



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. 2024. hal-04815308v1

HAL Id: hal-04815308

<https://utt.hal.science/hal-04815308v1>

Preprint submitted on 2 Dec 2024 (v1), last revised 5 Jan 2025 (v2)

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

^aLaboratory of Computer Science and Digital Society, Université de Technologie de Troyes, Troyes 10300, France

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 handle the problem of outliers, this paper aims to develop two different option pricing prediction strategies mainly based on the idea of robust linear regression. The proposed robust techniques are evaluated on numerical data, and 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 ([7]) can be defined as the price of foreign currency with respect to (w.r.t.) 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 w.r.t. the domestic one. For example, it has been observed that the Euro has shown a depreciation of about more than 18% against the United States dollar since 2013 and especially just after the start of the Russian-Ukraine war on Feb 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 Aug 04, 2020.

Dealing with such a huge fluctuations of the exchange rate can be alleviated via buying call European options ([12]). 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 which is well known as the 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 ([2]) 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 ([18], [14], [1], [4], [15], [16], [11]). 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 model that accurately prices options¹ ([3]). 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 thus, is not robust.

To move on from parametric to robust European option pricing, the one-period discrete-time binomial non-arbitrage pricing model was developed ([6]). It mainly assumes that the exchange rate at maturity (denoted as

¹See the online conversation ([17]) with Myron Scholes who briefly explained his work with Fisher Black.

X_1), can only go up or down w.r.t. its value at initial date (denoted as X_0). After creating a self-financing hedging portfolio and equating it to 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, 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 w.r.t. to X_0 . More precisely, X_1 is assumed to behave in a three possible directions w.r.t. X_0 : goes up (with a factor e^u , $u > 0$), remains the same, or goes down (with a factor e^d , $d < 0$). This is well known as the one-period trinomial option pricing model ([5, 20, 9]), and by which equating the self-financing hedging portfolio to the call option price at maturity results an over-determined linear system of three equations with two unknowns, that correspond again, 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 (ex. 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 ([13]) or finding a hedging strategy that is close to being perfect. The latter scenario means to find an approximation to the hedging portfolio and which can simply be done via the least squares machine learning technique ([19, 9]). 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)², and therefore, often fails to provide good fits to the bulk of the data. In addition, the least squares is derived under

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

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 replicates the option price. As the Least squares technique fails with the presence of outliers, we alleviate this challenge via robust linear regression techniques. More precisely, we first aim to detect the nature of the outliers and 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 exploit some linear regression M -estimators capable of automatically identifying the outliers during the estimation of the delta-hedging portfolio coefficients.

It is worthy to further note that the two different possible positive factors e^u and e^d by which the exchange rate X_1 goes up or down with respect to X_0 , respectively, are usually not known in the real market, and thus, a single specific values for both the parameters u and d can be chosen manually in order 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 (ex. three) might be not sufficient to accurately predict the option price under any regression model. To ensure reliable prediction results, generating enough observations may alleviate this challenge. However, it is not surprising that the option price greatly depends on the selection of both u and d . Hence, instead of selecting a single specific value for each of these two parameters, we aim to price options under a set of n possible values for both u and d , with $n \geq 1$. This will help us to generate $2n + 1$ observations to build the linear regression model and which 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 both the two parameters u and d .

This paper is structured along the following lines. First comes in section 2.1 a brief presentation of how can the standard trinomial option pricing model on exchange rates be applied and how is the delta-hedging portfolio coefficients estimated via the least squares technique. A numerical example is also presented to evaluate the least squares under the presence of outliers. The proposed two 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 European call option pricing 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 depict 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 norms on vectors will be used. For instance, $\boldsymbol{\delta}$ is a vector, and δ_j is the j^{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 w.r.t. 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 European call option price on exchange rates. Third, we briefly overview in subsection 2.3 how can the option price be predicted via the least squares machine learning technique. Finally, in subsection 2.4, we demonstrate via a simple numerical example, the suffer of the least squares to accurately predict the European call option price 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 [9]). 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 well a probability if and only if condition (1) is satisfied, and thus, $u > r_d$.

2.2. The self-financing hedging portfolio

The replicating portfolio can be defined as follows³:

At initial date :

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

At maturity :

$$V_1 = \delta_1 e^{r_f} X_1 + \delta_2 B_1$$

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

where δ_1 is the number of units of the foreign currency bought at initial date, δ_2 is the number of units of the 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 goes up, remains the same, or goes down 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 (δ_1 and δ_2). Based on both the values of X_0 and K , one can assume the following three possible different scenarios

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

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 exhibits 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 in estimating 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 the matrix of independent variables $\mathbf{A} = \begin{bmatrix} - & - & \mathbf{a}_1^T & - & - \\ - & - & \mathbf{a}_2^T & - & - \\ - & - & \mathbf{a}_3^T & - & - \end{bmatrix} =$

$$\begin{bmatrix} e^{rf} X_0 e^u & B_0 e^{rd} \\ e^{rf} X_0 & B_0 e^{rd} \\ e^{rf} X_0 e^d & B_0 e^{rd} \end{bmatrix} \in \mathbb{R}_+^{3 \times 2}, \text{ and hence, a solution for } \boldsymbol{\delta} = [\delta_1, \delta_2]^T \in \mathbb{R}^2 \text{ does}$$

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 can not be uniquely defined. Hence, the standard European option may not be replicable by only investing in the underlying exchange rate and the risk-free bond. As it is generally impossible to solve the linear system $\mathbf{h} = \mathbf{A}\boldsymbol{\delta}$, it is possible, however, to project \mathbf{h} into the column space of \mathbf{A} . By assuming 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 3×3 , the regression coefficient vector $\boldsymbol{\delta} \in \mathbb{R}^2$ can simply be

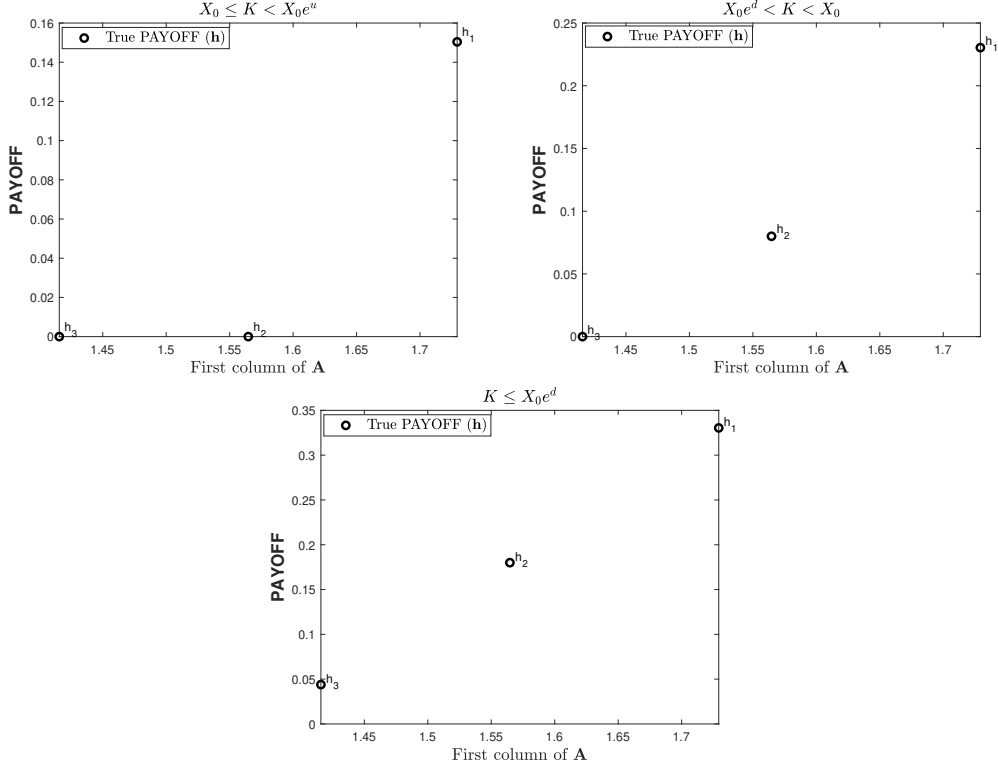


Figure 1: Example of the three PAYOFF scenarios when $X_0 = 1.43$, $u = 0.1$, and $d = -u$.

estimated by maximizing the log-likelihood under the Gaussian distribution 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:

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

2.4. European option pricing via the least squares

Once the delta-hedging coefficient vector $\boldsymbol{\delta}$ 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:

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

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 such that the sum of the squares of the differences between the original and 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 outliers.

As the option price should always be positive, thus, 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 $\boldsymbol{\delta}$ 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 efficient to predict the option price?

We study, via a simple numerical example, the performance of the least squares on predicting the option price at maturity (that is, the PAYOFF) as well as at initial date (that is, V_0) for different values of $K \in]X_0 e^d, X_0 e^u[$. Let fix $X_0 = 1.43$, $u = 0.1$, and $d = -u$. The value for X_1 under the trinomial model will behave as follows:

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

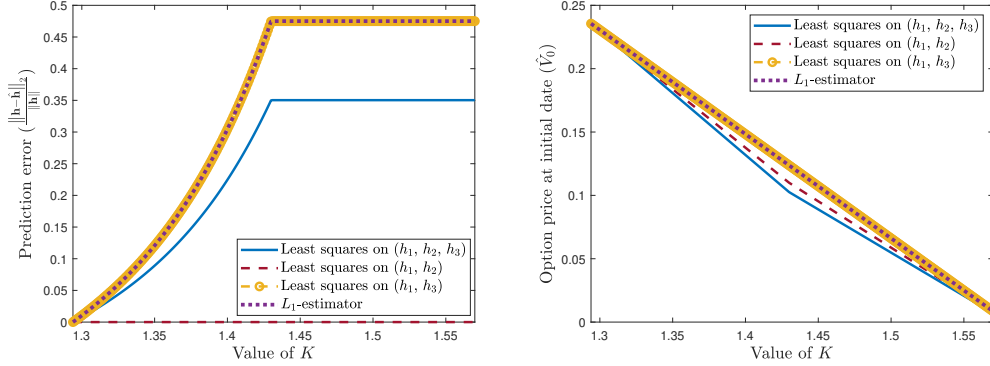


Figure 2: From left to right: PAYOFF prediction error $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\hat{\mathbf{h}}\|_2}$, the predicted option price \hat{V}_0 .

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 $\boldsymbol{\delta}$ is estimated as follows:

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

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,

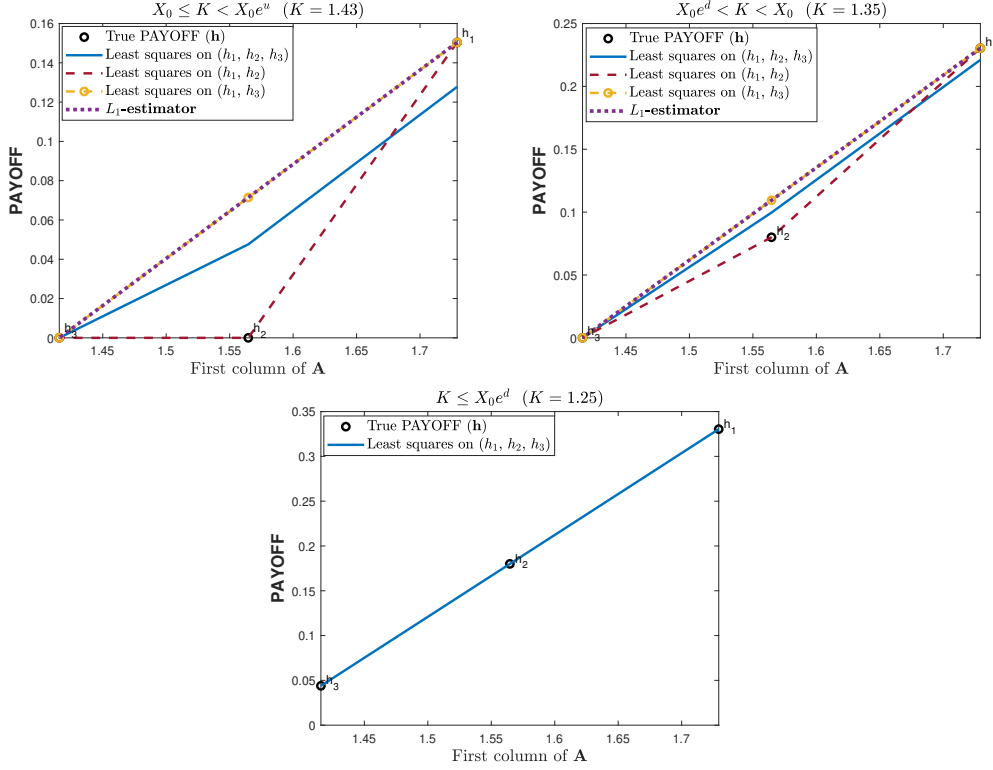


Figure 3: The difference between \mathbf{h} and $\hat{\mathbf{h}}$ when $X_0 \leq K < X_0e^u$, $X_0e^d < K < X_0$, and $K \leq X_0e^d$.

one can find recursive algorithms that are able to solve it (e.g., the 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 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 δ yields a precise prediction of the option price at initial date.

Figure 3 exhibits 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 coincides with the true ones). However, when $X_0 \leq K < X_0 e^u$ and $X_0 e^d < K < X_0$, it is the model “Least squares on (h_1, h_2) ” that has the lowest PAYOFF prediction error and its predicted PAYOFF observations coincides with the true ones. We can then consider that applying the least squares after removing the observation h_3 may potentially decrease the prediction error compared to the case if the least squares is applied on the whole 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 (e.g., 3) is odd, the derivative of the objective function in (4) with respect to $\mathbf{a}_i^T \boldsymbol{\delta}$ exists and which 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 indeed not differentiable. However, it is a continuous function and is decreasing when $\mathbf{a}_i^T \boldsymbol{\delta} < h_{\frac{3+1}{2}}$ and increasing when $\mathbf{a}_i^T \boldsymbol{\delta} > h_{\frac{3+1}{2}}$. Hence, the minimizer is given by $\mathbf{a}_i^T \hat{\boldsymbol{\delta}} = h_{\frac{3+1}{2}}$. 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 bulk 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 from 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 on a regression with $2n + 1$, $n \geq 1$, observations in order 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 apply the least squares on the remaining part of the data, we aim to exploit in subsection 3.2 the idea behind the robust M -estimators for linear regression analysis ([10, 8]). This will allow to accurately estimate the (unknown) delta-hedging coefficient vector δ 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

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

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

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 exhibits three different examples of those two possible different scenarios for the PAYOFF values when $X_0 = 1.43$, $n = 5$, and $K \in \{0.50, 1.43, 3.1\}$. We assume $r_d = 8\%$ and a set of 5 possible evenly spaced values $\{u_l\}_{l \in [1, 5]}$, with $u_1 > \dots > u_5 > r_d$. We choose $u_5 = r_d + \epsilon$, $\forall \epsilon > 0$, and $u_1 = 1$. We also fix $\{d_l\}_{l \in [1, 5]} = \{-u_l\}_{l \in [5, 1]}$. 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, $K \leq X_0e^{d_5}$).

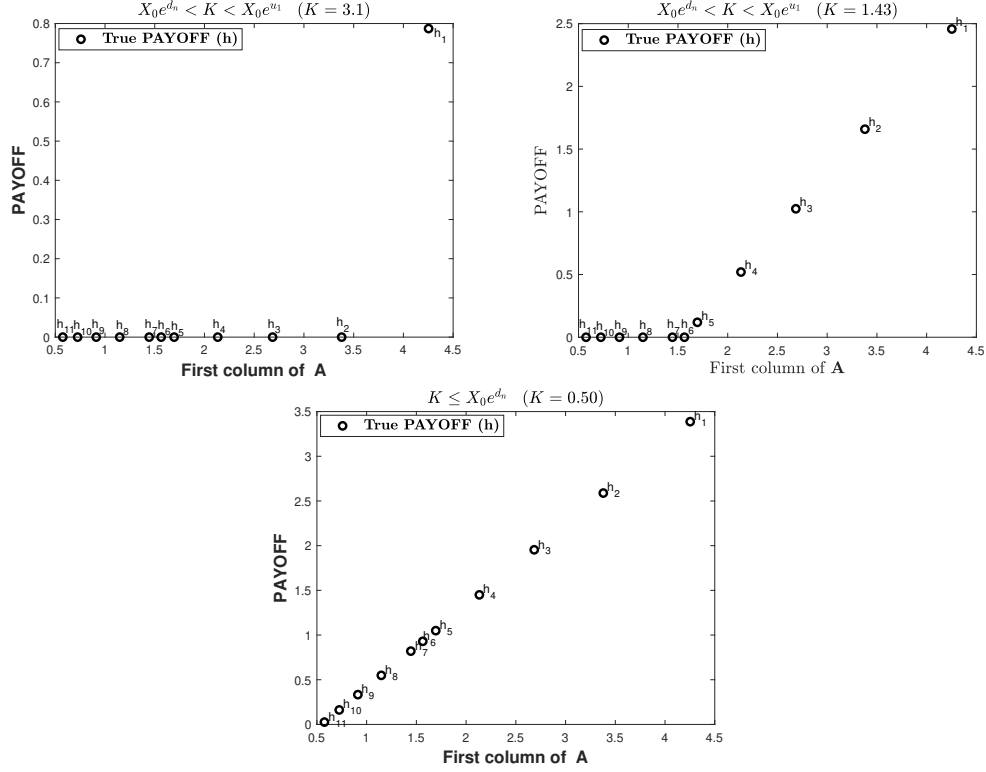


Figure 4: Example of the three PAYOFF scenarios when $X_0 = 1.43$ and $n = 5$. From left to right: $K = 3.1$, $K = 1.32$, $K = 0.50$.

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

By pursuing 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 option pricing prediction models:

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 on the two observations h_1 and h_2 .

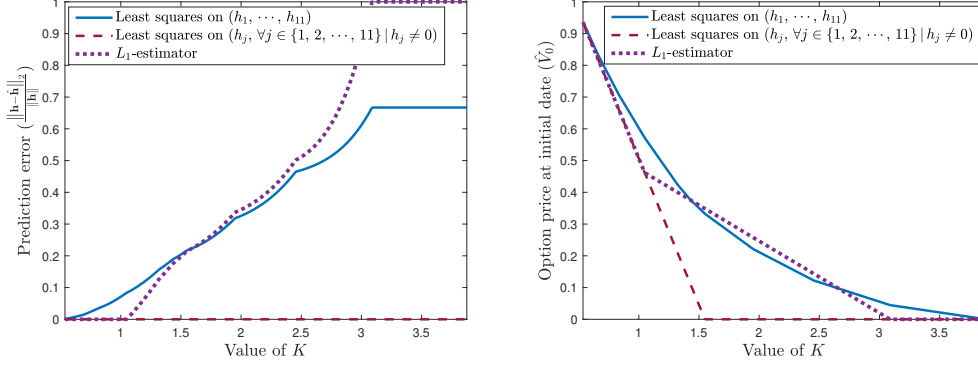


Figure 5: From left to right: PAYOFF prediction error $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$, the predicted option price \hat{V}_0 .

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 thus, its estimated delta-hedging coefficient vector δ 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 exhibits 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 coincides 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 coincides with the

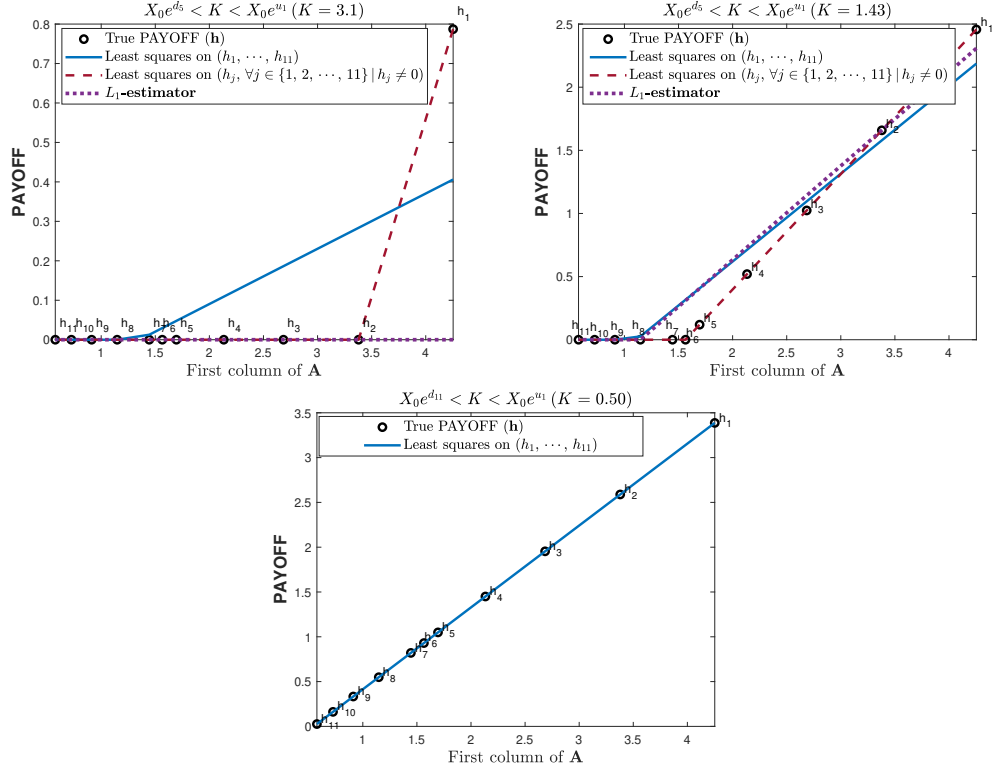


Figure 6: The difference between \mathbf{h} and $\hat{\mathbf{h}}$ when $X_0 e^{d_n} < K < X_0 e^{u_1}$ and $K \leq X_0 e^d$ when $n = 5$.

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 $\boldsymbol{\delta}$. Recall that the least squares is derived under the Gaussian assumption which does not take into consideration the outliers. In order to not 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 $\frac{1}{\sigma} f_0 \left(\frac{h_i - \mathbf{a}_i^T \boldsymbol{\delta}}{\sigma} \right)$. The likelihood

under this density function can be written as ([10, 8]):

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

Maximizing the log-likelihood w.r.t. $(\boldsymbol{\delta}, \sigma)$ 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 σ and set it to zero, one obtains ([10]):

$$\begin{aligned} & \sum_{i=1}^{2n+1} \rho'\left(\frac{h_i - \mathbf{a}_i^T \hat{\boldsymbol{\delta}}}{\hat{\sigma}}\right) \mathbf{a}_i = \mathbf{0} \\ \text{and} \quad & \frac{1}{(2n+1)} \sum_{i=1}^{2n+1} \left(\frac{h_i - \mathbf{a}_i^T \hat{\boldsymbol{\delta}}}{\hat{\sigma}}\right) \rho'\left(\frac{h_i - \mathbf{a}_i^T \hat{\boldsymbol{\delta}}}{\hat{\sigma}}\right) = 1, \end{aligned}$$

respectively.

Note that the function $\rho(\cdot)$ must be chosen to respect the following conditions ([10]):

- $\rho(x)$ 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 ρ 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) option price at 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 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

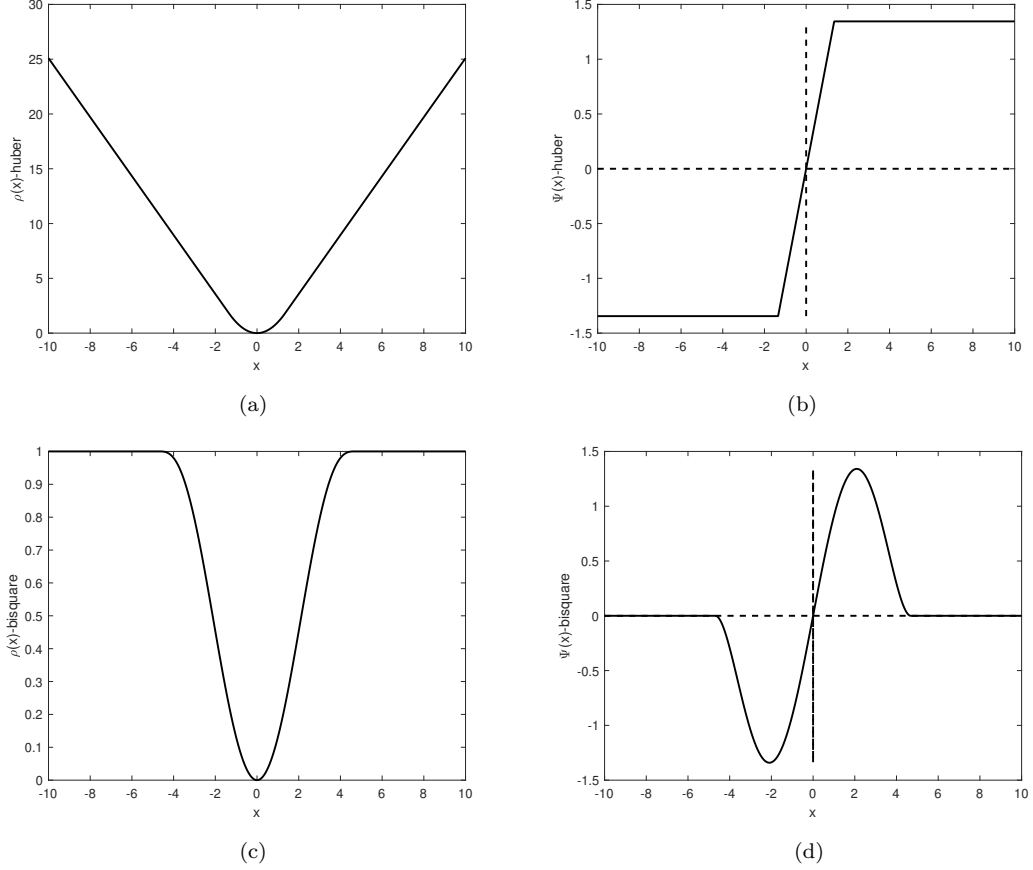


Figure 7: (a) and (b) the Huber ρ - and $\Psi = \rho'$ - functions (with $\lambda = 1.345$), respectively; (c) and (d) the Bisquare ρ - and $\Psi = \rho'$ - functions (with $\lambda = 4.685$), respectively.

([8]) 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, we mainly aim to compare between six different option pricing prediction models on exchange rates. More precisely, the six models

are the following:

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

Note that we use the *robustfit* package in MATLAB to simultaneously estimate both δ and σ for the Huber, Bisquare, and Talwar robust linear regression techniques. This package uses an iteratively reweighted least squares to compute the delta-hedging coefficients.

To choose the best parameter value of λ for every robust regression estimator (e.g. *Huber*, *Bisquare*, and *Talwar*), we decide to vary λ between 1 and 5 by step of 10^{-4} . The parameter value λ that minimizes the criterion $\frac{\|\mathbf{h} - \hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$ will be selected as the optimal value. All our experiments are conducted on MATLAB 2021a⁴.

⁴The MATLAB code of the proposed work is available upon request.

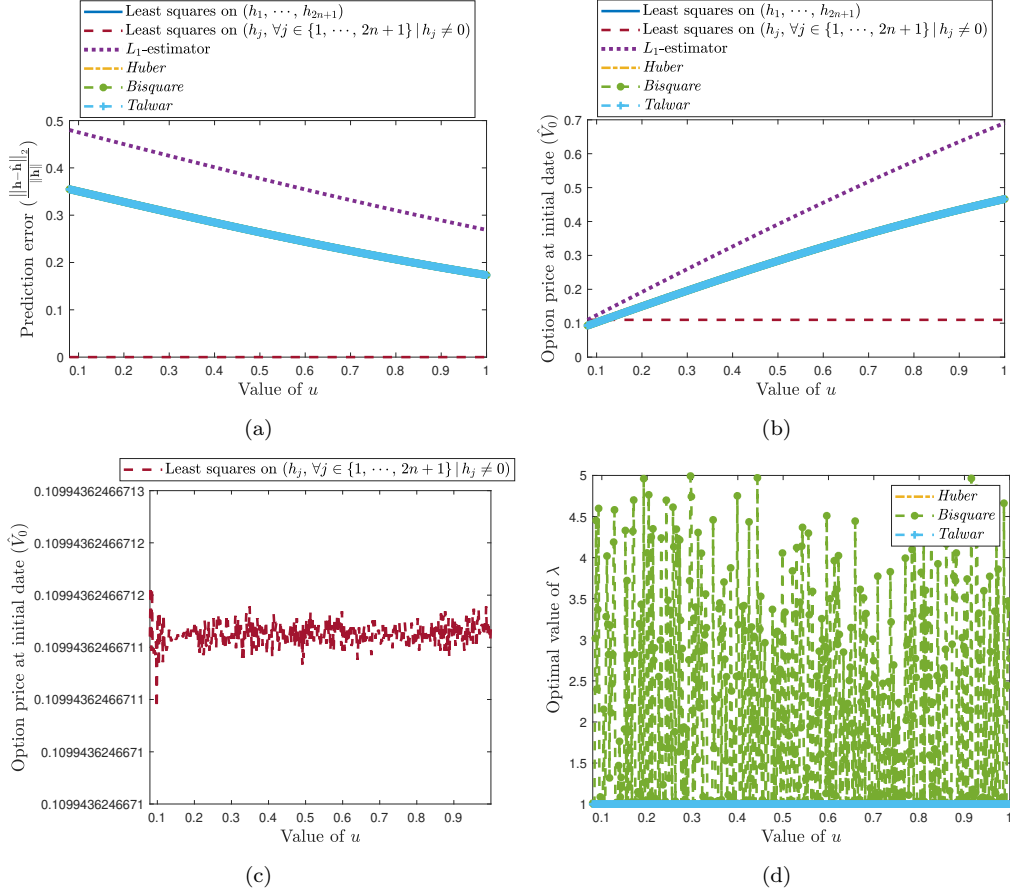


Figure 8: Evaluations when $n = 1$: (a) the prediction error $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$, (b) the predicted European call option pricing at initial date (\hat{V}_0) , (c) the predicted option price for “Least squares on $(h_j, \forall j \in \{1, 2, \dots, 2n+1\} | h_j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

4.1. European option pricing on exchange rates

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

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 exhibits, 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 obtained

optimal values of λ for the $\{Huber, Bisquare, \text{ and } Talwar\}$ -prediction models. As we can observe from figure 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 exhibits, 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 λ for the $\{Huber, Bisquare, \text{ and } Talwar\}$ -prediction models. From figure 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 1.5877×10^{-15} and 0.1099, respectively, and they 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, 4.7532×10^{-6} , 0.0113, 6.3428×10^{-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 figure 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 exhibits the difference between \mathbf{h} and $\hat{\mathbf{h}}$ for the different six 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*.

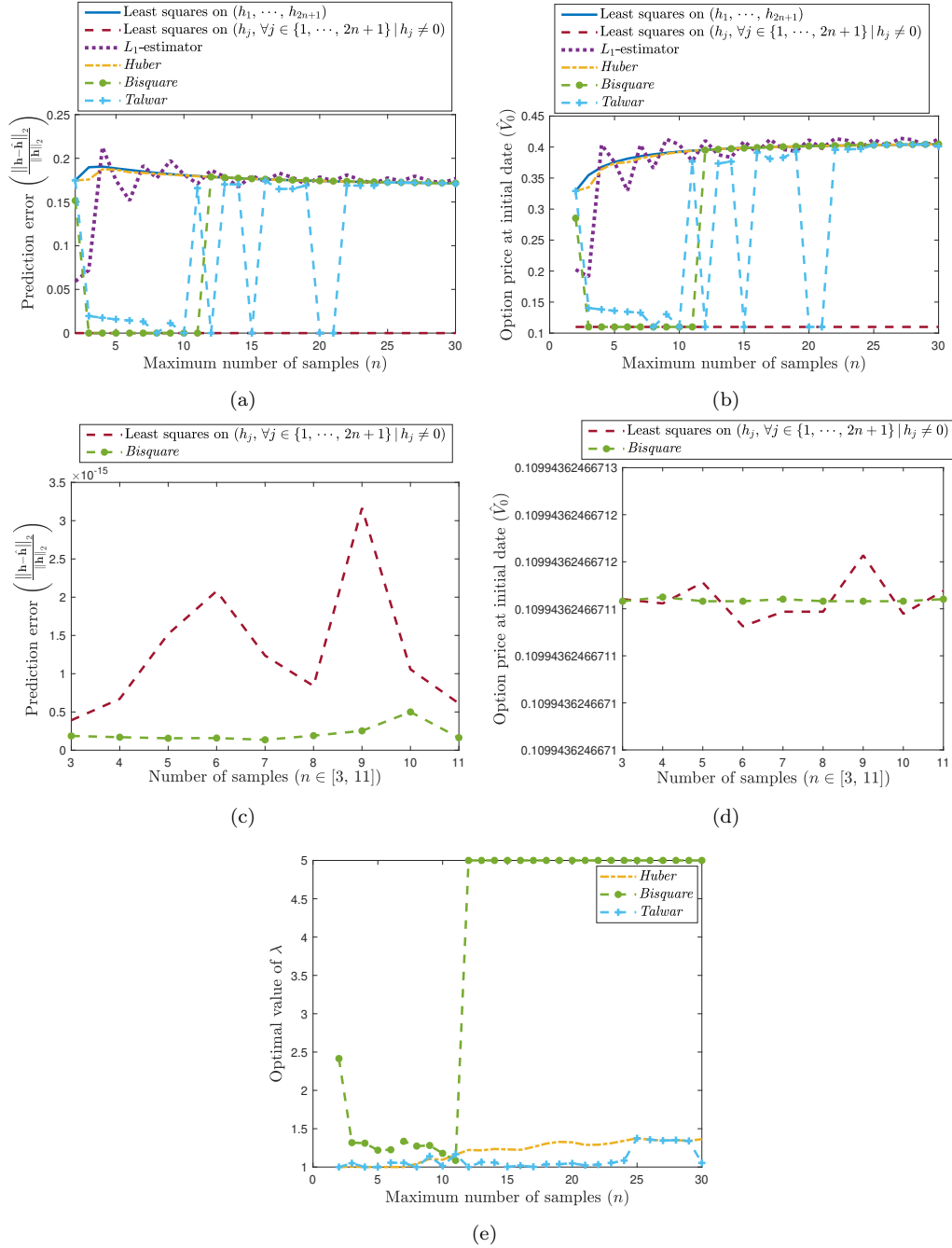


Figure 9: Evaluations when $n > 1$: (a) the PAYOFF prediction error $\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2}{\|\mathbf{h}\|_2}$, (b) the predicted European call option pricing at initial date (\hat{V}_0), (c) the PAYOFF prediction error when $n \in [3, 11]$, (d) the predicted option price \hat{V}_0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

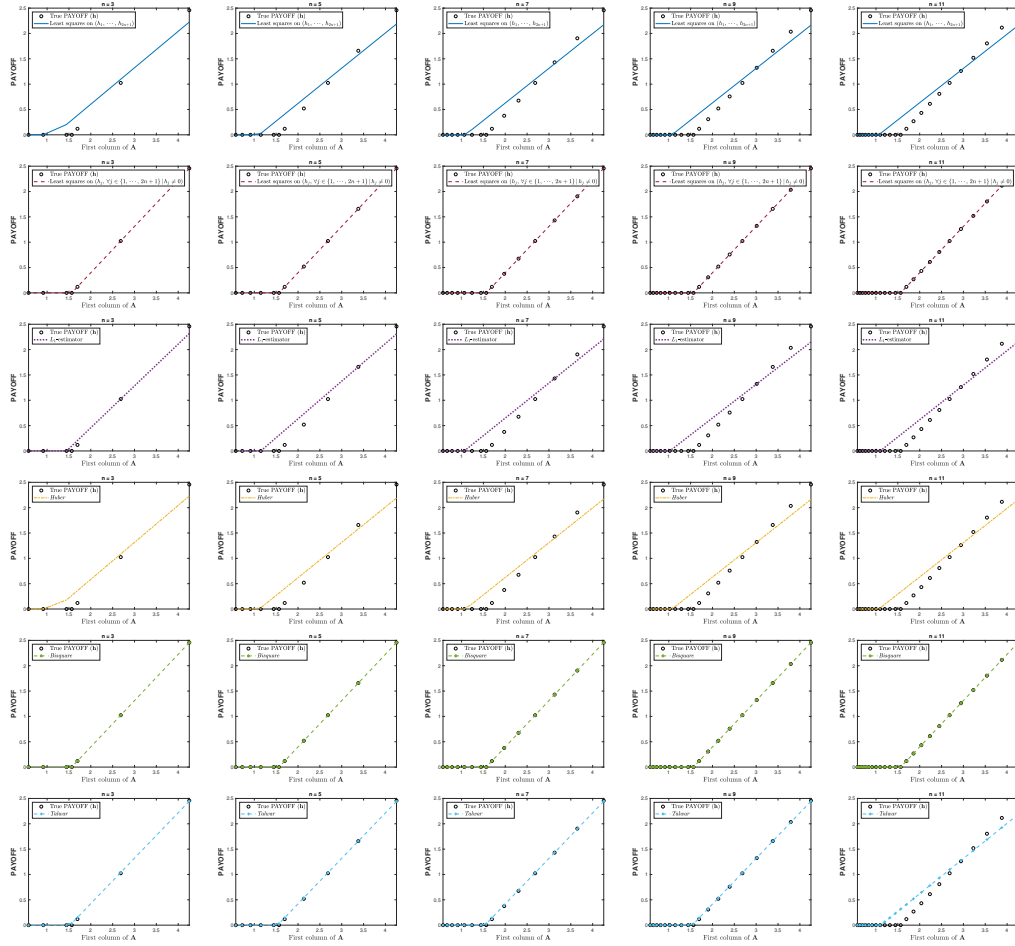


Figure 10: The difference between \mathbf{h} and $\hat{\mathbf{h}}$ for the six option pricing prediction models when $n \in \{3, 5, 7, 9, 11\}$. From top row to bottom row: Least squares on (h_1, \dots, h_{2n+1}) , Least squares on $(h_j, \forall j \in \{1, 2, \dots, 2n+1\} | h_j \neq 0)$, L_1 -estimator,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

5. Conclusion and future work

The trinomial European call option pricing model is well known to be incomplete, and thus, the least squares linear regression technique can simply be used to build a self-financing hedging portfolio that does not perfectly replicates 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. Whereas the second strategy exploits some robust linear regression M -estimators. Both the 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 the PAYOFF when n is sufficiently large. However, the big challenge remains of how to develop a method that can benefit from the estimated neural network weights to efficiently predict the option price at initial date.

References

- [1] Ayres, A., 1967. Risk aversion in the warrants market. *Indus. Management Review* , 497–505.
- [2] Bachelier, L., 1900. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure 3e série*, 17, 21–86. URL: <http://www.numdam.org/articles/10.24033/asens.476/>, doi:10.24033/asens.476.
- [3] Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654. URL: <http://www.jstor.org/stable/1831029>.

- [4] Boness, A.J., 1964. Elements of a theory of stock-option value. *Journal of Political Economy* 72, 163–175. URL: <http://www.jstor.org/stable/1828962>.
- [5] Boyle, P.P., 1986. Option valuation using a three jump process. URL: <https://api.semanticscholar.org/CorpusID:150764610>.
- [6] Cox, J.C., Ross, S.A., Rubinstein, M., 1979. Option pricing: A simplified approach. *Journal of Financial Economics* 7, 229–263. doi:[https://doi.org/10.1016/0304-405X\(79\)90015-1](https://doi.org/10.1016/0304-405X(79)90015-1).
- [7] Ehrmann, M., Fratzscher, M., 2005. Exchange rates and fundamentals: new evidence from real-time data. *Journal of International Money and Finance* 24, 317–341. URL: <https://www.sciencedirect.com/science/article/pii/S0261560604001196>, doi:<https://doi.org/10.1016/j.jimonfin.2004.12.010>. exchange Rate Economics.
- [8] Huber, P.J., 1964. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics* 35, 73 – 101. URL: <https://doi.org/10.1214/aoms/1177703732>, doi:10.1214/aoms/1177703732.
- [9] Johan, B., 2016. The trinomial asset pricing model. GIstitutionen for matematiska vetenskaper, Goteborgs universitet, Thesis report .
- [10] Maronna, R.A., 2019. *Robust Statistics: Theory and Methods (with R)*. Wiley Series in Probability and Statistics, John Wiley Sons. URL: <https://doi.org/10.1002/9781119214656>.
- [11] Merton, R., 1973. Continuous-time speculative processes: Appendix to paul a. samuelson’s ”mathematics of speculative price”. *SIAM Rev.* 15, 34–38.
- [12] Orlin Grabbe, J., 1983. The pricing of call and put options on foreign exchange. *Journal of International Money and Finance* 2, 239–253. URL: <https://www.sciencedirect.com/science/article/pii/S0261560683800023>, doi:[https://doi.org/10.1016/S0261-5606\(83\)80002-3](https://doi.org/10.1016/S0261-5606(83)80002-3).

- [13] Pascucci, A., Runggaldier, W.J., 2012. Financial Mathematics : Theory and Problems for Multi-period Models. La Matematica per il 3+2. 1st ed. 2012. ed., Springer Milan, Milano.
- [14] Rosett, R., 1963. Estimating the utility of wealth from call options transactions. Doctoral dissertation, University of Rochester .
- [15] Samuelson, P.A., 2015. Rational Theory of Warrant Pricing. Springer International Publishing, Cham. pp. 195–232. doi:10.1007/978-3-319-22237-0_11.
- [16] Samuelson, P.A., Merton, R.C., 1969. A complete model of warrant pricing that maximizes utility. URL: <https://api.semanticscholar.org/CorpusID:153453511>.
- [17] Scholes, M., . A conversation with myron scholes.
- [18] Sprengle, C.M., 1961. Warrant prices as indicators of expectations and preferences. Yale economic essays 1, 178–231.
- [19] Takahashi, H., 2000. A note on pricing derivartives in an incomplete market. Graduate School of Economics, Hitotsubashi University, Discussion Papers .
- [20] 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. p. 104–106. URL: <https://doi.org/10.1145/3507485.3507502>, doi:10.1145/3507485.3507502.