



**HAL**  
open science

# Linear regression for currency European call option pricing in incomplete markets

Ahmad W. Bitar

► **To cite this version:**

Ahmad W. Bitar. Linear regression for currency European call option pricing in incomplete markets. 2025. hal-04815308v2

**HAL Id: hal-04815308**

**<https://utt.hal.science/hal-04815308v2>**

Preprint submitted on 5 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Linear regression for currency European call option pricing in incomplete markets

Ahmad W. Bitar

Laboratory of Computer Science and Digital Society, Université de Technologie de Troyes, 12 Rue Marie Curie, Troyes, 10300, France.

Corresponding author(s). E-mail(s): [ahmad.bitar@utt.fr](mailto:ahmad.bitar@utt.fr);

## Abstract

The Least squares is the traditional regression technique for pricing European options in incomplete markets. However, the least squares is quite sensitive to even a single outlier in the data, and thus the predicted option price may potentially deviate from the true unknown one. To alleviate the problem of outliers, this paper aims to develop two different option pricing prediction strategies based mainly on the idea of robust linear regression. The proposed robust techniques are evaluated on numerical data, the results of which demonstrate their effectiveness for European call option pricing on exchange rates.

**Keywords:** Exchange rates, Trinomial model, European option pricing, Least squares, Robust linear regression

## 1 Introduction

The exchange rate ([Ehrmann and Fratzscher \(2005\)](#)) can be defined as the price of the foreign currency relative to the domestic one. It mainly affects the competitiveness of any country's goods and services and thus has a huge impact on its foreign trade, as well as its citizens' purchasing power. Since 1976, most of the top economies in the world have adopted a floating exchange rate regime, by which it is determined by supply and demand in the market. When the exchange rate increases (or decreases), it implies an appreciation (or depreciation) of the foreign currency with respect to the domestic one. For example, it has been observed that the euro has shown a depreciation of about more than 18% against the US dollar since 2013 and especially just after the start of the Russian-Ukraine war on February 24, 2022. In addition, the Lebanese

pound had a dramatic depreciation against the United States dollar mainly due to several political factors especially after the COVID-19 pandemic and the Beirut port non-nuclear explosion on August 04, 2020.

Dealing with such huge fluctuations of the exchange rate can be alleviated by buying call European options (Orlin Grabbe (1983)). The latter is a contract that gives a domestic party the right, but not the obligation, to buy a defined amount of a particular foreign currency only at maturity and at a price (called strike) specified at the initial date of the contract. At maturity, there is no need for any mathematical formulation to price a call option since its value would simply be equal to  $(\text{Exchange rate at maturity} - \text{Strike})_+$  and this is well known as PAYOFF. However, how to price options at any time before maturity? The birth of mathematical finance has essentially marked with the Ph.D. thesis of Louis Bachelier in 1900 (Bachelier (1900)) in which he developed an option pricing closed-form formula under the hypothesis that the asset prices increments are independent and follow an arithmetic Brownian motion, which can allow negative prices. The Bachelier option pricing model has been forgotten as a part of history until it gained attention again especially in the twenty-first century due to several serious economic crises (e.g., the negative interest rates are observed in some developed countries after the 2008 global financial crisis, the COVID-19 pandemic which led to lockdown worldwide, and an extremely sharp drop in the global demand for oil).

Since then, several notable contributions for pricing options, under the difficulty of estimating certain non-observable parameters, have been developed (Sprenkle (1961), Rosett (1963), Ayres (1967), Boness (1964), Samuelson (2015), Samuelson and Merton (1969), Merton (1973)). However, one of the biggest influence finance had on operation markets started when Myron Scholes met Fisher Black in the fall of 1968 and began working together to discover one of the most influential mathematical models that accurately prices options<sup>1</sup> (Black and Scholes (1973)). More precisely, they have developed a parametric non-arbitrage option pricing model with a closed-form formula that accurately prices options under the assumption that the market is complete. However, this model assumes a very specific dynamics of the exchange rates and is thus not robust.

To move from parametric to robust European option pricing, the one-period discrete-time binomial non-arbitrage pricing model was developed (Cox et al (1979)). Basically, it assumes that the exchange rate at maturity (denoted as  $X_1$ ) can only increase or decrease with respect to its value at the initial date (denoted as  $X_0$ ). After creating a self-financing hedging portfolio and equating it with the price of the call option at maturity, this yields a linear system of two equations with two unknowns that correspond to the (unknown) delta hedging portfolio coefficients, which can simply be calculated with a closed-form formula. However, the binomial model does not reflect the real market, since it considers only two possible directions of the exchange rate (that is, it either goes up or down).

For a more realistic option pricing model in the market, it would be more important to further consider the possibility that the exchange rate,  $X_1$ , may remain the same

---

<sup>1</sup>See the online conversation (Scholes (2018)) with Myron Scholes, who briefly explained his work with Fisher Black.

with respect to  $X_0$ . More precisely,  $X_1$  is assumed to behave in three possible directions with respect to  $X_0$ : increases (with a factor  $e^u$ ,  $u > 0$ ), remains the same or decreases (with a factor  $e^d$ ,  $d < 0$ ). This is well known as the one-period trinomial option pricing model (Boyle (1986); Zhu et al (2022); Johan (2016)), and by which equating the self-financing hedging portfolio to the call option price at maturity results in an overdetermined linear system of three equations with two unknowns, which again correspond to the delta hedging coefficients that need to be estimated very efficiently. As the solution of the constructed linear system can not be guaranteed in general (e.g. more equations than unknowns), the trinomial option pricing model is said to be incomplete, that is, there exist an infinite number of equivalent martingale measures.

Pricing European call options under this incompleteness can be achieved in different scenarios, including adding a second risky asset to the trinomial market model and which can help to replicate the option perfectly (Pascucci and Runggaldier (2012)) or finding a hedging strategy that is close to being perfect. The latter scenario means finding an approximation to the hedging portfolio and can be done simply using the least squares machine learning technique (Takahashi (2000); Johan (2016)). The latter is derived with the idea of minimizing the sum of the squared “errors”, that is, to adjust the unknown delta hedging coefficients such that the sum of the squares of the differences between the original and the predicted PAYOFF values is minimized. As a result, the least squares estimates is the regression equivalent to the sample mean, which is well known to be very sensitive to the influential observations (outliers)<sup>2</sup>, and therefore often fail to provide good fits to most of the data. In addition, the least squares is derived under the Gaussian assumption, and this ensures its lack of robustness against the outliers.

In this paper, we only focus on the second scenario for pricing European options in incomplete markets, that is, by building a self-financing hedging portfolio which does not perfectly replicate the option price. As the least squares technique fails with the presence of outliers, we alleviate this challenge via robust linear regression techniques Bitar (2024). More precisely, we first aim to detect the nature of the outliers which help us to manually remove them from the data and then simply apply the least squares on the remaining observations (that is, after the outliers have been removed) to estimate the (unknown) delta hedging portfolio coefficients. However, an alternative option pricing strategy would be to take advantage of some linear regression  $M$ -estimators capable of automatically identifying outliers during the estimation of the delta hedging portfolio coefficients.

It should also be noted that the two different possible positive factors  $e^u$  and  $e^d$  by which the exchange rate  $X_1$  increases or decreases with respect to  $X_0$ , respectively, are usually not known in the real market, and thus a single specific value for both parameters  $u$  and  $d$  can be chosen manually to price the corresponding option. As a result, only three possible observations were always used to construct the linear regression model, and thus using a very little number of observations (for example, three) might not be sufficient to accurately predict the option price under any regression model. To ensure reliable prediction results, generating enough observations may alleviate this

---

<sup>2</sup>By an influential observation we mean the observations (outliers) that deviate from the general pattern of the data.

challenge. However, it is not surprising that the price of the option greatly depends on the selection of both  $u$  and  $d$ . Hence, instead of selecting a single specific value for each of these two parameters, our aim is to price options under a set of  $n$  possible values for both  $u$  and  $d$ , with  $n \geq 1$ . This will help us generate  $2n + 1$  observations to build the linear regression model and can lead to an accurate estimation of the delta hedging coefficients especially with such a robust linear regression  $M$ -estimator. In addition, pricing the call option will be much less sensitive to such a single specific selection of the two parameters  $u$  and  $d$ .

This paper is structured in the following manner. In Section 2.1 a brief presentation is provided on how the standard trinomial option pricing model on exchange rates can be applied and how the delta hedging portfolio coefficients are estimated using the least squares technique. A numerical example is also presented to evaluate the least squares in the presence of outliers. The two proposed robust option pricing prediction strategies are presented in Section 3. Section 4 presents the experiments to gauge the effectiveness of the proposed prediction strategies for the pricing of the European call option on exchange rates. The paper ends with a summary of the work and some directions for future work.

### Summary of main notations

Throughout this paper, we represent vectors in lowercase boldface letters and matrices in uppercase boldface letters. The notation  $(\cdot)^T$ ,  $(\cdot)_+$ ,  $|\cdot|$ , and  $\widehat{(\cdot)}$  stand for the transpose, positive part, absolute value, and estimated/predicted value, respectively. A variety of vector norms will be used. For example,  $\boldsymbol{\delta}$  is a vector, and  $\delta_j$  is the  $j^{\text{th}}$  element. The vector  $l_2$  and  $l_1$ - norms are defined by  $\|\boldsymbol{\delta}\|_2 = \sqrt{\sum_j \delta_j^2}$  and  $\|\boldsymbol{\delta}\|_1 = \sum_j |\delta_j|$ , respectively.

## 2 The trinomial option pricing model on exchange rates

In this section, we first overview in Subsection 2.1 the one-period standard trinomial behavior of the exchange rate  $X_1$  with respect to  $X_0$ , as well as the corresponding arbitrage-free condition. Second, in Subsection 2.2, we present the self-financing replicating portfolio strategy that can be considered for the prediction of the price of the European call option on exchange rates. Third, in Subsection 2.3 we briefly describe how the option price can be predicted via the least squares machine learning technique. Finally, in Subsection 2.4, we demonstrate by a simple numerical example the difficulty of accurately predicting the price of the European call option on exchange rates.

## 2.1 The one-period standard trinomial model

The one-period trinomial pricing model assumes that the exchange rate  $X_1$  is characterized by three jump behavior as follows:

$$X_1 = \begin{cases} X_0 e^u & \text{with probability } p_u > 0 \\ X_0 & \text{with probability } p_0 = 1 - p_u - p_d > 0, \\ X_0 e^d & \text{with probability } p_d > 0 \end{cases}$$

where  $u > 0$ ,  $d < 0$ ,  $p_u, p_d \in (0, 1)$ , and  $p_u + p_d < 1$ .

Assume a domestic cash capitalization process  $B_1 = B_0 e^{r_d}$ , where  $B_0 > 0$  and  $r_d > 0$  is a domestic risk-free interest rate. To prove that  $p_u > 0$  and  $p_d > 0$ , the condition:

$$\frac{e^u - e^{r_d}}{e^u - 1} > p_0 > 0, \quad (1)$$

should be satisfied (for more details, see theorem 2.1.2 in Johan (2016)). As  $p_u + p_0 + p_d = 1$ , one can directly conclude that  $p_u + p_d < 1$ . Hence, the couple  $(p_u, p_d)$  defines a probability if and only if the condition (1) is satisfied, and thus  $u > r_d$ .

## 2.2 The self-financing hedging portfolio

The replicating portfolio can be defined as follows<sup>3</sup>:

At initial date :

$$V_0 = \delta_1 X_0 + \delta_2 B_0,$$

At maturity :

$$\begin{aligned} V_1 &= \delta_1 e^{r_f} X_1 + \delta_2 B_1 \\ &= \begin{cases} \delta_1 e^{r_f} (X_0 e^u) + \delta_2 (B_0 e^{r_d}) = (X_0 e^u - K)_+ = h_1 \\ \delta_1 e^{r_f} (X_0) + \delta_2 (B_0 e^{r_d}) = (X_0 - K)_+ = h_2, \\ \delta_1 e^{r_f} (X_0 e^d) + \delta_2 (B_0 e^{r_d}) = (X_0 e^d - K)_+ = h_3 \end{cases} \end{aligned}$$

where  $\delta_1$  is the number of units of the foreign currency bought at initial date,  $\delta_2$  is the number of units of risk-free investment in the domestic currency,  $r_f > 0$  is a foreign risk-free interest rate,  $K > 0$  is the strike price,  $h_1$ ,  $h_2$ , and  $h_3$ , represent the PAYOFF value when the exchange rate  $X_1$  increases, remains the same, or decreases with respect to  $X_0$ , respectively. As we can observe, the portfolio at maturity yields an over-determined linear system of three equations with two unknowns ( $\delta_1$  and  $\delta_2$ ). Based on both the values of  $X_0$  and  $K$ , one can assume the following three possible

---

<sup>3</sup>For more details, see the class notes in <https://web.ma.utexas.edu/users/mcudina/m339w-jan-27-2020.pdf>.

different scenarios for the PAYOFF values  $h_1$ ,  $h_2$ , and  $h_3$ :

$$\begin{aligned} \text{if } X_0 \leq K < X_0 e^u : \\ h_1 > 0, h_2 = 0, h_3 = 0. \end{aligned}$$

$$\begin{aligned} \text{if } X_0 e^d < K < X_0 : \\ h_1 > 0, h_2 > 0, h_3 = 0. \end{aligned}$$

$$\begin{aligned} \text{if } K \leq X_0 e^d : \\ h_1 > 0, h_2 > 0, h_3 > 0. \end{aligned}$$

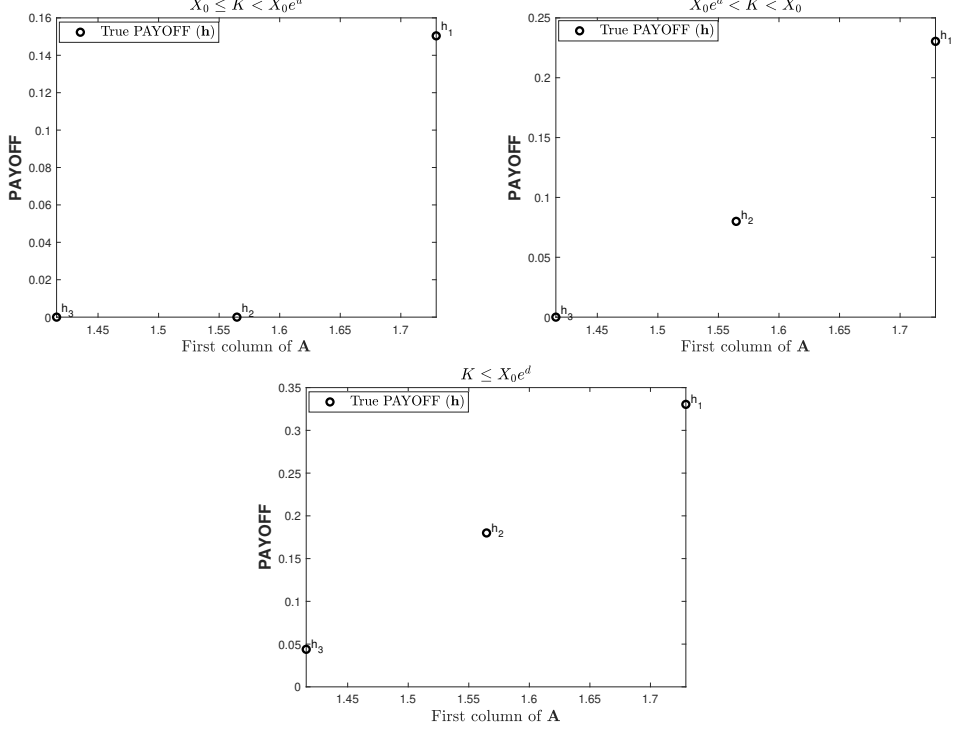
Figure 1 shows three plots of the three observations  $h_1$ ,  $h_2$ , and  $h_3$  when  $X_0 = 1.43$ ,  $u = 0.1$ , and  $d = -u$ . It is important to note that when  $X_0 e^d < K < X_0$ , the observation  $h_2$  will be strictly positive and will approach the line that joins the two observations  $h_1$  and  $h_3$  until they coincide all together once  $K$  becomes lower than or equal to  $X_0 e^d$  (that is,  $K \leq X_0 e^d$ ). In the latter situation, the least squares technique will obviously work very well to estimate the (unknown) delta hedging coefficient vector  $\boldsymbol{\delta}$ .

### 2.3 delta hedging via the least squares

We denote the corresponding vector of dependent variables by  $\mathbf{h} = [h_1, h_2, h_3]^T \in \mathbb{R}_+^3$ . In general, the vector  $\mathbf{h}$  does not span the column space of

$$\text{the matrix of independent variables } \mathbf{A} = \begin{bmatrix} - & - & \mathbf{a}_1^T & - & - \\ - & - & \mathbf{a}_2^T & - & - \\ - & - & \mathbf{a}_3^T & - & - \end{bmatrix} = \begin{bmatrix} e^{rf} X_0 e^u & B_0 e^{rd} \\ e^{rf} X_0 & B_0 e^{rd} \\ e^{rf} X_0 e^d & B_0 e^{rd} \end{bmatrix} \in \mathbb{R}_+^{3 \times 2},$$

and hence, a solution for  $\boldsymbol{\delta} = [\delta_1, \delta_2]^T \in \mathbb{R}^2$  does not exist if one forces  $\mathbf{h}$  to be exactly equal to  $\mathbf{A}\boldsymbol{\delta}$ , that is,  $\mathbf{h} = \mathbf{A}\boldsymbol{\delta}$ . As a result, the trinomial model is said to be incomplete, that is, a fair price of European options cannot be uniquely defined. Hence, the standard European option may not be replicable by investing only in the underlying exchange rate and the risk-free bond. However, since it is generally impossible to solve the linear system  $\mathbf{h} = \mathbf{A}\boldsymbol{\delta}$ , it is possible to project  $\mathbf{h}$  into the column space of  $\mathbf{A}$ . Taking the assumption that  $\mathbf{h} \sim \mathcal{N}(\mathbf{A}\boldsymbol{\delta}, \sigma^2 \mathbf{I}_3)$  with  $\sigma > 0$  and  $\mathbf{I}_3$  is an identity matrix of size  $3 \times 3$ , the regression coefficient vector  $\boldsymbol{\delta} \in \mathbb{R}^2$  can be simply estimated by maximizing the log-likelihood under the Gaussian distribution



**Fig. 1:** Example of the three PAYOFF scenarios when  $X_0 = 1.43$ ,  $u = 0.1$ , and  $d = -u$ .

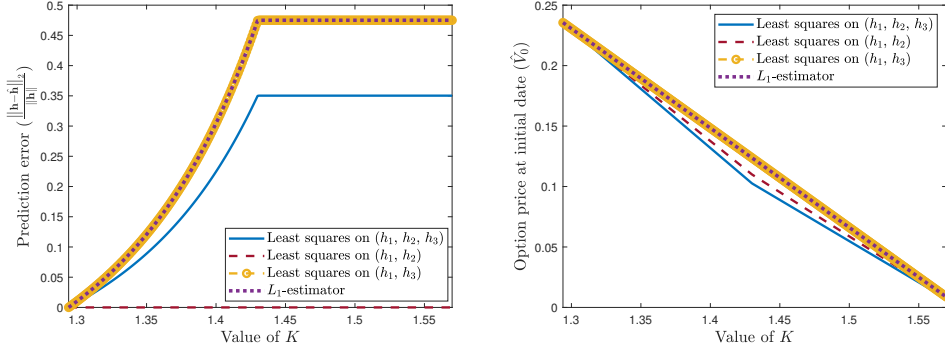
assumption. More precisely, we have:

$$\begin{aligned}
\hat{\boldsymbol{\delta}} &= \underset{\boldsymbol{\delta}}{\operatorname{argmax}} \left\{ -\frac{6}{2} \log(2\pi) - \frac{3}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^3 (h_i - \mathbf{a}_i^T \boldsymbol{\delta})^2 \right\} \\
&\equiv \underset{\boldsymbol{\delta}}{\operatorname{argmin}} \left\{ \sum_{i=1}^3 (h_i - \mathbf{a}_i^T \boldsymbol{\delta})^2 \right\} \\
&= \underset{\boldsymbol{\delta}}{\operatorname{argmin}} \left\{ \|\mathbf{h} - \mathbf{A}\boldsymbol{\delta}\|_2^2 \right\}.
\end{aligned} \tag{2}$$

It is well known that the domain  $\mathbb{R}^2$  is closed and convex. In addition, the objective function to minimize in (2),  $\|\mathbf{h} - \mathbf{A}\boldsymbol{\delta}\|_2^2$ , is inf-compact (since it is coercive and lower semi-continuous) and strictly convex. As a result, the estimation of  $\boldsymbol{\delta}$ , that is,  $\hat{\boldsymbol{\delta}}$ , exists and is unique with an analytical formula defined as:

$$\hat{\boldsymbol{\delta}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{h} = \begin{bmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \end{bmatrix} \in \mathbb{R}^2.$$





**Fig. 2:** From left to right: PAYOFF prediction error  $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$ , the predicted option price  $\hat{V}_0$ .

## 2.4 European option pricing via the least squares

Once the delta hedging coefficient vector  $\delta$  is estimated via the least squares technique (see subsection 2.3), the true (known) option price at maturity (that is, the PAYOFF vector  $\mathbf{h}$ ) can thus be predicted as:

$$\hat{\mathbf{h}} = \mathbf{A}\hat{\delta}.$$

Based on the optimization (minimization) problem in (2), we can obviously observe that the least squares is derived with the idea of minimizing the sum of the squared “errors”, that is, to adjust the unknown delta hedging coefficients so that the sum of the squares of the differences between the original and predicted PAYOFF values is minimized. As a result, the least squares estimate is the regression equivalent to the sample mean, which is well known to be very sensitive to outliers.

As the option price should always be positive, it would be more realistic to consider the rectified linear unit version of  $\hat{\mathbf{h}}$ . That is:

$$\hat{\mathbf{h}} = \left( \hat{\mathbf{h}} \right)_+.$$

Once  $\delta$  is estimated, the true (unknown) option price at initial date (that is,  $V_0$ ) can be predicted as:

$$\hat{V}_0 = \hat{\delta}_1 X_0 + \hat{\delta}_2 B_0. \quad (3)$$

### Does the least squares have efficiency in predicting the option price?

We study, using a simple numerical example, the performance of the least squares in predicting the option price at maturity (that is, PAYOFF) as well as at the initial

date (that is,  $V_0$ ) for different values of  $K \in ]X_0e^d, X_0e^u[$ . Fix  $X_0 = 1.43$ ,  $u = 0.1$ , and  $d = -u$ . The value for  $X_1$  under the trinomial model will behave as follows:

$$X_1 = \begin{cases} X_0e^u \approx 1.58 & \text{with probability } p_u > 0 \\ X_0 = 1.43 & \text{with probability } p_0 > 0. \\ X_0e^d \approx 1.29 & \text{with probability } p_d > 0 \end{cases}$$

We consider  $r_f = 9\%$ ,  $r_d = 8\%$ , and  $B_0 = 1$ . We aim to compare between four different option pricing prediction models:

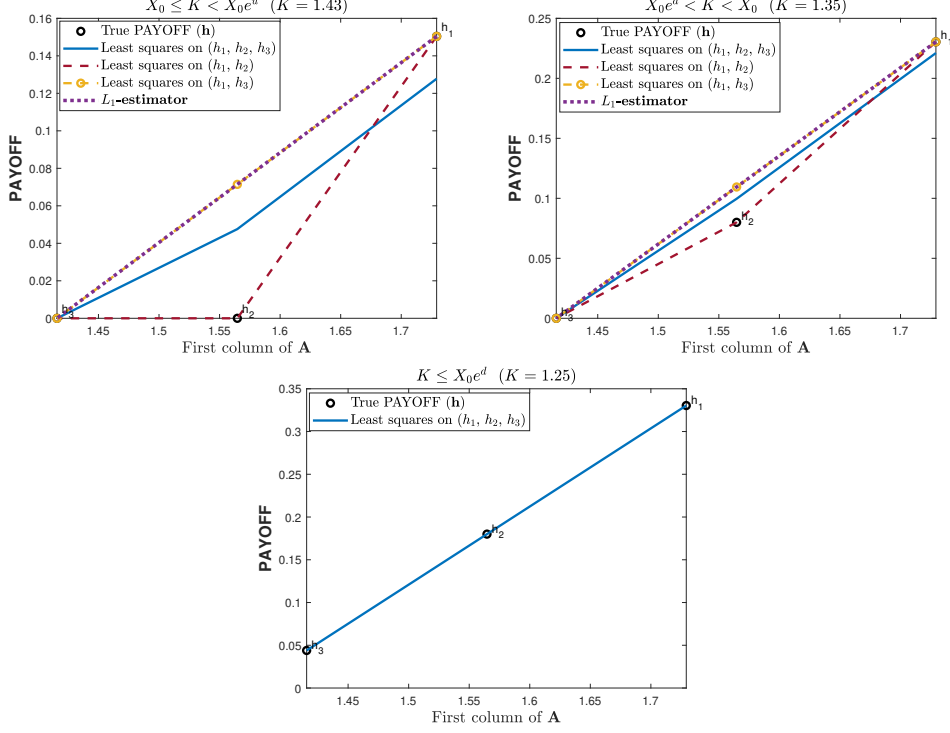
1. **Least squares on  $(h_1, h_2, h_3)$ :** the least squares is applied on the whole three observations  $h_1$ ,  $h_2$ , and  $h_3$ ;
2. **Least squares on  $(h_1, h_2)$ :** we assume that the observation  $h_3$  badly affects the performance of the least squares on predicting the true PAYOFF vector  $\mathbf{h}$ . Thus, we propose to manually delete the observation  $h_3$ , and then apply the least squares on the remaining part of data (that is,  $h_1$  and  $h_2$ );
3. **Least squares on  $(h_1, h_3)$ :** We assume that the observation  $h_2$  badly affects the performance of the least squares on predicting the true PAYOFF vector  $\mathbf{h}$ . Thus, we propose to manually delete the observation  $h_2$ , and then apply the least squares on the remaining part of data (that is,  $h_1$  and  $h_3$ );
4.  **$L_1$ -estimator:** the  $l_2$ -norm in the least squares objective function in (2) is replaced by the  $l_1$ -norm. More precisely, the vector  $\delta$  is estimated as follows:

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}} \left\{ \sum_{i=1}^3 |h_i - \mathbf{a}_i^T \delta| \right\} = \underset{\delta}{\operatorname{argmin}} \{ \|\mathbf{h} - \mathbf{A}\delta\|_1 \}. \quad (4)$$

Unlike the least squares, there is no explicit expression for the minimizer of (4) since the  $l_1$ -norm is not differentiable at zero. However, recursive algorithms can solve it (e.g., Matlab software for disciplined convex programming, CVX Matlab).

Figure 2 evaluates the four aforementioned option pricing prediction methods in terms of the PAYOFF prediction error  $\frac{\|\mathbf{h} - \hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$  and the predicted option price at the initial date  $\hat{V}_0$  for different values of  $K \in ]1.29, 1.58[$  in steps of  $10^{-2}$ . We can observe that the “Least squares on  $(h_1, h_2)$ ” achieves the lowest PAYOFF prediction error for all values of  $K$ , and thus, we expect that its estimated delta hedging coefficient vector  $\delta$  yields a precise prediction of the option price at the initial date.

Figure 3 shows the difference between  $\mathbf{h}$  and  $\hat{\mathbf{h}}$  when  $K = 1.43$  ( $X_0 \leq K < X_0e^u$ ),  $K = 1.35$  ( $X_0e^d < K < X_0$ ) and  $K = 1.25$  ( $K \leq X_0e^d$ ), respectively. We can obviously observe that when  $K \leq X_0e^d$ , the three observations  $h_1$ ,  $h_2$ , and  $h_3$  will always be located on the same line, and so the “least squares on  $(h_1, h_2, h_3)$ ” will behave perfectly (that is, the predicted PAYOFF observations coincide with the true ones). However, when  $X_0 \leq K < X_0e^u$  and  $X_0e^d < K < X_0$ , it is the model “Least squares on  $(h_1, h_2)$ ” that has the lowest PAYOFF prediction error and its predicted PAYOFF observations coincide with the



**Fig. 3:** The difference between  $\mathbf{h}$  and  $\hat{\mathbf{h}}$  when  $X_0 \leq K < X_0 e^u$ ,  $X_0 e^d < K < X_0$ , and  $K \leq X_0 e^d$ .

true ones. We can then consider that applying the least squares after removing the observation  $h_3$  can potentially decrease the prediction error compared to the case if the least squares is applied on all three observations. If we further assume that the bulk of the data is only formed by  $h_1$  and  $h_2$ , the regression line for the model “Least squares on  $(h_1, h_2, h_3)$ ” seems to be skewed towards the observation  $h_3$ , and thus the least squares when applied on the whole three observations has considered  $h_3$  as an outlier.

In addition, it is well known that the optimization (minimization) problem in (4) is obtained after maximizing the log-likelihood under the Laplace distribution, which is considered a heavy-tailed distribution (that is, having a density function with tails that tend more slowly to zero compared to the normal density tails). As the total number of observations (for example, three) is odd, the derivative of the objective function in (4) with respect to  $\mathbf{a}_i^T \boldsymbol{\delta}$  exists and is everywhere except for the observations  $h_1, h_2, h_3$ , and it is equal to  $-\sum_{i=1}^3 \text{sign}(h_i - \mathbf{a}_i^T \boldsymbol{\delta})$  which can never be equal to zero. Thus, the minimizer must occur at one of the points  $h_1, h_2, h_3$  where the objective function in (4) is not differentiable. However, it is a continuous function and decreases when  $\mathbf{a}_i^T \boldsymbol{\delta} < h_{\frac{3+i}{2}}$  and increases when  $\mathbf{a}_i^T \boldsymbol{\delta} > h_{\frac{3+i}{2}}$ . Hence, the minimizer is given by  $\mathbf{a}_i^T \hat{\boldsymbol{\delta}} = h_{\frac{3+i}{2}}$ . As a result, replacing the  $l_2$ -norm in the objective function in (2)

by the  $l_1$ -norm is the regression equivalent to the sample median. The latter is well known to be robust to outliers, but this is paid for by an increase in variance (“a loss in efficiency”) in the exact Gaussian distribution. However, when applying the model “ $L_1$ -estimator”, we can observe from Figure 3 that its regression line is equivalent to the one of the model “Least squares on  $(h_1, h_3)$ ”. Hence, it seems that model “ $L_1$ -estimator” has considered that the general pattern of the data is only formed by the observations  $h_1$  and  $h_3$ , whereas  $h_2$  is the observation that deviates from it (that is, it is considered as an outlier). As a result, both the models “Least squares on  $(h_1, h_3)$ ” and “ $L_1$ -estimator” obtained the largest PAYOFF prediction error, as can be seen in Figure 2.

### 3 Main contribution

In this section, we first present in subsection 3.1 a methodology that allows us to construct an option pricing model that considers  $n \geq 1$  possible values for each of the parameters  $u > 0$  and  $d < 0$ , and hence allows us to construct a linear regression model with  $2n + 1$  observations. Recall that our main purpose for pricing options under a large set of values for  $u$  and  $d$  is that we want to dramatically decrease the sensitivity of the option price on any such a single specific selection of the parameters as has been detailed in the standard trinomial option pricing model (see Subsection 2.1). In addition, we evaluate the least squares in a regression with  $2n + 1$ ,  $n \geq 1$ , observations to detect the nature of the influential observations (that is, the outliers). Once these outliers are identified, one can simply delete them and then re-apply the least squares on the remaining part of the data. Finally, instead of manually deleting the observations that are identified as outliers and then applying the least squares on the remaining part of the data, in Subsection 3.2 we aim to exploit the idea behind the robust  $M$ -estimators for linear regression analysis (Maronna (2019); Huber (1964)). This will allow for an accurate estimate of the (unknown) delta hedging coefficient vector  $\delta$  with an automatic deletion of the identified outliers.

#### 3.1 Increasing the number of observations from 3 to $2n + 1$ , $n \geq 1$

In order to predict the option price at initial date, that is,  $V_0$ , independently from any such a single specific selection value for both  $u > 0$  and  $d < 0$ , we propose to replace  $u$  and  $d$  by a set of  $n \geq 1$  possible values  $\{u_l\}_{l \in [1, n]}$  with  $u_1 > u_2 > \dots > u_n > r_d$ , and  $\{d_l\}_{l \in [1, n]}$  with  $d_n < d_{n-1} < \dots < d_1 < 0$ , respectively, with  $n \in \mathbb{N}$ . As a result, one has to expect  $2n + 1$  different values for  $X_1$ , and thus, the possibility to construct an over-determined linear system of  $2n + 1$  equations with only 2 unknowns. More

precisely, we have:

$$X_1 = \begin{cases} X_0 e^{u_1} & \text{with probability } p_{u_1} > 0 \\ \vdots \\ X_0 e^{u_n} & \text{with probability } p_{u_n} > 0 \\ X_0 & \text{with probability } p_0 > 0, \\ X_0 e^{d_1} & \text{with probability } p_{d_1} > 0 \\ \vdots \\ X_0 e^{d_n} & \text{with probability } p_{d_n} > 0 \end{cases}$$

with

$$p_0 = 1 - p_{u_1} - \cdots - p_{u_n} - p_{d_1} - \cdots - p_{d_n} > 0.$$

$$V_1 = \begin{cases} \delta_1 e^{r_f} (X_0 e^{u_1}) + \delta_2 B_0 e^{r_d} = (X_0 e^{u_1} - K)_+ = h_1 \\ \delta_1 e^{r_f} (X_0 e^{u_2}) + \delta_2 B_0 e^{r_d} = (X_0 e^{u_2} - K)_+ = h_2 \\ \vdots \\ \delta_1 e^{r_f} (X_0 e^{u_n}) + \delta_2 B_0 e^{r_d} = h_n = (X_0 e^{u_n} - K)_+ = h_n \\ \boxed{\delta_1 e^{r_f} X_0 + \delta_2 B_0 e^{r_d} = (X_0 - K)_+ = h_{n+1}} \\ \delta_1 e^{r_f} (X_0 e^{d_1}) + \delta_2 B_0 e^{r_d} = (X_0 e^{d_1} - K)_+ = h_{n+2} \\ \delta_1 e^{r_f} (X_0 e^{d_2}) + \delta_2 B_0 e^{r_d} = (X_0 e^{d_2} - K)_+ = h_{n+3} \\ \vdots \\ \delta_1 e^{r_f} (X_0 e^{d_n}) + \delta_2 B_0 e^{r_d} = (X_0 e^{d_n} - K)_+ = h_{2n+1} \end{cases}$$

We can observe that when  $n = 1$ , that is, a single choice for  $u > r_d > 0$  and  $d < 0$ , we return back to the standard (one-period) trinomial option pricing model already described in subsection 2.1. Based on both the values of  $X_0$  and  $K$ , one can expect the following two possible different scenarios for the PAYOFF values  $h_1, h_2, \dots, h_{2n+1}$ :

if  $X_0e^{d_n} < K < X_0e^{u_1}$  :

$$h_1 > 0, h_2 \geq 0, \dots, h_n \geq 0, h_{n+1} \geq 0, h_{n+2} \geq 0, \dots, \\ h_{2n} \geq 0, h_{2n+1} = 0.$$

if  $K \leq X_0e^{d_n}$  :

$$h_1 > 0, h_2 \geq 0, \dots, h_n \geq 0, h_{n+1} \geq 0, h_{n+2} \geq 0, \dots, \\ h_{2n} \geq 0, h_{2n+1} = 0.$$

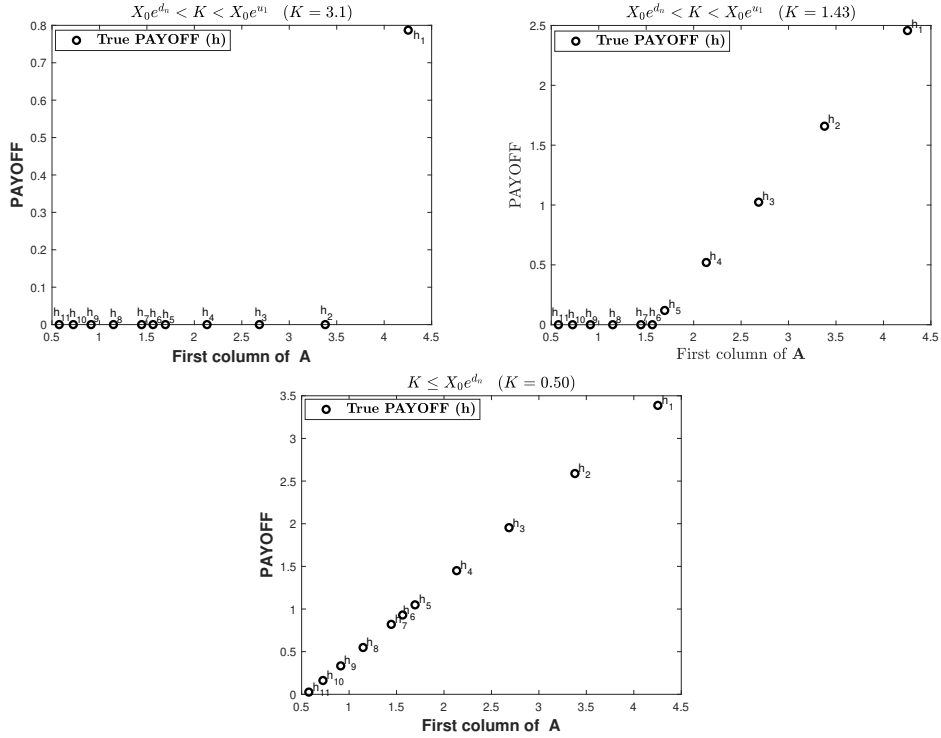
Figure 4 shows three different examples of these two different possible scenarios for the PAYOFF values when  $X_0 = 1.43$ ,  $n = 5$ , and  $K \in \{0.50, 1.43, 3.1\}$ . We assume  $\epsilon = 10^{-6}$ ,  $r_d = 8\%$  and a set of equally spaced 5 possible values  $\{u_i\}_{i \in [1, 5]}$ , with  $u_1 > \dots > u_5 > r_d$ . We choose  $u_5 = r_d + \epsilon$ , and  $u_1 = 1$ . We also fix  $d_5 = -u_1$ ,  $d_4 = -u_2$ ,  $d_3 = -u_3$ ,  $d_2 = -u_4$ , and  $d_1 = -u_5$ . The value for  $X_1$  will then behave as follows:

$$X_1 = \begin{cases} X_0e^{u_1} = 3.8871 & \text{with probability } p_{u_1} > 0 \\ X_0e^{u_2} = 3.0884 & \text{with probability } p_{u_2} > 0 \\ X_0e^{u_3} = 2.4538 & \text{with probability } p_{u_3} > 0 \\ X_0e^{u_4} = 1.9496 & \text{with probability } p_{u_4} > 0 \\ X_0e^{u_5} = 1.5491 & \text{with probability } p_{u_5} > 0 \\ X_0 = 1.43 & \text{with probability } p_0 > 0. \\ X_0e^{d_1} = 1.3200 & \text{with probability } p_{d_1} > 0 \\ X_0e^{d_2} = 1.0488 & \text{with probability } p_{d_2} > 0 \\ X_0e^{d_3} = 0.8333 & \text{with probability } p_{d_3} > 0 \\ X_0e^{d_4} = 0.6621 & \text{with probability } p_{d_4} > 0 \\ X_0e^{d_5} = 0.5260 & \text{with probability } p_{d_5} > 0 \end{cases}$$

We can observe from Figure 4 that as  $K$  decreases in value, less are the PAYOFF observations that are zeros, and hence, the whole observations  $h_1, h_2, \dots, h_{11}$  will locate on the same line once  $K$  becomes less than or equal to  $X_0e^{d_5}$  (that is,  $K \leq X_0e^{d_5}$ ).

### What happens to the least squares with $2n + 1$ , $n \geq 1$ , observations?

Using the same settings as in the example just above (that is, when  $n = 5 \implies 11$  observations in total), we aim to compare between three options pricing prediction models:

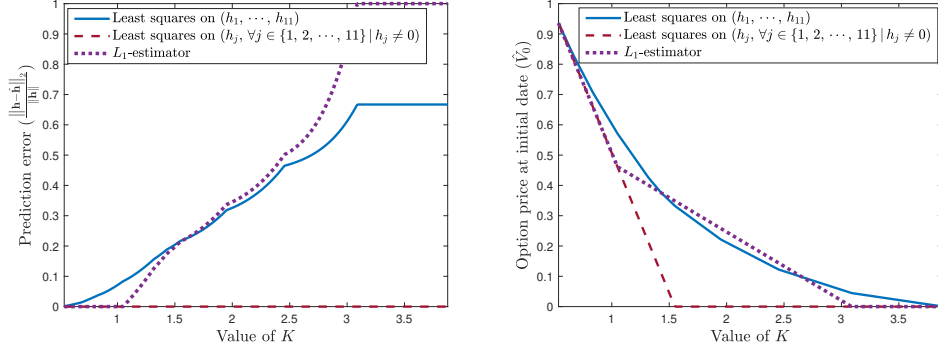


**Fig. 4:** Example of the three PAYOFF scenarios when  $X_0 = 1.43$  and  $n = 5$ . From left to right:  $K = 3.1$ ,  $K = 1.32$ ,  $K = 0.50$ .

1. **Least squares on  $(h_1, \dots, h_{11})$ :** we directly apply the least squares on the whole 11 observations;
2. **Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 11\} | h_j \neq 0)$ :** we assume that all the observations which are equal to zero correspond to the influential observations (outliers), and thus, the least squares is only applied on the remaining part of the data (that is, on the observations  $h_j \neq 0$ ). Note that when only one observation is not equal to zero (see Figure 4 when  $K = 3.1$ ), that is,  $h_1 > 0$  and  $h_j = 0 \forall j \in [2, 11]$ , the least squares will then only be applied to the two observations  $h_1$  and  $h_2$ .
3.  **$L_1$ -estimator:** it is the same model that we have already described in Subsection 2.4. More precisely, the  $l_2$ -norm in the least squares objective function in (2) is replaced by the  $l_1$ -norm.

Figure 5 evaluates the three aforementioned option pricing prediction models in terms of the PAYOFF prediction error  $\frac{\|\mathbf{h} - \hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$  and the predicted option price at initial date  $\hat{V}_0$  for different values of  $K \in ]0.5260, 3.8871[$  in steps of  $10^{-2}$ .

We can observe that the “Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 11\} | h_j \neq 0)$ ” achieves the lowest PAYOFF prediction error for all values of  $K$ , and



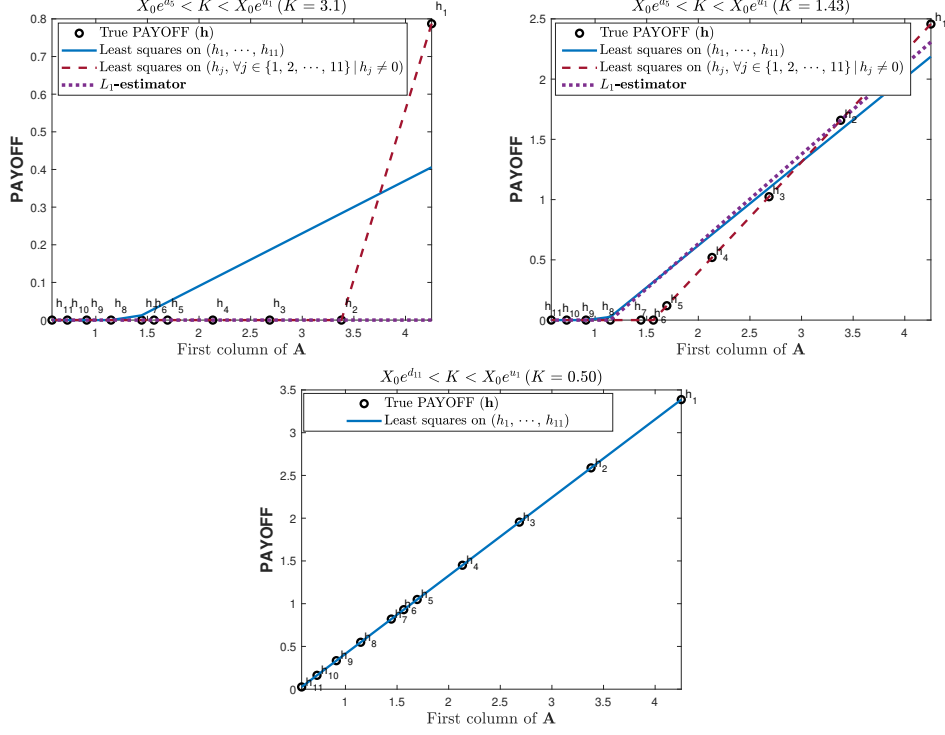
**Fig. 5:** From left to right: PAYOFF prediction error  $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\hat{\mathbf{h}}\|_2}$ , the predicted option price  $\hat{V}_0$ .

thus, its estimated delta hedging coefficient vector  $\delta$  should achieve a precise prediction of the option price at initial date. We can also observe that when  $K \in ]0.5260, 1.0560[$ , the PAYOFF prediction error as well as  $\hat{V}_0$  for the “ $L_1$ -estimator” coincides with those of the “Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 11\} | h_j \neq 0)$ ”. Figure 6 shows the difference between  $\mathbf{h}$  and  $\hat{\mathbf{h}}$  when  $K = 3.1$  ( $X_0e^{d_5} < K < X_0e^{u_1}$ ),  $K = 1.43$  ( $X_0e^{d_5} < K < X_0e^{u_1}$ ), and  $K = 0.50$  ( $K \leq X_0e^{d_5}$ ), respectively. We can obviously observe that when  $K \leq X_0e^{d_5}$ , the whole 11 observations  $h_1, h_2, \dots, h_{11}$ , will always be located on the same line and so the “Least squares on  $(h_1, \dots, h_{11})$ ” will behave perfectly (that is, the predicted PAYOFF observations coincide with the true ones). However, when  $X_0e^{d_5} < K < X_0e^{u_1}$ , it is the model “Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 11\} | h_j \neq 0)$ ” that has the lowest PAYOFF prediction error, and its predicted PAYOFF observations coincide with the true ones.

### 3.2 delta hedging via robust linear regression $M$ -estimators

Instead of manually removing the influential observations (outliers) from the data and then apply the least squares on the remaining part of the data, it seems that there exists other strategies that can automatically remove those outliers while estimating the (unknown) delta hedging coefficient vector  $\delta$ . Recall that the least squares is derived under the Gaussian assumption which does not take into consideration the outliers. In order not to consider any specific kind of distribution (ex. the Gaussian assumption), we are going to assume that the  $h_i$ 's,  $i \in [1, 2n + 1]$ ,  $n \geq 1$ , are independent and not identically distributed with any density of the form  $\frac{1}{\sigma} f_0\left(\frac{h_i - \mathbf{a}_i^T \delta}{\sigma}\right)$ . The likelihood under this density function can be written as (Maronna (2019); Huber





**Fig. 6:** The difference between  $\mathbf{h}$  and  $\hat{\mathbf{h}}$  when  $X_0 e^{d_n} < K < X_0 e^{u_1}$  and  $K \leq X_0 e^d$  when  $n = 5$ .

(1964)):

$$\mathcal{L}(h_1, \dots, h_{2n+1}; \boldsymbol{\delta}, \sigma) = \prod_{i=1}^{2n+1} \frac{1}{\sigma} f_0 \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right).$$

Maximizing the log-likelihood with respect to  $(\boldsymbol{\delta}, \sigma)$  is thus equivalent to the following (minimization) optimization problem:

$$\begin{aligned} \max_{\boldsymbol{\delta}, \sigma} & \left\{ -(2n+1) \log(\sigma) + \sum_{i=1}^{2n+1} \log \left( f_0 \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right) \right) \right\} \\ & \equiv \min_{\boldsymbol{\delta}, \sigma} \left\{ (2n+1) \log(\sigma) + \sum_{i=1}^{2n+1} \rho \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right) \right\}, \end{aligned} \quad (5)$$

where  $-\log \left( f_0 \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right) \right) = \rho \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right)$  for any non-constant function  $\rho(\cdot)$ .

One can clearly observe that the least squares and the  $L_1$ -estimator are a special case when  $\rho \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right) = \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right)^2$  and  $\rho \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right) = \left| \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right|$ , respectively. By taking

the derivative of the objective function in (5) with respect to  $\boldsymbol{\delta}$  and  $\sigma$  and set it to zero, one obtains (Maronna (2019)):

$$\sum_{i=1}^{2n+1} \rho' \left( \frac{h_i - \mathbf{a}_i^T \hat{\boldsymbol{\delta}}}{\hat{\sigma}} \right) \mathbf{a}_i = \mathbf{0}$$

and

$$\frac{1}{(2n+1)} \sum_{i=1}^{2n+1} \left( \frac{h_i - \mathbf{a}_i^T \hat{\boldsymbol{\delta}}}{\hat{\sigma}} \right) \rho' \left( \frac{h_i - \mathbf{a}_i^T \hat{\boldsymbol{\delta}}}{\hat{\sigma}} \right) = 1,$$

respectively.

Note that the function  $\rho(\cdot)$  must be chosen to respect the following conditions (Maronna (2019)):

- $\rho(x)$  should be a non-decreasing function of  $|x|$  ;
- $\rho(0) = 0$  ;
- $\rho(x)$  is increasing for  $x > 0$  such that  $\rho(x) < \rho(\infty)$  ;
- if  $\rho$  is bounded, it is also assumed that  $\rho(\infty) = 1$  .

### Choices for the function $\rho(\cdot)$

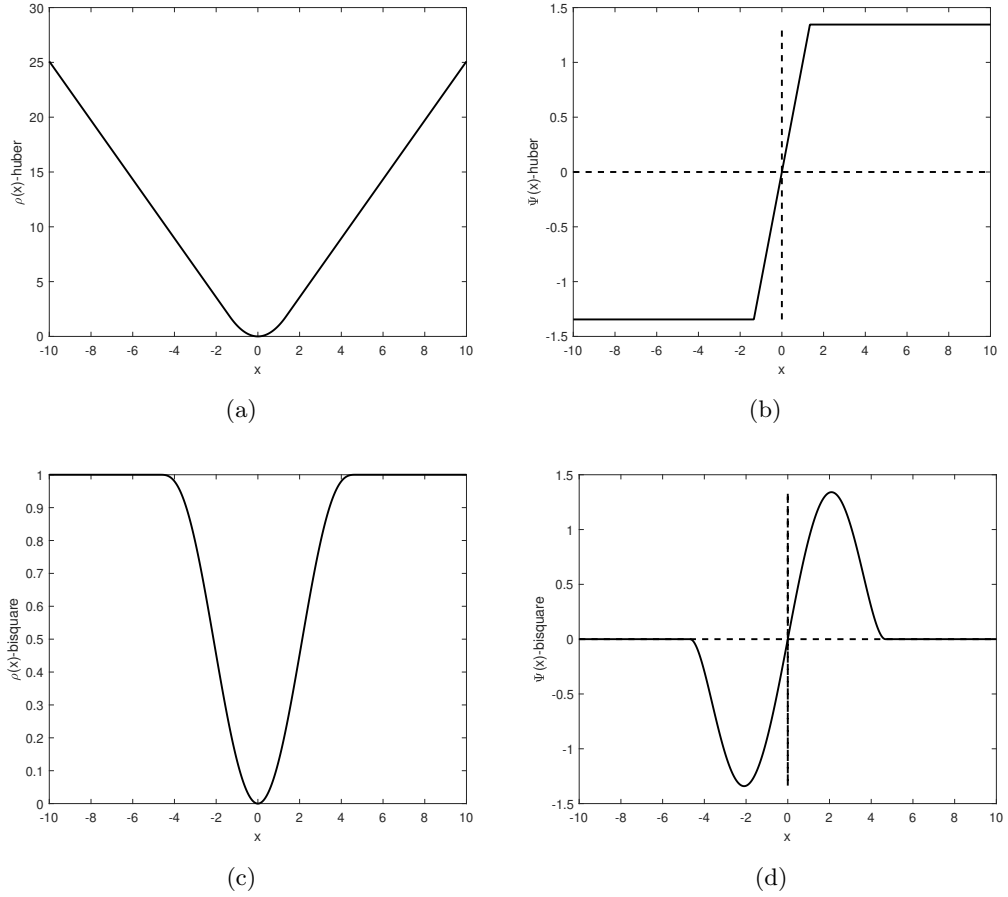
To estimate the delta hedging coefficient vector  $\boldsymbol{\delta}$ , we can assume three types of  $\rho(\cdot)$ -functions to estimate  $\boldsymbol{\delta}$ , and thus the prediction of the PAYOFF as well as the true (unknown) price of the option at the initial date  $V_0$ . More precisely, the following  $\rho(\cdot)$ -functions are considered:

1. The *Huber*- $\rho(\cdot)$  function: For the first impression, one can directly think to propose an estimator that combines the robustness of the median (under the non-Gaussian assumption) and the low-variance of the mean (under exact normality). Hence, for a fair compromise between the mean and median, the function  $\rho(\cdot)$  can be chosen to belong to the family of Huber functions. More precisely, the function  $\rho(\cdot)$  can be as follows:

$$\rho \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right) = \begin{cases} \left( \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right)^2 & \text{if } \left| \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right| \leq \lambda \\ 2\lambda \left| \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right| - \lambda^2 & \text{if } \left| \frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right| > \lambda \end{cases}.$$

Figure 7a and 7b exhibit both the  $\rho(\cdot)$ - and  $\rho'(\cdot)$ - functions under this family of Huber functions, respectively. Note that the solution  $\mathbf{a}_i^T \hat{\boldsymbol{\delta}}$  is closely related to Winsorizing.

2. The *Bisquare*- $\rho(\cdot)$  function: From figure 7a, we can observe that the function  $\rho(\cdot)$  aims to tend to infinity at infinity, whereas its derivative (figure 7b), that is, the function  $\rho'(\cdot)$ , is continuous but always shifted by  $\pm\lambda = 1.345$ . These findings demonstrate that both the gross as well as the moderately large outliers are going to be treated in the same manner. In this regard, thinking about using a function  $\rho'(\cdot)$  that re-descend smoothly to zero at infinity (see figure 7d) may be quite helpful since in this case the gross outliers will be completely rejected and the



**Fig. 7:** (a) and (b) the Huber  $\rho$ - and  $\Psi = \rho'$ - functions (with  $\lambda = 1.345$ ), respectively; (c) and (d) the Bisquare  $\rho$ - and  $\Psi = \rho'$ - functions (with  $\lambda = 4.685$ ), respectively.

moderately large outliers will be completely ignored. That is why, one can consider the following  $\rho(\cdot)$ -function:

$$\rho\left(\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma}\right) = \begin{cases} 1 - \left(1 - \left(\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma \lambda}\right)^2\right)^3 & \text{if } \left|\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma}\right| \leq \lambda \\ 1 & \text{if } \left|\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma}\right| > \lambda \end{cases}.$$

3. The *Talwar*- $\rho(\cdot)$  function: In addition to the sample median and winsorized mean, one can think of discarding a proportion of the largest and smallest values. This is well-known as the trimming procedure (Huber (1964)) and which can be much more robust to outliers than the sample mean and sample median. Hence, one can

consider the following  $\rho(\cdot)$ -function:

$$\rho\left(\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma}\right) = \begin{cases} \frac{1}{2} \left(\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma}\right)^2 & \text{if } \left|\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma}\right| \leq \lambda \\ \frac{1}{2} \lambda^2 & \text{if } \left|\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma}\right| > \lambda \end{cases}.$$

## 4 Experiments

In this section, our main objective is to compare six different option pricing prediction models on exchange rates.

1. Least squares on  $(h_1, \dots, h_{2n+1})$ : we directly apply the least squares on the whole  $2n + 1$  observations;
2. Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 2n + 1\} | h_j \neq 0)$ : the least squares is applied on the observations that are not equal to zero. When only one observation is different than zero (that is,  $h_1 > 0$ ), the least squares will then only be applied on the two observations  $h_1$  and  $h_2$ ;
3.  $L_1$ -estimator;
4. *Huber*;
5. *Bisquare*;
6. *Talwar*.

All our evaluations are performed on two different scenarios for  $n$ :

- $n = 1$  (the standard one-period trinomial option pricing (see Subsection 2.1)): a single value for  $u > 0$  and  $d < 0$  are specified to price the option. In our experiments, the value for  $d$  is always fixed to be equal to  $-u$ ;
- $n > 1$  (our proposed extension of the one-period trinomial option pricing model (see subsection 3.1)): a large set of  $n$  values for both  $u$  and  $d$  are considered at once. More precisely, the set  $\{u_l\}_{l \in [1, n]}$  and  $\{d_l\}_{l \in [1, n]}$ , with  $u_1 > u_2 > \dots > u_n > r_d$  and  $d_n < d_{n-1} < \dots < d_1 < 0$  are specified to price the European option. In our experiments, we consider  $d_n = -u_1, d_{n-1} = -u_2, \dots, d_1 = -u_n$ .

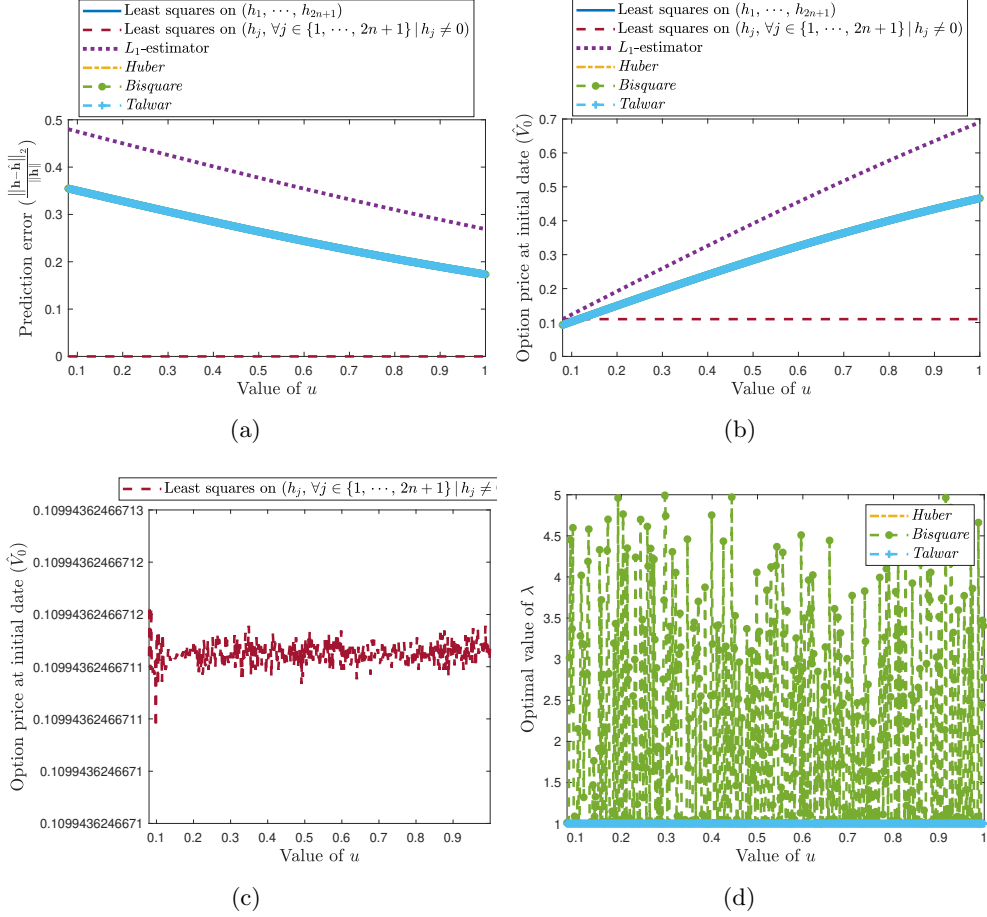
In all experiments, we fix  $r_d = 8\%$ ,  $r_f = 9\%$ ,  $B_0 = 1$ , and we assume that the exchange rate at the initial date is  $X_0 = 1.43$ . A call on the exchange rate is available with a strike price  $K = X_0$ , which expires at maturity.

Note that we use the *robustfit* package in MATLAB to simultaneously estimate both  $\boldsymbol{\delta}$  and  $\sigma$  for the Huber, Bisquare and Talwar robust linear regression techniques. This package uses iteratively reweighted least squares to compute the delta hedging coefficients.

To choose the best parameter value of  $\lambda$  for every robust regression estimator (e.g. *Huber*, *Bisquare*, and *Talwar*), we decide to vary  $\lambda$  between 1 and 5 by step of  $10^{-4}$ . The parameter value  $\lambda$  that minimizes the criterion  $\frac{\|\mathbf{h} - \hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$  will be selected as the optimal value. All of our experiments are conducted on MATLAB 2021a <sup>4</sup>.

---

<sup>4</sup>The MATLAB code of the proposed work is available upon request.



**Fig. 8:** Evaluations when  $n = 1$ : (a) the prediction error  $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$ , (b) the predicted European call option pricing at initial date  $(\hat{V}_0)$ , (c) the predicted option price for “Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 2n+1\} | h_j \neq 0)$ ”, (d) the optimal values of  $\lambda$  for  $\{Huber, Bisquare, \text{ and } Talwar\}$ -prediction models.

## 4.1 European option pricing on exchange rates

Let  $\epsilon = 10^{-6}$ .

### 4.1.1 When $n = 1$

We choose to evaluate the aforementioned six option pricing prediction models on a set of 500 single evenly spaced values for  $u$  between  $r_d + \epsilon$  and 1. Figure 8 shows, as a function of  $u$ , the error results in the prediction of  $\mathbf{h}$ , the prediction results of  $V_0$ , as well as the different optimal values of  $\lambda$  for the  $\{Huber, Bisquare, \text{ and } Talwar\}$ -prediction models. As we can observe from figures 8a and 8b, only the “Least squares

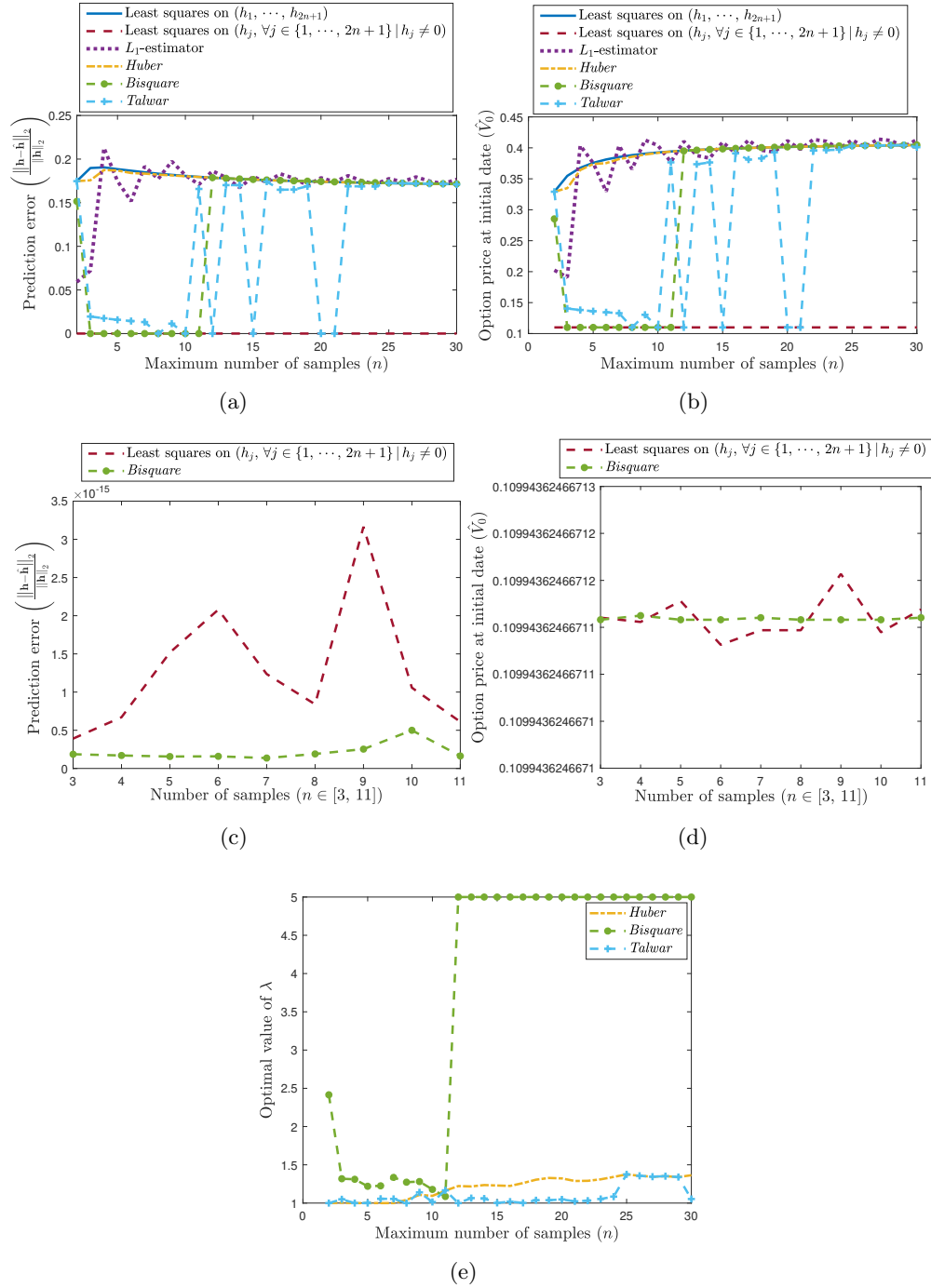
on  $(h_j, \forall j \in \{1, 2, 3\} | h_j \neq 0)$ " (that is, when applying the least squares after the manual removal of the outliers) achieves the lowest prediction error for all the values of  $u$  and presents a stable  $\hat{V}_0$  always varying around 0.1099 (see figure 8c). However, the other prediction models behave poorly for all the different values of  $u$ .

#### 4.1.2 When $n > 1$

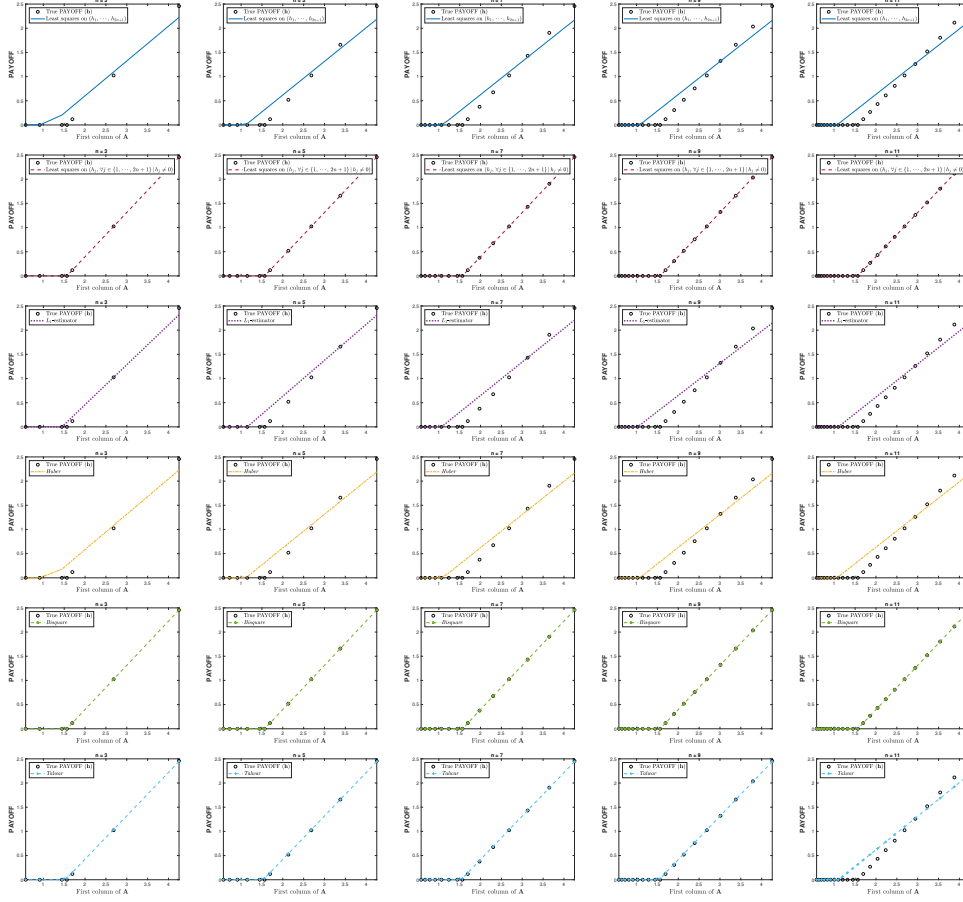
For any  $u_n \geq r + \epsilon$ , the six option pricing prediction models are evaluated on a large set of  $\{u_l\}_{l \in [1, n]}$  and  $\{d_l\}_{l \in [1, n]}$ , both containing  $n > 1$  evenly spaced values. Figure 9 shows, as a function of  $n \in \{2, \dots, 30\}$  in steps of 1, and for  $u_1 = 1$ , the error results in the prediction of  $\mathbf{h}$ , the prediction results of  $V_0$ , as well as the different obtained optimal values of  $\lambda$  for the *{Huber, Bisquare, and Talwar}*-prediction models. From figures 9a and 9b, it is important to note that when  $n \in [3, 11]$ , the PAYOFF prediction error and the European call option price for the *Bisquare* is about  $1.5877 \times 10^{-15}$  and 0.1099, respectively, and both coincide in values with that of the "Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 2n + 1\} | h_j \neq 0)$ ".

The *Talwar* achieves a low PAYOFF prediction error when  $n \in [3, 10]$ . More precisely, for every value of  $n \in \{3, 4, 5, 6, 7, 8, 9, 10\}$ , it achieves respectively a PAYOFF prediction error of about (0.0195, 0.0174, 0.0157, 0.0143, 0.0131,  $4.7532 \times 10^{-6}$ , 0.0113,  $6.3428 \times 10^{-6}$ ) and a European call option price of about (0.1403, 0.1380, 0.1360, 0.1343, 0.1328, 0.1099, 0.1303, 0.1099).

It is true that "Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 2n + 1\} | h_j \neq 0)$ " always achieve a very low PAYOFF prediction error for all the possibilities of  $n$ , it can be seen from figures 9c and 9d, however, that it presents slightly higher prediction error as well as higher variability in the call option price than *Bisquare* when  $n \in [3, 11]$ . Figure 10 shows the difference between  $\mathbf{h}$  and  $\hat{\mathbf{h}}$  for the six different option pricing prediction models when  $n \in \{3, 5, 7, 9, 11\}$ . As we can observe, the PAYOFF vector  $\mathbf{h}$  can be predicted very efficiently for the two prediction models "Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 2n + 1\} | h_j \neq 0,)$ " and *Bisquare*.



**Fig. 9:** Evaluations when  $n > 1$ : (a) the PAYOFF prediction error  $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$ , (b) the predicted European call option pricing at initial date ( $\hat{V}_0$ ), (c) the PAYOFF prediction error when  $n \in [3, 11]$ , (d) the predicted option price  $\hat{V}_0$  when  $n \in [3, 11]$ , (e) the optimal values of  $\lambda$  for  $\{Huber, Bisquare, \text{and Talwar}\}$ -prediction models.



**Fig. 10:** The difference between  $\mathbf{h}$  and  $\hat{\mathbf{h}}$  for the six option pricing prediction models when  $n \in \{3, 5, 7, 9, 11\}$ . From top row to bottom row: Least squares on  $(h_1, \dots, h_{2n+1})$ , Least squares on  $(h_j, \forall j \in \{1, 2, \dots, 2n+1\} | h_j \neq 0)$ ,  $L_1$ -estimator, *Huber*, *Bisquare*, *Talwar*.



## 5 Conclusion and future work

The trinomial European call option pricing model is well known to be incomplete, and therefore, the least squares linear regression technique can simply be used to build a self-financing hedging portfolio that does not perfectly replicate the call option. However, the least squares is quite sensitive to even a single influential observation, and thus, the prediction of the PAYOFF and the option price at initial date can potentially deviate from their true values. This paper has alleviated this challenge via robust linear regression. More precisely, two option pricing prediction strategies have been developed: the first strategy simply applies the least squares after the outliers are removed. The second strategy takes advantage of some robust linear regression  $M$ -estimators. Both prediction strategies have been evaluated on numerical experiments, and the results of which demonstrate their effectiveness for European call option pricing on exchange rates.

For future enhancements, a likely first step would be to extend the proposed work to the multi-period case. Other promising avenues for further research include the pricing of European call options on exchange rates via deep learning (e.g., neural networks). More precisely, by extending the trinomial option pricing model to the case of  $2n + 1$ ,  $n \geq 1$ , observations, the use of deep learning can be quite efficient in predicting PAYOFF when  $n$  is sufficiently large. However, the big challenge remains of how to develop a method that can benefit from the estimated weights of the neural network to efficiently predict the option price at the initial date.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available upon request.

## References

- Ayres A (1967) Risk aversion in the warrants market. *Indus Management Review* pp 497–505
- Bachelier L (1900) Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure 3e série*, 17:21–86. <https://doi.org/10.24033/asens.476>, URL <http://www.numdam.org/articles/10.24033/asens.476/>
- Bitar AW (2024) Robust European Call Option Pricing via Linear Regression, URL <https://utt.hal.science/hal-04754957>, working paper or preprint
- Black F, Scholes M (1973) The pricing of options and corporate liabilities. *Journal of Political Economy* 81(3):637–654. URL <http://www.jstor.org/stable/1831029>
- Boness AJ (1964) Elements of a theory of stock-option value. *Journal of Political Economy* 72(2):163–175. URL <http://www.jstor.org/stable/1828962>
- Boyle PP (1986) Option valuation using a three jump process. URL <https://api.semanticscholar.org/CorpusID:150764610>
- Cox JC, Ross SA, Rubinstein M (1979) Option pricing: A simplified approach. *Journal of Financial Economics* 7(3):229–263. [https://doi.org/https://doi.org/10.1016/0304-405X\(79\)90015-1](https://doi.org/https://doi.org/10.1016/0304-405X(79)90015-1)
- Ehrmann M, Fratzscher M (2005) Exchange rates and fundamentals: new evidence from real-time data. *Journal of International Money and Finance* 24(2):317–341. <https://doi.org/https://doi.org/10.1016/j.jimonfin.2004.12.010>, URL <https://www.sciencedirect.com/science/article/pii/S0261560604001196>, exchange Rate Economics
- Huber PJ (1964) Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics* 35(1):73 – 101. <https://doi.org/10.1214/aoms/1177703732>, URL <https://doi.org/10.1214/aoms/1177703732>
- Johan B (2016) The trinomial asset pricing model. GIstitutionen for matematiska vetenskaper, Goteborgs universitet, Thesis report
- Maronna RA (2019) *Robust Statistics: Theory and Methods (with R)*, Wiley Series in Probability and Statistics, John Wiley Sons. URL <https://doi.org/10.1002/9781119214656>
- Merton R (1973) Continuous-time speculative processes: Appendix to paul a. samuelson's "mathematics of speculative price". *SIAM Rev* 15:34–38
- Orlin Grabbe J (1983) The pricing of call and put options on foreign exchange. *Journal of International Money and Finance* 2(3):239–253. [https://doi.org/https://doi.org/10.1016/S0261-5606\(83\)80002-3](https://doi.org/https://doi.org/10.1016/S0261-5606(83)80002-3), URL <https://www.sciencedirect.com/science/>

[article/pii/S0261560683800023](#)

- Pascucci A, Runggaldier WJ (2012) *Financial Mathematics : Theory and Problems for Multi-period Models*, 1st edn. *La Matematica per il 3+2*, Springer Milan, Milano
- Rosett R (1963) *Estimating the utility of wealth from call options transactions*. Doctoral dissertation, University of Rochester
- Samuelson PA (2015) *Rational Theory of Warrant Pricing*, Springer International Publishing, Cham, pp 195–232. [https://doi.org/10.1007/978-3-319-22237-0\\_11](https://doi.org/10.1007/978-3-319-22237-0_11)
- Samuelson PA, Merton RC (1969) A complete model of warrant pricing that maximizes utility. URL <https://api.semanticscholar.org/CorpusID:153453511>
- Scholes M (2018) A conversation with myron scholes. URL [https://www.youtube.com/watch?v=zooUe\\_oDSpo](https://www.youtube.com/watch?v=zooUe_oDSpo)
- Sprenkle CM (1961) Warrant prices as indicators of expectations and preferences. *Yale economic essays* 1(2):178–231
- Takahashi H (2000) A note on pricing derivartives in an incomplete market. Graduate School of Economics, Hitotsubashi University, Discussion Papers
- Zhu S, Zhang J, Chen Z (2022) The maximum and minimum price in trinomial model. In: *Proceedings of the 2021 5th International Conference on Software and E-Business*. Association for Computing Machinery, New York, NY, USA, ICSEB '21, p 104–106, <https://doi.org/10.1145/3507485.3507502>, URL <https://doi.org/10.1145/3507485.3507502>