8-17-2022 10:00 AM

An Analysis of Weighted Least Squares Monte Carlo

Xiaotian Zhu, *The University of Western Ontario*

Supervisor: Reesor, Mark R., *Wilfrid Laurier University*
: Stentoft, Lars, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Statistics and Actuarial Sciences

ABSTRACT

Since Longstaff and Schwartz [2001] brought the amazing Regression-based Monte Carlo (LSMC) method in pricing American options, it has received heated discussion. Based on the research done by Fabozzi et al. [2017] that applies the heteroscedasticity correction method to LSMC, we further extend the study by introducing the methods from Park [1966] and Harvey [1976]. Our work shows that for a single stock American Call option modelled by GBM with two exercise opportunities, WLSMC or IRLSMC provides better estimates in continuation value than LSMC. However, they do not lead to better exercise decisions and hence have little to no effect on option price estimates. Our work finally indicates that in terms of real-life options pricing modelled by univariate GBM, bivariate GBM, and univariate GARCH, WLSMC or IRLSMC are not effective at producing more efficient price estimates .

Key Words: [LSMC, Regression, Heteroscedasticity Correction].

Summary for Lay Audience

As financial derivatives commonly trade in the market, options and their pricing have a long history. Back in the time when the computational power is weak, people do replicate portfolios to treat the price of an option as a combination of risk-free and risky assets. However, after the development of computational techniques and the study on continuous-time stochastic process, pricing options with certain formulas become possible.

However, there are many types of options trading in the market. The most famous two are the European option and the American option. The major difference between the two options is the European option allows option holders to exercise the option only on the maturity date while option holders can exercise the American option on any dates before maturity.

Obviously, pricing American options is tricky as the decision made by option holders to hold or exercise before maturity is unclear, and thus it has no closed-form solution. However, researchers have put a lot of effort into American options valuation techniques and come up with many fabulous ideas. The method introduced by Longstaff and Schwartz [2001] that uses regressions (denoted as LSMC) to estimate the continuation value of American options has shown great success. As one of the most commonly used regression, ordinary least-squares requires errors to be homoscedastic. The study by Fabozzi et al. [2017] showed heteroscedasticity exists in LSMC and suggested a method for correcting this, indicating a marginal improvement in price estimator efficiency. Here we investigate other methods to correct the heteroscedasticity in LSMC and find out to what extent the correction affects the option prices. Our work is a critical assessment of the work by Fabozzi et al. [2017].

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Dr. R.Mark Reesor and co-supervisor Dr. Lars Stentoft for their selfless help, support and guidance throughout my thesis year. When I got lost in the research, it was the weekly meeting with you that helps me to set a goal and rebuild my confidence. I am extremely looking forward to starting my Ph.D. career under your supervision.

Special thanks to my colleague Dr. Francois-Michel Boire, my former instructors Dr. Simon Bonner and Dr. Hyukjun Gweon for sharing their precious knowledge and answering my questions.

I would like to extend my sincere thanks to my parents Yajun Ni, Guobin Zhu and grandparents Luan Wang, Zhenqin Teng for their continuous love, understanding and support. Thank you for bringing me up and I love you all.

Lastly, I would be remiss in not mentioning my friends, my snooker coach and everyone who work in WSC 191 for the entertainment and enjoyable chat.

Contents

List of Figures

List of Tables

List of Abbreviations

LSMC Least Squares Monte Carlo

WLSMC Weighted Least Squares Monte Carlo

IRLSMC Iteratively Reweighted Least Squares Monte Carlo

BM Brownian Motion

GBM Geometric Brownian Motion

ITM In The Money

OOTM Out Of The Money

ATM At The Money

SDE Stochastic Differential Equation

BLUE Best Linear Unbiased Estimate

i.i.d Independent and Identically Distributed

c.d.f Cumulative Density Function

RMSE Root Mean Squares Error

MRE Mean Relative Error

List of Notations

W Brownian motion

T maturity time of a stock/option

t any time between $[0, T]$

S A stochastic process of stock

μ drift

τ stopping time

r interest rate

q dividend yield

Z random samples from $N(0, 1)$

σ volatility/variance

ρ correlation

h volatility in GARCH/NGARCH model

$\alpha_G, \alpha_{NG}, \beta_G, \beta_{NG}, \omega, \gamma, \zeta$ parameters of GARCH/NGARCH model

K strike price

d Black-Scholes formula

U random variables

u i.i.d random samples

y, g some functions

V price/pricing process of an option

$\boldsymbol{Y}$ response variables

$\boldsymbol{X}$ predictors

$\boldsymbol{\beta}$ regression coefficients

$\boldsymbol{\epsilon}$ residuals

$\boldsymbol{e}$ fitted residuals

$\boldsymbol{W}$ weight matrix

$\boldsymbol{P}$ diagonal matrix

λ parameter in WLS method 1

$\boldsymbol{Z}$ predictor matrix for WLS method 2

$\boldsymbol{v}$ error in WLS regression

N number of time step/exercise opportunities

n number of simulated stocks in a sample

m number of correlated stock paths in a set, a sample has n sets

i, j counting numbers

l an in-the-money path

L a set of in-the-money paths in a sample

$\mathscr{L}$ a set of in-the-money paths across all samples

$\boldsymbol{x}$ a column vector

k number of observations

$M, \mathscr{M}$ polynomial orders used in regression

Q number of replicates/samples

C continuation value

E Error

Chapter 1

Introduction

1.1 Background

Unlike the traditional European options that only exercise on the expiration date and can be priced using Black & Scholes formula, pricing American options is challenging, as it allows option holders to exercise the option at any time before expiration. Due to the uncertainty in pricing American options, there is been no closed-form solution for all but a few special cases.

Assuming time can be discretized into small sub-intervals, several approaches were invented. Brennan and Schwartz [1977] first introduced the finite difference method based on the Black & Scholes differential equation, where this method can provide standard benchmark prices for American options. Then, Cox et al. [1979] used a binomial tree to price American options, which is very easy to implement in practice. Due to technical constraints, it took many years for simulation approaches to appear. Broadie and Glasserman [1997] proposed the stochastic tree simulation approach that produces biased high and biased low estimators, both converging to the true price as the sample size increases. Longstaff and Schwartz [2001] proposed the Regression-based Monte Carlo method (denoted as LSMC) which quickly become popular, and Stentoft [2004] furthermore proved its convergence in probability.

However, LSMC is not a perfect method with many problems that need to be solved, and many improvements have been proposed. For example, Longstaff and Schwartz [2001] raised the concern of choosing the most suitable basis function. For the bias associated with estimators in LSMC, Kan and Reesor [2012] constructed a bias-corrected estimator by subtracting estimated bias from the uncorrected estimator at each exercise opportunity. Stentoft [2019] showed that the pricing of the American Call option could be improved significantly by using Put-Call Symmetry. Boire et al. [2021] used Importance Sampling as a variance reduction technique to reduce the variance in the estimates and also correct the bias.

1.2 Previous Study

The LSMC method is based on a regression at each time step to find estimates for continuation values of each path. However, there are not many works focused on the regression assumptions and Fabozzi et al. [2017] is an exception that first discussed the heteroscedasticity in LSMC. By deriving a heuristic proof, their work demonstrated that heteroscedasticity exists in LSMC regression and proved that error terms in the LSMC American Put option are heteroscedastic. Their results showed that pricing American basket put option in multiple stocks can be improved using Weighted Least Squares Monte Carlo (WLSMC) with less heteroscedasticity.

1.3 Our work

This paper continues the study on WLSMC by introducing two other methods from Park [1966] and Harvey [1976], and furthermore applying Iteratively Reweighted Least Squares Monte Carlo (IRLSMC) to reduce heteroscedasticity in the regressions. The models we consider for underlying assets in the thesis are Geometric Brownian Motion (denoted as GBM), two dimensional GBM, GARCH and its extension, the NGARCH model.

This thesis is constructed as follows: Section 2 reviews some mathematical materials that are highly correlated to this thesis. Section 3 shows the detailed LSMC, WLSMC and IRLSMC algorithms, respectively. Section 4 provides our preliminary simulation results with two exercise opportunities that mainly compare the estimated continuation values to true continuation values. Section 5 includes our results of in-sample American options' pricing with multiple exercise opportunities by applying LSMC/WLSMC/IRLSMC. Finally Section 6 gives our conclusion and suggestions for future works. In Appendix A, we will discuss the pros and cons of the WLS methods with some makeup data.

Chapter 2

Mathematical Background

In this section, we will review the two commonly used processes in stock path generation, GBM and GARCH, European/American options with their pricing, Monte Carlo method, OLS, homoscedasticity and heteroscedasticity and WLS. [1]

2.1 GBM and GARCH Models

2.1.1 GBM Model

Since 1827 the first time Brownian Motion, BM, was used to describe irregular moves of particles in water by Robert Brown, it has been applied to many different areas in science. In Mathematics, assuming a probability space $(\mathbf{\Omega}, \mathscr{F}, \mathscr{P})$, a BM is a random Process $\{W_t : t \in [0, \infty)\}$ with following properties

- $W_0 = 0$ almost surely.
- W_t has independent increments.
- W_t has stationary Gaussian increments. That is, let $s \leq t$, W_t - $W_s \sim N(0, t-s)$.

[1] Skip this section if you are familiar with every concepts.

- W_t has continuous paths.

Based on BM, a stochastic process, S_t, is a GBM having dynamics given by the following SDE

$$dS_t = (r-q)S_t dt + \sigma S_t dW_t, \tag{2.1.1}$$

where S_t is a stock price, r indicates the continuous risk-free rate, q is the continuous dividend yield, W_t represents BM, and σ is the volatility.

For stock paths generation, assume $t \in [0, T]$ and $[0, T]$ can be divided into n uniform sub-intervals, $\Delta t = \frac{T}{n}$, for each path at time $t + \Delta t$

$$S_{t+\Delta t} - S_t = (r-q)S_t \Delta t + \sigma S_t \sqrt{\Delta t} Z_t, \tag{2.1.2}$$

where $Z_t \overset{i.i.d}{\sim} N(0,1)$ is a random sample chosen from standard normal distribution. This method is known as an Euler discretization. However, it is an approximation of equation 2.1.1 and loops are required to generate a full path in software, leading to slightly inaccurate results and high computational cost. To solve both problems, a numerical solution is needed.

To give a numerical solution of this SDE, knowledge of Itô's lemma is necessary. Let $f(t, S_t)$ represents a function with two variables t and S_t, and S_t satisfies the following SDE

$$dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t, \tag{2.1.3}$$

where μ and σ are the drift and diffusion functions respectively.

Thus, Itô's lemma is defined as

$$df(t, S_t) = \frac{\partial f(t, S_t)}{\partial t} dt + \frac{\partial f(t, S_t)}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 f(t, S_t)}{\partial S_t^2} (dS_t)^2. \tag{2.1.4}$$

Let $f(t, S_t) = \ln S_t$ with $S_0 \in [0, \infty)$, the dynamics of $f(t, S_t)$ by equation 2.1.4 is

$$d \ln S_t = \frac{1}{S_t} dS_t - \frac{1}{2S_t^2} (dS_t)^2, \tag{2.1.5}$$

combining equation 2.1.5 with equation 2.1.1, and after some simple elimination and calculation, it is easy to derive

$$S_t = S_0 e^{(r-q-\frac{1}{2}\sigma^2)t + \sigma W_t}. \tag{2.1.6}$$

Based on equation 2.1.6 and same setting in equation 2.1.2, let k be an integer such that $1 \leq k \leq n$. Now for each path at time $k\Delta t$

$$S_{k\Delta t} = S_0 e^{(r-q-\frac{1}{2}\sigma^2)k\Delta t+\sigma\sqrt{k\Delta t}\sum_{i=1}^{k} Z_i}, \tag{2.1.7}$$

$Z_i \overset{i.i.d}{\sim}$ N(0,1) are random samples chosen from Standard Normal Distribution.

To generate multiple stock paths using GBM, let $S^1, ..., S^m$ represent the simulated correlated stocks. The correlation matrix $\mathbf{\Sigma} = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1} & \rho_{m2} & \cdots & \rho_{mm} \end{pmatrix}$ with $\rho_{ii} = 1$ and $\forall i \neq j$, $\rho_{ij} = \rho_{ji}$. The algorithm is given by, as stated in Goudenège et al. [2019]:

1. Set $t = 0$ with $S_0^1, ..., S_0^k$ to be the initial value.
2. For each stock S^i, $1 \leq i \leq m$, simulate k Z^i random samples from $N(\mathbf{0}, \mathbf{\Sigma})$ where $\mathbf{0}$ is a $m \times m$ null matrix and $\mathbf{\Sigma}$ is defined above. Use Equation 2.1.7 as the stock price for S^i at time $k\Delta t$.

2.1.2 GARCH Model

As a Generalized ARCH model that was firstly introduced by Engle [1982], the idea of GARCH(p,q) model was suggested by Bollerslev [1986]. In comparison to the GBM model, GARCH models have their own advantage, as the volatility terms are based on past volatilities and some random terms instead of a fixed number. Thus, GARCH models reproduce features of financial time series, such as volatility clustering, and typically provide a better fit to option prices than GBM. Duan [1995] provided the method of using the GARCH(1,1) model in option pricing and followed by Stentoft [2005], who combined the GARCH(1,1) option pricing model with LSMC in pricing the American Put option. In this thesis, we are interested in whether WLSMC or IRWLSMC can help to remove heteroscedasticity effectively with data from GARCH models and improve their pricing. This thesis implemented two models: the ordinary GARCH model and the NGARCH model from Engle and Ng [1993].

The stock price process for the ordinary GARCH and NGARCH models shares the same formula. Let S_t represent a stochastic process (underlying asset) with the following equations, listed in Stentoft [2011]

$$S_t = S_{t-1} e^{r - \frac{1}{2}h_t + \sqrt{h_t} Z_t^*}, \tag{2.1.8}$$

where $Z_t^* \overset{i.i.d}{\sim}$ N(0,1) and h_t is the volatility term. h_t is a slightly different term for the ordinary GARCH and NGARCH models. For the ordinary GARCH model

$$h_t = \omega + \beta_G h_{t-1} + \alpha_G h_{t-1} (\tilde{Z}_{t-1}^*)^2, \tag{2.1.9}$$

and for NGARCH model

$$h_t = \omega + \beta_{NG} h_{t-1} + \alpha_{NG} h_{t-1} (\tilde{Z}_{t-1}^* + \gamma)^2, \tag{2.1.10}$$

where $\tilde{Z}_{t-1}^* = Z_{t-1}^* - \zeta$, $\alpha_G, \alpha_{NG}, \beta_G, \beta_{NG}, \omega, \gamma, \zeta$ are parameters to be specified[1].

Thus, for stock path generation, we assume h_t^G and h_t^{NG} represent the volatility terms for ordinary GARCH and NGARCH models, respectively. The initial values h_0^G and h_0^{NG} are set to the unconditional level of variance, specifically.

To generate GARCH paths, the following steps are used.

1. Set t = 0 with $h_0^G = \frac{\omega}{1-\beta_G-\alpha_G}$ to be the initial value.
2. Simulate $Z_{t+\Delta t}^G \overset{i.i.d}{\sim} N(0,1)$.
3. Compute $h_{t+\Delta t}^G = \omega + \beta_G h_t^G + \alpha_G h_t^G (Z_t^G - \zeta)^2$, where Z_t^G is simulated in the previous time step.
4. Compute $S_{t+\Delta t}^G = S_t e^{r - \frac{1}{2}h_{t+\Delta t}^G + \sqrt{h_{t+\Delta t}^G} Z_{t+\Delta t}^G}$.
5. Set $t = t + \Delta t$.
6. Repeat step 2 to 5 until $t = T$, the time horizon of interest.

To generate stock paths for the NGARCH model, straightforward modifications to the above algorithm apply, specifically.

[1] As stated in Stentoft [2011], these parameters need to be empirically plausible, here $\zeta = \lambda$

1. Set t = 0 with $h_0^{NG} = \frac{\omega}{1-\beta_{NG}-\alpha_{NG}(1+\gamma)^2}$ to be the initial value.

2. Simulate $Z_{t+\Delta t}^{NG} \overset{i.i.d}{\sim} N(0,1)$.

3. Compute $h_{t+\Delta t}^{NG} = \omega + \beta_{NG} h_t^{NG} + \alpha_{NG} h_t^{NG}(Z_t^{NG} - \zeta + \gamma)^2$, where Z_t^{NG} is simulated in the previous time step.

4. Compute $S_{t+\Delta t}^{NG} = S_t e^{r - \frac{1}{2} h_{t+\Delta t}^{NG} + \sqrt{h_{t+\Delta t}^{NG}} Z_{t+\Delta t}^{NG}}$.

5. Set $t = t + \Delta t$.

6. Repeat step 2 to 5 until $t = T$, the time horizon of interest.

2.2 Options and Their Pricing

2.2.1 European Option

A European option gives the option holder the right to exercise the option at the expiration date. In mathematics, assume a European option with expiration time T, stock price S_t with $t \leq T$, strike K, and payoff function $f(T, S_T)$ is defined as

$$\text{European Call option} : f_C(T, S_T) = \max(S_T - K, 0), \text{ and} \tag{2.2.1}$$

$$\text{European Put option} : f_P(T, S_T) = \max(K - S_T, 0). \tag{2.2.2}$$

Pricing of a European option using risk-neutral valuation is given by

$$V = \mathbb{E}^{\mathbb{Q}}[e^{-rT} f(T, S_T)], \tag{2.2.3}$$

where V is the price of the European option and the expectation is done under a risk-neutral probability measure $\mathbb{Q}$.

Black and Scholes [1973] first derived the closed-form solutions to the European option, where

$$V_C^{EU} = S_0 e^{-qt} N(d_1) - K e^{-rt} N(d_2), \text{ and} \tag{2.2.4}$$

$$V_P^{EU} = K e^{-rt}(1 - N(d_2)) - S_0 e^{-qt}(1 - N(d_1)), \tag{2.2.5}$$

V_C^{EU} and V_P^{EU} are the prices of European Call and European Put respectively. $N(.)$ calculates the c.d.f of $N(0,1)$. d_1 and d_2 are calculated as

$$d_1 = \frac{1}{\sigma\sqrt{t}}\left[\ln\left(\frac{S_t}{K}\right) + t\left(r - q + \frac{\sigma^2}{2}\right)\right], \text{and} \tag{2.2.6}$$

$$d_2 = d_1 - \sigma\sqrt{t}, \tag{2.2.7}$$

where r is the interest rate, q is the dividend yield and σ is the volatility.

2.2.2 American Option

An American option gives the option holder the right to exercise the option any time before expiration date. In mathematics, an American option with expiration time T, stock price S_t with $t \leq T$, strike K, payoff function $f(t, S_t)$ is defined as

$$\text{American Call option} : f_C(t, S_t) = \max(S_t - K, 0), \text{and} \tag{2.2.8}$$

$$\text{American Put option} : f_P(t, S_t) = \max(K - S_t, 0). \tag{2.2.9}$$

As stated in Myneni [1992] in section 3, the price process $V(t, S_t)$ of a American style option under risk-neutral evaluation is given by

$$V(t, S_t) = \sup_{\tau\in[t,T]} \mathbb{E}_t^{\mathbb{Q}}[e^{-r\tau} f(\tau, S_\tau)], \tag{2.2.10}$$

where τ is a stopping time and $\mathbb{E}_t^{\mathbb{Q}}$ denotes the risk-neutral expectation conditional on information known at time t.

The American option value, $V(t, S_t)$ satisfies the property $V(t, S_t) \geq f(t, S_t)$, otherwise there is arbitrage opportunity. Also Myneni [1992] outlined in section 4 that before expiration date T, $V(t, S_t)$ satisfies the following Black-Scholes Differential Equation. Let $S = S_t$, $V(t, S) = V(t, S_t)$, we have

$$\frac{\partial V(t,S)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2} + (r-q)S\frac{\partial V(t,S)}{\partial S} = (r-q)V(t,S), \tag{2.2.11}$$

with boundary conditions

$$V(T, S_T) = f(T, S_T). \tag{2.2.12}$$

Specifically, for an American put option, we have

$$\lim_{S \to \infty} V(T, S) = 0, \text{and} \tag{2.2.13}$$

$$\lim_{S \to 0} V(T, S) = K. \tag{2.2.14}$$

2.2.3 Monte Carlo Method

As stated in James [1980], let U be a continuous random variable with support on (a, b) and probability density function $f_U(u) = y(u)$. Assume $g(U)$ to be some function of U. Now suppose we are interested in calculating

$$V = E(g(U)) = \int_a^b g(u)y(u)du. \tag{2.2.15}$$

Suppose that $g(u)y(u)$ are complicated functions such that the antiderivative is not derivable. In this case, an estimate $\hat{V}$ is required to replace V. One approach to estimate V is the Monte Carlo Estimator $\tilde{V}_{MC}$.

Assume $U_1, ..., U_n$ have the same i.i.d distribution with U and $Var(g(U)) = \sigma^2$. The Monte Carlo Estimator $\tilde{V}_{MC}$ is defined as

$$\tilde{V}_{MC} = \frac{1}{n}\sum_{i=1}^{n} g(U_i). \tag{2.2.16}$$

It is easy to show that $E(\tilde{V}_{MC}) = V$ and $Var(\tilde{V}_{MC}) = \frac{\sigma^2}{n}$. Based on Equation 3.3.2, let $u_1, ..., u_n$ be i.i.d random samples generated from $y(u)$, we have the MC estimates to be

$$\hat{V}_{MC} = \frac{1}{n}\sum_{i=1}^{n} g(u_i) \quad \text{with} \quad \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(g(u_i) - \hat{V}_{MC})^2}{n-1}}. \tag{2.2.17}$$

The Law of Large number tells us $\hat{V}_{MC} \to V$ as $n \to \infty$ and the Central Limit Theorem allows for easy construction of a 95% confidence interval for V by

$$[\hat{V}_{MC} - 1.96 * \frac{\hat{\sigma}}{\sqrt{n}}, \quad \hat{V}_{MC} + 1.96 * \frac{\hat{\sigma}}{\sqrt{n}}]. \tag{2.2.18}$$

2.3 OLS and WLR

2.3.1 Ordinary Least Squares

As stated in Greene [2012], consider the following equation

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon} \quad \text{with} \quad E(\boldsymbol{\epsilon}) = \boldsymbol{0}, \quad Var(\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{I}. \tag{2.3.1}$$

$\boldsymbol{Y} = (y_1, y_2, ..., y_n)^T$ is a $n \times 1$ vector that represents the response variable. $\mathbf{1}$ is the $n \times 1$ vector $(1, ..., 1)^T$, $\boldsymbol{X} = (\boldsymbol{1}, \boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_p})$, where $\boldsymbol{x_i} = (x_{i1}, x_{i1}, ..., x_{in})^T$ with $1 \le i \le p$, is defined to be a $n \times (p+1)$ matrix. $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, ..., \beta_p)^T$ is a $(p+1) \times 1$ vector of regression coefficients, $\boldsymbol{\epsilon} = (\epsilon_0, \epsilon_1, \epsilon_2, ..., \epsilon_n)^T$ is a $n \times 1$ vector as the error term. σ is a constant and $\boldsymbol{I}$ is a $n \times n$ identity matrix. With the notation listed above, for each y_i in $\boldsymbol{Y}$, there is a vector $(1, x_{i1}, ..., x_{ip})$ where

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + \epsilon_i, \tag{2.3.2}$$

and Ordinary Least Squares gives estimates of $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{Y}}$ and $\boldsymbol{e}$ as [1]

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X^T X})^{-1} \boldsymbol{X^T Y}, \tag{2.3.3}$$

$$\hat{\boldsymbol{Y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}} \quad \text{and} \quad \boldsymbol{e} = \boldsymbol{Y} - \hat{\boldsymbol{Y}}. \tag{2.3.4}$$

2.3.2 Homoscedasticity and Heteroscedasticity

By equation 2.3.1, $\boldsymbol{\epsilon}$ should have mean $\mathbf{0}$ and covariance matrix $\sigma^2\boldsymbol{I}$ where σ is a constant. As one of the assumptions for Ordinary Least Squares, homoscedasticity requires sigma to be a constant. However, there are some cases where σ is not a constant with $\sigma = \sigma_i$. This is known as heteroscedasicity and violates the assumption of linear regression.

There are two ways to detect evidence of heteroscedasticity graphical and statistical. Graphical evidence comes with observing the residual plots. If all residuals are

[1] Note that under the assumptions given above, $\hat{\boldsymbol{\beta}}$ is the BLUE of $\boldsymbol{\beta}$.

uniformly distributed with no apparent pattern, we consider this as no evidence of heteroscedasticity. Statistical evidence is provided via the Breusch-Pagan test (BPtest) first introduced by Breusch and Pagan [1979]. The test is conducted with null hypotheses H_0 = heteroscedasicity does not exist, and alternative hypotheses H_a = heteroscedasicity exists. Typically a p.value of 0.05 is used for the significance level. If the p.value is greater than 0.05 we do not reject H_0 and otherwise we reject H_0 in favour of H_a and conclude the existence of heteroscedasticity [1].

2.3.3 Weighted Linear Regression

As stated in section 2.3.2, once a residual plot/BPtest shows there is graphical/statistical evidence of heteroscedasicity in regression, we should consider applying Weighted Linear Regression (WLR). If the source of heteroscedasicity is known, suppose for some positive function f(x_{ik},y_i), where $1 \leq k \leq p$, and assume $\sigma_i^2 = \sigma^2$f(x_{ik},y_i). Let $\boldsymbol{P}$ = diag($\frac{1}{\sigma\sqrt{f(x_{1k},y_1)}}$, ..., $\frac{1}{\sigma\sqrt{f(x_{nk},y_n)}}$), $\boldsymbol{W} = \boldsymbol{P}^T\boldsymbol{P}$, where now we transform the original model by multiplying both sides of Equation 3.4.1 by $\boldsymbol{P}$ given

$$\boldsymbol{PY} = \boldsymbol{PX\beta} + \boldsymbol{P\epsilon} \quad \text{with} \quad E(\boldsymbol{P\epsilon}) = \boldsymbol{0}, \quad Var(\boldsymbol{P\epsilon}) = \boldsymbol{I}. \tag{2.3.5}$$

this results in the weighted least squares estimator and fitted value

$$\hat{\boldsymbol{\beta}}^{\boldsymbol{W}} = (\boldsymbol{X^TWX})^{-1}\boldsymbol{X^TWY}, \text{and} \tag{2.3.6}$$

$$\hat{\boldsymbol{Y}}^{\boldsymbol{W}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}^{\boldsymbol{W}}, \tag{2.3.7}$$

respectively.

However, in the real world, sources of heteroscedasicity are unknown for most of the cases. As stated in Greene [2012], let $\boldsymbol{W}$ = diag($\frac{1}{\sigma_1^2}$, ..., $\frac{1}{\sigma_n^2}$), then a consistent estimator $\tilde{\boldsymbol{\beta}}^{\boldsymbol{W}}$ is

$$\tilde{\boldsymbol{\beta}}^{\boldsymbol{W}} = (\boldsymbol{X^TWX})^{-1}\boldsymbol{X^TWY}. \tag{2.3.8}$$

The weight matrix is estimated by $\widehat{\boldsymbol{W}}$ = diag($\frac{1}{\hat{\sigma}_1^2}$, ..., $\frac{1}{\hat{\sigma}_n^2}$). This gives estimates of the

[1]For more technical details, see Breusch and Pagan [1979].

regression coefficients, fitted values and residuals

$$\hat{\hat{\boldsymbol{\beta}}}^{\boldsymbol{W}} = (\boldsymbol{X}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Y}, \tag{2.3.9}$$

$$\hat{\hat{\boldsymbol{Y}}}^{\boldsymbol{W}} = \boldsymbol{X}\hat{\hat{\boldsymbol{\beta}}}^{\boldsymbol{W}}, \text{ and} \tag{2.3.10}$$

$$\boldsymbol{e}^{\boldsymbol{W}} = \boldsymbol{Y} - \hat{\hat{\boldsymbol{Y}}}^{\boldsymbol{W}}, \tag{2.3.11}$$

respectively.

What remains is how to specify $\widehat{\boldsymbol{W}}$ or equivalently $\widehat{\sigma_i^2}$, $i = 1, ..., n$. Here we provide three methods to solve this problem.

Method 1: Park [1966] has proposed a model that assumes σ_i^2 is proportional to an unknown power of one or more predictors in the regression. Suppose σ_i^2 is proportional to the kth predictor $\boldsymbol{x_k}$. Let $\boldsymbol{\sigma^2} = (\sigma_1^2, ..., \sigma_n^2)^T$, we assume the following equation for the ith entry

$$\sigma_i^2 = \sigma^2(x_{ik})^\lambda, \tag{2.3.12}$$

where λ is an unknown parameter that needs to be estimated.

Furthermore, let $\boldsymbol{e^2} = (e_1^2, ..., e_n^2)^T$ and $\boldsymbol{\epsilon^2} = (\epsilon_1^2, ..., \epsilon_n^2)^T$. Heuristically since $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}$, $\boldsymbol{e^2}$ is approximately equal to $\boldsymbol{\epsilon^2}$ and as $E\boldsymbol{(\epsilon^2) = \sigma^2}$, we have approximately

$$e_i^2 = \sigma^2(x_{ik})^\lambda e^{v_1}, \tag{2.3.13}$$

where v_1 is the error term for method 1, and by taking natural logarithm on both sides, it is easy to get

$$\ln e_i^2 = \ln \sigma^2 + \lambda \ln x_{ik} + v_1. \tag{2.3.14}$$

Therefore, let $\psi_0 = \ln \sigma^2$ and $\psi_1 = \lambda$, an OLS is constructed with predictor variable $\ln x_{ik}$, response variable $\ln e_i^2$ and error term v_1. By taking exponential on $\ln e_i^2$, using $\widehat{e_i^2}$ as an estimate of σ_i^2 and constructing $\widehat{\boldsymbol{W}} = \text{diag}(\frac{1}{\widehat{e_1^2}}, ..., \frac{1}{\widehat{e_n^2}})$, we obtain the estimated $\hat{\hat{\boldsymbol{\beta}}}^{\boldsymbol{W}}$ from equation 3.7.5.

Method 2: This method is outlined in Greene [2012]. Let $\boldsymbol{Z} = (\boldsymbol{1, z_1, ..., z_h})$ be a set of variables that may or may not be a subset of $\boldsymbol{X}$, and $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_h)$ be the coefficient vector that corresponds to $\boldsymbol{Z}$. With the same notation in method 1, we can assume the following structure

$$\boldsymbol{\sigma^2} = \boldsymbol{Z\alpha}. \tag{2.3.15}$$

Therefore, we can obtain the following OLS regression

$$\boldsymbol{e^2} = \boldsymbol{Z\alpha} + \boldsymbol{v_2}, \tag{2.3.16}$$

where $\boldsymbol{v_2}$ is the error term for method 2.

Similar to method 1, let $\widehat{\boldsymbol{e^2}}$ be the fitted value and $\widehat{\boldsymbol{W}} = \text{diag}(\frac{1}{\widehat{e_1^2}}, \ldots, \frac{1}{\widehat{e_n^2}})$, we are able to get the updated $\hat{\hat{\boldsymbol{\beta}}}^{\boldsymbol{W}}$.

Method 3: This model is introduced by Harvey [1976]. Instead of modelling $\boldsymbol{\sigma^2}$ by equation 3.7.11, now suppose we have the following structure:

$$\boldsymbol{\sigma^2} = e^{\boldsymbol{Z\alpha}} \quad \text{or} \quad \ln \boldsymbol{\sigma^2} = \boldsymbol{Z\alpha}, \tag{2.3.17}$$

and $\sigma^2 = e^{\alpha_0}$.

Similarly,

$$\boldsymbol{e^2} = e^{\boldsymbol{Z\alpha} + \boldsymbol{v_3}} \quad \text{or} \quad \ln \boldsymbol{e^2} = \boldsymbol{Z\alpha} + \boldsymbol{v_3}, \tag{2.3.18}$$

where $\boldsymbol{v_3}$ is the error term for method 3.

With $\boldsymbol{z_i}$ to be predictors and $\ln \boldsymbol{e^2}$ to be response variable, we obtain $\hat{\boldsymbol{\alpha}}$. However, Harvey [1976] stated that the calculated $\hat{\alpha_0}$ is not a consistent estimator, and could be corrected by adding a constant 1.2704. In this case, with $\hat{\boldsymbol{\alpha}} = (\hat{\alpha_0} + 1.2704, \hat{\alpha_1}, \ldots, \hat{\alpha_h})$ and $\widehat{\boldsymbol{W}} = \text{diag}(\frac{1}{\widehat{e_1^2}}, \ldots, \frac{1}{\widehat{e_n^2}})$, we obtain $\hat{\hat{\boldsymbol{\beta}}}^{\boldsymbol{W}}$.

Chapter 3

LSMC and Its Extensions

3.1 Least Squares Monte Carlo

Longstaff and Schwartz [2001] introduced LSMC by showing a short but intuitive example. In general, by applying backward dynamic programming, the algorithm to price an American put option with paths simulated from univariate GBM model using LSMC is described as

1. Discretize time $[0, T]$ into N time steps, generate a set of stock paths with a total number n from equation 2.1.7 and record them as S_i^j, where i represents a simulated path and j is the time index. [1]

2. Calculate payoffs P_i^N at T (time index N) by equation 2.2.9.

3. Discount P_i^N back to time step $N-1$ as P_i^{N-1}. Since only the stock paths that are below the strike would be considered to exercise, thus, record all ITM paths with a set L and other paths with a set $\bar{L}$. Suppose for an ITM path $l \in L$, use S_l^{N-1} as predictors and corresponding P_l^{N-1} as response variable. Construct a

[1] For example, S_{10}^{10} means the 10^{th} path at time index 10.

regression with polynomial of degree M, where

$$P_l^{N-1} = \sum_{i=0}^{M} \beta_i^{N-1} (S_l^{N-1})^i + \epsilon_l^{N-1}. \quad (3.1.1)$$

4. Calculate $\hat{\beta}_i^{N-1}$, fitted values $\hat{C}_l^{N-1}$ using equation 3.4.2 and 3.4.3 and their corresponding payoff H_l^{N-1}. Let $\hat{C}_l^{N-1}$ be continuation value and H_l^{N-1} be exercise value. Here, a decision $\hat{P}_l^{N-1}$ is made by

$$\hat{P}_l^{N-1} = \begin{cases} H_l^{N-1}, & H_l^{N-1} \geq \hat{C}_l^{N-1} \\ P_l^{N-1}, & H_l^{N-1} < \hat{C}_l^{N-1}. \end{cases} \quad (3.1.2)$$

5. If the path $i \in L$, substitute P_i^{N-1} with $\hat{P}_i^{N-1}$. Do nothing to P_i^{N-1} if $i \in \bar{L}$. Discount P_i^{N-1} to time step $N-2$ to be P_i^{N-2}, repeat step 3 and 5 until time step 1. The final estimated price $\hat{V}$ is given by, let $\Delta t = \frac{1}{N}$

$$\hat{V} = \frac{e^{-r\Delta t}}{n} \sum_{i=1}^{n} P_i^1. \quad (3.1.3)$$

Similarly, for an American Call option, we change the payoff function in steps 2 and 4. The selected paths used for regression are above the strike price.

However, the LSMC algorithm can be slightly improved at time step $N-1$. Instead of using fitted values from regression as continuation values for all ITM paths, we use their Black & Scholes European prices stated in section 2.2.1. This technique removes possible errors from analyzing regression coefficients.

For the LSMC algorithm of an American put option with paths simulated from univariate GARCH/NGARCH model, we first generate n stock paths and record their volatility terms using the algorithms listed in section 2.1.2, denoting as S_i^j and h_i^j. For the regression construction in step 3, unlike the GBM model with constant volatility σ, GARCH models' volatility terms change with time; thus they should be considered predictors. In this case, with P_l^{N-1} as response variable, (S_l^{N-1}, h_l^{N-1}) and their cross-sectional interaction terms as predictors, we construct a regression with a full basis M[1] to substitute equation 4.1.1 and proceed analogously as before.

[1]For example, a basis of 3 includes $(1, S_l, S_l^2, S_l^3, h_l, h_l^2, h_l^3, S_l * h_l, S_l^2 * h_l, S_l * h_l^2)$ as predictors.

For the LSMC algorithm of an American option with paths simulated from multivariate GBM model, we generate n sets of stocks, where each set contains m paths simulated with the correlation matrix $\mathbf{\Sigma}$. Record each path as $S^j_{i(k)}$, $1 \leq k \leq m$. A regression with full basis M is constructed to estimate the continuation values, similar to the case of GARCH model.

3.2 WLSMC in GBM Model

Fabozzi et al. [2017] proved the existence of heteroscedasticity in the regressions of LSMC for the American Put option. This violates the assumption of OLS and thus, may cause the estimated coefficients from OLS to be non-BLUE and lead to poor estimates of the continuation value. To correct the heteroscedasticity brought by the underlying assets, they implemented method 2 with quadratic mean function in predictors. Here we extend another two models stated in the second part of section 3.7, which are method 1 and method 3, respectively.

Suppose at a time step between 1 and $N-2$ to price an American put option, we use equation 3.5.1 as the initial unweighted regression and can describe all three methods with a two-step algorithm.

For Method 1 step 1, by equation 3.7.9, since stock prices are positive, we assume the following structure:

$$(e_l)^2 = \sigma^2 e^{u_l} \prod_{i=1}^{n} (S_l)^{i\lambda_i} \quad n \leq M \tag{3.2.1}$$

, and

$$\ln{(e_l)^2} = \ln \sigma^2 + \sum_{i=1}^{n} (i\lambda_i) \ln S_l + v_l, \tag{3.2.2}$$

where M is the polynomial order used to fit the continuation value.

There is only one predictor in equation 4.2.2, thus we can drop those terms with orders more than 1. Then followed by the instruction listed in Section 3.7 Method one as step 2 to conduct WLS. Since all entries in the diagonal matrix $\widehat{\boldsymbol{W}}$ are positive, it factors into two diagonal matrices $\widehat{\boldsymbol{P}}$ and $\widehat{\boldsymbol{P}}^{\boldsymbol{T}}$ such that $\widehat{\boldsymbol{W}} = \widehat{\boldsymbol{P}}^{\boldsymbol{T}}\widehat{\boldsymbol{P}}$ and use equation 3.6.5 to calculate $\hat{\hat{\boldsymbol{\beta}}}^{\boldsymbol{W}}$. To prevent possible numerical issues brought by substantial values from predictors, we use $\frac{S_l}{K}$ instead of S_l for both LSMC and WLSMC, where K is the strike price.

For Method 2 step 1, by equation 3.6.11, we can assume the following structure

$$(e_l)^2 = \alpha_0 + \alpha_1 S_l + v_l. \tag{3.2.3}$$

As stated above, Fabozzi et al. [2017] also suggests using this method. However, we find some numerical issues relating to this method by observing that the fitted values for estimated error terms are not always positive. Intuitively, as $\widehat{(e_l)^2}$ is treated as a proxy of $(\sigma_l)^2$, it should not be negative. Furthermore, consider the Least Square Method in estimating $\boldsymbol{\beta}$ is calculated by minimizing $S(\boldsymbol{\beta})$ where

$$S(\boldsymbol{\beta}) = \boldsymbol{(Y - X\beta)^T(Y - X\beta)}, \tag{3.2.4}$$

The solution, $\hat{\boldsymbol{\beta}}$, satisfies

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \boldsymbol{0}, \tag{3.2.5}$$

with the partial derivative calculated as

$$\frac{\partial^2 S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^{\boldsymbol{T}}} = 2\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}. \tag{3.2.6}$$

In order for $\hat{\boldsymbol{\beta}}$ to minimize $S(\boldsymbol{\beta})$, $\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}$ needs to be positive definite. However, suppose $\boldsymbol{W}$ is a diagonal matrix with some negative elements, $\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{X}$ may no longer be a positive definite matrix and thus, the Least Squares Method breaks down.

Therefore, we choose to run a log-transform regression by taking the natural logarithm on the response variable instead. Observing that the log-transform regression has the same format as Method 3, we combine Method 2 and Method 3, where

$$\ln{(e_l)^2} = \alpha_0 + \alpha_1 S_l + v_l. \tag{3.2.7}$$

Let $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0 + 1.2704, \hat{\alpha}_1)$ be the estimated coefficients, method 2 step 2 is the same as Method 1 step 2. We use the same predictor $\frac{S_l}{K}$ to prevent numerical issues.

Finally, the new estimated continuation values are calculated as $\hat{C}_l = \sum_{i=0}^{M} \hat{\hat{\beta}}_i^w (S_l)^i$, which are used in place of C_l in LSMC step 4.

In terms of multivariate GBM model, for method 1, assume that

$$\ln{(e_l)^2} = \ln\sigma^2 + \sum_{k=1}^{m} \lambda_k \ln S_{l(k)} + v_l, \tag{3.2.8}$$

where m represents the the number of stock paths used in LSMC.

And for method 2

$$\ln{(e_l)^2} = \alpha_0 + \sum_{k=1}^{m} \alpha_k S_{l(k)} + v_l, \tag{3.2.9}$$

with $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0 + 1.2704, \hat{\alpha}_1, ..., \hat{\alpha}_n)$.

3.3 WLSMC in GARCH single stock model

By construction, stock prices generated using the GARCH model are heteroscedastic. It is natural to consider whether this translates into additional heteroscedasticity in LSMC and to investigate effect of LSMC. Thus, we consider applying WLSMC with the two-step algorithms stated in section 5.1. For Method 1, similar to equation 5.1.2, we have

$$\ln{(e_l)^2} = \ln\sigma^2 + \lambda_1 \ln S_l + \lambda_2 \ln h_l + v_l. \tag{3.3.1}$$

And for Method 2

$$\ln{(e_l)^2} = \alpha_0 + \alpha_1 S_l + \alpha_2 h_l + v_l, \tag{3.3.2}$$

with $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0 + 1.2704, \hat{\alpha}_1, \hat{\alpha}_2)$.

3.4 IRLSMC

As stated in Greene [2012], the procedure of IRLS is described as recomputing the residual from equation 3.7.7 and reentering the computation. It has the same asymptotic properties with WLS estimators. Although he also states that IRLS may provide little additional benefit, we wonder to what extent IRLS will affect the results compared to WLS.

The algorithm of IRLS is: From equation 3.7.7, starting at $\boldsymbol{e^W}$, let $\boldsymbol{e_i^W}$, $\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{i}}^{\boldsymbol{W}}$ represent the residuals and coefficients in i^{th} iteration. We further recompute $\boldsymbol{e_i^W}$ until $||\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{i+1}}^{\boldsymbol{W}} - \hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{i}}^{\boldsymbol{W}}||$ is less than a tolerant value or a maximum iteration number is reached, the second criterion is for those that do not converge.

Note that the algorithm listed above does not provide maximum likelihood estimates for method 2. The reason we did not use the algorithm suggested by Harvey [1976] for maximum likelihood estimation is due to non-convergence in some cases.

Chapter 4

Results with Two Exercise Opportunities

OLS regressions in LSMC are heteroscedastic as shown by Fabozzi et al. [2017] and results in Chapter 6 show that the problem of heteroscedasticity can lead to more biased regression coefficient estimates. Thus in this Chapter, by applying WLSMC/IRLSMC with two methods to correct for heteroscedasticity, we investigate the effect of these corrections on the estimated continuation values, the options pricing, and the problem of heteroscedasticity. The model we consider is univariate GBM for underlying assets.

Since in the context of option pricing, WLSMC/IRLSMC modifies the estimated continuation values at each exercise opportunity before maturity, we wonder whether WLSMC or IRLSMC gives more accurate estimated continuation values compare to the true continuation values. However, there is no formula to calculate the true continuation values except when there are only two exercise opportunities, where the maturity time is at time step 2. The true continuation values at the intermediate time step can be calculated by the Black and Scholes formula, as stated in section 2.2.1.

To compare the difference between the estimated continuation values from LSMC to WLSMC or IRLSMC, we set up three criteria. Take an example of American put option, first, we generate Q samples and for the i^{th} sample such that $1 \leq i \leq Q$, it contains

n_i simulated ITM paths. This gives $\sum_{i=1}^{Q} n_i$ ITM paths in total and we put them in a set $\mathscr{L}$. Since for each ITM path there will be a true continuation value obtained by the B-S formula and a continuation value estimated from LSMC/WLSMC/IRLSMC. We then calculate the RMSE and MRE of them, and plot the error against stock price. For the j^{th} ITM paths in $\mathscr{L}$, let C_j represents the true continuation value and $\hat{C}_j$ is the estimated continuation value, the RMSE and MRE can be calculated as

$$RMSE = \sqrt{\frac{\sum_{j=1}^{|\mathscr{L}|}(C_j - \hat{C}_j)^2}{|\mathscr{L}|}}, \text{ and} \tag{4.0.1}$$

$$MRE = \frac{\sum_{j=1}^{|\mathscr{L}|}\left|\frac{C_j - \hat{C}_j}{C_j}\right|}{|\mathscr{L}|}. \tag{4.0.2}$$

The error E_j between the true continuation value and estimated continuation value at the j^{th} ITM path is obtained by

$$E_j = C_j - \hat{C}_j. \tag{4.0.3}$$

Additionally, the true continuation values lead to the correct exercise/hold decision at time step 1. We construct 2×2 frequency tables detailing the frequency of correct and incorrect decisions. This criterion is more important than the previous one as in term of option pricing, the improvement depends on better decision making. Since there are Q samples in total, the tables displayed below are the average frequencies over the Q samples.

Moreover, we split all ITM paths in $\mathscr{L}$ into 100 bins and plot the RMSE of the estimated continuation values to the true values against mean stock price in each bin. The calculation of RMSE is similar the one that introduced in criteria 1. To compare the decision made by LSMC, WLSMC and IRLSMC, using the same bins as above, we also make some 2×2 plots where each graph plots the percentage of same decision made by two methods[1] against mean stock price in each bin.

Besides these analysis, we also include the prices of the American options used above with the price estimates, standard errors, RMSE and MRE[2] compare to the true

[1] LSMC vs WLSMC, LSMC vs IRLSMC.

[2] See chapter 5.1 for the calculation of RMSE and MRE.

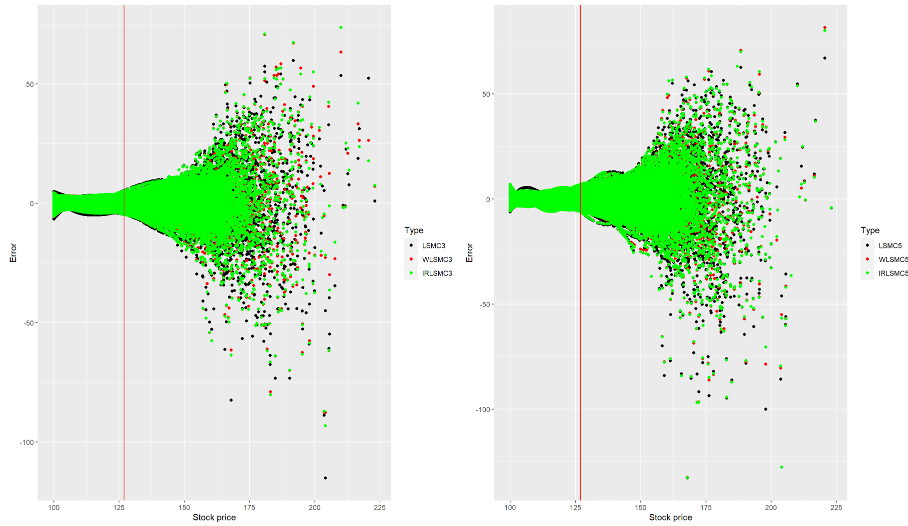

Figure 4.1: Plots of errors between true continuation values and estimated continuation values using LSMC(black), Method 1 WLSMC(red) and Method 1 IRLSMC(green) with 1000 simulated paths and 1000 repetitions in pricing American call option, same parameters used in Table 4.3 with $N = 2$. Left and right panels correspond to 3^{rd} and 5^{th} order polynomials. Red line is the true exercise boundary.

price calculated by the explicit finite difference method. Residual plots are used to determine whether WLSMC/IRLSMC works on correcting heteroscedasticity.

Stopping creteria for IRLSMC is a maximum number of iterations of 10 and a tolerant value of 0.01. We fit all LSMC, WLSMC and IRLSMC with 3^{rd} and 5^{th} order polynomials. In the tables and figures listed below, we denote them with the polynomial order attached to LSMC, WLSMC or IRLSMC[1].

4.1 Compare LSMC/WLSMC/IRLSMC with Criteria I

Table 4.1 shows that compared to LSMC, WLSMC using method 1 provides a lower RMSE and MRE hence providing a more accurate estimation of the continuation values, particularly for samples with a small number of simulated paths. Focusing on LSMC,

[1] For example, WLSMC3 means a third order polynomial is used in the WLSMC method.

Table 4.1: American Call option continuation values comparison using LSMC, WLSMC and IRLSMC with true values

Type	Method	npath	MRE	RMSE
LSMC3	-	1000	0.072(0.001)	1.878(0.024)
LSMC5	-	1000	0.083(0.001)	2.361(0.026)
WLSMC3	1	1000	0.069(0.001)	1.834(0.024)
WLSMC5	1	1000	0.081(0.001)	2.342(0.026)
IRLSMC3	1	1000	0.069(0.001)	1.838(0.024)
IRLSMC5	1	1000	0.081(0.001)	2.343(0.026)
WLSMC3	2	1000	0.069(0.001)	1.852(0.024)
WLSMC5	2	1000	0.081(0.001)	2.349(0.026)
IRLSMC3	2	1000	0.069(0.001)	1.861(0.024)
IRLSMC5	2	1000	0.081(0.001)	2.353(0.026)
LSMC3	-	10000	0.024(0.001)	0.629(0.023)
LSMC5	-	10000	0.025(0.001)	0.800(0.026)
WLSMC3	1	10000	0.022(0.001)	0.607(0.023)
WLSMC5	1	10000	0.024(0.001)	0.788(0.025)
IRLSMC3	1	10000	0.022(0.001)	0.607(0.023)
IRLSMC5	1	10000	0.024(0.001)	0.788(0.025)
WLSMC3	2	10000	0.022(0.001)	0.621(0.024)
WLSMC5	2	10000	0.024(0.001)	0.794(0.025)
IRLSMC3	2	10000	0.022(0.001)	0.621(0.024)
IRLSMC5	2	10000	0.024(0.001)	0.795(0.025)

Note: Numbers in Type represents the polynomial orders for regression, the parameters used for stock paths generation and option are $S_0 = 100$, $K = 100$, $r = q = 0.06$, $T = 1$, $N = 2$, $\sigma = 0.25$, The total computational cost is set to be 1,000,000 (For example, when paths is 1000, we run 1000 different simulations), numbers in the bracket are the standard error.

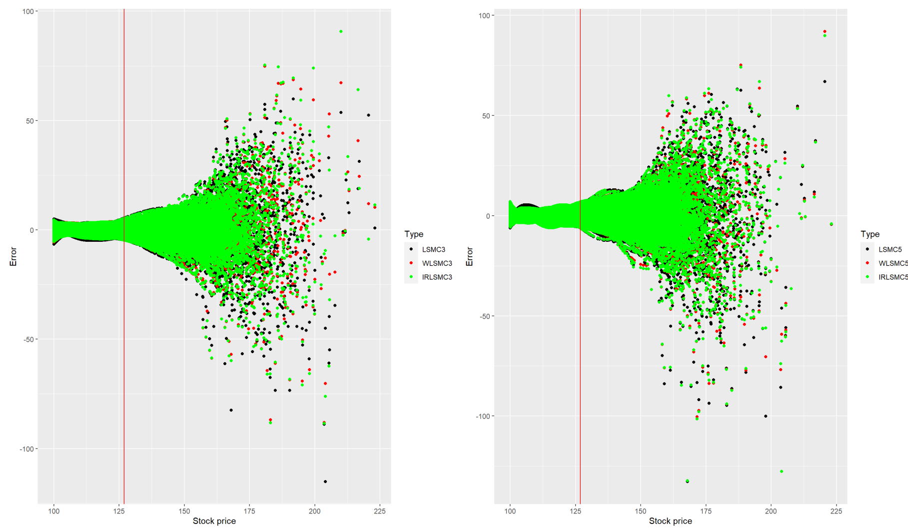

Figure 4.2: Plots of errors between true continuation values and estimated continuation values using LSMC(black), Method 2 WLSMC(red) and Method 2 IRLSMC(green) with 1000 simulated paths and 1000 repetitions in pricing American call option, same parameters used in Table 4.3 with $N = 2$. Left and right panels correspond to 3^{rd} and 5^{th} order polynomials. Red line is the true exercise boundary.

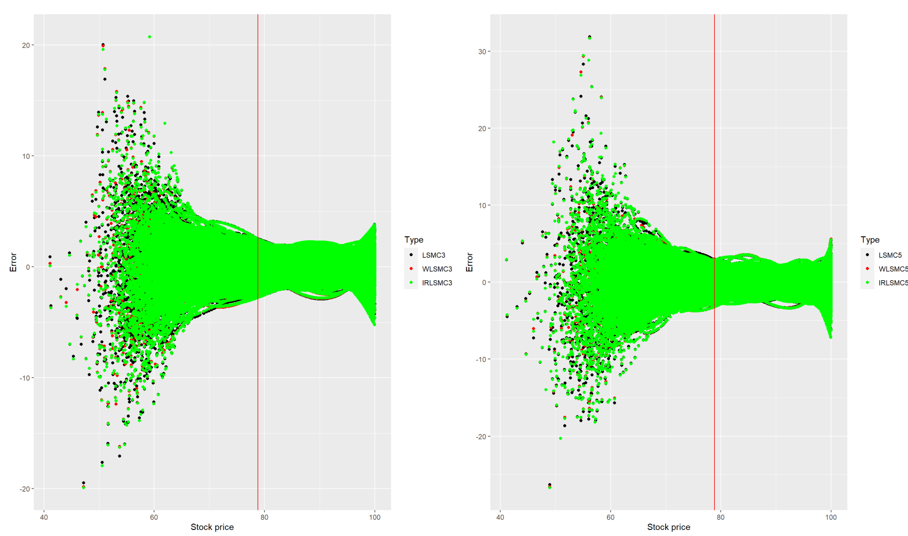

Figure 4.3: Plots of errors between true continuation values and estimated continuation values using LSMC(black), Method 1 WLSMC(red) and Method 1 IRLSMC(green) with 1000 simulated paths and 1000 repetitions in pricing American put option. Same parameters used in Table 4.1 with N = 2. Left and right panels correspond to 3^{rd} and 5^{th} order polynomials. Red line is the true exercise boundary.

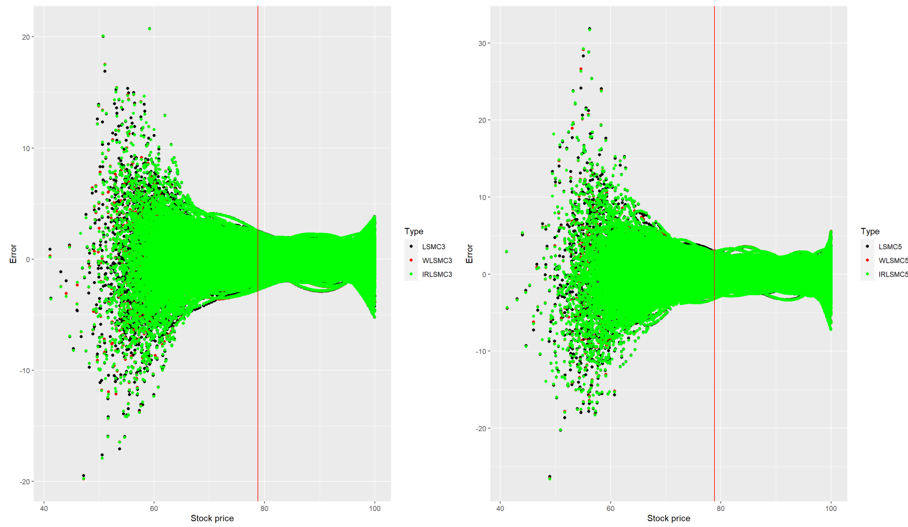

Figure 4.4: Plots of errors between true continuation values and estimated continuation values using LSMC(black), Method 2 WLSMC(red) and Method 2 IRLSMC(green) with 1000 simulated paths and 1000 repetitions in pricing American put option. Same parameters used in Table 4.1 with N = 2. Left and right panels correspond to 3^{rd} and 5^{th} order polynomials. Red line is the true exercise boundary.

WLSMC with third order polynomials and a thousand simulated paths, the RMSE drops from 1.878 to 1.834 and MRE drops from 0.072 to 0.069. IRLSMC however, does not give additional benefit in any of the cases. We do not observe too much difference in the standard error of these estimates among LSMC, WLSMC and IRLSMC.

WLSMC using method 2 also gives closer estimates to the true value compared to LSMC, but not as good as method 1 when using the RMSE as a criteria. With the same example described in the previous paragraph, RMSE using method 2 is 1.852, that is slightly higher than the RMSE using method 1. Data from MRE is relatively the same using both methods. Similar to method 1, IRLSMC does not furthermore improve the results.

Results from Figure 4.1 and Figure 4.2 are consistent with Table 4.3, if we focus on the stock range from [100, 150], estimated continuation values from WLSMC and IRLSMC with both methods perform better than LSMC. Smaller errors are observed from the plots. For large stock prices, WLSMC or IRLSMC does not seem to make further improvement.

Table 4.2: American Put continuation values comparison using LSMC, WLSMC and IRLSMC with true values using RMSE and MRE

Type	method	npath	MRE	RMSE
LSMC3	-	1000	0.051(0.001)	0.879(0.010)
LSMC5	-	1000	0.060(0.001)	1.097(0.011)
WLSMC3	1	1000	0.051(0.001)	0.880(0.010)
WLSMC5	1	1000	0.060(0.001)	1.098(0.011)
IRLSMC3	1	1000	0.051(0.001)	0.880(0.010)
IRLSMC5	1	1000	0.060(0.001)	1.098(0.011)
WLSMC3	2	1000	0.051(0.001)	0.880(0.010)
WLSMC5	2	1000	0.060(0.001)	1.098(0.011)
IRLSMC3	2	1000	0.051(0.001)	0.880(0.010)
IRLSMC5	2	1000	0.060(0.001)	1.098(0.011)
LSMC3	-	10000	0.017(0.001)	0.287(0.010)
LSMC5	-	10000	0.019(0.001)	0.342(0.010)
WLSMC3	1	10000	0.017(0.001)	0.287(0.010)
WLSMC5	1	10000	0.019(0.001)	0.342(0.010)
IRLSMC3	1	10000	0.017(0.001)	0.287(0.010)
IRLSMC5	1	10000	0.019(0.001)	0.342(0.010)
WLSMC3	2	10000	0.017(0.001)	0.287(0.010)
WLSMC5	2	10000	0.019(0.001)	0.342(0.010)
IRLSMC3	2	10000	0.017(0.001)	0.287(0.010)
IRLSMC5	2	10000	0.019(0.001)	0.342(0.010)

Note: Same parameters as in Table 4.1

Turning to the ATM American Put option, RMSE and MRE listed in Table 4.2 does not indicate that WLSMC or IRLSMC is a better approach compared to LSMC. Both estimates and standard error using WLSMC are close to those from LSMC. No improvement observed from IRLSMC. Similar results are obtained from Figure 4.3 and Figure 4.4, we do not see much difference in errors among LSMC, WLSMC and IRLSMC in both methods. Thus, we conclude that WLSMC or IRLSMC does not provide more accurate estimated continuation values compared to LSMC.

4.2 Compare LSMC/WLSMC/IRLSMC with Criteria II

Results from the previous section show that WLSMC applied to an ATM American call option leads to more accurate estimated continuation values. In this section, Table 4.3 and Table 4.4 show the results of comparing the decisions made by true continuation values and estimated continuation values from LSMC, WLSMC and IRLSMC, respectively. Focus on Table 4.3, around 93.9% of the time that LSMC3 can make correct decisions. WLSMC3 and IRLSMC3 with both methods do not make better decisions. Turning to LSMC5, 93.2% of the estimates make decisions consistent with the true continuation value. The data increased to 93.3% using WLSMC5 in both methods. In the second table when there are ten thousand paths in a sample. It still turns out that WLSMC5 makes more consistent decisions with the true continuation value. The accuracy increases from 96.5% to 96.7% in both methods, IRLSMC does not make any further improvement.

Table 4.3: American Call 2×2 contingency table of Exercise decided by True Continuation Value and Estimated Continuation Value. Average frequency per 1000 paths.

Type	Exercise by True Cont.Value	Exercise by Estimated Cont.Value	
		Yes	No
LSMC3	Yes	45.4	29.2
	No	32.1	893.3
WLSMC3 Method 1	Yes	45.1	29.5
	No	31.7	893.7
IRLSMC3 Method 1	Yes	45.1	29.4
	No	31.8	893.7
WLSMC3 Method 2	Yes	45.0	29.6
	No	31.6	893.7
IRLSMC3 Method 2	Yes	45.0	29.6
	No	31.6	893.8
LSMC5	Yes	43.3	31.2
	No	36.7	888.8
WLSMC5 Method 1	Yes	43.4	31.2
	No	35.9	889.5
IRLSMC5 Method 1	Yes	43.4	31.2
	No	36.0	889.4
WLSMC5 Method 2	Yes	43.3	31.2
	No	36.0	889.5
IRLSMC5 Method 2	Yes	43.3	31.3
	No	35.9	889.5

Note: Same parameters as in Table 4.1 and run 1000 different simulations.

Table 4.4: American Call 2×2 contingency table of Exercise decided by True Continuation Value and Estimated Continuation Value. Average frequency per 10000 paths.

Type	Exercise by True Cont.Value	Exercise by Estimated Cont.Value	
		Yes	No
LSMC3	Yes	602.4	150.5
	No	134.9	9112.2
WLSMC3 Method 1	Yes	590.2	162.8
	No	133.4	9113.6
IRLSMC3 Method 1	Yes	590.3	162.7
	No	133.5	9113.5
WLSMC3 Method 2	Yes	587.0	165.9
	No	132.5	9114.6
IRLSMC3 Method 2	Yes	587.3	165.7
	No	132.5	9114.5
LSMC5	Yes	552.7	200.3
	No	151.3	9095.7
WLSMC5 Method 1	Yes	557.0	196.0
	No	138.9	9108.1
IRLSMC5 Method 1	Yes	557.0	196.0
	No	139.0	9108.0
WLSMC5 Method 2	Yes	557.8	195.1
	No	139.5	9107.6
IRLSMC5 Method 2	Yes	558.1	194.8
	No	139.7	9107.4

Note: Same parameters as in Table 4.1 and run 100 different simulations.

4.3 Compare LSMC/WLSMC/IRLSMC with Criteria III

Figure 4.5 ,Figure 4.7, Figure 4.9, and Figure 4.11 display the results of RMSE in each bin for ATM American call option and put option, method 1 and 2, respectively. They behave similarly to those error plots. For call option, smaller RMSE are obtained for stock price less than 120 and for put option, RMSE are relatively the same everywhere. Moreover, we do not see improvement in RMSE around the exercise boundary for all plots. Thus we think that WLSMC/IRLSMC may not help in making better decision than LSMC. Figure 4.6 ,Figure 4.8, Figure 4.10, and Figure 4.12 confirm our thought. By plotting the percentage of same decisions made by two methods in each bin, we find that: For call option at any stock price, over 90% of the time LSMC and WLSMC/IRLSMC make the same decision, this data goes to 95% when the stock price approaches the early exercise boundary. For put option, over 98% of the time LSMC and WLSMC/IRLSMC make the same decision. In this case, we conclude that the improvement in RMSE by WLSMC/IRLSMC does not translate into making better decisions.

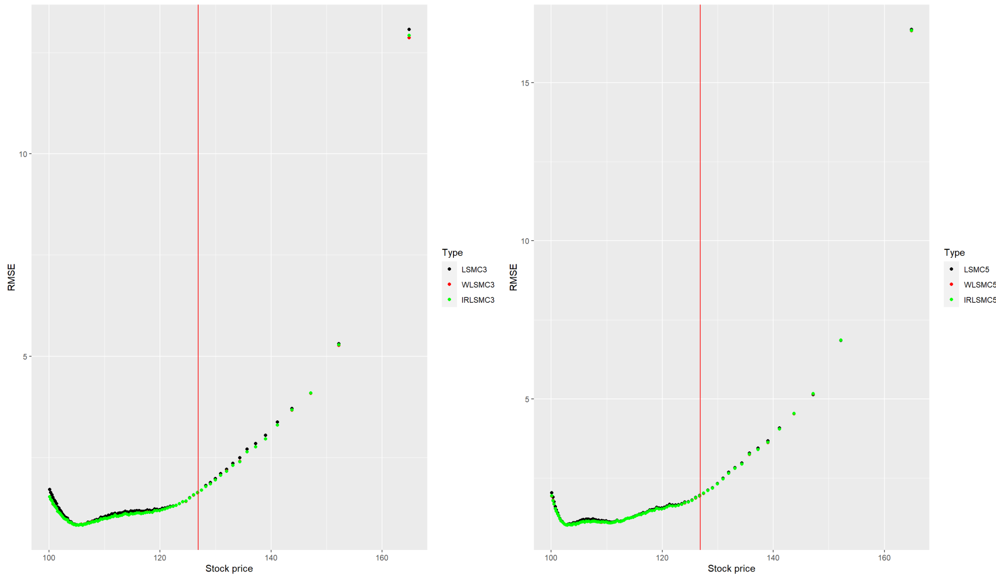

Figure 4.5: Plots of RMSE between true continuation values and estimated continuation values using LSMC3, Method 1 WLSMC3 and Method 1 IRLSMC3 on the left and 5^{th} order polynomial on the right against mean stock price in each bin. 1000 simulated paths with 1000 repetitions are used in pricing American call option. Same parameters used in Table 4.1, red line is the early exercise boundary.

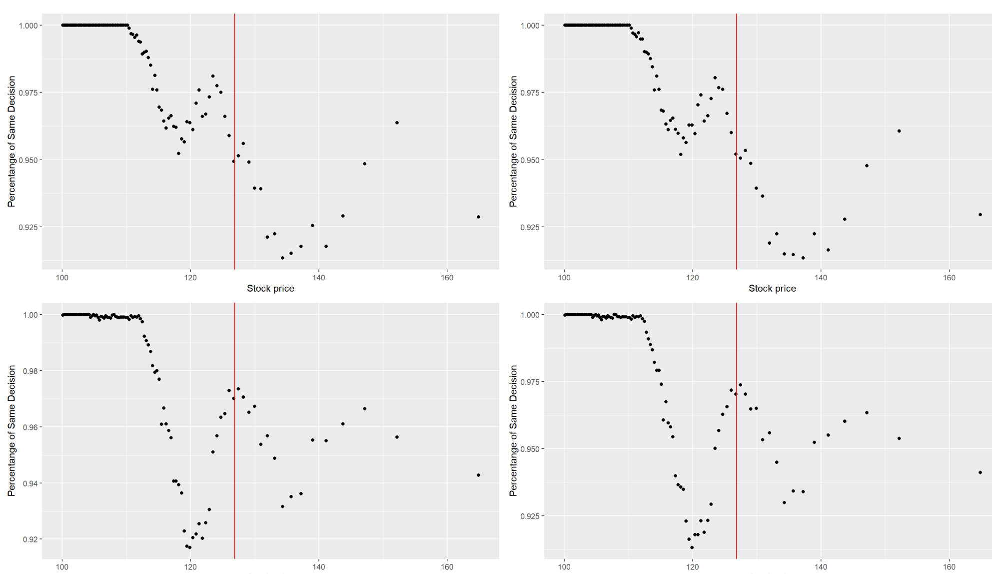

Figure 4.6: Plots of percentage of same decisions made by LSMC3 vs WLSMC3(top left), LSMC3 vs IRLSMC3(top right), LSMC5 vs WLSMC5(bottom left), LSMC5 vs IRLSMC5(bottom right), same bins and same method of WLSMC/IRLSMC as in Figure 4.5. 1000 simulated paths with 1000 repetitions were used in pricing ATM American Call option, same parameters used in Table 4.1, red line is the early exercise boundary.

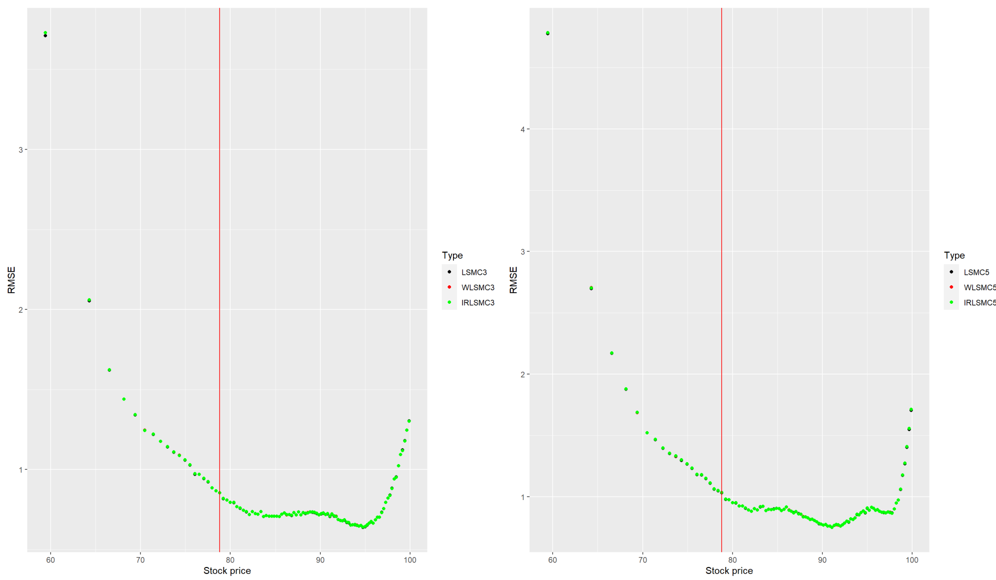

Figure 4.7: Plots of RMSE between true continuation values and estimated continuation values using LSMC3, Method 1 WLSMC3 and Method 1 IRLSMC3 on the left and 5^{th} order polynomial on the right against mean stock price in each bin. 1000 simulated paths with 1000 repetitions are used in pricing American put option. Same parameters used in Table 4.1, red line is the early exercise boundary.

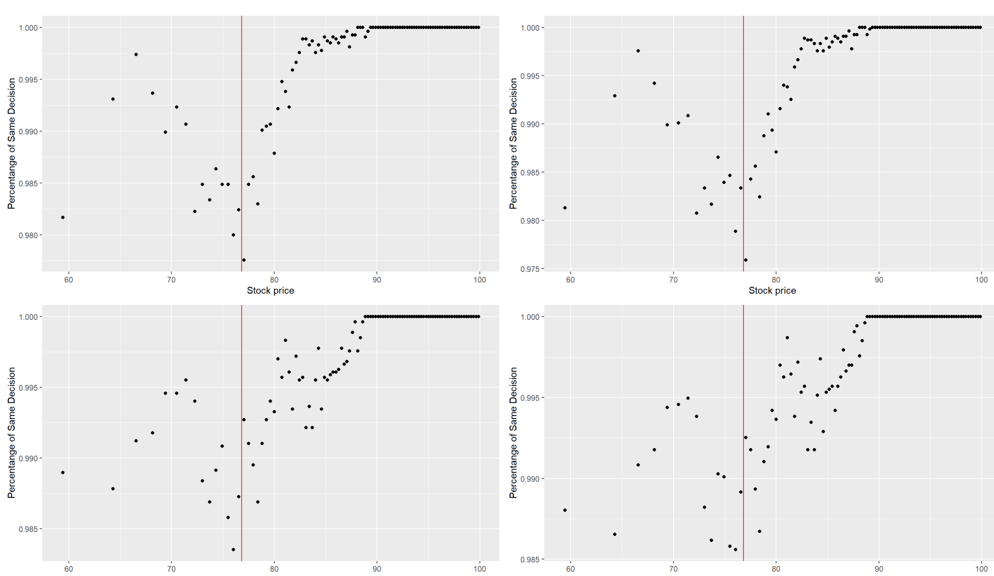

Figure 4.8: Plots of percentage of same decision made by LSMC3 vs WLSMC3(top left), LSMC3 vs IRLSMC3(top right), LSMC5 vs WLSMC5(bottom left), LSMC5 vs IRLSMC5(bottom right), same bins and same method of WLSMC/IRLSMC as in Figure 4.7. 1000 simulated paths with 1000 repetitions were used in pricing ATM American put option, same parameters used in Table 4.1, red line is the early exercise boundary.

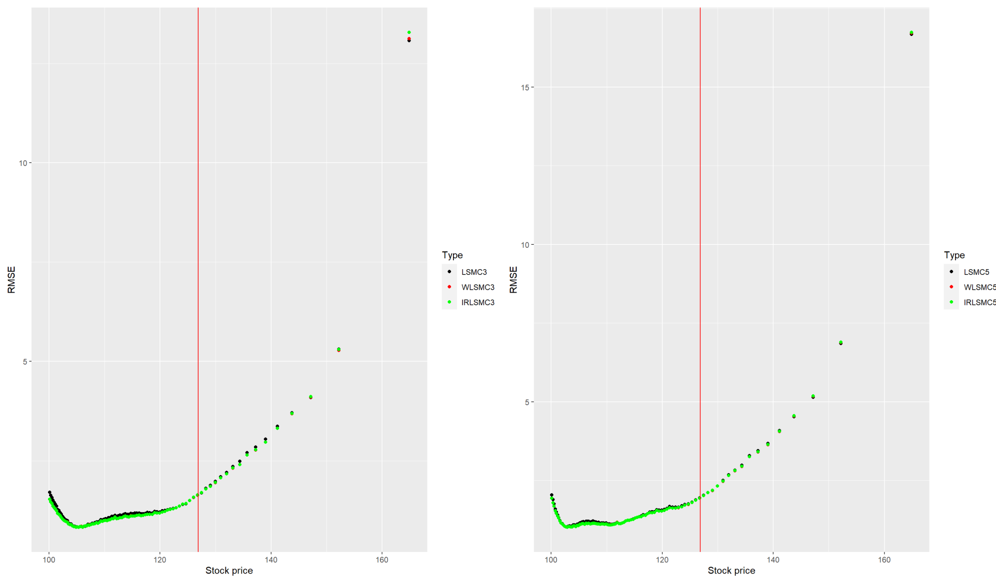

Figure 4.9: Plots of RMSE between true continuation values and estimated continuation values using LSMC3, Method 2 WLSMC3 and Method 2 IRLSMC3 on the left and 5^{th} order polynomial on the right against mean stock price in each bin. 1000 simulated paths with 1000 repetitions are used in pricing American call option. Same parameters used in Table 4.1, red line is the early exercise boundary.

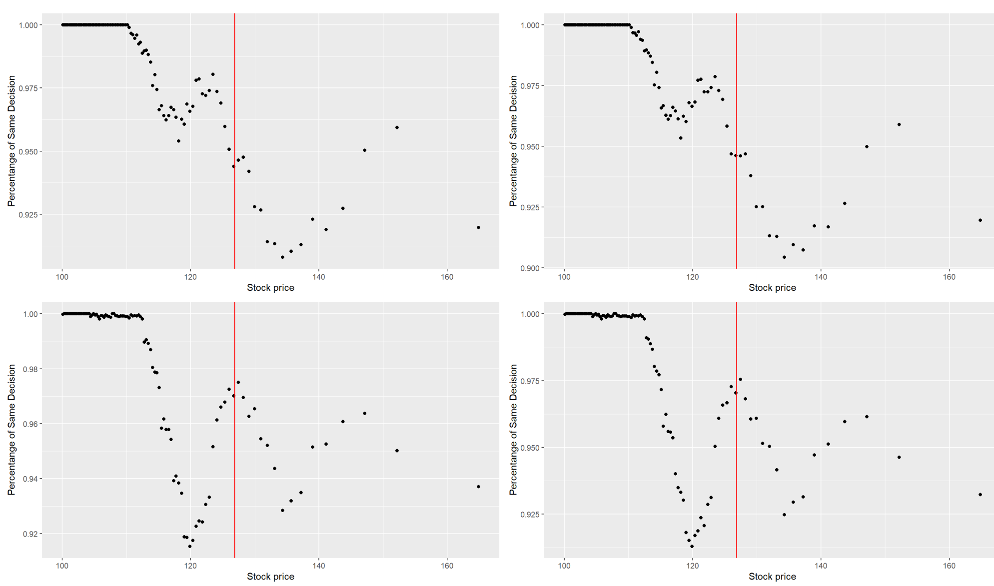

Figure 4.10: Plots of percentage of same decision made by LSMC3 vs WLSMC3(top left), LSMC3 vs IRLSMC3(top right), LSMC5 vs WLSMC5(bottom left), LSMC5 vs IRLSMC5(bottom right), same bins and same method of WLSMC/IRLSMC as in Figure 4.9. 1000 simulated paths with 1000 repetitions were used in pricing ATM American call option, same parameters used in Table 4.1, red line is the early exercise boundary.

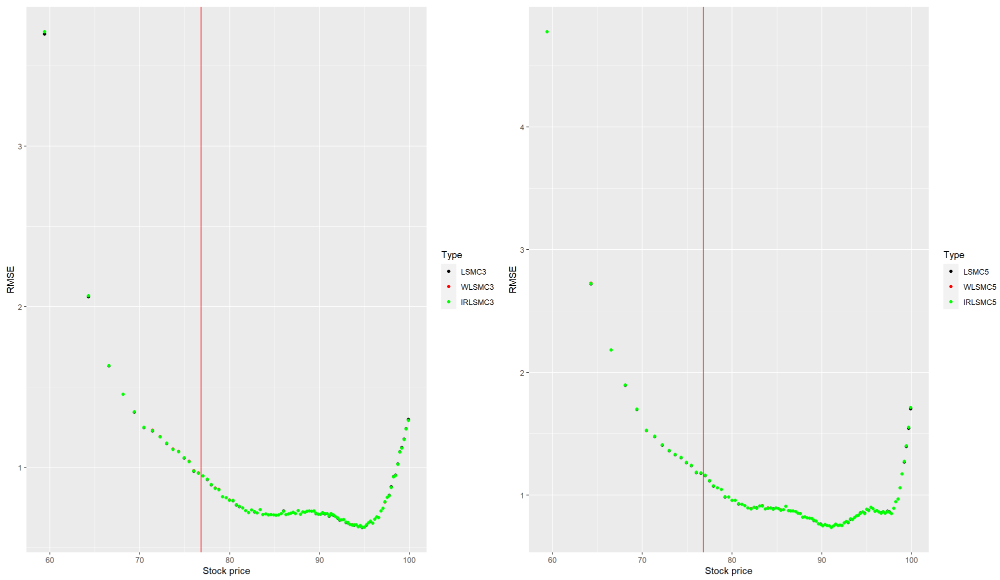

Figure 4.11: Plots of RMSE between true continuation values and estimated continuation values using LSMC3, Method 2 WLSMC3 and Method 2 IRLSMC3 on the left and 5^{th} order polynomial on the right against mean stock price in each bin. 1000 simulated paths with 1000 repetitions are used in pricing American put option. Same parameters used in Table 4.1, red line is the early exercise boundary.

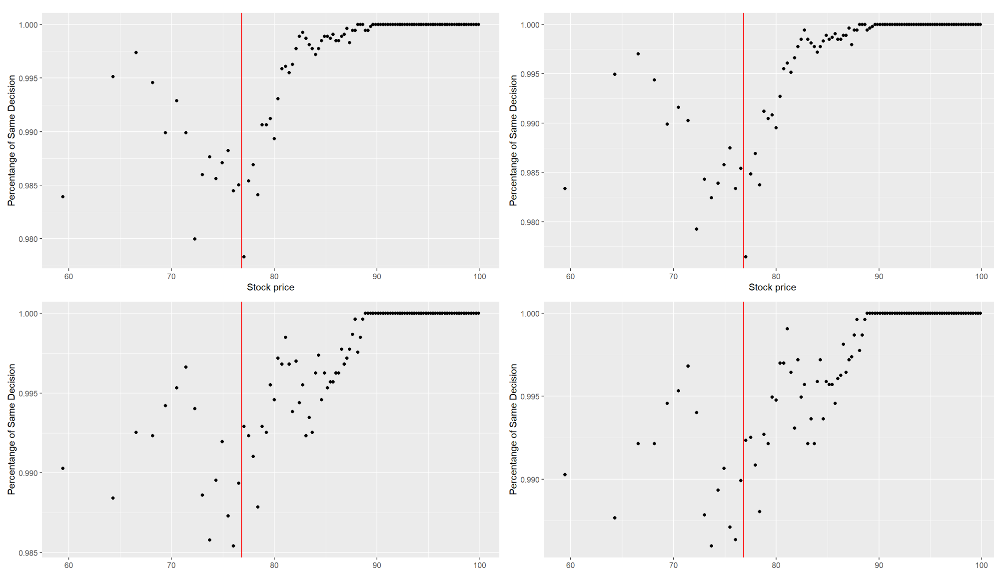

Figure 4.12: Plots of percentage of same decision made by LSMC3 vs WLSMC3(top left), LSMC3 vs IRLSMC3(top right), LSMC5 vs WLSMC5(bottom left), LSMC5 vs IRLSMC5(bottom right), same bins and same method of WLSMC/IRLSMC as in Figure 4.11. 1000 simulated paths with 1000 repetitions were used in pricing ATM American put option, same parameters used in Table 4.1, red line is the early exercise boundary.

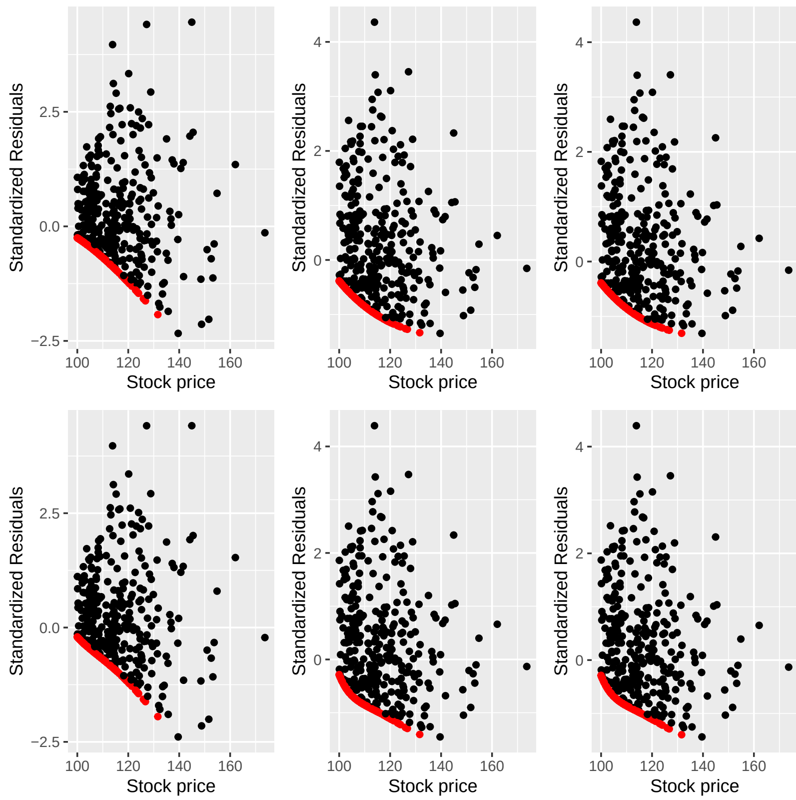

Figure 4.13: Standardized residual plots of LSMC(left), Method 1 WLSMC(middle) and Method 1 IRLSMC(right) with 1000 simulated paths in pricing ATM American Call option, top and bottom rows correspond to polynomial orders 3 and 5 used in regressions, same parameters used in Table 4.1

4.4 Compare LSMC/WLSMC/IRLSMC with Price and Heteroscedasticity Correction

Figure 4.13 and Figure 4.14 include the standardized residual plots of LSMC, WLSMC and IRLSMC in two methods under polynomial of 3 and 5 in pricing ATM American Call option. Figure 4.15 and Figure 4.16 are the correspond residuals plots in pricing ATM American Put option. Red points in the graphs are those with 0 in the response variables, which means for these particular stock paths at this time step, they have not been exercised. This leads to a particular trend formed in the residual plots, which is apparent in every plots. In this case, we think using BPtest may not be a good idea to

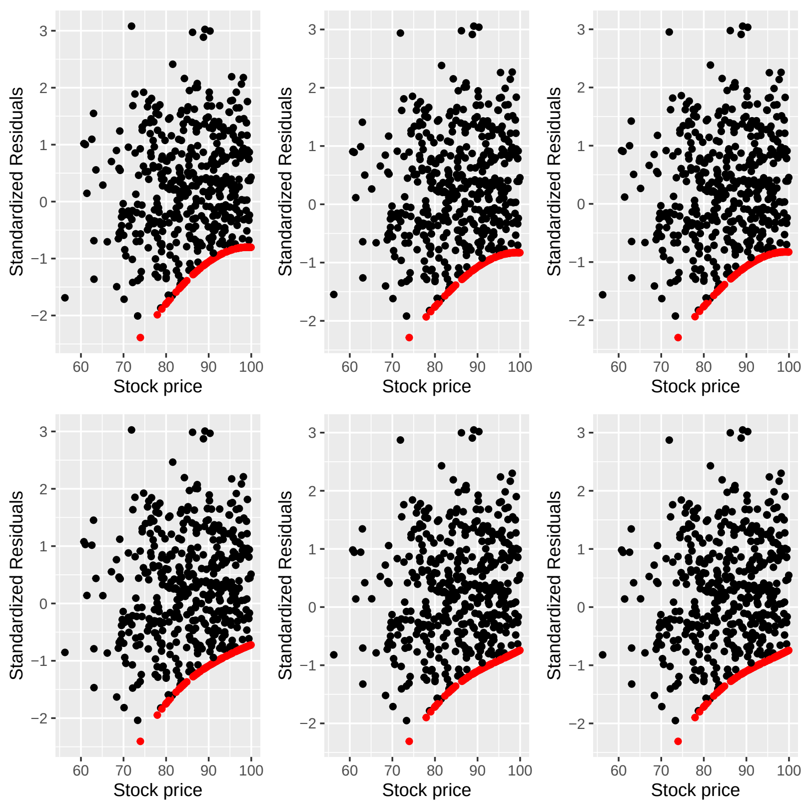

Figure 4.14: Standardized residual plots of LSMC(left), Method 1 WLSMC(middle) and Method 1 IRLSMC(right) with 1000 simulated paths in pricing ATM American Put option, top and bottom rows correspond to polynomial orders 3 and 5 used in regressions, same parameters used in Table 4.1

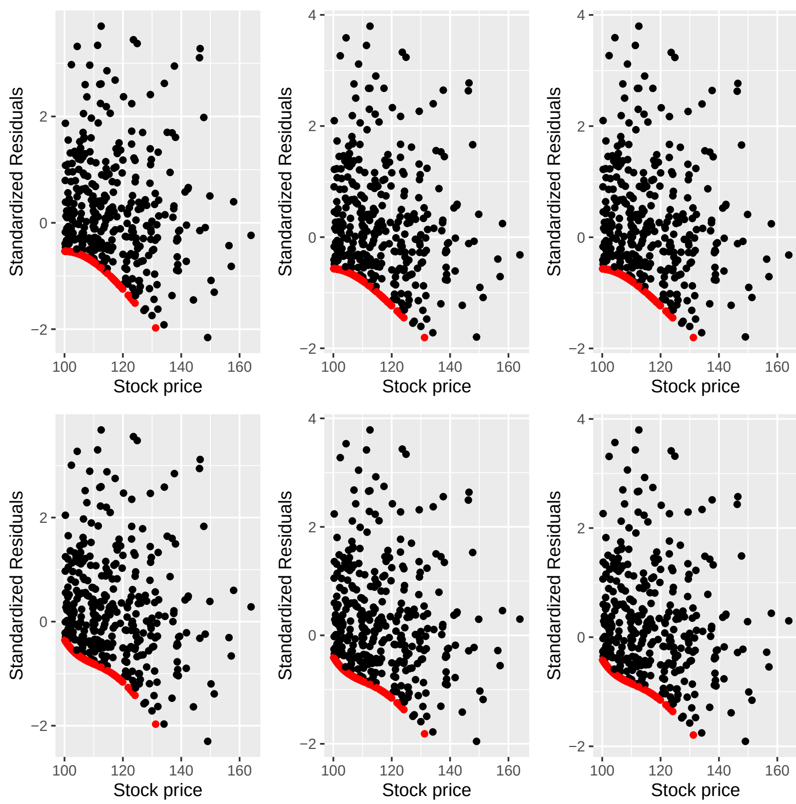

Figure 4.15: Standardized residual plots of LSMC(left), Method 2 WLSMC(middle) and Method 2 IRLSMC(right) with 1000 simulated paths in pricing ATM American Call option, top and bottom rows correspond to polynomial orders 3 and 5 used in regressions, same parameters used in Table 4.1

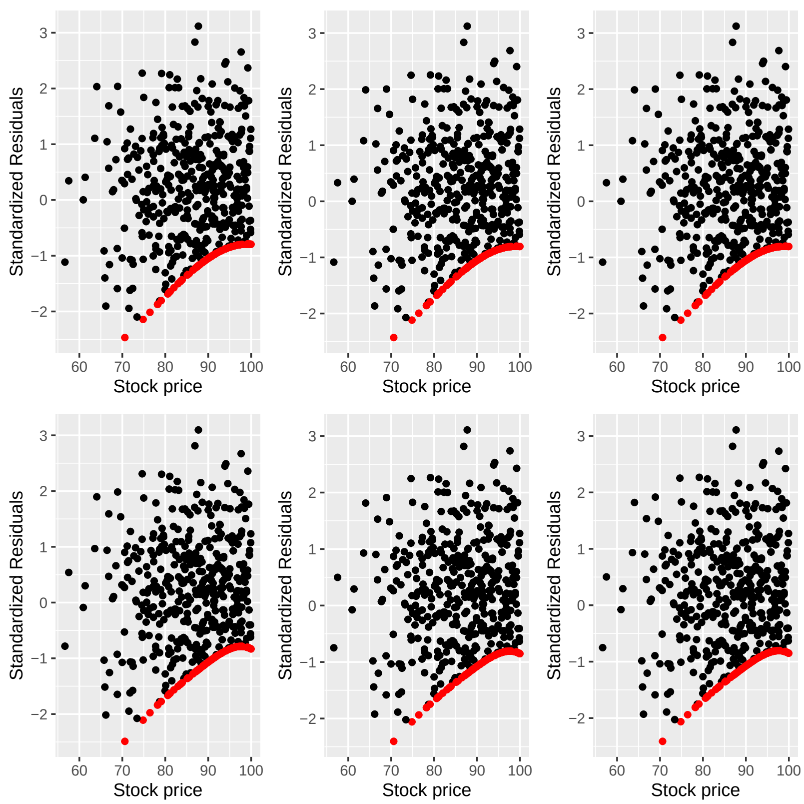

Figure 4.16: Standardized residual plots of LSMC(left), Method 2 WLSMC(middle) and Method 2 IRLSMC(right) with 1000 simulated paths in pricing ATM American Put option, top and bottom rows correspond to polynomial orders 3 and 5 used in regressions, same parameters used in Table 4.1

test for heteroscedasticity as this trend has already violated the assumption of errors being homoscedastic in OLS.

Focused on the standardized residual plots, we find WLSMC cannot remove the trend as described in the previous paragraph and do not observe too much improvement over WLSMC made by IRLSMC. In this case, we consider that WLSMC/IRLSMC does not help with heteroscedasticity correction.

Table 4.5 and Table 5.1 display the prices of the American options using LSMC, WLSMC, and IRLSMC, respectively. We hardly see any improvement made by applying WLSMC compare to LSMC in both methods, especially for American put option.

Table 4.5: Pricing performance of ATM American Call with 2 exercise opportunities using LSMC, WLSMC and IRLSMC.

Type	Path	Method	Price	se	MRE	RMSE
LSMC3	1000	-	9.5281	0.0153	0.0416	0.4941
LSMC5	1000	-	9.5601	0.0151	0.0419	0.4975
WLSMC3	1000	1	9.5272	0.0151	0.0415	0.4903
WLSMC5	1000	1	9.5606	0.0151	0.0419	0.4958
IRLSMC3	1000	1	9.5276	0.0152	0.0414	0.4904
IRLSMC5	1000	1	9.5611	0.0151	0.0419	0.4960
WLSMC3	1000	2	9.5287	0.0151	0.0414	0.4898
WLSMC5	1000	2	9.5606	0.0151	0.0420	0.4963
IRLSMC3	1000	2	9.5288	0.0151	0.0413	0.4890
IRLSMC5	1000	2	9.5608	0.0151	0.0419	0.4962
LSMC3	10000	-	9.4347	0.0135	0.0117	0.1347
LSMC5	10000	-	9.4424	0.0133	0.0114	0.1335
WLSMC3	10000	1	9.4343	0.0133	0.0114	0.1328
WLSMC5	10000	1	9.4385	0.0132	0.0114	0.1326
IRLSMC3	10000	1	9.4345	0.0133	0.0114	0.1325
IRLSMC5	10000	1	9.4386	0.0132	0.0115	0.1327
WLSMC3	10000	2	9.4338	0.0133	0.0113	0.1324
WLSMC5	10000	2	9.4390	0.0132	0.0114	0.1326
IRLSMC3	10000	2	9.4340	0.0133	0.0113	0.1327
IRLSMC5	10000	2	9.4394	0.0132	0.0114	0.1329

Note: Same parameters used for stock generations and option as in Table 4.1 with N = 2, exercise once every 126 days, the true price is calculated from explicit finite difference method to be 9.422.

Table 4.6: Pricing performance of ATM American Put with 2 exercise opportunities using LSMC, WLSMC and IRLSMC.

Type	Path	Method	Price	se	MRE	RMSE
LSMC3	1000	-	9.4548	0.0112	0.0302	0.3555
LSMC5	1000	-	9.4692	0.0111	0.0300	0.3524
WLSMC3	1000	1	9.4546	0.0112	0.0301	0.3544
WLSMC5	1000	1	9.4693	0.0110	0.0299	0.3521
IRLSMC3	1000	1	9.4546	0.0112	0.0301	0.3546
IRLSMC5	1000	1	9.4692	0.0110	0.0299	0.3522
WLSMC3	1000	2	9.4545	0.0112	0.0301	0.3550
WLSMC5	1000	2	9.4691	0.0110	0.0300	0.3522
IRLSMC3	1000	2	9.4544	0.0112	0.0301	0.3548
IRLSMC5	1000	2	9.4690	0.0110	0.0300	0.3524
LSMC3	10000	-	9.4169	0.0112	0.0094	0.1119
LSMC5	10000	-	9.4205	0.0113	0.0093	0.1127
WLSMC3	10000	1	9.4171	0.0112	0.0094	0.1120
WLSMC5	10000	1	9.4205	0.0113	0.0093	0.1125
IRLSMC3	10000	1	9.4171	0.0112	0.0094	0.1120
IRLSMC5	10000	1	9.4205	0.0113	0.0093	0.1125
WLSMC3	10000	2	9.4170	0.0112	0.0094	0.1120
WLSMC5	10000	2	9.4205	0.0113	0.0093	0.1125
IRLSMC3	10000	2	9.4170	0.0112	0.0094	0.1120
IRLSMC5	10000	2	9.4204	0.0113	0.0093	0.1125

Note: Same parameters used for stock generations and option as in Table 4.1 with N = 2, exercise once every 126 days, the true price is calculated from explicit finite difference method to be 9.422.

Chapter 5

General Results

In this section, we extend the exercise opportunities from two to many. The models we consider are univariate GBM, univaraite GARCH and bivariate GBM, respectively, for underlying assets. The criteria we set to compare the performance of LSMC, WLSMC and IRLSMC are the estimated option values compared with the true value, standard error, RMSE and MRE. The true option value is calculated from the Explicit Finite Difference method brought by Brennan and Schwartz [1977]. Recall we generate q samples stated in section 6.1, thus, suppose V represents the true price and $\hat{V}_i$ represents the estimated price from the i^{th} sample such that $1 \leq i \leq Q$, RMSE and MRE are calculated as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{Q}(\hat{V}_i - V)^2}{Q}}, \text{ and} \tag{5.0.1}$$

$$MRE = \frac{\sum_{i=1}^{Q}\left|\frac{\hat{V}_i - V}{V}\right|}{Q}. \tag{5.0.2}$$

5.1 Univariate GBM

Table 5.1 and Table 5.2 display the prices of ATM American Call and Put options with 50 exercise opportunities using LSMC, two methods WLSMC and IRLSMC respectively. Since we set the interest rate equal to the dividend rate in our examples, the true price

Table 5.1: Pricing performance of ATM American Call with 50 exercise opportunities using LSMC, WLSMC and IRLSMC.

Type	Paths	Method	Price	se	MRE	RMSE
LSMC3	1000	-	9.884	0.014	0.050	0.595
LSMC5	1000	-	10.055	0.014	0.063	0.715
WLSMC3	1000	1	9.861	0.015	0.050	0.592
WLSMC5	1000	1	10.035	0.014	0.061	0.702
IRLSMC3	1000	1	9.866	0.014	0.050	0.587
IRLSMC5	1000	1	10.042	0.014	0.062	0.709
WLSMC3	1000	2	9.853	0.015	0.050	0.595
WLSMC5	1000	2	10.037	0.014	0.062	0.705
IRLSMC3	1000	2	9.857	0.015	0.050	0.586
IRLSMC5	1000	2	10.037	0.014	0.062	0.707
LSMC3	10000	-	9.539	0.013	0.011	0.138
LSMC5	10000	-	9.582	0.014	0.013	0.161
WLSMC3	10000	1	9.530	0.014	0.011	0.140
WLSMC5	10000	1	9.576	0.014	0.013	0.161
IRLSMC3	10000	1	9.531	0.014	0.011	0.141
IRLSMC5	10000	1	9.573	0.014	0.013	0.159
WLSMC3	10000	2	9.517	0.015	0.012	0.150
WLSMC5	10000	2	9.569	0.014	0.013	0.154
IRLSMC3	10000	2	9.518	0.015	0.012	0.146
IRLSMC5	10000	2	9.568	0.014	0.013	0.161

Note: Same parameters used for stock generations and option as in Table 4.1 with $N = 50$, exercise once every 5 days, the true price is calculated from explicit finite difference method to be 9.497.

Table 5.2: Pricing performance of ATM American Put with 50 exercise opportunities using LSMC, WLSMC and IRLSMC.

Type	Paths	Method	Price	se	MRE	RMSE
LSMC3	1000	-	9.733	0.0110	0.036	0.419
LSMC5	1000	-	9.822	0.0107	0.040	0.469
WLSMC3	1000	1	9.723	0.0114	0.036	0.424
WLSMC5	1000	1	9.822	0.0107	0.040	0.470
IRLSMC3	1000	1	9.728	0.0111	0.036	0.420
IRLSMC5	1000	1	9.825	0.0108	0.041	0.473
WLSMC3	1000	2	9.737	0.0110	0.036	0.422
WLSMC5	1000	2	9.832	0.0107	0.041	0.476
IRLSMC3	1000	2	9.736	0.0109	0.036	0.420
IRLSMC5	1000	2	9.831	0.0108	0.041	0.477
LSMC3	10000	-	9.547	0.0104	0.010	0.115
LSMC5	10000	-	9.570	0.0102	0.011	0.125
WLSMC3	10000	1	9.542	0.0106	0.010	0.114
WLSMC5	10000	1	9.564	0.0102	0.010	0.121
IRLSMC3	10000	1	9.542	0.0105	0.010	0.114
IRLSMC5	10000	1	9.566	0.0101	0.011	0.122
WLSMC3	10000	2	9.547	0.0105	0.010	0.116
WLSMC5	10000	2	9.571	0.0100	0.011	0.124
IRLSMC3	10000	2	9.545	0.0105	0.010	0.115
IRLSMC5	10000	2	9.571	0.0100	0.011	0.123

Note: Same parameters used for stock generations and option as in Table 4.1 with $N = 50$, exercise once every 5 days, the true price is calculated from explicit finite difference method to be 9.497.

of ATM American Call option is the same as ATM American Put option, and the true price is 9.497. Using the true price as the standard, for method 1 in pricing Call option, compared with LSMC, WLSMC brings low biased estimates across all samples with 1000 or 10000 paths. However, the standard error from WLSMC is slightly higher than LSMC. In terms of the MRE and RMSE calculated by equation 7.2.1 and 7.2.2, with a sample size of 1000 WLSMC method 1 makes lower or relatively the same results compared to LSMC. With a sample size of 10000, WLSMC method 1 using third-order polynomial achieves higher RMSE. Although IRLSMC gets higher biased prices compared to WLSMC for most cases, the standard error decreases or keeps the same across all samples. It leads to lower RMSE for a polynomial of order 3 and 1000 paths and order 5 with 10000 paths. In terms of method 1 when pricing the put option, we could hardly see any benefit from using WLSMC or IRLSMC.

Focused on method 2 WLSMC and use the true price 9.497 as the standard, it provides lower biased estimates compared to LSMC and method 1 WLSMC for most of the time in Table 5.1. However, the major disadvantage of using method 2 WLSMC comes from its high standard error in pricing ATM American Call options across all cases listed in the table. Method 2 IRLSMC helps to reduce the standard error when the sample size is 1000. Method 2 WLSMC or IRLSMC does not perform better than LSMC in pricing ATM American Put option.

Table 5.3 shows the prices when there are 126 exercise opportunities for ATM American Call options. As the number of exercise opportunities increases to 126, the true price rises as well to 9.499. Using the new benchmark price 9.499 as the standard, method 1 WLSMC exerts more effect on the prices of the option, especially for third-order polynomial. The difference in estimates is 23 cents when the sample size is 1000, much larger than the difference of 2.3 cents in the previous table. However, it also brings much more volatility across all samples. Due to the high standard error brought by Method 1 WLSMC, higher RMSE and MRE were observed. Method 1 IRLSMC helps to cut down the standard error compared to WLSMC, reflecting on the RMSE and MRE calculation with lower results.

Turning to method 2 WLSMC with the same benchmark price 9.499 as the standard.

Table 5.3: Pricing performance of ATM American Call with 126 exercise opportunities using LSMC, WLSMC and IRLSMC .

Type	Paths	Method	Price	se	MRE	RMSE
LSMC3	1000	-	9.915	0.014	0.052	0.609
LSMC5	1000	-	10.110	0.014	0.067	0.753
WLSMC3	1000	1	9.686	0.024	0.062	0.768
WLSMC5	1000	1	10.031	0.016	0.065	0.736
IRLSMC3	1000	1	9.823	0.016	0.052	0.602
IRLSMC5	1000	1	10.048	0.015	0.065	0.727
WLSMC3	1000	2	9.602	0.027	0.066	0.859
WLSMC5	1000	2	10.014	0.017	0.066	0.737
IRLSMC3	1000	2	9.791	0.017	0.051	0.603
IRLSMC5	1000	2	10.031	0.015	0.064	0.719
LSMC3	10000	-	9.530	0.013	0.012	0.137
LSMC5	10000	-	9.596	0.013	0.013	0.159
WLSMC3	10000	1	9.492	0.016	0.012	0.161
WLSMC5	10000	1	9.529	0.023	0.015	0.232
IRLSMC3	10000	1	9.501	0.014	0.012	0.142
IRLSMC5	10000	1	9.559	0.014	0.012	0.152
WLSMC3	10000	2	9.365	0.036	0.024	0.378
WLSMC5	10000	2	9.452	0.038	0.021	0.380
IRLSMC3	10000	2	9.456	0.017	0.014	0.176
IRLSMC5	10000	2	9.505	0.019	0.014	0.191

Note: Same parameters used for stock generations and option as in Table 4.1 with $N = 126$, exercise once every 2 days. The true price is 9.499.

It behaves similarly to Method 1 WLSMC. Although Method 2 WLSMC reduces the bias in the price estimates when the sample size is 1000, the standard error increases dramatically. When the sample size is 10000, Method 2 WLSMC gives poor estimates with high standard error. Method 2 IRLSMC lessens the standard error to a large extent.

5.2 Univariate GARCH

In this section, we focus on GARCH models including the ordinary GARCH and NGARCH models discussed in section 2. Similar to the previous section, we test the performance of LSMC, WLSMC and IRLSMC via option price compare with true price, standard error, RMSE and MRE. The option type is ATM American put option in one dimension.

Table 5.4 exhibits the pricing estimates of American put option under GARCH model using LSMC, WLSMC and IRLSMC in both methods. The true price is selected from Stentoft [2011]. Obviously except the case of Method 2 WLSMC5 when the number of paths is 1000, WLSMC provides lower biased estimates compared to LSMC and higher standard error, similar to the GBM American call option as described in the previous section. Checking the results from IRLSMC, it appears that method 1 IRLSMC makes extra contribution to WLSMC by reducing the standard error effectively, leading to smaller MRE and RMSE. However, method 2 IRLSMC produces relatively the same results with WLSMC. Turn to the NGARCH model, results from Table 5.5 is pretty consistent with Table 5.4.

5.3 Bivariate GBM

In this section, the option type we discussed is the two dimensional ATM American basket put option. Suppose the option is expired at T. Let S^1 and S^2 represent the two correlated stock paths generated from section 2.1, the payoff function at $t \leq T$ is

Table 5.4: Pricing performance of ATM American Put with 63 exercise opportunities using LSMC, WLSMC and IRLSMC in GARCH model.

Type	Paths	Method	Price	se	MRE	RMSE
LSMC3	1000	-	4.547	0.0056	0.067	0.331
LSMC5	1000	-	4.735	0.0056	0.109	0.500
WLSMC3	1000	1	4.520	0.0064	0.065	0.324
WLSMC5	1000	1	4.728	0.0058	0.108	0.495
IRLSMC3	1000	1	4.513	0.0064	0.063	0.317
IRLSMC5	1000	1	4.719	0.0059	0.106	0.488
WLSMC3	1000	2	4.518	0.0062	0.063	0.318
WLSMC5	1000	2	4.738	0.0059	0.110	0.506
IRLSMC3	1000	2	4.526	0.0061	0.064	0.322
IRLSMC5	1000	2	4.713	0.0063	0.105	0.487
LSMC3	10000	-	4.324	0.0054	0.016	0.078
LSMC5	10000	-	4.363	0.0055	0.023	0.110
WLSMC3	10000	1	4.282	0.0076	0.014	0.076
WLSMC5	10000	1	4.343	0.0064	0.020	0.099
IRLSMC3	10000	1	4.288	0.0066	0.013	0.069
IRLSMC5	10000	1	4.341	0.0062	0.019	0.095
WLSMC3	10000	2	4.304	0.0062	0.014	0.071
WLSMC5	10000	2	4.352	0.0059	0.022	0.103
IRLSMC3	10000	2	4.305	0.0061	0.014	0.071
IRLSMC5	10000	2	4.351	0.0059	0.021	0.102

Note: We use the GARCH parameters outlined in Stentoft [2011] on page 889 with $N = 63$, exercise once every day. The benchmark price is 4.268.

Table 5.5: Pricing performance of ATM American Put with 63 exercise opportunities using LSMC, WLSMC and IRLSMC in NGARCH model.

Type	npaths	Method	Price	se	MRE	RMSE
LSMC3	1000	-	4.715	0.0063	0.075	0.379
LSMC5	1000	-	4.925	0.0062	0.121	0.568
WLSMC3	1000	1	4.672	0.0072	0.070	0.360
WLSMC5	1000	1	4.912	0.0064	0.118	0.558
IRLSMC3	1000	1	4.672	0.0071	0.070	0.358
IRLSMC5	1000	1	4.899	0.0067	0.116	0.549
WLSMC3	1000	2	4.679	0.0070	0.070	0.362
WLSMC5	1000	2	4.922	0.0064	0.121	0.567
IRLSMC3	1000	2	4.688	0.0070	0.072	0.369
IRLSMC5	1000	2	4.896	0.0066	0.115	0.546
LSMC3	10000	-	4.454	0.0060	0.017	0.086
LSMC5	10000	-	4.503	0.0060	0.026	0.126
WLSMC3	10000	1	4.395	0.0089	0.015	0.089
WLSMC5	10000	1	4.473	0.0069	0.021	0.106
IRLSMC3	10000	1	4.403	0.0075	0.014	0.076
IRLSMC5	10000	1	4.472	0.0066	0.020	0.103
WLSMC3	10000	2	4.431	0.0064	0.014	0.075
WLSMC5	10000	2	4.491	0.0063	0.024	0.117
IRLSMC3	10000	2	4.432	0.0064	0.014	0.075
IRLSMC5	10000	2	4.490	0.0063	0.023	0.116

Note: We use the NGARCH parameters outlined in Stentoft [2011] on page 889 with $N = 63$, exercise once every day. The benchmark price is 4.392.

given by

$$f_{BP} = \max(K - \frac{S_t^1 + S_t^2}{2}, 0) \tag{5.3.1}$$

where BP denotes for American basket put option.

Table 5.5 displays the pricing estimates using LSMC, WLSMC and IRLSMC in both methods. The correlation between two stocks is 0.5. The results from the table behave similarly to Table 5.3, There is no benefit in using WLSMC or IRLSMC for pricing this 2-dimensional American style option.

Table 5.6: Pricing performance of ATM American Basket Put with 13 exercise opportunities using LSMC, WLSMC and IRLSMC in GBM Two Stocks Model.

Type	npaths	Method	Price	se	MRE	RMSE
LSMC3	1000	-	3.267	0.0042	0.048	0.185
LSMC5	1000	-	3.355	0.0041	0.071	0.255
WLSMC3	1000	1	3.267	0.0041	0.048	0.184
WLSMC5	1000	1	3.356	0.0042	0.071	0.255
IRLSMC3	1000	1	3.266	0.0042	0.048	0.185
IRLSMC5	1000	1	3.355	0.0042	0.071	0.255
WLSMC3	1000	2	3.267	0.0042	0.048	0.185
WLSMC5	1000	2	3.356	0.0042	0.071	0.256
IRLSMC3	1000	2	3.267	0.0042	0.048	0.185
IRLSMC5	1000	2	3.355	0.0042	0.071	0.255
LSMC3	10000	-	3.162	0.0034	0.011	0.042
LSMC5	10000	-	3.179	0.0033	0.014	0.054
WLSMC3	10000	1	3.162	0.0034	0.011	0.042
WLSMC5	10000	1	3.180	0.0033	0.015	0.055
IRLSMC3	10000	1	3.163	0.0035	0.011	0.043
IRLSMC5	10000	1	3.180	0.0034	0.014	0.054
WLSMC3	10000	2	3.164	0.0034	0.011	0.044
WLSMC5	10000	2	3.180	0.0034	0.015	0.055
IRLSMC3	10000	2	3.164	0.0035	0.011	0.044
IRLSMC5	10000	2	3.180	0.0034	0.014	0.055

Note: We use the parameters outlined in Fabozzi et al. [2017] on page 702 with $N = 13$, approximately exercise once every 5 days. The benchmark price is 3.137.

Chapter 6

Conclusion

In this thesis, based on the work from Fabozzi et al. [2017], we introduce two methods of heteroscedasticity correction in the regression part of LSMC, which are Method 1 and 2 WLSMC from Park [1966] and Harvey [1976]. Moreover, by recomputing the residuals from WLSMC and running several iterations, we take IRLSMC into account. To test the performance of the two methods, we create some corresponding make-up heteroscedastic data and apply two methods separately. By varying the polynomial order used for the mean function, we find that WLS provide better coefficient estimates but is sometimes insufficient for heteroscedasticity correction, and IRLS can help to further correct the heteroscedasticity. In the results part, we first do an analysis when there are only two exercise opportunities. Comparing true continuation values to estimated continuation values from LSMC, WLSMC and IRLSMC respectively, we find that WLSMC and IRLSMC are able to give more accurate continuation values than LSMC when pricing ATM American call options. However, this increased in accuracy does not translate into better exercise decisions and hence there is no effect on option prices. Finally, by comparing their residual plots, the general pattern of the residuals in LSMC is not changed by applying WLSMC/IRLSMC. IRLSMC does not perform better than WLSMC.

Next we apply LSMC, WLSMC and IRLSMC with more realistic number of exercise opportunities in both methods in pricing ATM American call and put option under

GBM model. As expected, WLSMC works better with call option and IRLSMC helps to reduce the standard error. However, as the standard errors from WLSMC and IRLSMC are higher than LSMC, we conclude that they are not effective at improving price estimator efficiency in realistic settings. Last but not least, we test WLSMC and IRLSMC on GARCH, NGARCH and the bivariate GBM models to price ATM American put option[1], the conclusion is the same with the GBM single stock case.

6.1 Suggestions on Future Work

Based on our analysis, we think some interesting problems could be worked on in the future. Here we list some of these.

1. The reason we abandon the method from Greene [2012] is due to negative weights, is there a way to avoid them?

2. Recall we mention that Harvey [1976] suggested using Maximum Likelihood Estimation [2] in the iterative process for method 2. However in our test, it fails in a lot of cases due to non-convergence. Is the problem solvable?

3. We set up the maximum number of iterations for IRLSMC to be 10, should it be higher to ensure convergence or lower to save computational costs?

4. Notice that we only test our WLSMC and IRLSMC algorithms on one and two dimensional underlying assets American options, future work may consider in increasing the number of underlying assets.

[1] American basket put for GBM two stocks model.

[2] For more details, see Harvey [1976].

Appendix A

Tests with Makeup Data

In this chapter, we introduce some simple makeup data to test the pros and cons of the methods described in section 4. By conducting OLS, WLS and IRLS with the same, less or higher order polynomial as that used to fit the mean function, we investigate whether the estimated coefficients are close to the makeup values and use the standardized residual plots to determine whether WLS and IRLS help correct for heteroscedasticity.

A.1 Method 1 Test

Let $\boldsymbol{x}$ be a $k \times 1$ column vector that uniformly discretizes the interval $[0, 1]$ with k points and $\boldsymbol{x}^M$ be a $k \times 1$ vector such that each element in $\boldsymbol{x}^M$ is to the M^{th} polynomial order of the corresponding element in $\boldsymbol{x}$[1]. Then we construct the matrix $\boldsymbol{X} = (\mathbf{1}, \boldsymbol{x}, ..., \boldsymbol{x}^M)$ as predictors with $\boldsymbol{\beta}$ as coefficients to be determined. The noise term $\boldsymbol{\epsilon}$ is constructed to be a $k \times 1$ column vector. For the j^{th} entry in $\boldsymbol{\epsilon}$, we have

$$\epsilon_j = Z_j \sqrt{(x_j)^\lambda}, \tag{A.1.1}$$

where x_j is the j^{th} entry of $\boldsymbol{x}$, Z_j is a random sample generated from $N(0, 1)$ and λ is a parameter to be specified.

[1] In this case, $\boldsymbol{x} = \boldsymbol{x}^1$

Table A.1: Estimated β and λ of cubic mean function in predictors, method 1.

k	Type	β_0	β_1	β_2	β_3	2-norm	λ
500	OLS	1.04(**)	1.39(**)	4.69(**)	2.84(**)	2.14	0
500	WLS	1.00(**)	1.98(**)	3.08(**)	3.98(**)	0.08	2.08(*-)
500	IRLS	1.00(**)	2.00(**)	2.97(**)	4.08(**)	0.09	3.95(**)
5000	OLS	1.00(**)	1.92(**)	3.31(**)	3.73(**)	0.42	0
5000	WLS	1.00(**)	2.00(**)	3.06(**)	3.93(**)	0.09	3.06(*-)
5000	IRLS	1.00(**)	2.00(**)	3.03(**)	3.97(**)	0.04	3.98(**)
50000	OLS	1.00(**)	1.94(**)	3.19(**)	3.87(**)	0.24	0
50000	WLS	1.00(**)	2.00(**)	3.00(**)	4.01(**)	0.01	3.11(*-)
50000	IRLS	1.00(**)	2.00(**)	3.00(**)	4.01(**)	0.01	4.00(**)

Note: β_0 to β_3 are the estimated coefficients, 2-norm computes the second norm between estimated coefficients and true values. λ represents the estimated value according to equation 5.1.2. The symbols in brackets are the hypothesis testings on estimated coefficients at a significance level of 0.05, with (**: different from 0 but not true value, *-: different from 0 and true value, -*: not different from both 0 and true value, –: not different from 0 and different from true value).

Finally the response variable $\boldsymbol{Y}$ is computed as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}. \tag{A.1.2}$$

The parameters we select for $k = \{500, 5000, 50000\}$, $M = 3$, $\boldsymbol{\beta} = (1, 2, 3, 4)$, $\lambda = 4$. In the IRLS algorithm, the stopping rule uses a tolerant value of 0.01 and a maximum number of iterations 10. We first consider applying a cubic mean function to predictors, which is the same polynomial orders with $\boldsymbol{X}$.

Table A.1 shows the results of estimating parameters. It is clear that WLS gives better coefficient estimates compared to OLS, especially for $k = 500$, the value of the second norm drops from 2.14 to 0.09. For λ estimates, although for all k WLS gives estimates significantly different from 0, they are different from the true value. However, IRLS can provide more accurate coefficient estimates than OLS and generates better

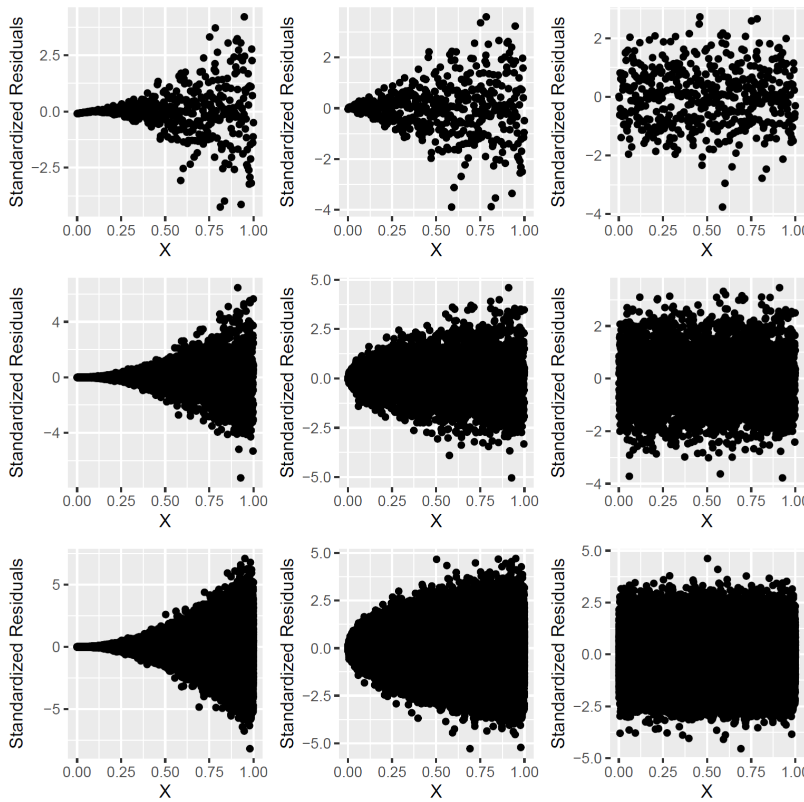

Figure A.1: Standardized residual plots of OLS(left), WLS(middle) and IRLS(right) for k = 500, 5000 and 50000 top, middle and bottom rows with cubic mean function on predictors.

estimates in λ compared to WLS. Furthermore, IRLS algorithm converges before 10 iterations for every k. Thus IRLS produces better results in this case.

Figure A.1 are the residual plots correspond to OLS, WLS and IRLS. For OLS, since the variance of residuals increases with $\boldsymbol{x}$, and p.values from BPtest are less than 0.05, we conclude that heteroscedasticity exists. In contrast with OLS, a trend with smaller variance change in residuals is observed after using WLS. However, the p.values from BPtest are still less than 0.05. Thus, there is statistical evidence of heteroscedasticity in residuals of WLS. IRLS gives the best residual plots where residuals are uniformly distributed with p.values greater than 0.05, no evidence of heteroscedasticity. This result is convincing since IRLS provides the best λ estimates.

Next, we consider the case of applying the linear mean function to predictors. That is, we use a lower polynomial order and misspecified mean function when the true mean

Table A.2: Estimated β and λ for linear mean function on predictors, method 1

k	Type	β_0	β_1	λ
500	OLS	-0.31(+)	8.65(+)	0
500	WLS	-0.36(+)	8.73(+)	-0.04(−)
500	IRLS	-1.50(+)	10.43(+)	-0.99(*-)
5000	OLS	-0.30(+)	8.59(+)	0
5000	WLS	-0.40(+)	8.75(+)	-0.08(*-)
5000	IRLS	-1.41(+)	10.27(+)	-0.94(*-)
50000	OLS	-0.31(+)	8.61(+)	0
50000	WLS	-0.39(+)	8.75(+)	-0.07(*-)
50000	IRLS	-1.41(+)	10.27(+)	-0.91(*-)

Note: (+) means the result is statistically different from 0 and (-) is the opposite, other symbols in brackets are consistent with Table A.1.

function is cubic.

Table A.2 presents that all estimating coefficients are significantly different from 0. Unlike the cubic mean function, WLS and IRLS have trouble estimating the true λ, and IRLS does not converge before 10 iterations for all k. Thus when the fitted mean function is significantly misspecified, WLS and IRLS are not effective.

As expected, we do not find any improvement from WLS and IRLS compared to OLS in Figure A.2. Since they all have p.values smaller than 0.05 from Bptest, we conclude that statistical evidence of heteroscedasticity exists in all 9 cases.

Finally, we want to check the performance of the quartic mean function on predictors, with the expected coefficient value on the quartic term to be 0, since the data is generated using a cubic mean function.

Similar to Table A.1, Table A.3 shows that WLS gives better coefficient estimates than OLS, and lambda is different from the true value. The quartic term is not different from 0, as expected. IRLS is able to achieve both goals and converges before 10

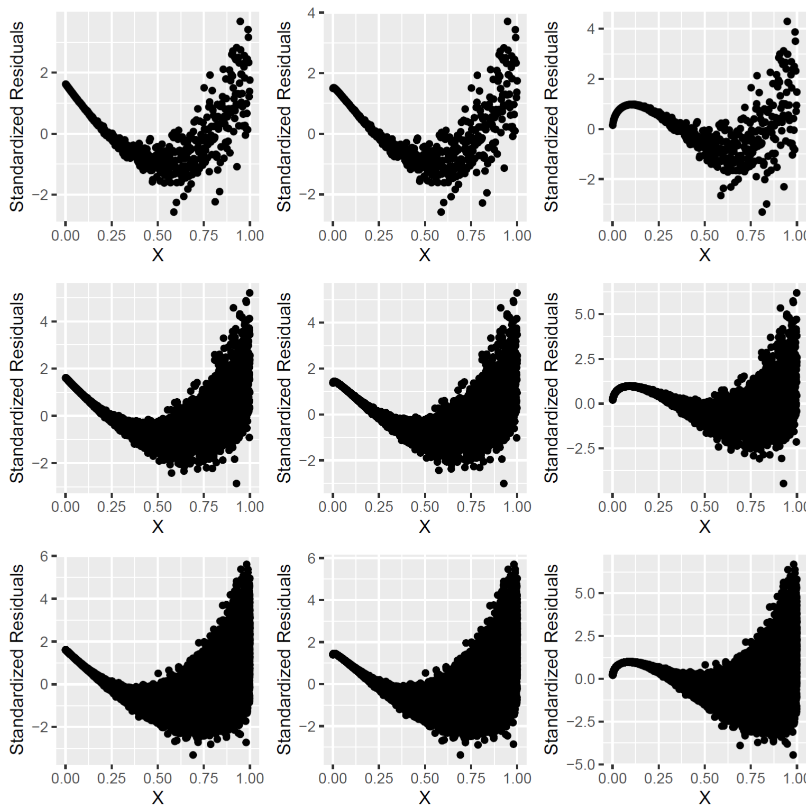

Figure A.2: Standardized residual plots of OLS(left), WLS(middle) and IRLS(right) for $k = 500$, 5000 and 50000 top, medium and bottom rows with linear mean function on predictors.

Table A.3: Estimated β and λ with quartic mean function on predictors, method 1

k	Type	β_0	β_1	β_2	β_3	β_4	λ
500	OLS	0.94(**)	3.37(**)	-4.23(-*)	16.71(**)	-6.93(-)	0
500	WLS	1.00(**)	2.06(**)	2.20(**)	6.12(**)	-1.42(-)	1.98(*-)
500	IRLS	1.00(**)	2.00(**)	2.97(**)	4.07(**)	0.01(-)	3.96(**)
5000	OLS	1.00(**)	2.02(**)	2.85(-*)	4.45(-*)	-0.36(-)	0
5000	WLS	1.00(**)	2.00(**)	2.99(**)	4.17(**)	-0.02(-)	3.79(*-)
5000	IRLS	1.00(**)	2.00(**)	2.99(**)	4.17(**)	-0.02(-)	3.98(**)
50000	OLS	1.00(**)	2.06(**)	2.63(**)	4.73(**)	-0.43(-)	0
50000	WLS	1.00(**)	2.00(**)	2.98(**)	4.07(**)	-0.05(-)	3.34(*-)
50000	IRLS	1.00(**)	2.00(**)	3.01(**)	3.98(**)	0.02(-)	4.00(**)

Note: Same symbols in the brackets with Table A.1

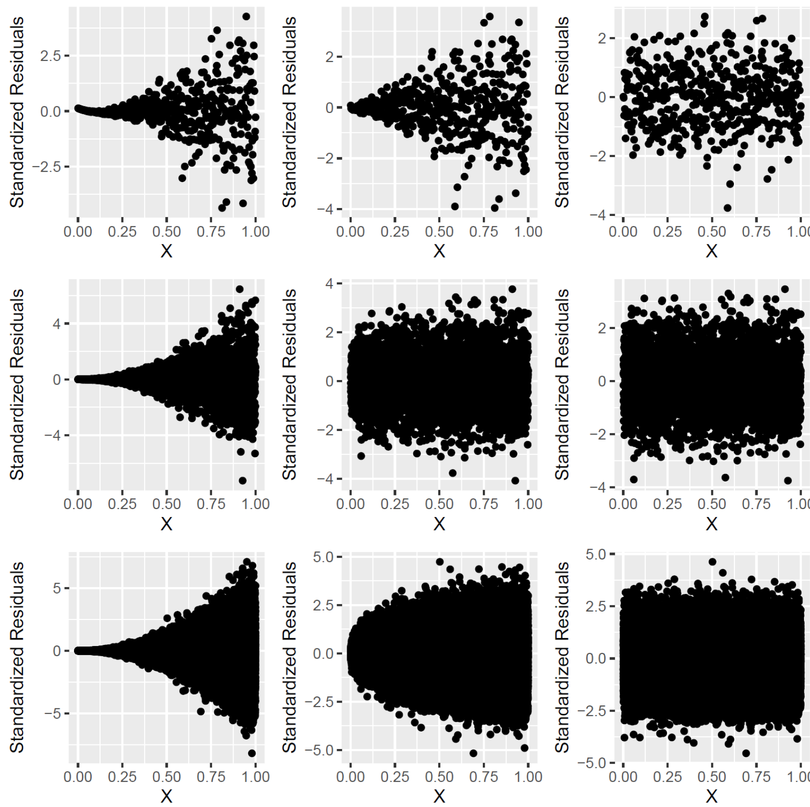

Figure A.3: Standardized residual plots of OLS(left), WLS(middle) and IRWLS(right) for $k = 500$, 5000 and 50000 top, medium and bottom rows with quartic mean function on predictors.

iterations. We have the same results with the case of cubic mean function.

Figure A.3 also looks similar to Figure A.1. The most significant difference comes with the case of WLS, $k = 5000$. We could hardly see variance change in residuals with $\boldsymbol{x}$. However, results from Bptest demonstrate that statistical evidence of heteroscedasticity exists in OLS and WLS, but IRLS corrects the heteroscedasticity pretty well.

A.2 Method 2 Test

With the same setting as in method 1, we make some modifications to the noise term $\boldsymbol{\epsilon}$. Let $\boldsymbol{Z} = (\mathbf{1}, \boldsymbol{x}, ..., \boldsymbol{x}^{\mathscr{M}})$ be a $k \times (\mathscr{M}+1)$ matrix where $\mathscr{M} \leq M$ and $\boldsymbol{\alpha}$ be a $(\mathscr{M}+1)\times 1$ vector corresponds to $\boldsymbol{Z}$. For the j^{th} row of $\boldsymbol{Z}$, we have $\boldsymbol{Z_j} = (1, x_j, ..., x_j^{\mathscr{M}})$. Thus, the

j^{th} entry of $\boldsymbol{\epsilon}$ is computed as

$$\boldsymbol{\epsilon}_j = Z_j\sqrt{e^{\boldsymbol{Z_j\alpha}}}, \tag{A.2.1}$$

where Z_j is a random sample generated from $N(0,1)$.

The parameters we select for $k = \{500, 5000, 50000\}$, $m = 3$, $\boldsymbol{\beta} = (1, 2, 3, 4)$, $\boldsymbol{\alpha}$ $= (0.01, 1, 1)$. For IRLS, We use the same tolerant value and maximum iteration of iterations as used for method 1. Since the error term is quadratic exponential affine heteroscedastic, thus besides quadratic heteroscedasticity correction[1], we are also interested in the performance of linear heteroscedasticity correction[2]. In this case, we first consider the case of cubic mean function on predictors.

Table A.4 presents the results of estimated coefficients for both linear and quadratic heteroscedasticity correction. Starting at the linear case, it is evident that WLS1 gives better coefficient estimates compared to OLS. Especially for $k = 500$, the second norm decreases from 6.17 to 3.35. IRLS provides slight benefit in the case of $k = 500$, but no significant improvement is found for $k = 5000$ or 50000. Compared to the linear case and OLS, WLS2 has better performance in estimating coefficients for $k = 500$. IRLS2 performs similarly to WLS2 in large k. Turn to α estimation, more accurate and statistically significant results are obtained as k increases.

Look at Figure A.4, it is clear that OLS has the problem of heteroscedasticity. As $\boldsymbol{x}$ increases, the variance of residuals increases with an exponential trend. WLS1 removes the trend pretty well for $k = 500$. However, small concave curves were observed at the higher and lower bound of the residual plots for $k = 5000$ and 50000. And BPtest gives results with p.value less than 0.05. Therefore we conclude statistical evidence of heteroscedasticity exists for $k = 5000$ and 50000 for all OLS/WLS/IRLS. IRLS1 provides little benefit above WLS1.

Turning to quadratic heteroscedasticity correction plotted in Figure A.5, it looks very similar to Figure A.4. Compared to OLS and WLS1, the trend is removed completely for WLS2 and IRLS2. Since BPtest gives results with p.values greater than 0.05,

[1] Denoted as WLS2 and IRLS2 in the following tables and context.

[2] Denoted as WLS1 and IRLS1 in the following tables and context.

Table A.4: Estimated β and α with cubic mean function in predictors, method 2 using both linear ($\alpha_2 = 0$) and quadratic heteroscedasticity correction.

k	Type	β_0	β_1	β_2	β_3	2-norm	α_0	α_1	α_2
500	OLS	1.06(**)	0.35(-*)	7.88(-*)	0.61(-*)	6.17	0	0	0
500	WLS1	1.00(**)	1.20(-*)	5.67(-*)	2.13(-*)	3.35	0.15(+)	1.44(+)	0
500	WLS2	1.00(**)	1.22(-*)	5.58(-*)	2.21(-*)	3.24	0.21(-*)	1.07(-*)	0.36(-*)
500	IRLS1	1.00(**)	1.23(-*)	5.56(-*)	2.21(-*)	3.22	0.07(-)	1.52(+)	0
500	IRLS2	1.00(**)	1.26(-*)	5.45(-*)	2.31(-*)	3.06	0.19(-*)	0.91(-*)	0.60(-*)
5000	OLS	0.98(**)	2.12(**)	3.14(-*)	3.66(**)	0.38	0	0	0
5000	WLS1	0.97(**)	2.25(**)	2.78(-*)	3.91(**)	0.35	-0.12(-)	1.99(+)	0
5000	WLS2	0.97(**)	2.25(**)	2.78(-*)	3.92(**)	0.35	0.04(-*)	1.04(**)	0.95(**)
5000	IRLS1	0.97(**)	2.25(**)	2.78(-*)	3.91(**)	0.35	-0.12(-)	1.98(+)	0
5000	IRLS2	0.97(**)	2.25(**)	2.78(-*)	3.92(**)	0.35	0.02(-*)	1.15(**)	0.84(**)
50000	OLS	1.03(**)	1.64(**)	3.94(**)	3.38(**)	1.18	0	0	0
50000	WLS1	1.03(**)	1.72(**)	3.72(**)	3.54(**)	0.90	-0.16(+)	2.00(+)	0
50000	WLS2	1.03(**)	1.70(**)	3.76(**)	3.51(**)	0.95	0.03(**)	0.87(**)	1.13(**)
50000	IRLS1	1.03(**)	1.72(**)	3.72(**)	3.54(**)	0.90	-0.17(+)	2.01(+)	0
50000	IRLS2	1.03(**)	1.70(**)	3.76(**)	3.51(**)	0.95	0.02(**)	0.88(**)	1.12(**)

Note: WLS1 and IRLS1 represent linear heteroscedasticity correction and WLS2, IRLS2 are for quadratic heteroscedasticity correction, other forms and symbols in brackets follow Table A.1 and Table A.2.

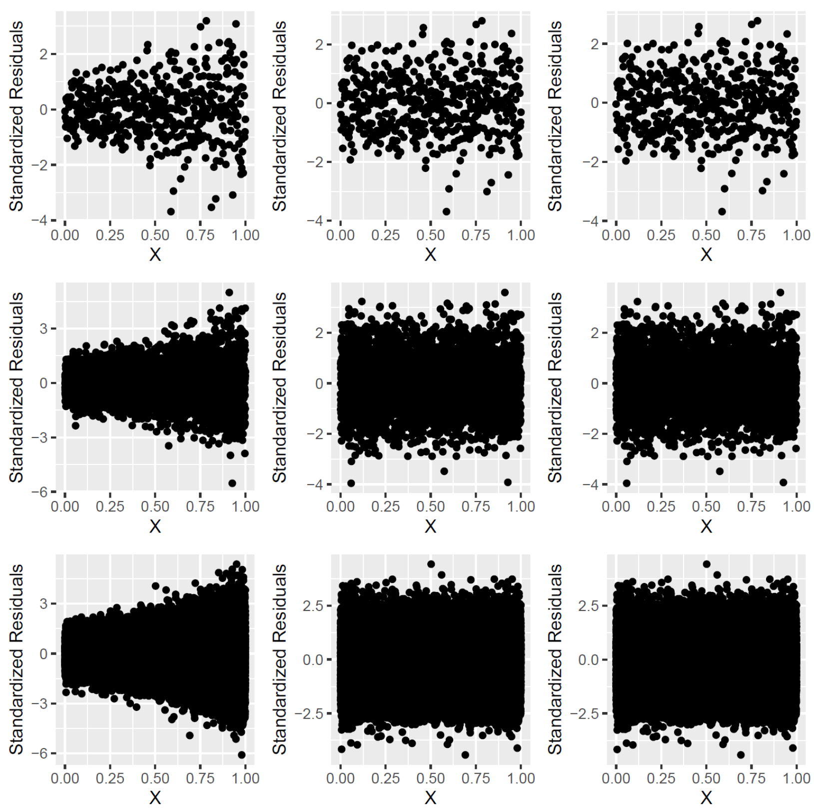

Figure A.4: Residual Plots of Method 2 with OLS(left), WLS(middle) and IR-WLS(right) for k = 500, 5000 and 50000 top, medium and bottom rows with cubic mean function on predictors, linear heteroscedasticity correction.

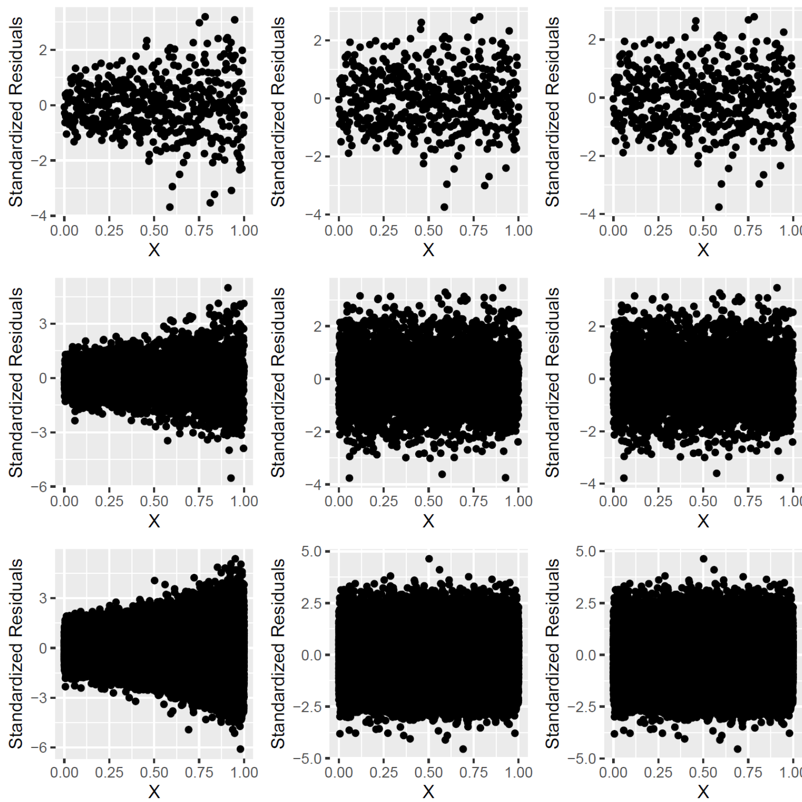

Figure A.5: Standardized residual plots of Method 2 with OLS(left), WLS(middle) and IRLS(right) for $k = 500$, 5000 and 50000 top, medium and bottom rows with cubic mean function on predictors, quadratic heteroscedasticity correction.

Table A.5: Estimated β and α of linear mean function on predictors, method 2 using both linear ($\alpha_2 = 0$) and quadratic heteroscedasticity correction.

k	Type	β_0	β_1	α_0	α_1	α_2
500	OLS	-0.37(+)	8.78(+)	0	0	0
500	WLS1	-0.02(-)	7.99(+)	0.41(+)	1.34(+)	0
500	WLS2	-0.05(-)	7.98(+)	0.54(-*)	0.56(-*)	0.78(-*)
500	IRLS1	0.12(-)	7.62(+)	0.04(-)	1.98(+)	0
500	IRLS2	0.12(-)	7.42(+)	0.20(-*)	0.50(-*)	1.73(-*)
5000	OLS	-0.27(+)	8.56(+)	0	0	0
5000	WLS1	0.13(+)	7.64(+)	0.22(+)	1.63(+)	0
5000	WLS2	0.05(-)	7.57(+)	0.64(*-)	-0.92(*-)	2.54(*-)
5000	IRLS1	0.18(+)	7.50(+)	0.09(-)	1.87(+)	0
5000	IRLS2	0.12(+)	7.39(+)	0.44(*-)	-0.52(*-)	2.47(*-)
50000	OLS	-0.30(+)	8.62(+)	0	0	0
50000	WLS1	0.13(+)	7.61(+)	0.20(+)	1.70(+)	0
50000	WLS2	0.05(+)	7.55(+)	0.63(*-)	-0.87(*-)	2.56(*-)
50000	IRLS1	0.23(+)	7.34(+)	-0.06(+)	2.16(+)	0
50000	IRLS2	0.16(+)	7.25(+)	0.39(*-)	-0.69(*-)	2.89(*-)

Note: Same format as in Table A.4.

no statistical evidence of heteroscedasticity exists for all cases in WLS2 and IRLS2.

Next, we focus on the case of linear mean function on predictors.

Most of the estimated coefficients in Table A.5 are statistically different from 0. Since the makeup data is made with cubic polynomial order, comparing them with true values is useless. Focus on α estimations for WLS and IRLS, unlike the previous case, they have significant differences for different k. Thus we think they may affect the results of heteroscedasticity correction.

In Figure A.6, as the variance of residuals increases with $\boldsymbol{x}$ in OLS, and BPtest gives results of p.values less than 0.05. We conclude that heteroscedasticity exists in OLS.

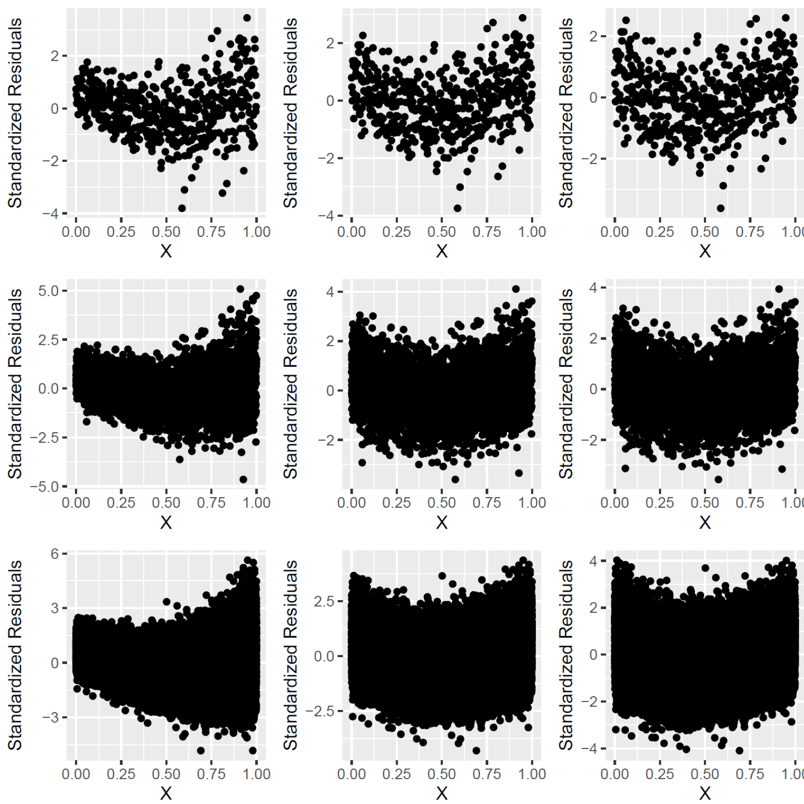

Figure A.6: Standardized residual plots of Method 2 with OLS(left), WLS(middle) and IRLS(right) for k = 500, 5000 and 50000 top, medium and bottom rows with linear mean function on predictors, linear heteroscedasticity correction.

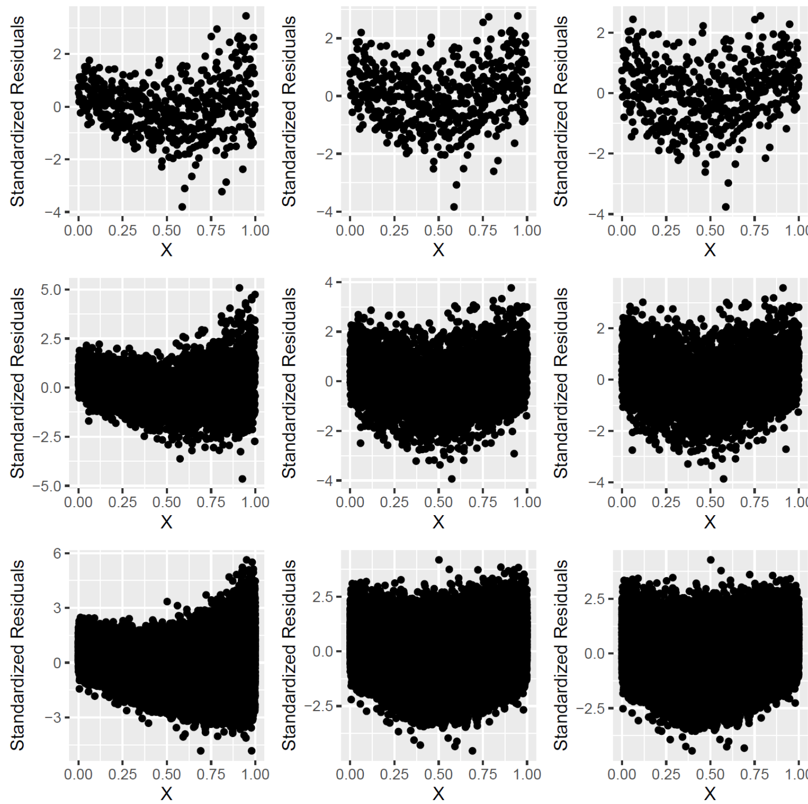

Figure A.7: Standardized residual plots of Method 2 with OLS(left), WLS(middle) and IRLS(right) for k = 500, 5000 and 50000 top, medium and bottom rows with linear mean function on predictors, quadratic heteroscedasticity correction.

There are two advantages in conducting WLS1. 1: Focus on the Y-axis, points with high residuals are reduced, especially for the case of k = 50000. 2: Removed the trend to some extent. However, statistical evidence of heteroscedasticity exists as BPtest gives results of p.values less than 0.05. IRLS1 furthermore slightly fixes the trend and for $k = 500$ and 5000, results from BPtest show no evidence of heteroscedasticity.

Residual plots in Figure A.7 looks similar to Figure A.6. For $k = 50000$, the upper bound from WLS2 and IRLS2 plots seems to be less concave. Focused on IRLS2, since BPtest provides results of p.values greater than 0.05 for all k, no statistical evidence of heteroscedasticity.

Finally, we consider the case of quartic mean function in predictors.

In Table A.6, all estimated β_4 are not statistically different from 0. This makes sense

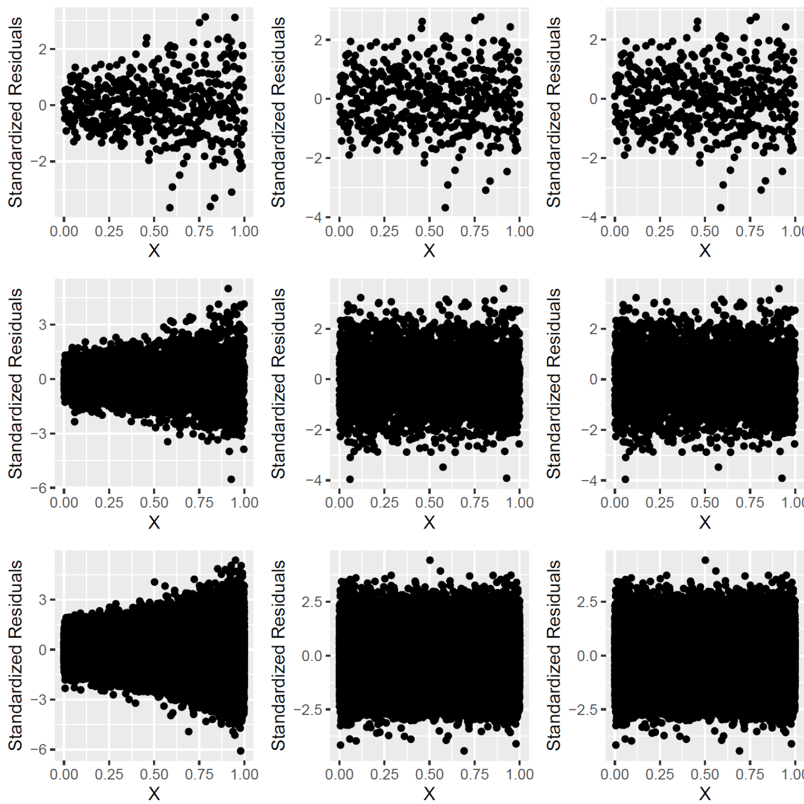

Figure A.8: Standardized residual plots of Method 2 with OLS(left), WLS(middle) and IRLS(right) for $k = 500$, 5000 and 50000 top, medium and bottom rows with quadratic mean function on predictors, linear heteroscedasticity correction.

as the makeup data is with cubic mean function. For α estimation, closer estimates were obtained when k increases compared to Table A.4.

In terms of Figure A.8 and Figure A.9, they have the same results with Figure A.4 and Figure A.5. Thus we consider that heteroscedasticity correction with a higher mean function makes no difference with the same mean function.

A.3 Findings from test examples

For method 1, in the cases of cubic mean function and quartic mean function in predictors, WLS can achieve better coefficient estimates, especially for small k but is insufficient for heteroscedasticity correction. IRLS can fix both problems and thus, is

Table A.6: Estimated β and α for quartic mean function in predictors, method 2 using both linear ($\alpha_2 = 0$) and quadratic heteroscedasticity correction.

k & Type	β_0	β_1	β_2	β_3	β_4	α_0	α_1	α_2
500 OLS	0.79(**)	5.95(-*)	-17.37(-*)	39.92(-*)	-19.65(-)	0	0	0
500 WLS1	0.85(**)	4.49(-*)	-10.55(-*)	29.00(-*)	-14.08(-)	0.26(-)	1.38(+)	0
500 WLS2	0.86(**)	4.42(-*)	-10.11(-*)	28.15(-*)	-13.58(-)	0.34(-*)	0.94(-*)	0.45(-*)
500 IRWLS1	0.86(**)	4.48(-*)	-10.50(-*)	28.92(-*)	-14.04(-)	0.23(-)	1.40(+)	0
500 IRWLS2	0.86(**)	4.40(-*)	-9.97(-*)	27.89(-*)	-13.44(-)	0.32(-*)	0.87(-*)	0.53(-*)
5000 OLS	0.96(**)	2.65(-*)	0.76(-*)	7.37(-*)	-1.85(-)	0	0	0
5000 WLS1	0.96(**)	2.67(**)	0.67(-*)	7.51(-*)	-1.92(-)	-0.13(+)	1.99(+)	0
5000 WLS2	0.95(**)	2.71(**)	0.48(-*)	7.81(-*)	-2.08(-)	0.02(-*)	1.13(**)	0.86(**)
5000 IRWLS1	0.96(**)	2.67(**)	0.67(-*)	7.51(-*)	-1.92(-)	-0.13(+)	1.99(+)	0
5000 IRWLS2	0.95(**)	2.71(**)	0.48(-*)	7.80(-*)	-2.07(-)	0.01(-*)	1.15(**)	0.85(**)
50000 OLS	1.02(**)	1.94(**)	2.58(-*)	5.50(-*)	-1.06(-)	0	0	0
50000 WLS1	1.01(**)	2.03(**)	2.11(-*)	6.27(**)	-1.46(-)	-0.16(+)	2.01(+)	0
50000 WLS2	1.02(**)	1.99(**)	2.33(-*)	5.93(**)	-1.29(-)	0.03(-*)	0.87(**)	1.13(**)
50000 IRWLS1	1.01(**)	2.03(**)	2.11(-*)	6.27(**)	-1.46(-)	-0.16(+)	2.01(+)	0
50000 IRWLS2	1.02(**)	1.98(**)	2.33(-*)	5.93(**)	-1.29(-)	0.03(-*)	0.86(**)	1.15(**)

Note: Same format as in Table A.4.

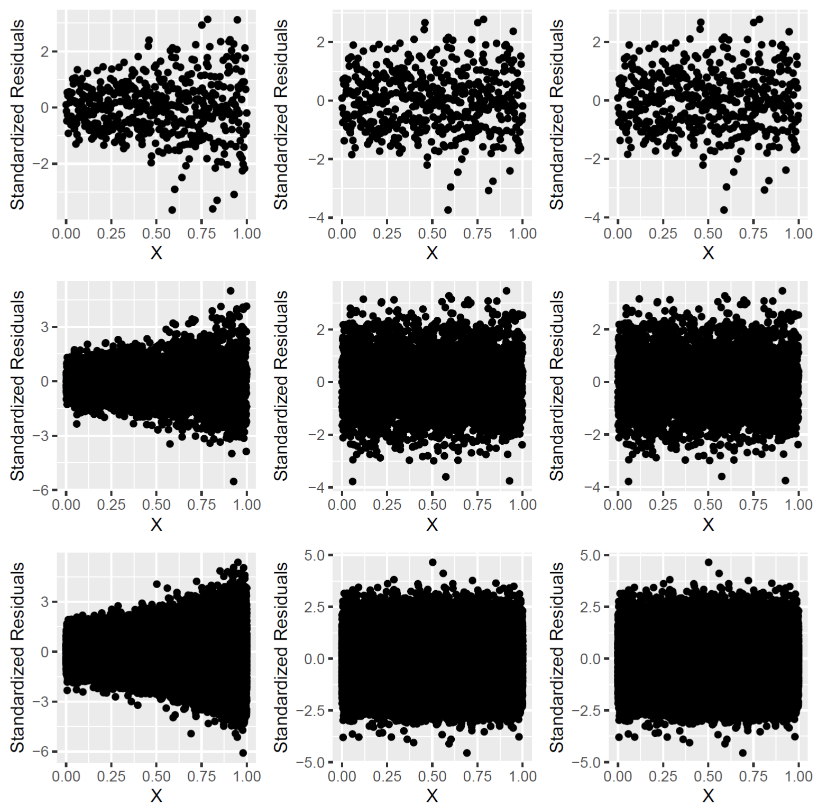

Figure A.9: Standardized residual plots of Method 2 with OLS(left), WLS(middle) and IRLS(right) for $k = 500$, 5000 and 50000 top, medium and bottom rows with quadratic mean function on predictors, quadratic heteroscedasticity correction.

considered the best approach. However, for linear mean function in predictors, WLS and IRLS are unable to correct heteroscedasticity and cannot obtain the correct value for lambda.

In terms of method 2, the significant difference compared to method 1 is that IRLS performs better in both linear and quadratic heteroscedasticity correction under linear mean function in predictors.

Since we do not know the actual parameters in LSMC for continuation values estimates, thus, when conducting WLSMC with two methods, we should not use linear mean function in predictors. IRLSMC is necessary to implement. We also expect that WLSMC or IRLSMC will have a larger effect on the option price when there are fewer stock paths in the sample.

Bibliography

Francis A Longstaff and Eduardo S Schwartz. Valuing American options by simulation: a simple least-squares approach. *The Review of Financial Studies*, 14(1):113–147, 2001.

Frank J Fabozzi, Tommaso Paletta, and Radu Tunaru. An improved least squares Monte Carlo valuation method based on heteroscedasticity. *European Journal of Operational Research*, 263(2):698–706, 2017.

Rolla E Park. Estimation with heteroscedastic error terms. *Econometrica*, 34(4):888, 1966.

Andrew C Harvey. Estimating regression models with multiplicative heteroscedasticity. *Econometrica*, pages 461–465, 1976.

Michael J Brennan and Eduardo S Schwartz. Finite difference methods and jump processes arising in the pricing of contingent claims: a synthesis. *Journal of Financial and Quantitative Analysis*, 12:659, 1977.

John C Cox, Stephen A Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229–263, 1979.

Mark Broadie and Paul Glasserman. Pricing American-style securities using simulation. *The Journal of Economic Dynamics and Control*, 21(8-9):1323–1352, 1997.

Lars Stentoft. Convergence of the least squares Monte Carlo approach to American option valuation. *Management Science*, 50(9):1193–1203, 2004.

Kin Hung Kan and R Mark Reesor. Bias reduction for pricing American options by least-squares Monte Carlo. *Applied Mathematical Finance*, 19(3):195–217, 2012.

Lars Stentoft. Efficient numerical pricing of American call options using symmetry arguments. *Journal of Risk and Financial Management*, 12(2):59, 2019.

Francois-Michel Boire, R Mark Reesor, and Lars Stentoft. American option pricing with importance sampling and shifted regressions. *Journal of Risk and Financial Management*, 14(8):340, 2021.

Ludovic Goudenège, Andrea Molent, and Antonino Zanette. Variance reduction applied to machine learning for pricing Bermudan/American options in high dimension. *arXiv:1903.11275*, 2019.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, pages 987–1007, 1982.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

Jin-Chuan Duan. The GARCH option pricing model. *Mathematical Finance*, 5(1): 13–32, 1995.

Lars Stentoft. Pricing American options when the underlying asset follows GARCH processes. *Journal of Empirical Finance*, 12(4):576–611, 2005.

Robert F Engle and Victor K Ng. Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1749–1778, 1993.

Lars Stentoft. American option pricing with discrete and continuous time models: An empirical comparison. *Journal of Empirical Finance*, 18(5):880–902, 2011.

Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

Ravi Myneni. The pricing of the American option. *The Annals of Applied Probability*, pages 1–23, 1992.

Frederick James. Monte Carlo theory and practice. *Reports on progress in Physics*, 43 (9):1145, 1980.

William H Greene. *Econometric Analysis*. Prentice Hall, 2012.

Trevor S Breusch and Adrian R Pagan. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, pages 1287–1294, 1979.

Curriculum Vitae

Name

Xiaotian Zhu

Education and Degrees

Bachelor of Mathematics (2016 - 2020)

Univeristy of Waterloo, ON, Canada

Major: Actuarial Science & Statistics

Minor: Applied Mathematics

Master of Science (2020 - 2022)

Western University, ON, Canada

Major: Financial Modelling

Related Work Experience

Teaching Assistant (2021-2022)

-