Learning Minimum Variance Discrete Hedging Directly from Market

Ke Nian[a,1], Thomas Coleman[b,2], Yuying Li[a,1,*]

[a] *Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada N2L 3G1*
[b] *Combinatorics and Optimization, University of Waterloo, Waterloo, ON, Canada N2L 3G1*

Abstract

Option hedging is a critical risk management problem in finance. In the Black-Scholes model, it has been recognized that computing hedging position from the sensitivity of the calibrated model option value function is inadequate in minimizing variance of the option hedge risk, as it fails to capture the model parameter's dependence on the underlying price, see e.g., [16, 37]. In this paper we demonstrate that this issue can exist generally when determining hedging position from the sensitivity of the option function, either calibrated from a parametric model from current option prices or estimated nonparametricaly from historical option prices. Consequently the sensitivity of the estimated model option function typically does not minimize variance of the hedge risk, even instantaneously. We propose a data driven approach to directly learn a hedging function from the market data by minimizing variance of the local hedge risk. Using the S&P 500 index daily option data for more than a decade ending in August 2015, we show that the proposed method outperforms the parametric minimum variance hedging method proposed in [37], as well as minimum variance hedging corrective techniques based on stochastic volatility or local volatility models. Furthermore, we show that the proposed approach achieves significant gain over the implied BS delta hedging for weekly and monthly hedging.

Keywords: machine learning, dynamic hedging, risk management, kernels, regularized network

---

1. Introduction

Options hedging is a critical problem in financial risk management. The prevailing approach in financial derivative pricing and hedging has been to first assume a parametric model describing the underlying price dynamics. An option model function V is then calibrated to current market option prices and various sensitivities are computed and used to hedge the option risk. For example, the sensitivity of the option value function to the underlying price is used in delta hedging. When the Black-Scholes (BS) model is assumed with the implied volatility calibrated to the market price at the rebalancing time, this is referred to as the practitioner's BS delta hedging. Unfortunately this widely established option risk management practice suffers a few fundamental challenges.

Firstly it is widely acknowledged that most proposed parametric underlying price models fail to capture the option market accurately. This inadequacy is clearly illustrated in the widely documented volatility smile for the BS model [10, 43]. Specifically there is strong empirical evidence that the BS model tends to overestimate option prices of deeply out-of-money options [29]. The BS model fails to capture the negative relationship between the implied volatility and the underlying price [26, 11]. A number of alternative parametric models have been proposed as improvement to the BS model, including the Stochastic Volatility (SV) models [32, 35, 36, 7], Local Volatility Function (LVF) models

---

*Corresponding author.

Email addresses: `knian@uwaterloo.ca` (Ke Nian), `tfcoleman@uwaterloo.ca` (Thomas Coleman), `yuying@uwaterloo.ca` (Yuying Li)

[1] All authors acknowledge funding from the National Sciences and Engineering Research Council of Canada. The authors also acknowledge referees' comments which have led to improved presentation of the paper.

[2] This author acknowledges funding from the Ophelia Lazaridis University Research Chair. The views expressed herein are solely from the authors.

[16, 21, 47, 23], and jump models, e.g., [41, 2, 33]. Unfortunately all models have been shown to have limitations in adequately modeling the option market.

Errors in the option value model have significant implications in hedging. As a consequence, one naturally cannot expect partial derivatives of the option value function to accurately represent option value sensitivity. In the context of a parametric option model, particularly the BS model, using the partial derivative $\frac{\partial V}{\partial S}$ as a hedging function, i.e., delta, is suboptimal, since this hedging function fails to minimize variance of the option risk, even infinitesimally, see, e.g., [16, 37]. Furthermore, using delta as the hedging position does not minimize the discrete option hedging risk, noting that option hedging can only be done discretely in practice.

Even for instantaneous hedging, $\frac{\partial V}{\partial S}$ fails to minimize variance of the hedging error because the calibrated model parameters depend on the underlying price S, when the assumed model is erroneous. For the BS model, it has been recognized that implied volatility depends on the underlying [16, 37, 1]. In addition, accounting for this parameter dependence is not straightforward. Methods have been proposed to correct the practitioner's BS delta hedging to minimize variance of the hedging error [37, 1, 4, 5, 30], based on analysis of some parametric models, e.g., LVF and SV. Recently Hull and White [37] propose a parametric model for minimum variance delta hedging based on approximations to the BS vega and analysis of the empirical option market data. The proposed parametric model, however, is based on instantaneous hedging analysis.

The dependence of implied volatility on the underlying arises because the BS model mis-specifies the underlying price dynamics for option pricing. Even when this mis-specification error is acceptable with respect to the option value, it can result in dependence of model parameters on the underlying. Consequently using the partial derivative of the option value function to the underlying as the hedging position can fail to eliminate the option risk completely, even infinitesimally. Given that it is unlikely that any parametric model describes the underlying dynamics perfectly for option pricing, a parametric model option value function calibrated from the option market inevitably will always have the model parameter dependence issue, which is often not explicitly acknowledged in the option hedging literature.

Recognizing these challenges in the parametric financial modeling approach, nonparametric option pricing has also been studied. The nonparametric option value modeling approach has the distinctive advantage of not relying on specific assumptions about the underlying asset price dynamics. Hutchinson et al. [39] first propose a nonparametric data-driven approach to price and hedge European options using neural networks, radial basis functions, and projection pursuit regressions. Many other neural network methods for European option pricing have also been proposed, see, e.g., [52, 9, 31, 28, 42].

Although there are quite a few studies on nonparametric option pricing models, to our knowledge, there has been little research specifically focusing on discrete hedging using a nonparametric method. Even when the hedging problem is considered, e.g., [39], it is treated as a byproduct of obtaining a nonparametric pricing function: The hedging position is obtained from the partial derivative of the option value function. Hutchinson et al. [39] show that, based on hedging errors on some simulated paths, this indirect data-driven hedging approach can potentially be an effective alternative to the traditional parametric delta hedging methods.

In this paper, we study the discrete option hedge problem by explicitly focusing on the issues arising from model specification errors. We illustrate that the inability to minimize variance of the hedging error, when determining a hedging position using an option value function determined from a parametric model, is also shared by an option model estimated from a nonparametric method. Although a nonparametric modeling approach to the option value can potentially lead to smaller mis-specification error, we illustrate that non-parametric model parameters can similarly depend on the underlying. Consequently the sensitivity of the optimally estimated option value function will not lead to minimization of option hedge risk. Furthermore, the estimated pricing function inevitably has errors, due to both model mis-specification, discretization, and numerical roundoff. The error in the value function can potentially be substantially magnified when computing partial derivatives as hedging positions.

We explore a direct market data driven approach to bypass this challenge to achieve effective hedging performance. We propose a nonparametric data-driven approach to learn a local risk minimization hedging model directly from the market data. Using a regularized network [27, 24] with a spline kernel [50], we learn a hedging function from the market data by minimizing the empirical local hedge risk with a suitable regularization. The local risk corresponds directly to the variance of the hedging error in the discrete rebalancing period.

Specifically, the main contributions of this paper include:

- We analyze and discuss implications from model mis-specification in the option value function for discrete option hedging. We illustrate challenges in accounting for dependence of the calibrated model parameters on the underlying, which arises due to model mis-specification.

- We formalize a more general relation between the partial derivatives of the implied volatility for the call and put, under the assumption of put and call symmetry. This relation can potentially be useful in accounting for the implied volatility dependence on the underlying.
- We analyze a regularized kernel network for option value estimation and illustrate that the partial derivative of the estimated value function with respect to the underlying similarly does not minimize variance of the hedging risk in general, even infinitesimally.
- We propose a data driven approach to learn a hedging position function directly by minimizing the variance of the local hedge risk. Specifically we implement a regularized spline kernel method to nonparametrically estimate the hedge function from the market data.
- Using synthetic data sets, we compare daily, weekly, and monthly hedging performance using the proposed direct data-driven hedging approach with the performance of the indirect approach where hedging positions are computed from the sensitivity of the nonparametric option value function. In particular, we present computational results which demonstrate that the direct spline kernel hedging position learning outperforms the hedging position computed from the sensitivity of the spline kernel option value function.
- Using S&P 500 index option market data for more than a decade ending in August 31, 2015, we demonstrate that the daily hedging performance of the direct spline kernel hedging function learning method significantly surpasses that of the minimum variance quadratic hedging formula proposed in [37], as well as corrective methods based on LVF and SABR implemented in [37].
- We also present weekly and monthly hedging results using the S&P 500 index option market data and demonstrate significant enhanced performance over the BS implied volatility hedging.

The structure of the subsequent presentation is as follows. We first review parametric option models, model mis-specification, and its implication in option hedging in section 2. In section 3, we discuss the proposed nonparametric data-driven approach, regularized network, and spline kernels. Computational performance comparison, based on the synthetic and real market data, is reported in section 6. The concluding remarks are given in section 7.

2. Hedging Challenges from the Parametric Option Modeling Approach

We first discuss challenges in option hedging in practice, including dependence of the model parameters on the underlying price, accounting for this dependence in minimum variance hedging, and minimizing discrete hedging risk.

2.1. Dependence of the Implied Volatility on the Underlying and Minimum Variance Hedging

To see how calibrating a mis-specified model to market option prices leads to dependence of the calibrated parameters on the underlying, we consider a BS underlying price model

$$dS_t = (r-q)S_t dt + \sigma S_t dW_t \tag{1}$$

where W_t is a standard Brownian motion, $r > 0$ is the constant risk free rate, $q \geq 0$ is dividend yield, and $\sigma > 0$ is a constant volatility. To hedge an option with a fixed strike K and time to expiry $T-t$ at time t, one first computes the implied volatility $\check{\sigma}$ from the market option price $\check{V}$, i.e., $\check{\sigma}$ is the solution to

$$V(S,\sigma) = \check{V} \tag{2}$$

where $V(S,\sigma)$ is BS model option value function and the hedging position in the underlying is $\frac{\partial V(S,\cdot)}{\partial S}$. Note that the current underlying price S, the constant risk free rate r, and dividend yield q are observed. This is frequently referred to as the "practitioner Black-Scholes delta". Since the underlying price following a BS model with a constant

σ cannot match market option prices exactly, the calibration equation links the implied volatility to the underlying and consequently the implied volatility $\breve{\sigma}$ also depends on the underlying price S. Consequently the sensitivity of the model function $V(S,\breve{\sigma})$ on the underlying should include the sensitivity of the implied volatility $\breve{\sigma}$ on S.

As an improvement over the BS model, the local volatility function (LVF) model,

$$\frac{dS_t}{S_t} = (r-q)dt + \sigma(S_t,t)dW_t \tag{3}$$

has been considered, see e.g., [22, 19, 20], and LVF and its extensions remain widely popular in practice. Many methods have been proposed to calibrate a local volatility function $\sigma(S,t)$ from the traded market option prices, see e.g., [40, 3, 15].

To determine the delta hedge $\frac{\partial V(S,\cdot)}{\partial S}$, typically an option model $V(S,\cdot)$ is first calibrated to the current market option prices, see, e.g., implied volatility [10], local volatility function [16, 21, 47, 23], or stochastic volatility [32, 35, 36, 7]. Hedging using a more accurate option value model can often significantly outperforms hedging from a less accurately option value function, see, e.g., LVF hedging compared to the BS implied volatility delta hedging in [16]. In addition, Coleman et al. [16] discuss a relationship between the partial derivatives of calls and puts in the context of the LVF model, under which a call and put symmetry relation holds, see, e.g., [12, 13]. This relationship is found useful in correcting the dependence of the implied volatility on the underlying for BS delta hedging.

In Theorem 1 below, we formalize this relation [16] to any call and put functions satisfying the call-put-symmetry. We note that, for this relationship to be useful in correcting for MV hedging in practice, hence the relevant price functions correspond to the market option prices, not model option values.

Theorem 1. *Let $C(S,K,T,r,q)$ and $P(S,K,T,r,q)$ be the call option price and put option price with underlying price S, strike K, time to expiry T, interest rate r and dividend yield q. Assume further that there exists a unique implied volatility calibrating to $C(S,K,T,r,q)$ and $P(S,K,T,r,q)$ respectively and the call-put-symmetry below holds:*

$$C(S,K,T,r,q) = P(K,S,T,q,r). \tag{4}$$

Then

$$\frac{\partial \breve{\sigma}_c(S,K,T,r,q)}{\partial S} = \frac{\partial \breve{\sigma}_p(K,S,T,q,r)}{\partial S}. \tag{5}$$

where $\breve{\sigma}_c(\cdot)$ is the BS implied volatility calibrated to $C(\cdot)$ and $\breve{\sigma}_p(\cdot)$ is the BS implied volatility calibrated to $P(\cdot)$ respectively.

Proof. Let $C_{\mathrm{BS}}(\cdot)$ and $P_{\mathrm{BS}}(\cdot)$ denote the BS model option value functions. Put and call symmetry under the BS model implies that

$$C_{\mathrm{BS}}(S,K,T,r,q,\sigma) = P_{\mathrm{BS}}(K,S,T,q,r,\sigma), \tag{6}$$

where σ is any constant volatility. Let $\breve{\sigma}_c(S,K,T,r,q)$ and $\breve{\sigma}_p(K,S,T,q,r)$ be the BS implied volatilities calibrated to $C(S,K,T,r,q)$ and $P(K,S,T,q,r)$ respectively. Then

$$\begin{aligned} C(S,K,T,r,q) &= C_{\mathrm{BS}}(S,K,T,r,q,\breve{\sigma}_c(S,K,T,r,q)) \\ P(K,S,T,q,r) &= P_{\mathrm{BS}}(K,S,T,q,r,\breve{\sigma}_p(K,S,T,q,r)). \end{aligned} \tag{7}$$

From (4) and above, it follows

$$C_{\mathrm{BS}}(S,K,T,r,q,\breve{\sigma}_c(S,K,T,r,q)) = P_{\mathrm{BS}}(K,S,T,q,r,\breve{\sigma}_p(K,S,T,q,r)) \tag{8}$$

Using (6) with $\breve{\sigma}_c(\cdot)$,

$$C_{\mathrm{BS}}(S,K,T,r,q,\breve{\sigma}_c(S,K,T,r,q)) = P_{\mathrm{BS}}(K,S,T,q,r,\breve{\sigma}_c(S,K,T,r,q)),$$

From above and (8), it follows

$$P_{\mathrm{BS}}(K,S,T,q,r,\breve{\sigma}_c(S,K,T,r,q)) = P_{\mathrm{BS}}(K,S,T,q,r,\breve{\sigma}_p(K,S,T,q,r))$$

Assuming that there is a unique implied volatility from the BS formula, we have

$$\breve{\sigma}_c(S,K,T,r,q) = \breve{\sigma}_p(K,S,T,q,r) \tag{9}$$

Taking derivative with respect to S, we have

$$\frac{\partial \breve{\sigma}_c(S,K,T,r,q)}{\partial S} = \frac{\partial \breve{\sigma}_p(K,S,T,q,r)}{\partial S},$$

i.e., (5) holds. This completes the proof. □

Assume that the market option prices satisfy put-call symmetry. Then the relevance of Theorem 1 in accounting for dependence of the implied volatility on the underlying can be appreciated as follows. On the left hand side of (5), the derivative is with respect to the underlying price (the first argument). On the right hand side, the derivative is with respect to the strike price (the second argument). When $q \approx r$ (as in the futures options market), $\frac{\partial \breve{\sigma}_p(K,S,T,q,r)}{\partial S}$ can be estimated from the observed implied volatility surface. Consequently the sensitivity of the implied volatility to the underlying $\frac{\partial \breve{\sigma}_c(S,K,T,r,q)}{\partial S}$ can be estimated. Recently Hull and White [37] implement a corrective formula for minimum variance hedging based on (5), referred to as the LVF minimum variance hedging. Other minimum variance delta hedge methods have also been proposed to correct practitioner BS delta explicitly, see, e.g., [7, 18, 8, 46, 6].

Although a more complex parametric underlying price model, e.g., LVF and SV, can fit option market prices more accurately and generate better hedging performance than the simple BS model, it is unlikely that this will eliminate all dependence of the model parameters on the underlying. This can be seen from the following arguments.

Assuming that the calibration equation holds for all S and the partial derivative of the market option price with respect to S exists, taking derivative with respect to S from the both sides of (2),

$$\frac{\partial V}{\partial S} + \frac{\partial V}{\partial \sigma}\frac{\partial \sigma}{\partial S} = \frac{\partial \breve{V}}{\partial S}$$

Consequently, $\frac{\partial \sigma}{\partial S} \neq 0$ unless the model partial derivative $\frac{\partial V}{\partial S}$ equals the market option delta $\frac{\partial \breve{V}}{\partial S}$, which is very unlikely in general since any assumed model only attempts to calibrate to option prices (and typically cannot even match all option prices). Therefore, the dependence of the model parameters on the underlying price exists in general when calibrating option value function. Hence using $\frac{\partial V}{\partial S}$ does not generally capture the market option delta, even instantaneously.

2.2. *A Quadratic Formula Driven by the Market Price*

Hull and White [37] propose a quadratic MV delta hedge formula based on the BS vega and empirical option data analysis. Here we discuss it briefly as it is the method which is most connected to the approach proposed in this paper.

The objective of the minimum variance delta hedging is to minimize the variance of the following hedging error

$$dV - \delta_{\mathrm{MV}} \cdot dS \tag{10}$$

where δ_{MV} is the hedging position in the underlying and dV and dS denote the differential of the market option price and the market underlying price respectively. It is important to note that the hedging performance is assessed here based on the market prices, not the model prices. In particular, dS is different from that in the risk neutral model (1) or (3) used in option valuation. Indeed, for hedging performance, the risk neutral model for option valuation is irrelevant. In the traditional hedging computation based on parametric option pricing model, hedging position computation requires a risk neutral model, e.g., (1) or (3), the hedging approach proposed in this paper uses the market option and underlying prices directly to determine hedging positions, avoiding inevitable errors in the assumed risk neutral model.

Subsequently we denote the BS formula as $V_{\mathrm{BS}}(S, \breve{\sigma})$ explicitly, where $\breve{\sigma}$ denotes the implied volatility calibrated to the corresponding market option price. Then the minimum variance delta hedge needs to consider both the sensitivity of the model option value to the underlying price, as well as the expected change in the implied volatility conditional on the change in the underlying price. In [37], the MV delta, δ_{MV}, is assumed to be

$$\delta_{\mathrm{MV}} = \frac{\partial V_{\mathrm{BS}}}{\partial S} + \frac{\partial V_{\mathrm{BS}}}{\partial \breve{\sigma}} \frac{\partial \mathbf{E}(\breve{\sigma})}{\partial S}, \tag{11}$$

where $\mathbf{E}(\breve{\sigma})$ denotes expectation.

Hull and White [37] analyze the market option data and consider the BS vega for European options and propose that $\frac{\partial \mathbf{E}(\breve{\sigma})}{\partial S}$ can be reasonably approximated by assuming a quadratic function of the BS delta. Specifically,

$$\frac{\partial \mathbf{E}(\breve{\sigma})}{\partial S} = \frac{a + b\,\delta^{\mathrm{BS}} + c\,(\delta^{\mathrm{BS}})^2}{S\,\sqrt{T}}$$

where a, b and c are the parameters to be fitted using market option data. This is equivalent to assuming that

$$\delta_{MV} = \delta^{\mathrm{BS}} + \frac{a + b\delta^{\mathrm{BS}} + c(\delta^{\mathrm{BS}})^2}{S\,\sqrt{T}}\,\mathrm{vega}^{BS}. \tag{12}$$

where vega^{BS} is the BS vega using implied volatility.

2.3. Risk Minimization Discrete Hedging

Hedging error (10) as well as the corrective formula (12) are based on assuming instantaneous hedging, which is not optimal in the context of discrete hedging. When rebalancing is done discretely, optimal local risk and global risk minimizing hedging can be computed by considering hedge risk in finite rebalancing time periods, see e.g., [48, 49, 25, 34, 17, 14]. Instantaneous hedging using the partial derivative of the option value function determines trading positions indirectly while risk minimization computes optimal hedging positions directly to minimize discrete hedge risk. Local and total risk minimization are typically based on an assumed parametric stochastic model for the underlying price; hence the option market data is not explicitly incorporated in the hedging position computation.

3. Data-driven Option Value Function Learning and Dependence Problem in Hedging

In §2, we have seen various challenges for hedging when the position is computed from the calibrated option value function. Since nonparametric option function estimation does not make any specific assumptions and can potentially match option prices more accurately, it is not unreasonable to expect that this challenge in hedging can potentially be addressed by reducing mis-specification using a data driven approach to first learn an option value function. Here we discuss this approach and analyze its challenge for option hedging.

Assume that a set of market option prices $\{\check{V}_1, \cdots, \check{V}_m\}$ on the same underlying and corresponding vector of dependent observable attributes $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$ are given. For options associated with a single underlying, the dependent attributes $\mathbf{x}_i$ may include the underlying price, volatility, time to expiry, strike price, and any other attributes which affect the market option price $\check{V}_i$. The objective is to learn a hedging position function $f(\mathbf{x})$ which can be used to directly hedge options.

Similar to discrete hedging under a parametric model, the hedging position can be computed by learning a nonparametric option value model and then determining the hedging position from the partial derivatives. Indeed this is the approach adopted in [39]. In this paper we propose to learn a nonparametric hedging function $f(\mathbf{x})$ directly from the change in market prices in discrete hedging periods. In particular we focus on regularized kernel learning methods. Subsequently we first review key components to the regularized kernel methods in §3.1. We then discuss in §3.2 potential issues associated with computing hedging positions from partial derivatives of the optimal regularized kernel functions. The proposed direct kernel hedging learning method is presented in §4.

3.1. Regularized Kernel Methods

Assume that a set of m training points $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_m, y_m)\}$ are given, where $(\mathbf{x}_i, y_i) \in \mathscr{X} \times \mathbb{R}$. A regularized kernel method, e.g., [24], learns a nonlinear function $f(\mathbf{x})$ to capture the dependence between the target y and feature $\mathbf{x}$. Assume that we are given a positive definite kernel similarity

$$k(\mathbf{x},\mathbf{x}') : \mathscr{X} \times \mathscr{X} \to \mathbb{R},$$

which captures similarity between data instances $\mathbf{x}$ and $\mathbf{x}'$ implicitly in a high dimensional feature space. Assume that $\mathscr{H}_K$ is the *Reproducing Kernel Hilbert Space* (RKHS) induced by the symmetric positive definite kernel function $k(\mathbf{x},\mathbf{y})$ and $\|f\|_K$ is the norm in RKHS.

A regularized kernel regression problem can be formulated as

$$\min_{f\in\mathscr{H}_K} \left(\sum_{i=1}^{m} L(y_i, f(\mathbf{x}_i)) + \lambda \|f\|_K^2 \right) \tag{13}$$

where $L(\cdot,\cdot)$ is a loss function, e.g., a p-th norm $L(y_i, f(\mathbf{x}_i)) = (y_i - f(\mathbf{x}_i))^p$, $0 < p < +\infty$ and $\lambda > 0$ is a regularization parameter which can be determined using cross validation. Following the Representor Theorem, e.g., [51], a solution of (13) has the form

$$f(\mathbf{x}) = \sum_{i=1}^{m} \alpha_i k(\mathbf{x},\mathbf{x}_i) \tag{14}$$

and the regularization term is given by

$$\|f\|_K^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i\alpha_j k(\mathbf{x_i},\mathbf{x}_j). \tag{15}$$

3.2. Hedging Positions as Partial Derivatives of Estimated Kernel Functions

Given a set of market option prices $\{\check{V}_1,\cdots,\check{V}_m\}$ on the same underlying S and a set of features $\{\mathbf{x}_1,\cdots,\mathbf{x}_m\}$,where $\mathbf{x}_i$ is a vector of the attributes affecting option price $\check{V}_i$, we can estimate an option value function $V(\mathbf{x})$ based on the regularized kernel estimation (13).

Using (14) and (15), assuming the quadratic loss and letting $y_i = \check{V}_i,\ i = 1,\cdots,m$, the option value function $V(\mathbf{x};\alpha) = \sum_{i=1}^{m} \alpha_i k(\mathbf{x},\mathbf{x}_i)$ can be computed by solving

$$\min_{\alpha} \left(\sum_{i=1}^{m} \left(\check{V}_i - \sum_{j=1}^{m} \alpha_j k(\mathbf{x}_i,\mathbf{x}_j) \right)^2 + \lambda \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i\alpha_j k(\mathbf{x}_i,\mathbf{x}_j) \right). \tag{16}$$

For standard options, the universal RBF kernel

$$k(\mathbf{x},\tilde{\mathbf{x}}) = e^{-\frac{\|\mathbf{x}-\tilde{\mathbf{x}}\|_2^2}{2\rho^2}} \tag{17}$$

is a reasonable kernel choice, since the option value function is very smooth, and a suitable bandwidth ρ is typically problem dependent and can be determined using cross validation.

Assume that the first component of the d-dimensional attribute vector $\mathbf{x} = (x_1,\cdots,x_d)$ corresponds to the underlying price S. Then the delta hedging function is typically determined as

$$\frac{\partial V(\mathbf{x};\alpha)}{\partial S} = \sum_{i=1}^{m} \alpha_i \frac{\partial k(\mathbf{x},\mathbf{x}_i)}{\partial S} \tag{18}$$

Hutchinson et al. [39] demonstrate that this nonparametric hedging approach, using the partial derivative of a nonparametric pricing function learned from historical market data, can be a useful alternative for option hedging.

However, using (18) as the hedging position does not minimize variance of the hedge risk in general and the challenge in accounting for parameter dependence on the underlying remains. We can see that the following arguments can be made, similar to those made before when calibrating a parametric model. Assume that the estimated kernel function $V(\mathbf{x};\alpha)$ matches the target market option price exactly, i.e.,

$$V(\mathbf{x};\alpha) = \check{V}$$

Then

$$\frac{\partial V}{\partial S}(\mathbf{x};\alpha) = \sum_{i=1}^{m} \alpha_i \frac{\partial k}{\partial S}(\mathbf{x},\mathbf{x}_i) + \sum_{i=1}^{m} \frac{\partial \alpha_i}{\partial S} k(\mathbf{x},\mathbf{x}_i) = \frac{\partial \check{V}}{\partial S}$$

Hence in general

$$\frac{\partial \alpha}{\partial S} \neq 0$$

unless

$$\frac{\partial V}{\partial S}(\mathbf{x};\alpha) = \frac{\partial \check{V}}{\partial S}. \tag{19}$$

However there is no reason that a solution of the regression problem (16) should satisfy (19). Consequently it is similarly difficult to account for all dependence on S, even infinitesimally, in the estimated kernel model.

Furthermore, error magnification can happen by deriving the hedging position from an estimated kernel function. To see this, let $y(x)$ be the target function and $f(x)$ be the corresponding approximation from the kernel regression, which generally contains errors. Let ε denote the total option error at x and $x+\Delta x$ (either roundoff error and/or model error), i.e., $\varepsilon = f(x+\Delta x) - y(x+\Delta x) - f(x) + y(x)$. Then the error in the derivative approximation by $\frac{df}{dx}$ become

$$\begin{aligned} \frac{dy}{dx} - \frac{df}{dx} &= \frac{dy}{dx} - \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\ &= \frac{dy}{dx} - \lim_{\Delta x \to 0} \frac{y(x+\Delta x) - y(x) + \varepsilon}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{\varepsilon}{\Delta x} \end{aligned}$$

which can become arbitrarily large when $\varepsilon \neq 0$ does not converge to zero.

4. Learning Option Hedging Functions Directly from the Market

Assume that both the underlying and option market prices are available, at a fixed time period Δt, e.g., 1-day or 1-week, for a reasonably long time horizon. Specifically we have both market option price observations $\{\check{V}_1, \cdots, \check{V}_m\}$ on the same underlying and attributes $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$, with $\mathbf{x}_i$ corresponding to the option $\check{V}_i$. Let market prices, after a time period Δt, be denoted by $\{\check{V}_1^{+\Delta t}, \cdots, \check{V}_m^{+\Delta t}\}$. The goal of the option hedging is to determine a hedging position function (in the underlying), $\delta(\mathbf{x})$, to minimize the variance of hedging error for a fixed time length Δt.

We choose the empirical loss function to correspond to the variance in option hedging error. Let $\Delta S_i = S_i^{+\Delta t} - S_i$ denote the change in the underlying price and $\Delta \check{V}_i = \check{V}_i^{+\Delta t} - \check{V}_i$ denote the change of option value, with the corresponding option attributes $\mathbf{x}_i$, for the data instance i. We use the following variance hedge risk as the empirical loss function, i.e.,

$$L(y_i, \delta(\mathbf{x}_i)) = (\Delta \check{V}_i - \Delta S_i \delta(\mathbf{x}_i))^2. \tag{20}$$

Note that this corresponds to the minimum variance hedging (10) except the instantaneous changes dV and dS are replaced by discrete change $\Delta \check{V}$ and ΔS. As an alternative to the loss function (20), one can also consider

$$L(y_i, f(\mathbf{x}_i)) = |\Delta \check{V}_i - \Delta S_i f(\mathbf{x}_i)|^p,$$

where $1 \leq p < +\infty$. Using the quadratic loss (20), the hedging position function $\delta(\mathbf{x})$ can be estimated from the regularized optimization below:

$$\min_{\delta \in \mathscr{H}_K} \left\{ \sum_{i=1}^{m} \left(\Delta \check{V}_i - \Delta S_i \delta(\mathbf{x}_i) \right)^2 + \lambda \|\delta\|_K^2 \right\} \tag{21}$$

Following the Representor Theorem, (14) and (15), the solution can be computed by solving the following convex quadratic minimization

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^m}\left\{\sum_{i=1}^{m}\left(\Delta\check{V}_i-\Delta S_i\left(\sum_{j=1}^{m}\alpha_j k(\mathbf{x}_i,\mathbf{x}_j)\right)\right)^2+\lambda\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j k(\mathbf{x}_i,\mathbf{x}_j)\right\} \tag{22}$$

The loss function for the hedging function in the proposed formulation (21) directly corresponds to the variance of hedge risk for a discrete Δt-time period. In addition the hedge function is the solution of the optimization problem and there is no potential error magnification through partial derivative computation.

Since the delta of the payoff function at the expiry is a discontinuous step function, the delta hedging function of an option changes quickly as the underlying changes near the expiry. Consequently we choose to use a spline kernel function [50] for the hedging function estimation. Specifically we use kernel generating splines with an infinite number of knots, which maps, in the 1-D case, a variable x defined on an interval, to splines of order $d \geq 0$ with infinite number of knots $\{t_i\},\ i=1,2,\ldots,$

$$x \to \phi(x) = (1, x, \ldots, x^d, (x-t_1)_+^d, \ldots, (x-t_i)_+^d, \ldots)$$

An explicit expression for the spline kernel can be derived, see e.g., [50] for more details. For multidimensional data, the spline kernel is the product of one-dimensional spline kernel functions.

5. Cross Validation

For the proposed data-driven formulation (21), we need to select an appropriate penalty λ to control model complexity. Cross-validation (CV) is a commonly used method for the model selection of an learning algorithm, which can be computationally expensive in general. Fortunately, for the regularized kernel method with a quadratic loss, the CV error can be computed efficiently without retraining the model in each CV round.

Let $\check{\boldsymbol{V}} = \{\check{V}_1, \check{V}_2, \cdots, \check{V}_m\}$ be the market option price observations and $\boldsymbol{K}$ be the kernel matrix with $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j),\ i = 1, \cdots, m,\ j = 1, \cdots, m$. For the indirect data-driven approach, (16) can be rewritten in matrix form:

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^m}(\boldsymbol{K\alpha}-\check{\boldsymbol{V}})^T(\boldsymbol{K\alpha}-\check{\boldsymbol{V}})+\lambda\boldsymbol{\alpha}^T\boldsymbol{K\alpha} \tag{23}$$

and the solution is,

$$\boldsymbol{\alpha} = (\boldsymbol{K}+\lambda\boldsymbol{I})^{-1}\check{\boldsymbol{V}}.$$

It can be shown, see e.g., [45, 51], that, given the eigen-decomposition $\boldsymbol{K} = \boldsymbol{Q\Lambda Q}^T$, the computational complexity for the leave-one-out cross-validation (LOOCV) on λ is $O(m^2)$ and the computational complexity for the n-fold cross-validation (nFCV) is $O(m^3/n)$.

Let $\Delta\check{\boldsymbol{V}} = \{\Delta\check{V}_1, \Delta\check{V}_2, \cdots, \Delta\check{V}_m\}$ and $\boldsymbol{D}$ be a diagonal matrix with $D_{ii} = \Delta S_i,\ i = 1, \cdots, m$, on its diagonal. We can similarly rewrite (22), for the direct data driven hedging, in a matrix form:

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^m}(\boldsymbol{DK\alpha}-\Delta\check{\boldsymbol{V}})^T(\boldsymbol{DK\alpha}-\Delta\check{\boldsymbol{V}})+\lambda\boldsymbol{\alpha}^T\boldsymbol{K\alpha} \tag{24}$$

Let $\widetilde{\boldsymbol{K}} = \boldsymbol{DK}$, the solution to (24) is given below:

$$\boldsymbol{\alpha} = (\widetilde{\boldsymbol{K}}^T\widetilde{\boldsymbol{K}}+\lambda\boldsymbol{K})^{-1}\widetilde{\boldsymbol{K}}^T\Delta\check{\boldsymbol{V}}$$

Unfortunately, the fast cross validation algorithms from [45, 51] for the minimization problem (24) cannot be applied for (24), since the Hessian of the regularization term is no longer diagonal. However if the regularization term for (24) is changed from $\lambda\boldsymbol{\alpha}^T\boldsymbol{K\alpha}$ to $\lambda\boldsymbol{\alpha}^T\boldsymbol{\alpha}$, the minimization problem becomes:

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^m}(\boldsymbol{DK\alpha}-\Delta\check{\boldsymbol{V}})^T(\boldsymbol{DK\alpha}-\Delta\check{\boldsymbol{V}})+\lambda\boldsymbol{\alpha}^T\boldsymbol{\alpha} \tag{25}$$

The solution to (25) has the form:

$$\boldsymbol{\alpha} = (\widetilde{\boldsymbol{K}}^T \widetilde{\boldsymbol{K}} + \lambda \boldsymbol{I})^{-1} \widetilde{\boldsymbol{K}}^T \Delta \check{\boldsymbol{V}}$$

Utilizing the ideas from [45, 51], it is easy to show that, given the singular value decomposition $\widetilde{\boldsymbol{K}} = \boldsymbol{U\Sigma V}^T$, for (25), the computational complexity for the LOOCV remains $O(m^2)$ and the computational complexity for the nFCV remains $O(m^3/n)$. In practice, changing the regularization term from $\lambda \boldsymbol{\alpha}^T \boldsymbol{K\alpha}$ to $\lambda \boldsymbol{\alpha}^T \boldsymbol{\alpha}$ does not seem to affect the performance significantly. Therefore to improve the computation efficiency for the direct data driven approach, we solve problem (25) instead of (24) when training data size is large.

6. Hedging Performance Comparisons

Using both synthetic data and S&P 500 market data, we demonstrate and compare hedging performance using the proposed direct hedging formulation (22) with the indirect approach which uses the hedging position given by the estimated option function from (23), as well as minimum variance hedging methods. Specifically,

- In §6.1, we demonstrate that the direct hedging function estimation (22) using the spline kernel outperforms the approach of determining hedging position indirectly as the partial derivative (18) of the estimated option value function from (23). This comparative study is done using a synthetic data set. We demonstrate that the direct spline kernel hedging function learning DKL_{SPL} outperforms the indirect approach and using the spline kernel yields better performance than using the RBF kernel.
- In §6.2, we compare the direct hedging method (25) using the spline kernel DKL_{SPL} with the minimum variance hedging methods using the S&P 500 market data. We demonstrate that the direct spline kernel hedging DKL_{SPL} performs significantly better than the compared minimum variance hedging methods.

We evaluate hedging performance using the gain ratio defined in [37], which measures the average quadratic hedging error (corresponding to variance when the mean is zero) of a hedging method over the average quadratic hedging error of the implied volatility BS delta hedging. Specifically, let m be the total number of testing data instances, δ_i^{BS} be the implied BS delta for the data instance i, δ_i^M be the hedging position calculated from the comparing method for data instance i, $\Delta \check{V}_i$ be the change in option price for data instance i, and ΔS_i be the change in the underlying price for data instance i. As in [37], the ratio Gain is the percentage reduction in the sum of the squared errors resulting from the one step hedge against the implied volatility BS delta hedge:

$$\text{GAIN} = 1 - \frac{\sum_{i=1}^m \left(\Delta \check{V}_i - \delta_i^M \, \Delta S_i\right)^2}{\sum_{i=1}^m \left(\Delta \check{V}_i - \delta_i^{\text{BS}} \, \Delta S_i\right)^2} \tag{26}$$

For additional comparison metrics, for comparisons using the synthetic data, we also report mean absolute value, standard deviation, VaR and CVaR of the hedging error $\Delta \check{V} - \delta^M \, \Delta S$.

6.1. Comparing learning hedging function directly vs indirectly using synthetic data

Following the S&P 500 market option specifications in [38], we first synthetically generate option data assuming that the underlying price follows a Heston model [35]. We compare hedging effectiveness of direct hedging function learning (21) and indirect hedging function estimation (16). In addition, we compare their performance to that of using analytic delta under the Heston model, which can be regarded as a best case benchmark, at least for daily hedging. Note that there is no model mis-specification for the underlying price in this case. Since the loss function is quadratic, solutions to (21) and (16) can be easily computed from linear equation solves. In addition we also compare hedging performance using a RBF kernel (17) versus a spline kernel. Specifically, using the generated synthetic data, we compare here hedging performance of the following hedging computation methods:

- δ^{BS}: implied volatility BS delta
- HESTON: analytic Heston delta

- DKL_{SPL} : direct learning a spline kernel hedging function based on (21)
- DKL_{RBF} : directly learning a RBF kernel hedging function directly based on (21)
- IKL_{SPL}: determining hedging position indirectly as the partial derivative (18) of the option value function estimated from (16) using a spline kernel
- IKL_{RBF}: determining hedging position indirectly as the partial derivative (18) of the option value function estimated from (16) using a RBF kernel

Training data consists of simulated daily underlying price S_t for two years[3], $t = 1, \cdots, 2 \times 252$, assuming a Heston model below

$$
\begin{aligned}
dS_t &= rS_t dt + \sqrt{\upsilon_t} S_t dW_t \\
d\upsilon_t &= \kappa(\overline{\upsilon} - \upsilon_t)dt + \eta\sqrt{\upsilon_t} dZ_t \\
\mathbf{E}(dZ_t dW_t) &= \rho dt
\end{aligned}
$$

In addition, a vector of simulated (traded) market call option prices $\check{\mathbf{V}}_t$ are generated, on each day t, with different strikes and time to expiry, following the specifications of stock options described in [38]. The option prices $\check{\mathbf{V}}_t$ are computed using the analytic option formula under the Heston model [35], using the the parameters from [7], which are given in Table 1.

r	$\overline{\upsilon}$	κ	η	ρ	S_0	υ_0
0.02	0.04	1.15	0.39	-0.64	100	0.04

Table 1: Parameters for the Heston Model

Testing data consists of 100 daily underlying price paths and corresponding option prices, spanning a six month period. We report average performance measures over 10 random training-test-data sets generated as described. Table A.12 in Appendix A.1 describes the average sizes[4] of training and testing data sets.

We train a regularized option price kernel model for the indirect hedging IKL_{SPL} and IKL_{RBF}. For IKL_{SPL} and IKL_{RBF}, the hedging position is the partial derivative of the estimated model function against the moneyness, as in [39], since the ratio of the option price to the strike is a function of the moneyness and time to expiry under the Heston model due to homogeneity. We note that the simulation hedging analysis in [39] considers only the BS model and the results in [39] indicate that a lower hedging error can be achieved on a subset of simulated paths.

Following a common practice, we use the standard deviation of the pairwise Euclidean distance of the training data as the bandwidth ρ for the RBF kernel. The regularization parameter λ is however selected using a 5-fold cross-validation. We also consider weekly hedging and monthly hedging, which correspond to hedging over a 5-business-days period and 20-business-days period respectively.

To investigate impact of the feature choice, we evaluate hedging performance using

- Feature Set #1 = {MONEYNESS, TIME-TO-EXPIRY}.
- Feature Set #2 = {MONEYNESS, TIME-TO-EXPIRY, δ^{BS}}.

In the second feature set, the BS delta using the implied volatility from the option market is used as an additional feature in hedging.

6.1.1. Feature Set #1: {MONEYNESS, TIME-TO-EXPIRY}

For the synthetic data, it is known that the option price is a function of the moneyness $\frac{S}{K}$ and time to expiry τ, which are the attributes $\mathbf{x}$ in Feature Set #1. Table 2, 3 and 4 report results for daily, weekly, and monthly hedging respectively.

[3] We assume that there are 252 trading days in a year.

[4] Following CBOE option specification rules, the size of a training or testing data set can vary slightly for each simulation run.

Method	Gain (%)	$\mathbf{E}(\lvert\Delta\check{V} - \Delta S f(\mathbf{x})\rvert)$	Std	VaR	CVaR
δ^{BS}	0.0	0.185	0.286	0.380	0.574
IKL_{RBF}	-3.3	0.171	0.291	0.356	0.566
IKL_{SPL}	-183.3	0.291	0.482	0.669	1.105
DKL_{RBF}	63.1	0.120	0.174	0.251	0.352
DKL_{SPL}	**64.9**	**0.121**	**0.170**	**0.255**	**0.345**
HESTON	63.6	0.121	0.173	0.266	0.360

Table 2: Daily Hedging Comparison

[1] FS #1: $\mathbf{x} = \{\text{MONEYNESS, TIME-TO-EXPIRY}\}$
[2] Bold entry indicating best Gain

Method	Gain (%)	$\mathbf{E}(\lvert\Delta\check{V} - \Delta S f(\mathbf{x})\rvert)$	Std	VaR	CVaR
δ^{BS}	0.0	0.414	0.620	0.776	1.009
IKL_{RBF}	-197.6	0.406	1.070	0.741	1.254
IKL_{SPL}	-94.7	0.548	0.866	1.114	1.738
DKL_{RBF}	47.0	0.312	0.451	0.620	0.825
DKL_{SPL}	**50.8**	**0.312**	**0.435**	**0.622**	**0.797**
HESTON	45.7	0.319	0.456	0.651	0.840

Table 3: Weekly Hedging Comparison

[1] FS #1: $\mathbf{x} = \{\text{MONEYNESS, TIME-TO-EXPIRY}\}$
[2] Bold entry indicating best Gain

Method	Gain (%)	$\mathbf{E}(\lvert\Delta\check{V} - \Delta S f(\mathbf{x})\rvert)$	Std	VaR	CVaR
δ^{BS}	0.0	0.941	1.484	1.516	1.808
IKL_{RBF}	1.0	0.888	1.470	1.516	1.829
IKL_{SPL}	-36.9	1.135	1.729	2.033	2.894
DKL_{RBF}	33.6	0.860	1.181	1.612	1.949
DKL_{SPL}	35.4	0.858	1.165	1.610	1.922
HESTON	**38.7**	**0.814**	**1.136**	**1.544**	**1.829**

Table 4: Monthly Hedging Comparison

[1] FS #1: $\mathbf{x} = \{\text{MONEYNESS, TIME-TO-EXPIRY}\}$
[2] Bold entry indicating best Gain

Table 2 and 3 demonstrate that the direct hedging function learning DKL_{SPL} & DKL_{RBF} significantly outperform the indirect hedging learning IKL_{SPL} & IKL_{RBF} in Gain and different risk measures considered. Indeed, DKL_{SPL} & DKL_{RBF} slightly outperform the benchmark of using the analytic Heston delta. The indirect hedging function learning performs more poorly than the implied BS delta hedging. In addition, the spline kernel performs better than the RBF kernel (with the standard deviation as the bandwidth parameter). The RBF kernel yields larger risk measures and smaller Gain for both the direct and indirect hedging learning methods.

Table 4 reports hedging comparison for monthly hedging. We observe that DKL_{SPL} & DKL_{RBF} significantly outperform the indirect hedging learning IKL_{SPL} & IKL_{RBF} in Gain and various risk measures. In addition, DKL_{SPL} & DKL_{RBF} continue to achieve enhanced performance over δ^{BS}, with the spline kernel DKL_{SPL} yielding better results than DKL_{RBF}. Not surprisingly, hedging performance of each method also deteriorates as the length of the hedging period increases, with larger mean absolute hedging error and larger standard deviation for monthly hedging than for daily and weekly hedging.

Table 4 also illustrates that, unlike daily and weekly hedging, DKL_{SPL} & DKL_{RBF} slightly underperform the analytic Heston delta benchmark for monthly hedging. Given that the analytic delta is for instantaneous hedging while the direct hedging learning DKL_{SPL} minimizes quadratic hedging error, one would expect better performance from the direct hedging learning DKL_{SPL}. We suspect that this is due to the effect of the specific combination of choices of features and kernel. Next we show that, with a different feature set, performance of direct hedging is improved, which suggests the possibility of surpassing analytic Heston delta benchmark, using a more suitable feature set, for a longer period hedging.

6.1.2. Feature Set #2: $\{$MONEYNESS, TIME-TO-EXPIRY, $\delta^{BS}\}$

We add the BS delta using the implied volatility as an additional feature in the direct hedging learning, since Hull and White [37] indicate that a better minimum variance hedge can be calculated based on the implied volatility delta. Table 5, 6, and 7 present hedging results for DKL_{SPL} & DKL_{RBF} for $\mathbf{x} = \{$MONEYNESS, TIME-TO-EXPIRY, $\delta^{BS}\}$. For clarity, we also include the results for FS #1 $\mathbf{x} = \{$MONEYNESS, TIME-TO-EXPIRY$\}$ and Heston delta for ease of comparison.

Method		Gain (%)	$\mathbf{E}(\lvert\Delta\check{V} - \Delta S f(\mathbf{x})\rvert)$	Std	VaR	CVaR
DKL_{RBF}	FS #1	63.1	0.120	0.174	0.251	0.352
	FS #2	62.3	0.114	0.176	0.238	0.349
DKL_{SPL}	FS #1	64.9	0.121	0.170	0.255	0.345
	FS #2	**70.9**	**0.110**	**0.154**	**0.234**	**0.322**
HESTON		63.6	0.121	0.173	0.266	0.360

Table 5: Daily Hedging Comparison

1 FS #2: $\mathbf{x} = \{$MONEYNESS, TIME-TO-EXPIRY, $\delta^{BS}\}$

2 Bold entry indicating best Gain

Method		Gain (%)	$\mathbf{E}(\lvert\Delta\check{V} - \Delta S f(\mathbf{x})\rvert)$	Std	VaR	CVaR
DKL_{RBF}	FS #1	47.0	0.312	0.451	0.620	0.825
	FS #2	51.4	0.301	0.432	0.611	0.816
DKL_{SPL}	FS #1	50.8	0.312	0.435	0.622	0.797
	FS #2	**53.5**	**0.299**	**0.422**	**0.606**	**0.794**
HESTON		45.7	0.319	0.456	0.651	0.840

Table 6: Weekly Hedging Comparison

1 FS #2: $\mathbf{x} = \{$MONEYNESS, TIME-TO-EXPIRY, $\delta^{BS}\}$

2 Bold entry indicating best Gain

Method		Gain (%)	$\mathbf{E}(\lvert\Delta\check{V}-\Delta Sf(\mathbf{x})\rvert)$	Std	VaR	CVaR
DKL_{RBF}	FS #1	33.6	0.860	1.181	1.612	1.949
	FS #2	30.1	0.863	1.217	1.652	2.104
DKL_{SPL}	FS #1	35.4	0.858	1.165	1.610	1.922
	FS #2	36.6	0.836	1.156	1.609	1.953
HESTON		**38.7**	**0.814**	**1.136**	**1.544**	**1.829**

Table 7: Monthly Hedging Comparison

[1] FS #2: $\mathbf{x} = \{\text{MONEYNESS, TIME-TO-EXPIRY, } \delta^{\text{BS}}\}$
[2] Bold entry indicating best Gain

From Table 5, 6 and 7, we observe that, for daily and weekly hedging, including the BS delta further improves the performance of the direct hedging learning methods, which outperforms analytical delta hedging. For monthly hedging, however, the performance is similar to what we obtain with that of the feature set #1. Overall, including the BS delta as an attribute is beneficial for the direct hedging function learning.

6.2. Hedging Comparison in the S&P 500 Option Market

Using more than ten years of the S&P 500 market data, ending August 2015, we compare the direct spline kernel hedging DKL_{SPL} with the following minimum variance hedging methods considered in [37]:

- MV: Hull-White's minimum variance hedging based on formulation (12),
- LVF: MV hedging from Local Volatility Function,
- δ^{BS}: implied volatility BS delta,
- SABR : MV hedging from SABR.

Since we use the exactly same S&P 500 index option data in [37] from the OptionMetric Database [44], here we simply quote the results presented in [37] for MV quadratic model (12) by Hull and White [37], LVF, and SABR implemented in [37].

As stated in [37], a SABR model is calibrated daily, from which the hedge position is determined and applied for the next day. For LVF, the partial derivative of the expected implied volatility with respect to the underlying is calculated from the slope of the observed implied volatility daily. For MV, the model parameter, a, b and c in (12), are estimated using all options **traded** in a 36 months window and then applied to determine the hedging position every day in the next month.

We process the data following the same procedure described in [37]. From the data set, on each day, the closing bid and ask option prices, daily underlying price, daily risk-free interest rate, and Greeks from the BS models using implied volatilities are available. The mid price from the bid and ask is used as the price of the option on that day. The closing price for the underlying is regarded as the daily underlying price. Options with time to expiry less than 14 days are removed from the data set. The call options with the implied volatility delta less than 0.05 or greater than 0.95 and the put options with the implied volatility delta less than -0.95 or greater than -0.05 are also removed from the data set. For hedging performance comparison, the option data instances are also divided into nine buckets, as in [37], according to the implied volatility delta δ^{BS} rounded to the nearest tenth.

We use the proposed DKL_{SPL} direct hedging approach with the feature set $\mathbf{x} = \{\text{MONEYNESS, TIME-TO-EXPIRY, } \delta^{\text{BS}}\}$. We build a separate model for each delta bucket, since learning a single model will be computationally demanding, as training and testing need to be done frequently, e.g., daily, for more than a 10-year horizon. For call options, we use the date set from January 2, 2004 to August 31, 2015, which is the same as in [37]. For the call option, the model for each bucket is estimated using all options **traded** in a 36 months window. For the put option, there is much larger variation in the number of data instances in each bucket. To ensure a reasonable training data size, we have used a time window of 24 months for delta buckets -0.1 and -0.2, window of 36 months for delta buckets -0.3 and -0.4, window of 48 months for delta buckets -0.5 and -0.6, and window of 72 months for delta buckets -0.7, -0.8 and -0.9. Hence, the training data for puts ranges from Jan 2001 to August 31 2015. Table A.13 & A.14 in Appendix A.2 report average sizes of training and test data sets.

We train the direct hedging function model $\mathrm{DKL_{SPL}}$ using all options **traded** in a fixed time window, as described above, and determine the hedge each day in the following month using the trained model, following the same sliding training-testing window procedure in [37]. Since a model needs to be trained for more than a decade, we use the efficient LOOCV to choose the regularization parameter λ for $\mathrm{DKL_{SPL}}$.

6.2.1. Daily Hedging Comparison

For $\mathrm{DKL_{SPL}}$, we present out-of-sample daily hedging performance test result using either **traded** data or **all** data in the database, with each learned model applied to the next test month. [5] Table 8 and Table 9 present the out-of-sample daily hedging Gain ratios for the S&P 500 options from different hedging models from January 2007 to August 2015.

Delta	MV (%)	SABR(%)	LVF(%)	$\mathrm{DKL_{SPL}}$ (%) LOOCV [1] Traded	All
0.1	42.1	39.4	42.6	**47.1**	**48.6**
0.2	35.8	33.4	36.2	**37.8**	**40.0**
0.3	31.1	29.4	30.3	**34.1**	**35.1**
0.4	28.5	26.3	26.7	**32.3**	**32.0**
0.5	27.1	24.9	25.5	**29.3**	**29.4**
0.6	25.7	25.2	25.2	**29.9**	**28.4**
0.7	25.4	24.7	25.8	**29.0**	**26.8**
0.8	24.1	23.5	25.4	**25.9**	24.7
0.9	16.6	17.0	16.9	**17.7**	13.9
Overall	25.7	24.6	25.5	**31.3**	**26.0**

Table 8: S&P 500 Call Option Daily Hedging: bold entry indicating better Gain than methods proposed in [37]

[1] For each month, LOOCV is used to chooses a constant $\lambda \in \{10^7, 10^6, 10^5, 10^4, 10^3, 10^2, 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$.

Delta	MV (%)	SABR(%)	LVF(%)	$\mathrm{DKL_{SPL}}$ (%) LOOCV [1] Traded	All
-0.9	15.1	11.2	-7.4	8.6	13.6
-0.8	18.7	19.6	6.8	6.5	16.7
-0.7	20.3	17.7	9.1	10.6	19.8
-0.6	20.4	16.7	9.2	14.9	**21.0**
-0.5	22.1	16.7	10.8	**22.5**	**23.1**
-0.4	23.8	17.7	12.0	**24.2**	**25.2**
-0.4	27.1	21.7	16.8	**27.7**	**28.3**
-0.2	29.6	25.8	20.6	**30.1**	**30.8**
-0.1	27.5	26.9	17.7	**29.1**	**31.2**
Overall	22.5	19.0	10.2	**23.4**	**23.2**

Table 9: S&P 500 Put Option Daily Hedging: bold entry indicating better Gain than methods proposed in [37]

[1] For each month, LOOCV is used to chooses a constant $\lambda \in \{10^7, 10^6, 10^5, 10^4, 10^3, 10^2, 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$.

[5] Note that, in the OptionMetric Database [44], not all the option data instances are actually traded. For many data instances, the trading volumes are zero. It is not clear to us whether the reported test results in [37] are computed using all data or traded data only.

Table 8 presents hedging comparisons for call options. From the table, we observe that, using either **traded** data or **all** data to test, for the more liquid delta buckets between 0.1 and 0.7, the direct hedging method $\mathrm{DKL}_{\mathrm{SPL}}$ significantly outperforms the other minimum variance hedging methods proposed in [37] and the performance, for the less frequently traded delta buckets 0.9 and 0.8, is similar to methods proposed in [37]. Furthermore, the overall performance from the direct hedging $\mathrm{DKL}_{\mathrm{SPL}}$ method is also better, especially when we only use **traded** data to evaluate the performance. Table 9 presents hedging comparisons for put options. From Table 9, we observe that, using either **traded** data or **all** data to test, for more liquid delta buckets between -0.5 and -0.1, $\mathrm{DKL}_{\mathrm{SPL}}$ performs better than all minimum variance methods. For the less frequently traded delta buckets -0.9, -0.8, -0.7,-0.6, the test Gains for $\mathrm{DKL}_{\mathrm{SPL}}$ are slightly worse than those from the Hull-White MV method. In addition, for put options, the overall performance from the direct hedging $\mathrm{DKL}_{\mathrm{SPL}}$ method is still better using either **all** data or **traded** data to evaluate the performance.

From Table 8 and Table 9, we see that the direct hedging $\mathrm{DKL}_{\mathrm{SPL}}$ method performs better than the other methods for more liquid delta buckets. For the less frequently traded delta buckets, the direct hedging $\mathrm{DKL}_{\mathrm{SPL}}$ method is less effective. This is understandable since sufficiently large training data sets are required to build effective models which have reasonably good out-of-sample performance. For those buckets, the number of traded data instances is too small. In those cases, parametric models can potentially be better choices because the model calibration process for parametric models requires less data.

6.2.2. Performance on Weekly and Monthly Hedging from Directly Hedging Learning DKL_{SPL}

Since the minimum variance hedging methods are based mainly on instantaneous hedging analysis, only daily hedging performance results are presented in [37]. The proposed direct hedging $\mathrm{DKL}_{\mathrm{SPL}}$, in contrast, is suitable also for a longer hedging period, since it uses the variance of the discrete hedging error (assuming zero mean) as the empirical cost.

Table 10 presents weekly and monthly hedging performance from the direct hedging spline kernel learning $\mathrm{DKL}_{\mathrm{SPL}}$ for call and put options respectively. We assume a period of 5-business-day for weekly hedging and a period of 20-business-day for monthly hedging.

For call options, Table 10 shows that $\mathrm{DKL}_{\mathrm{SPL}}$ achieves significant Gain ratios in both weekly and monthly hedging, indicating that $\mathrm{DKL}_{\mathrm{SPL}}$ outperforms implied volatility BS delta hedging. For put options, Table 11 shows that $\mathrm{DKL}_{\mathrm{SPL}}$ similarly achieves significant improvement over the implied volatility BS delta hedging in all buckets.

Delta	$\mathrm{DKL}_{\mathrm{SPL}}$ (%)			
	Weekly		Monthly	
	Traded	All	Traded	All
0.1	38.9	38.3	22.7	24.8
0.2	29.0	26.9	23.5	25.5
0.3	23.5	25.3	24.0	24.6
0.4	20.8	24.3	21.0	20.7
0.5	19.9	22.8	13.5	12.7
0.6	17.3	19.5	14.3	13.5
0.7	16.8	17.7	6.1	7.0
0.8	12.5	12.3	5.3	4.1
0.9	6.2	5.1	4.1	2.3
Overall	20.2	17.1	16.3	12.5

Table 10: S&P 500 Call Options Weekly and Monthly Hedging

[1] For each month, LOOCV is used to chooses a constant $\lambda \in \{10^7, 10^6, 10^5, 10^4, 10^3, 10^2, 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$.

In addition, we observe that Gain ratios are smaller for monthly hedging than weekly hedging and daily hedging. We note that hedging performance for the synthetic dataset in §6.1 also deteriorates for a longer hedging period and introducing δ^{BS} as the additional feature in the feature set #2 has improved hedging performance. We conjecture

Delta	DKL_{SPL} (%)			
	Weekly		Monthly	
	Traded	All	Traded	All
-0.9	10.1	7.3	6.5	5.8
-0.8	18.3	11.5	6.1	7.8
-0.7	20.2	16.3	7.3	11.9
-0.6	20.8	18.4	10.3	9.5
-0.5	22.4	21.2	13.9	12.8
-0.4	21.0	23.9	15.6	16.7
-0.3	22.2	26.1	19.5	13.4
-0.2	20.8	29.7	20.6	18.4
-0.1	19.2	29.1	13.0	19.9
Overall	20.4	20.3	13.5	12.7

Table 11: S&P 500 Put Options Weekly and Monthly Hedging

[1] For each month, LOOCV is used to chooses a constant $\lambda \in \{10^7, 10^6, 10^5, 10^4, 10^3, 10^2, 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$.

that additional features and kernel combinations may lead to further improved hedging performance for longer period discrete hedging for put option and leave this for future investigation.

7. Concluding Remarks

The prevailing approach for option hedging is to first calibrate a parametric model to determine an option value function and the hedging position is computed as the partial derivative of the option function. For example implied volatility BS delta hedging calibrates the option price to the constant BS model. Unfortunately dependence of the model parameter on the underlying occurs from this process, e.g., in [16]. Failure to capture this dependence means that hedging based on the sensitivity of the calibrated option function fails to minimize the variance of the option hedging risk.

Unfortunately, accounting for this model parameter dependence on the underlying is challenging. Minimum variance hedging methods, proposed to correct this issue, are typically based on an option function corresponding to stochastic volatility models, local volatility functions and SABR, see, e.g., [37]. In addition Hull and White [37] study the option market prices and propose a quadratic minimum variance hedging formula based on the BS instantaneous hedging analysis. Specifically Hull and White [37] estimate the quadratic formula from the market option data and illustrate that it yields better performance than minimum variance hedging methods derived from LVF and SABR.

In this paper we argue that dependence of the option model parameters on the underlying generally exists for any parametric option modeling method, including stochastic volatility models. We further illustrate that, even in a nonparametric kernel approach to model the option value function, dependence on the underlying can exist for the estimated kernel parameters. Consequently, using the partial derivatives of the model option function, parametric or nonparametrically estimated, will fail to minimize variance in hedging error, even instantaneously.

In this paper we propose to directly learn nonparametric kernel hedging functions by minimizing variance of the discrete hedging error, bypassing the intermediate step of the option value function estimation. Using synthetic data, we first demonstrate that the proposed direct hedging function learning significantly outperforms hedging based on the sensitivity of the model option function learned nonparametrically. In addition, we demonstrate that spline kernel yields better hedging performance in comparison to that of the RBF kernel.

In addition, using S&P 500 index option data for more than a decade ending in August 31, 2015, we compare hedging performance of the direct spline kernel hedging learning DKL_{SPL} with minimum variance hedging methods, including empirical quadratic formula recently proposed by Hull and White [37], as well as methods based on LVF and SABR implemented in [37]. We demonstrate that, for daily hedging call options, the direct hedging learning method DKL_{SPL} significantly outperforms the quadratic MV method [37], local volatility function, and SABR minimum variance methods for the more liquid delta buckets between 0.1 and 0.7. We also illustrate that, for daily put option hedging in buckets with delta between -0.5 and -0.1, direct hedging spline kernel $\mathrm{DKL}_{\mathrm{SPL}}$ performs better than all minimum variance methods. For the less frequently traded delta buckets, e.g., -0.9, -0.8, -0.7, and -0.6, Gain ratios computed using all data for $\mathrm{DKL}_{\mathrm{SPL}}$ are very close to those from the Hull-White MV method.

We also present weekly and monthly hedging results using the proposed direct spline kernel learning $\mathrm{DKL}_{\mathrm{SPL}}$. We demonstrate that $\mathrm{DKL}_{\mathrm{SPL}}$ achieves significantly positive Gain ratios for weekly hedging and monthly hedging, thus outperforming implied volatility BS delta hedging. Positive Gain ratios are smaller for monthly hedging in comparison to daily and weekly hedging for both synthetic and S&P 500 index option put hedging studies. Since we observe hedging enhancement when using a different feature set which includes the implied volatility BS delta δ^{BS} as a feature, we plan to investigate different feature and kernel constructions to improve hedging performance for longer periods.

The exploratory research in this paper clearly demonstrates the potential role of a market data driven approach for financial derivative modeling and risk management. In addition we discuss a relation between the partial derivatives of implied volatility for put and call under a more general setting than discussion in [16], which can play a critical role in accounting for implied volatility parameter dependence for minimum variance hedging.

References

[1] C. Alexander, A. Kaeck, and L. M. Nogueira. Model risk adjusted hedge ratios. *Journal of Futures Markets*, 29 (11):1021–1049, 2009.

[2] L. Andersen and J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivative Research*, 4:231–262, 2000.

[3] Leif Andersen and Rupert Brotherton-Ratcliffe. The equity option volatility smile: an implicit finite-difference approach. *The Journal of Computational Finance*, 1(2):5–32, 1998.

[4] F. Angelini and S. Herzel. Measuring the error of dynamic hedging: a laplace transform approach. *Journal of Computational Finance*, 13:47–72, 2009.

[5] F. Angelini and S. Herzel. Explicit formulas for the minimal hedging strategy in a martingale case. *Decisions in Economics and Finance*, 33:63–79, 2010.

[6] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. *The Journal of finance*, 52(5):2003–2049, 1997.

[7] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models. *The Journal of finance*, 52(5):2003–2049, 1997.

[8] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. *The Journal of finance*, 52(5):2003–2049, 1997.

[9] Julia Bennell and Charles Sutcliffe. Black–scholes versus artificial neural networks in pricing ftse 100 options. *Intelligent Systems in Accounting, Finance and Management*, 12(4):243–260, 2004.

[10] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973.

[11] Tim Bollerslev, Julia Litvinova, and George Tauchen. Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, 4(3):353–384, 2006.

[12] P. Carr. European put call symmetry. pages 509–537. Cornell University working paper, 1994.

[13] P. Carr, K. Ellis, and V. Gupta. Static hedging of exotic options. *Journal of Finance*, 53:1165–1190, 1998.

[14] T. F. Coleman, Y. Li, and M. Patron. Total risk minimization using Monte-carlo simulations. In John R. Birge and Vadim Linetsky, editors, *Handbook of Financial Engineering*. Elsevier, 2007.

[15] Thomas F. Coleman, Yuying Li, and Arun Verma. Reconstructing the unknown local volatility function. *The Journal of Computational Finance*, 2(3):77–102, 1999.

[16] Thomas F Coleman, Yohan Kim, Yuying Li, and Arun Verma. Dynamic hedging with a deterministic local volatility function model. *The Journal of Risk*, 5(6):63–89, 2001.

[17] Thomas F. Coleman, Yuying Li, and M. Patron. Discrete hedging under piecewise linear risk minimization. *The Journal of Risk*, 5:39–65, 2003.

[18] S. Crépey. Delta-hedging vega risk. *Quantitative Finance*, 4(October):559–579, 2004.

[19] E. Derman and I. Kani. Riding on a smile. *Risk*, 7:32–39, 1994.

[20] E. Derman, I. Kani, and J. Zou. The local volatility surface: Unlocking the information in index option prices. *Financial Analysts Journal*, pages 25–36, 1996.

[21] Bernard Dumas, Jeff Fleming, and Robert E Whaley. Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6):2059–2106, 1998.

[22] B. Dupire. Pricing with a smile. *Risk*, 7:18–20, 1994.

[23] Bruno Dupire et al. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[24] Theodoros Evgeniou, Massimiliano Pontil, and Tomaso Poggio. Regularization networks and support vector machines. *Advances in computational mathematics*, 13(1):1–50, 2000.

[25] H. Föllmer and M. Schweizer. Hedging by sequential regression: An introduction to the mathematics of option trading. *The ASTIN Bulletin*, 1:147–160, 1989.

[26] Kenneth R French, G William Schwert, and Robert F Stambaugh. Expected stock returns and volatility. *Journal of financial Economics*, 19(1):3–29, 1987.

[27] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin, 2001.

[28] René Garcia and Ramazan Gençay. Pricing and hedging derivative securities with neural networks and a homogeneity hint. *Journal of Econometrics*, 94(1):93–115, 2000.

[29] Ramazan Gençay, Aslihan Salih, et al. Degree of mispricing with the black-scholes model and nonparametric cures. *Annals of Economics and Finance*, 4:73–102, 2003.

[30] S. Goutte, N. Oudjane, and F. Russo. Variance optimal hedging for discrete time processes with independent increments, application to electricity markets. *Journal of Computational Finance*, 17:71–111, 2013.

[31] Nikola Gradojevic, Ramazan Gençay, and Dragan Kukolj. Option pricing with modular neural networks. *IEEE transactions on neural networks*, 20(4):626–637, 2009.

[32] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *The Best of Wilmott*, page 249, 2002.

[33] C. He, J. S. Kennedy, T. F. Coleman, P. A. Forsyth, Y. Li, and K. Vetzal. Calibration and hedging under jump diffusion. *Review of Derivative Research*, 9:1–35, 2006.

[34] D. Heath, E. Platen, and M. Schweizer. Numerical comparison of local risk-minimisation and mean-variance hedging. In *Option pricing, interest rates and risk management*, pages 509–537. (ed. E. Jouini, J. Cvitanic and, M. Musiela), Cambridge Univ. Press, 2001b.

[35] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343, 1993.

[36] John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300, 1987.

[37] John Hull and Alan White. Optimal delta hedging for options. *Journal of Banking & Finance*, 82:180–190, 2017.

[38] John C Hull. *Options, futures, and other derivatives*. Pearson Education India, 2006.

[39] James M Hutchinson, Andrew W Lo, and Tomaso Poggio. A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, 49(3):851–889, 1994.

[40] Jens Jackwerth and Mark Rubinstein. Recovering probability distributions from option prices. *The Journal of Finance*, 51(5):1611–1631, 1996.

[41] A. Lewis. Fear of jumps. *Wilmott Magazine*, December:60–67, 2002.

[42] Mary Malliaris and Linda Salchenberger. A neural network model for estimating option prices. *Applied Intelligence*, 3(3):193–206, 1993.

[43] R. Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4: 141–183, 1973.

[44] LLC OptionMetrics. Ivy db file and data reference manual, 2008.

[45] Tapio Pahikkala, Jorma Boberg, and Tapio Salakoski. Fast n-fold cross-validation for regularized least-squares. In *Proceedings of the ninth Scandinavian conference on artificial intelligence (SCAI 2006)*, pages 83–90. Otamedia Oy, Espoo, Finland, 2006.

[46] K. Poulsen R., R. Schenk-Hoppé, and C.-O Ewald. Risk minimization in stochastic volatility models: model risk and empirical performance. *Quantitative Finance*, 9(6):693–704, 2009.

[47] Mark Rubinstein. Implied binomial trees. *The Journal of Finance*, 49(3):771–818, 1994.

[48] M. Schäl. On quadratic cost criteria for option hedging. *Mathematics of Operation Research*, 19(1):121–131, 1994.

[49] M. Schweizer. Variance-optimal hedging in discrete time. *Mathematics of Operation Research*, 20:1–32, 1995.

[50] Vladimir Naumovich Vapnik. *Statistical learning theory*, volume 1.

[51] Grace Wahba. *Spline models for observational data*, volume 59. Siam, 1990.

[52] Jingtao Yao, Yili Li, and Chew Lim Tan. Option price forecasting using neural networks. *Omega*, 28(4):455–466, 2000.

Appendix A. Further Details on Training and Testing Sets

Appendix A.1. Synthetic Data Sets

Method	Daily		Weekly		Monthly	
	Train	Test	Train	Test	Train	Test
IKL	6447	158934	6399	122802	6418	40952
DKL	6127	158934	4854	122802	1821	40952

Table A.12: Average Size of Training and Testing Synthetic Data Sets

Appendix A.2. S&P 500 Data Sets

Delta	Daily			Weekly			Monthly		
	Train	Test (Traded)	Test (All)	Train	Test (Traded)	Test (All)	Train	Test (Traded)	Test (All)
0.1	17350	649	1293	17105	688	1269	13549	495	1079
0.2	12116	426	918	11935	448	901	9744	332	780
0.3	10873	374	850	10715	395	834	8837	295	727
0.4	10947	373	861	10791	394	845	8941	295	739
0.5	13541	458	942	13334	483	924	11275	371	810
0.6	9668	312	1089	9532	233	1069	7655	238	934
0.7	6643	223	1334	6559	238	1309	4867	157	1142
0.8	5478	184	1728	5388	196	1693	3646	119	1459
0.9	4481	145	2486	4396	158	2434	2599	81	1970

Table A.13: Average Sizes of Training and Testing Data Sets for Call Options

Delta	Daily			Weekly			Monthly		
	Train	Test (Traded)	Test (All)	Train	Test (Traded)	Test (All)	Train	Test (Traded)	Test (All)
-0.9	5803	118	1445	5708	114	1412	4073	69	1167
-0.8	6089	121	895	6001	117	879	4676	75	756
-0.7	7885	156	832	7773	152	816	6571	106	708
-0.6	8133	234	850	8028	229	835	6253	171	728
-0.5	12476	455	926	15359	446	909	12929	364	795
-0.4	15100	531	1071	14903	521	1051	12584	427	919
-0.3	17052	622	1300	16823	610	1276	14026	494	1111
-0.2	14274	848	1742	13461	842	1707	12116	663	1477
-0.1	24038	1469	3235	22614	1462	3172	19597	1102	2696

Table A.14: Average Sizes of Training and Testing Data Sets for Put Options