \boldsymbol{X}_{t-} + \Delta \boldsymbol{X}_t) - V(t, \boldsymbol{X}_{t-}))\mathrm{d}N_t^i \\ & = \partial_t V(t, \boldsymbol{X}_t)\mathrm{d}t + \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet (\mathbf{D}_{\boldsymbol{X}_t} r_t \boldsymbol{\iota})\mathrm{d}t \\ & \quad + \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet (\mathbf{D}_{\boldsymbol{X}_t} \boldsymbol{\sigma}(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t^{\mathbb{Q}}) \\ & \quad - \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet (\mathbf{D}_{\boldsymbol{X}_{t-}} \boldsymbol{\theta}(t, \boldsymbol{X}_{t-}) \boldsymbol{\lambda}^{\mathbb{Q}} \circ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1])\mathrm{d}t \\ & \quad + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}^\top(t, \boldsymbol{X}_t)\mathbf{D}_{\boldsymbol{X}_t} \nabla^2_{\boldsymbol{xx}} V(t, \boldsymbol{X}_t) \mathbf{D}_{\boldsymbol{X}_t} \boldsymbol{\sigma}(t, \boldsymbol{X}_t)]\mathrm{d}t \\ & \quad + \sum_{i=1}^{d_y} (V(t, \boldsymbol{X}_{t-} \circ (\boldsymbol{\iota} + \boldsymbol{\theta}_{:,i}(t, \boldsymbol{X}_{t-}) Z^i_{N_t^i})) - V(t, \boldsymbol{X}_{t-}))\mathrm{d}N_t^i \\ & = \partial_t V(t, \boldsymbol{X}_t)\mathrm{d}t + \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet (\mathbf{D}_{\boldsymbol{X}_t} r_t \boldsymbol{\iota})\mathrm{d}t \\ & \quad - \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet (\mathbf{D}_{\boldsymbol{X}_{t-}} \boldsymbol{\theta}(t, \boldsymbol{X}_{t-}) \boldsymbol{\lambda}^{\mathbb{Q}} \circ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1])\mathrm{d}t \\ & \quad + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}^\top(t, \boldsymbol{X}_t)\mathbf{D}_{\boldsymbol{X}_t} \nabla^2_{\boldsymbol{xx}} V(t, \boldsymbol{X}_t) \mathbf{D}_{\boldsymbol{X}_t} \boldsymbol{\sigma}(t, \boldsymbol{X}_t)]\mathrm{d}t \\ & \quad + \sum_{i=1}^{d_y} \lambda_i^{\mathbb{Q}} \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}\big[(V(t, \boldsymbol{x} \circ (\boldsymbol{\iota} \\ & \qquad + \boldsymbol{\theta}_{:,i}(t, \boldsymbol{x}) Z_1^i)) - V(t, \boldsymbol{x}))\big]_{\boldsymbol{x} = \boldsymbol{X}_{t-}} \mathrm{d}t \\ & \quad + \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet (\mathbf{D}_{\boldsymbol{X}_t} \boldsymbol{\sigma}(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t^{\mathbb{Q}}) \\ & \quad + \sum_{i=1}^{d_y} \big\{(V(t, \boldsymbol{X}_{t-} \circ (\boldsymbol{\iota} + \boldsymbol{\theta}_{:,i}(t, \boldsymbol{X}_{t-}) Z^i_{N_t^i})) - V(t, \boldsymbol{X}_{t-}))\mathrm{d}N_t^i \\ & \quad - \lambda_i^{\mathbb{Q}} \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}\big[(V(t, \boldsymbol{x} \circ (\boldsymbol{\iota} + \boldsymbol{\theta}_{:,i}(t, \boldsymbol{x}) Z_1^i)) - V(t, \boldsymbol{x}))\big]_{\boldsymbol{x} = \boldsymbol{X}_{t-}} \mathrm{d}t\big\}. \end{aligned}$$

Under $\mathbb{Q}$, the expectations of the diffusion term and the *compensated* jump terms (the last three lines) vanish. Furthermore, since

$$V_t^* := e^{-\int_0^t r_u \mathrm{d}u} V(t, \boldsymbol{X}_t),$$

is a $\mathbb{Q}$ martingale; $\mathrm{d}V_t^*$ should be driftless. These facts jointly imply that

$$\begin{aligned} & - r_t V(t, \boldsymbol{X}_t) + \partial_t V(t, \boldsymbol{X}_t) + r_t \mathbf{D}_{\boldsymbol{X}_t} \boldsymbol{\iota} \bullet \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \\ & - \nabla_{\boldsymbol{x}} V(t, \boldsymbol{X}_t) \bullet (\mathbf{D}_{\boldsymbol{X}_{t-}} \boldsymbol{\theta}(t, \boldsymbol{X}_{t-}) \boldsymbol{\lambda}^{\mathbb{Q}} \circ \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}[\boldsymbol{Z}_1]) \\ & + \tfrac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}^\top(t, \boldsymbol{X}_t)\mathbf{D}_{\boldsymbol{X}_t} \nabla^2_{\boldsymbol{xx}} V(t, \boldsymbol{X}_t) \mathbf{D}_{\boldsymbol{X}_t} \boldsymbol{\sigma}(t, \boldsymbol{X}_t)] \\ & + \sum_{i=1}^{d_y} \lambda_i^{\mathbb{Q}} \mathbb{E}_{\boldsymbol{\nu}^{\mathbb{Q}}}\big[(V(t, \boldsymbol{x} \circ (\boldsymbol{\iota} + \boldsymbol{\theta}_{:,i}(t, \boldsymbol{x}) Z_1^i)) - V(t, \boldsymbol{x}))\big]_{\boldsymbol{x} = \boldsymbol{X}_{t-}} = 0, \end{aligned}$$

which essentially is what we wanted to show. □

Appendix 3. Data

Table C1. The first column lists the purchasing dates of the 36 contracts. Column two shows the ATM strikes at which the contracts are purchased and column three show the prices at which this happens. The fourth column gives the terminal P&L for each contract, when the hedge is performed with an ‘actual’ (EGARCH(1,1)) volatility forecast. Column five likewise but when the hedge is with the implied volatilities. Finally, columns six and seven give the quadratic variation, defined as $\sum_{i=1}^{N} |dP\&L_i|^2/N$, where $N = 63$ is the number of trading days, for the entire actual and implied paths, respectively.

Contract	ATM strike	Option price	$\text{P\&L}_T^{\text{actual}}$	$\text{P\&L}_T^{\text{implied}}$	$\text{Q.V.}^{\text{actual}}$	$\text{Q.V.}^{\text{implied}}$
07-Jul-2004	1118.3	36.5852	12.2045	15.1591	0.5269	0.2615
05-Oct-2004	1134.5	33.0392	5.8372	5.0520	0.1683	0.1386
05-Jan-2005	1183.7	34.7050	11.4080	13.6705	0.1975	0.1759
06-Apr-2005	1184.1	34.9985	7.3072	9.0917	0.3162	0.1693
06-Jul-2005	1194.9	34.4864	11.9818	10.5282	0.2974	0.0894
04-Oct-2005	1214.5	37.8141	7.2779	7.4261	0.5384	0.1670
06-Jan-2006	1285.4	37.1621	12.6952	12.4934	0.1539	0.1406
07-Apr-2006	1295.5	38.2703	0.0765	0.5022	0.2827	0.2444
07-Jul-2006	1265.5	45.5356	15.3714	13.7452	0.3655	0.1974
05-Oct-2006	1353.2	42.7682	12.6179	12.6400	0.0904	0.0945
08-Jan-2007	1412.8	45.4682	−5.0096	2.1741	2.4476	1.0569
09-Apr-2007	1444.6	47.0689	19.4885	7.4564	0.7699	0.0865
09-Jul-2007	1531.8	55.8378	−11.4976	−7.5524	1.8603	1.1396
05-Oct-2007	1557.6	63.1625	1.6451	−1.2115	1.2330	0.4542
09-Jan-2008	1409.1	74.2874	9.6117	9.6158	1.1975	0.6555
09-Apr-2008	1354.5	66.2276	17.3617	19.0049	0.8019	0.6270
09-Jul-2008	1244.7	62.8179	−56.9636	−47.0345	8.0872	10.4193
07-Oct-2008	996.2	83.8510	55.3847	51.8900	9.7721	6.4129
09-Jan-2009	890.3	69.9489	14.1892	3.2637	3.0947	0.4083
10-Apr-2009	856.6	62.9702	30.2400	27.2551	0.5701	0.4336
09-Jul-2009	882.7	49.8464	−12.4499	−9.8467	0.1245	0.1039
07-Oct-2009	1057.6	49.0640	17.0496	18.0507	0.2944	0.2135
07-Jan-2010	1141.7	42.2410	16.4989	16.4106	0.2595	0.1990
09-Apr-2010	1194.4	36.6784	−10.3121	−9.5031	0.5463	0.5578
09-Jul-2010	1078.0	52.2001	15.6833	17.6455	3.0501	0.3326
07-Oct-2010	1158.1	50.6050	20.7394	19.8607	0.1926	0.2166
06-Jan-2011	1273.8	43.6970	9.4015	11.7384	0.3762	0.2400
07-Apr-2011	1333.5	44.7866	13.3942	13.8116	0.3490	0.3055
08-Jul-2011	1343.8	43.0900	−3.8722	3.8883	0.1692	0.3196
06-Oct-2011	1165.0	73.4417	14.3245	16.8015	0.8601	0.7112
06-Jan-2012	1277.8	53.9770	−17.4158	−21.6739	0.3472	0.1853
09-Apr-2012	1382.2	48.9735	−9.7517	−9.9641	0.4760	0.4018
09-Jul-2012	1352.5	47.5814	15.8417	15.4475	0.3181	0.3184
05-Oct-2012	1460.9	42.9608	11.2648	9.0422	3.0925	0.8156
09-Jan-2013	1461.0	43.7355	17.6935	14.7094	0.2747	0.1360
11-Apr-2013	1593.4	39.2535	9.4000	6.6261	1.0037	0.7546