THE LOG-PERIODIC POWER LAW MODEL: AN EXPLORATION

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

MATHEMATICS

AUGUST 2016

By:
Marcus Jared Scott Takeshi Mukai

Dissertation Committee:
Laurence Thomas Ramsey, Chairperson
David A. Ross
Daisuke Takagi
Bjørn Kjos-Hanssen
Lynne Wilkens

Dedicated to those who came before me, lighting the path on which I now walk.

ACKNOWLEDGMENTS

Words cannot express how thankful I am for all the guidance and advice Dr. Ramsey shared with me throughout this project. The uncountable number of seminars, meetings, discussions, and email correspondence provided by Dr. Ramsey were instrumental in the completion of this dissertation. I am extremely grateful for all the dedication, attention, and patience Dr. Ramsey graciously bestowed upon me.

I extend thanks to my committee members: Dr. Ross, Dr. Takagi, Dr. Kjos-Hanssen, and Dr. Wilkens. Their feedback was welcomed and appreciated. With their input, this dissertation was shaped into a body of work of which I am immensely proud.

To my family: my parents, Marcus and Mimi, and my two brothers, Justin and Jason. Their encouragement and support over the years has been invaluable. To my colleagues, especially Jamal, Alex, Erica, and Doug, I most certainly would not have reached the finish line without their friendship and aid.

Finally, I thank all the teachers and professors at the following departments who inspired me to study mathematics and provided me my education: Curtis Senior High School, Pepperdine University, and the University of Hawaiʻi at Mānoa.

ABSTRACT

Over the last two decades, a new financial model has emerged that might explain some of the rare instances of truly extreme volatility in asset prices. This model, known as the Log-Periodic Power Law (LPPL) model or the Johansen-Ledoit-Sornette (JLS) model, attempts to diagnose, time, and predict the termination of these bubbles; we caution that there is no academic agreement about the existence or definition of a "bubble." The creators of the model provide a motivation built upon some natural assumptions including risk-neutral assets, rational expectations, local self-reinforcing imitation, and probabilistic critical times. The model has evolved over time, partially in response to some sound criticism. This dissertation is focused on two criticisms that have not been fully addressed. First, it is unknown whether there exist unique best fits of the JLS model to log-price data. In this dissertation we explore the first level JLS model and analyze its relationship with extreme boundary vectors which serve as the building blocks for increasing concave up log-price paths. Second, it is unknown whether the current method for locating local minima is sufficient. Using numerical analysis, this dissertation uses the Cauchy-Schwarz Theorem and Taylor's Theorem to find bounds for various moduli of continuity and first and second derivatives of the error of the JLS model fit to appropriate log-price paths. In addition to discussing these criticisms, the question of the applicability of the JLS model is considered. Without committing ourselves to a definition of a bubble, this dissertation also presents the results of an ongoing test attempting to determine whether the JLS model can be used to generate systematic profits. An experiment using the JLS model on specifically chosen stocks filtered from the New York Stock Exchange (NYSE) is detailed along with a step-by-step procedure to determine specific dates on which to utilize a trading strategy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

CHAPTER 1
INTRODUCTION

1.1 Preliminaries

Various forms of the Efficient Market Hypothesis are routinely taught in such places as graduate programs in finance or economics [16]. In particular, it is often hypothesized that stock prices should be modeled by Geometric Brownian Motion (GBM). That is, the logarithms of stock prices should, and can, be modeled as Brownian motion with a constant drift and volatility. More specifically, if $P(t)$ denotes the price of an asset at time t, then there are drift and volatility constants, α and $\beta > 0$, respectively, such that

$$d\left[\ln P(t)\right] = \alpha\, dt + \beta\, dW(t), \quad \text{where } W(t) \text{ is a standard Wiener process}$$

It is easy to criticize the GBM model as over-simplistic. Remarkably, the model is not as bad as intuition might dictate. In particular, the GBM model can be used to derive the Black-Scholes-Merton formula for pricing (certain) options [16]. The GBM hypothesis is used in this dissertation as a null hypothesis.

Researchers Johansen, Ledoit, and Sornette (JLS) essentially accept, most likely with some caveats, that the GBM hypothesis basically works most of the time, in most markets, for most assets. Their main quibble would be that the distribution of $\ln P(t+\delta) - \ln P(t)$ might be too skewed to be considered normal or has tails that are too heavy. While these are valid and frequently made criticisms, efforts to replace the null hypothesis with an equally specific one have been unsatisfactory. For example, no distribution has been determined to systematically characterize the heaviness of the tails [14]. However, JLS strongly believe that there are (rare) occasions when a completely different dynamic takes over, leading to the formation of bubbles and crashes. After much experimentation, they have settled on the following model for these admittedly rare episodes of market history:

$$\ln[P(t)] \approx A + B(T-t)^m \left[1 + C\cos[\omega \ln(T-t) + \theta]\right] \tag{1.1}$$

where $A > 0$, $B < 0$, $C \neq 0$, T is the critical time when the special dynamic ends and a transition must occur to a different governing behavior, m is the critical exponent, ω is the frequency of the oscillations during the bubble, and $\theta \in [0, 2\pi)$ is a phase parameter. In this dissertation we call Equation 1.1 the second level JLS model. It is also known as the Log Periodic Power Law (LPPL) model, partly inspired by (when $C \neq 0$) the systematic oscillations that occur ever more frequently until a discontinuity occurs.

More recently, the JLS model was rewritten to shift from 3 linear parameters and 4 non-linear parameters to 4 linear parameters and 3 non-linear parameters [4] via the application of the trigonometric identity

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y) \tag{1.2}$$

Using $x = \omega \ln(T-t)$ and $y = \theta$ in Equation 1.2, we have a version of the second level JLS model with 4 linear parameters

$$\ln[P(t)] \approx A' + (T-t)^m \left[B' + C' \cos[\omega \ln(T-t)] + D' \sin[\omega \ln(T-t)]\right]$$

where $A' = A$, $B' = B$, $C' = C\cos(\theta)$, $D' = -C\sin(\theta)$. Note that when $C = D = 0$, the JLS model becomes, what this dissertation considers, the first level JLS model as used in Chapter 2. The restriction of m to $[0.1, 0.9]$ and ω to $[6, 13]$ [4] is therefore a divergence from the null hypothesis. JLS and other researchers have fit this model to numerous episodes of bubbles in recent history and have, in some cases, predicted several crashes ex-ante [19]. Several examples have been verified [6, 10, 22].

In this dissertation, mathematical and practical aspects of the JLS model are explored. Just about everything is unknown.

- JLS determine the coefficients of the model by minimizing the sum of squared errors. Unfortunately, it is still unknown whether there is a unique minimum [1, 2]. Chapter 2 addresses a simpler version of the model and proves that the minimum is unique for the simpler model and some canonical cases. For practical reasons, JLS limit the parameter m to the real interval $[0.1, 0.9]$. Chapter 2 proves that, for log-prices that are truly concave up and increasing (with no noise), that (mathematically) m must be in $[0, 1]$. Finally, Chapter 2 uses limit arguments to establish the model appropriate for the value of $m = 0$. This extends the applicability of the JLS model as it supplies possible models for log-prices that exhibit greater concavity.

- The optimization over the non-linear parameters T, m, and ω is challenging. JLS settled on a strategy of approximately 20 random selections from which to start a minimization strategy (presumably converging to some local minimum) [20]. Chapter 3 gives some bounds on first and second order derivatives of the objective function with respect to T, m, and ω. These bounds govern a more exhaustive grid search for an approximation to the global minimum.

- Chapter 4 describes the results of an experiment that tries to utilize the JLS model to successfully trade bubbles and crashes. Note that there is no consensus whatsoever that bubbles even exist, nor is there an agreement about what constitutes a bubble. This dissertation proposes a specific and novel test to identify stock prices that might be considered (perhaps after employing additional criterion) to be bubbles (crashes): the log-prices must be increasing (decreasing) and concave up (down) to a degree that would be unusual under the GBM hypothesis. To these the JLS model is fit. If the JLS model fits unusually well, then that stock price example is tracked for possible changes in the dynamic and traded under specific rules.

The rest of Chapter 1 briefly discusses the existing narrative proposed by JLS for the LPPL model and also discusses a few competing ideas for what constitutes a bubble.

1.2 Motivation Behind the JLS Model

The majority of this section has appeared in multiple papers by JLS and associated authors. More specifically, we cite much of this section from [6] and [11].

The JLS model assumes the existence of two types of traders: rational and irrational. The former group trades with rational expectations, moves the majority of invested money, has more efficient access to information, and are fewer in number than their irrational counterparts. The latter group can also be described as noise traders who act irrationally, follow herd behavior, and depend more on the opinions of their trading neighbors. According to the model, traders are organized into networks and are always in one of two states: buy or sell. Each trader's state is allowed to depend on factors including the external influences on him/her by nearby traders and his/her internal idiosyncratic behaviors. Because of these interactions between traders, over time, groups can form with similar behavior that result in significant price movements. In this sense, the model regards extreme shifts in price as times of order, while normal everyday price movements are considered times of disorder [11].

With respect to rapidly increasing prices, the JLS model views massive drawdowns as having a probability density distribution that is different from the vast majority of distributions of smaller drawdowns. These large and rare (negative) price movements tend to occur following super-exponential price growth and have been coined [5] by Sornette as "dragon-kings": unusually large outliers whose behavior is governed by a distribution different from the typically assumed distribution. This perspective on bubbles opens up the possibility of not only modeling, but also predicting when these bubbles are going to burst. A typical situation in this binary trader scenario is as follows. Rational traders take a long position in an asset. Due to the time lag of information and response by the irrational traders, the asset begins to experience extreme price movements when the herding behavior of irrational traders begins. Since asset prices cannot grow to infinity, rational traders begin taking short positions and soon the large group of irrational agents follow, causing the price of the asset to crash. Notice here that the event of the crash is probabilistic and not certain. Traders are incentivized to stay invested despite a positive probability of a crash due to potentially additional asset price increases.

In terms of price dynamics, the JLS model makes a few more assumptions [11]. First, it considers an ideal market paying no dividends and where interest rates, risk aversion, and market liquidity are negligible. Therefore, the fundamental price of a particular asset is always $p(t) = 0$ and any positive value represents a bubble. In this scenario, rational traders are risk neutral and have rational expectations. Thus, $p(t)$ also follows a martingale process

$$E_t[p(t')] = p(t), \quad \forall t < t'$$

Assuming there is a non-zero probability a crash will occur, *"we can define a jump process j such that $j = 0$ leading up to the crash and $j = 1$ the moment the crash occurs at some critical time t_c. Since t_c is unknown, it is described by a stochastic variable with a density function $q(t)$, a cumulative distribution* *function $Q(t)$, and a hazard rate, $h(t) = \frac{q(t)}{1-Q(t)}$, which is the probability per unit of time that a crash will occur in the next instant, given the crash has not yet occurred."* [6] Assuming the price crash is some percentage drop $\kappa \in (0, 1)$ with $\mu(t)$ denoting a time-varying drift parameter at the price level, [6] and [11] view the price dynamic before the crash as

$$dp = \mu(t)p(t)dt - \kappa p(t)dj \tag{1.3}$$

Assuming $t < t_c$, then

$$\begin{aligned} E[dp] &= E[\mu(t)p(t)dt] - E[\kappa p(t)dj] \\ &= \mu(t)p(t)dt - \kappa p(t)\left(\mathbb{P}(dj = 0) \times 0 + \mathbb{P}(dj = 1) \times 1\right) \\ &= \mu(t)p(t)dt - \kappa p(t)h(t)dt \end{aligned}$$

By the no arbitrage condition and rational expectations, $E[dp] = 0$ and $\mu(t) = \kappa h(t)$. Substituting this into Equation 1.3, the equation before the crash (where $dj = 0$), becomes

$$\frac{dp}{dt} = \kappa h(t)p(t) \tag{1.4}$$

Solving Equation 1.4 yields

$$\ln\left[\frac{p(t)}{p(t_0)}\right] = \kappa \int_{t_0}^{t} h(t')dt' \tag{1.5}$$

At this point, an explicit description of the hazard rate function would be useful. Note that one of the most important aspects of the JLS model is the idea that asset price crashes are not consequences of large sweeping global influences, but rather of many small imitative local interactions among rational and irrational traders. [6] and [11] show this by discussing both the macroscopic and microscopic approaches and identifying similarities.

Macroscopically, *"according to the mean field theory from Statistical Mechanics (see e.g. Stanley (1971) and Goldenfeld, (1992)),* [1] *a simple way for describing an imitative process is by assuming the hazard rate, $h(t)$, can be described by the following equation:*

$$\frac{dh}{dt} = Ch^{\delta}$$

where $C > 0$ is a constant and $\delta > 1$ represents the average number of interactions among traders minus one." [6] Solving this differential equation,

$$h(t) = \left(\frac{h_0}{t_c - t}\right)^{\alpha}, \quad \alpha = \frac{1}{\delta - 1}$$

[1] Stanley (1971) and Goldenfeld (1992) are references [21] and [7], respectively in the bibliography of this dissertation.

where t_c is the critical time determined by the initial conditions at some t_0. Notice the exponent α is necessarily larger than 0 (so that $h(t)$ increases as $t \to t_c$) and less than 1 (to prevent the price from diverging at t_c in Equation 1.5).

From a microscopic point of view, JLS assume the group of irrational traders or agents are connected in a network [11]. *"Each agent is indexed by an integer $i = 1, \cdots, I$ and $N(i)$ represents the number of agents who are directly connected with agent i in the network. JLS (2000)* [2] *assume that each agent can have only two possible states s_i: "buy" ($s_i = +1$) or "sell" ($s_i = -1$). JLS (2000) suppose the state of agent i is determined by the following Markov process* [3]:

$$s_i = sign\left(K \sum_{k \in N(i)} s_k + \sigma\epsilon_i\right) \tag{1.6}$$

where the sign function, sign(x), is equal to +1 if $x > 0$ and to -1 if $x < 0$, K is a positive constant, and ϵ_i is an i.i.d. standard normal random variable." [6] In this model, K (called the coupling strength [11]) governs the tendency for imitation and σ governs the tendency for idiosyncratic behavior. As K increases, the network will tend to fall into order and as σ increases, the network will tend toward disorder. With regard to a rapidly increasing asset price, order reigning will lead to a price crash. It should be noted at this time that by [11] this equation belongs to a class of stochastic dynamical models of interacting particles [12, 13]. Of most interest to us is the Ising model, a mathematical model of ferromagnetic interacting particles whose magnetic dipole moments of atomic spins can be one of two states. For the Ising model, *"there exists a critical point K_c that determines the separation between regimes: when $K < K_c$, disorder reigns and the sensibility to a small global influence is low. When the imitation force K grows approaching K_c, a hierarchy of groups of agents acting collectively and with the same position is formed. As a consequence, the market becomes extremely sensitive to small global disturbances. Finally, for a large imitation force so that $K > K_c$, the tendency of imitation is so intense that there exists a strong predominance of one state/position among agents."* [6] A measure of the sensitivity of the network to global influence is called *susceptibility* [11]. It describes the probability that a large group of traders will be in the same state, conditioned on the existing influences on the network. JLS *"assume the existence of a term G which measures the of global influence, and add it to Equation 1.6:*

$$s_i = sign\left(K \sum_{k \in N(i)} s_k + \sigma\epsilon_i + G\right)$$

If we define the average state of the market as $M = \frac{1}{I}\sum_{i=1}^{I} s_i$, for $G = 0$ we have $E[M] = 0$ by symmetry. For $G > 0$, we have $M > 0$, while for $G < 0$, $M < 0$. Thus, it follows that $E[M] \times G \geq 0$. The

[2] JLS (2000) is reference [11] in the bibliography of this dissertation

[3] For the mathematical reader, Equation 1.6 has a hidden time dependence and can be recast as follows:

$$s_i(t+1) = \text{sign}\left(K \sum_{k \in N(i)} s_k(t) + \sigma\epsilon_i(t)\right)$$

susceptibility of the system is then defined as $\chi = \frac{dE[M]}{dG}\Big|_{G=0}$*."* [6]

JLS believe that *"the susceptibility correctly measures the ability of the system of agents to agree on an opinion."* [11] Thus, they believe susceptibility of the network and the corresponding hazard rate of the asset price behave similarly. In the context of the (2-dimensional) Ising model, the behavior of K as it tends to K_c, resembles that of "critical *phenomena. Formally, in this case the susceptibility* χ *of the system goes to infinity. The hallmark of criticality is the* power law*, and indeed the susceptibility goes to infinity according the power law:*

$$\chi \approx A(K_c - K)^{-\gamma} \tag{1.7}$$

where A *is a positive constant and* $\gamma > 0$ *is the* critical exponent*."* [11] While the dynamics of K are unknown, JLS assume K evolves smoothly and it is possible to generate a Taylor series around K_c and use its first order Taylor polynomial to approximate $K_c - K$. If $K(t_c) = K_c$, then for $t < t_c$,

$$K_c - K \approx \text{constant} \times (t_c - t)^{-\gamma}$$

Using the assumption that the susceptibility and hazard rate behave similarly [6],

$$h(t) \approx B(t_c - t)^{-\alpha}$$

where $\alpha \in (0, 1)$. Note that t_c is not the time of the crash, but the most probabilistic time for the crash. Integrating $h(t)$ (as in Equation 1.5), yields a formula for the log prices of the asset

$$\ln[p(t)] \approx \ln[p(t_0)] - \frac{\kappa B}{\beta}(t_c - t)^{\beta}$$

where $\beta = 1 - \alpha$. This is what this dissertation considers the first level JLS model. Also note that this version of the JLS model agrees with the macroscopic view of the imitation process associated with the hazard rate.

Unfortunately, the reality of trader networks is more sophisticated than the uniform interactions for which the 2-dimensional Ising model accounts. As such, to model reality more realistically, JLS turn to *hierarchical diamond lattices* [11]. *"A version of this model was solved by Derrida et al. [3]. The basic properties are similar to those of the rational imitation model using the bi-dimensional network. The only crucial difference is that the critical exponent* γ *of the susceptibility in Equation 1.7 can be a complex number."* [6] As such, taking the real part of the first order expansion of the general solution yields

$$\chi \approx A_0'(K_c - K)^{\gamma} + A_1'(K_c - K)^{\gamma}\cos[\omega \ln(K_c - K) + \psi] + \cdots$$

where $\gamma, \omega \in \mathbb{R}$. Again, assuming the susceptibility of the network and the hazard rate behave similarly, the hazard rate can be approximated as

$$h(t) \approx B_0(t_c - t)^{-\alpha} + B_1(t_c - t)^{-\alpha}\cos[\omega\ln(t_c - t) + \psi']$$

Integrating this in Equation 1.5 results in, what this dissertation considers, the second level JLS model:

$$\ln[p(t)] \approx \ln[p(t_0)] - \frac{\kappa}{\beta}\left[B_0(t_c - t)^{\beta} + B_1(t_c - t)^{\beta}\cos[\omega\ln(t_c - t) + \theta]\right]$$

1.3 Other Ideas of Bubbles

This review of the past literature is limited to two main ideas. First, in the early papers ([11, 18] among others) JLS proposed "super-exponential growth," meaning there are constants $\alpha > 0$ and $\beta > 1$ such that

$$\frac{d}{dt}P(t) \approx \alpha P(t)^{\beta}$$

Note that non-trivial solutions of this differential equation have a vertical asymptote at a finite time. No doubt any price that followed that path would constitute a bubble! As a practical matter, Sornette found this did not work well as stock prices were too noisy to reliably estimate α and β [14]. In addition, prices do not have vertical asymptotes. So while super-exponential growth remains at the core of what Sornette considers (intuitively) to be a bubble, he has opted for a different mathematical formalism.

A second theme is that the term bubble is simply an instance of a severe departure from the fundamental value. Most often the departure itself and its severity give weak information about when the bubble will break. Professor Schiller of Yale invented the Cyclically Adjusted Price Earnings (CAPE) ratios to do this for U.S. stock prices; this statistic is maintained [15]. British economist, Andrew Smithers, has extended Yale Professor Tobin's Q-ratio to estimate the true (replacement) value of all U.S. industrial companies [17]. Mutual fund manager, John Hussman, has extensively updated and refined these ideas [9]. In the management of his mutual funds, Hussman now includes a proprietary overlay that tries to capture the timing of bubbles and crashes; the prediction intervals have not been made public but seem to be at least a year long. Finally, Jeremy Grantham of GMO, Ltd. defines a bubble as a 2-sigma event [8], always in relation to some underlying measure of true value, where the 2-sigma-ness is related to an average over time. The following is an abbreviated description of the use of Grantham's definition. Let $\lambda(t)$ be, at time t, the average (over some universe of homes) of the ratio of the price of the home to the net rental income that could be generated by renting that home. Over a long enough history, collect regular instances $\mathcal{O} = \{\lambda(t_0 + i\delta)\}_{i=1}^{N}$ of this measurement. From $\mathcal{O}$, estimate an average $\widehat{\mu}$ and standard deviation $\widehat{\sigma}$ for λ. If $\lambda(t)$ is greater than $\widehat{\mu} + 2\widehat{\sigma}$, then house prices are judged to be in a bubble at time t.

CHAPTER 2
THEOREMS ABOUT THE OPTIMIZATION

For this chapter, we specifically deal with m as the lone independent non-linear variable for the first level JLS model. Let $L \geq 3$ be the length of the original data vector. Let $T \geq L+1$ be the critical time. We consider T to be a fixed constant. Given a price vector exhibiting bubble behavior, $P = \{p_i\}_{i=1}^{L}$, the first level JLS model approximates $Y = \ln P$:

$$y_i \approx \begin{cases} a + b(T-i)^m & : m \in (0,\infty) \\ \\ a + b\ln(T-i) & : m = 0 \end{cases}$$

Note that we have extended the scope of the parameter m to include the value 0; the models specified for $m = 0$ are derived by a limiting argument applied to the optimized models for $m \neq 0$. See Proposition 10 below for the details of this limit argument. Given m, there are unique parameters $a^*(m)$ and $b^*(m)$ that minimize the sum of squared errors. We denote this optimized first level JLS model as

$$F_T^*(m,i) := \begin{cases} a^*(m) + b^*(m)(T-i)^m & : m \in (0,\infty) \\ \\ a^*(0) + b^*(0)\ln(T-i) & : m = 0 \end{cases}$$

Throughout this chapter, we will often use the interpolation $x \mapsto a^*(m) + b^*(m)(T-x)^m$ of $F_T^*(m,i)$. We denote this continuous interpolation as $F_T^*(m,x)$. As noted in [20], even under strong assumptions such that Y is concave up and/or strictly increasing, it is not clear that optimization over m is unique. In this chapter, we provide five theorems: three limiting the interval of relevant values for m and two completely identifying the optimal m for two "extreme" vectors of log-price vectors.

Because the vector $\{(T-i)^m\}_{i=1}^{L}$ and the constant vector $\mathbf{1}$ are linearly independent in $\mathbb{R}^L$ (and the same applies to $\mathbf{1}$ and $\{\ln(T-i)\}_{i=1}^{L}$), we may apply the Gram-Schmidt orthogonalization process starting with $\mathbf{1}$ to obtain a pair of orthonormal vectors $Z_1 = \frac{1}{\sqrt{L}} \cdot \mathbf{1}$ and $Z_2(m)$ that generate the same vector subspace, where

$$Z_2(m) = \begin{cases} \frac{AX(m)}{\|AX(m)\|} & : m > 0 \\ \\ \frac{AW_1}{\|AW_1\|} & : m = 0 \end{cases}$$

with

$$X(m) = \{(T-i)^m\}_{i=1}^{L}, \qquad W_k = \{\ln^k(T-i)\}_{i=1}^{L}, \qquad A_{i,j} = \begin{cases} \frac{L-1}{L} & \text{if } i = j \\ \\ \frac{-1}{L} & \text{if } i \neq j \end{cases}$$

The role of the matrix A is that, for all vectors $V \in \mathbb{R}^L$, AV is the projection of V onto the orthogonal complement of the subspace of constant vectors. There is a unique closest vector $\widehat{Y}$ to Y in the subspace generated by $\mathbf{1}$ and $Z_2(m)$ given by

$$\widehat{Y} = (Y \cdot Z_1)Z_1 + (Y \cdot Z_2(m))Z_2(m)$$

Since the least sum of squared errors is equal to $\|Y - \widehat{Y}\|^2 = \|Y\|^2 - (Y \cdot Z_1)^2 - (Y \cdot Z_2(m))^2$, where Z_1 does not depend on m, we know that minimizing the sum of squared errors is equivalent to maximizing

$$\phi(m) := (Y \cdot Z_2(m))^2$$

Because $\widehat{Y} = \{F_T^*(m,i)\}_{i=1}^L$, it follows that

$$b^*(m) = \begin{cases} \frac{Y \cdot AX}{AX \cdot X} & : m > 0 \\ \\ \frac{Y \cdot AW_1}{AW_1 \cdot W_1} & : m = 0 \end{cases} \tag{2.1}$$

and

$$a^*(m) = \begin{cases} \frac{1}{L}\left(\sum_{i=1}^L y_i - b^*(m)\sum_{i=1}^L (T-i)^m\right) & : m > 0 \\ \frac{1}{L}\left(\sum_{i=1}^L y_i - b^*(0)\sum_{i=1}^L \ln(T-i)\right) & : m = 0 \end{cases} \tag{2.2}$$

In this chapter, we often replace $X(m)$ with X, suppressing the independent variable. We will also use prime notation to denote the partial derivatives of X with respect to m. Additionally, we will often use $\gamma_i = T - i$ for computational purposes. Throughout this chapter, we restrict Y to being concave up and/or increasing with linear functions considered both concave up and concave down, and constant functions as increasing. By this we mean that, viewing Y as a function on the real line $i \mapsto y_i$, the piece-wise linear interpolating function of Y on $[1, L]$ has the corresponding property. These conditions on Y define a cone in $\mathbb{R}^L$ which can be described as linear combinations of a finite set of extreme examples. For concave up increasing vectors, we chose these L extreme examples:

$$Y_1 = \{1\}_{i=1}^L \tag{2.3}$$

and for $k \in \{2, \cdots L\}$, set $Y_k = \{Y_k(i)\}_{i=1}^L$ where

$$Y_k(i) = \begin{cases} 0 & \text{if } k > i \\ \\ i + 1 - k & \text{if } k \leq i \end{cases} \tag{2.4}$$

In the following sections, we state and prove five theorems and provide corresponding proofs for necessary lemmas, propositions, and corollaries. Section 2.1 contains technical results that provide some needed details about how first level models interact. Some of these lemmas have helpful intuitive interpretations. In particular, Lemma 5 can be understood as saying that, as m increases over the interval $[0, 1]$, the first level models curve less in a global sense on the interval $[1, L]$ until becoming a linear function at $m = 1$.

2.1 Geometric Relations Between General First Level JLS Models

It should be emphasized we differentiate between the general first level JLS model and the optimized first level JLS model. The former allows for any real a and b for linear parameters while the latter uses the specific linear parameters, $a^*(m)$ and $b^*(m)$. In this section, we consider the general first level JLS model:

$$f_T(m, x) = a + b(T - x)^m, \quad \text{for all real } x \in [1, L], \quad \text{for } m > 0$$

with $f_T(0, x) = a + b\ln(T - x)$. Suppose we have fixed $x_1 < x_2$ in $[1, L]$ and real numbers y_1 and y_2. For every $m \geq 0$ we choose $a = a(m)$ and $b = b(m))$ so that $f_T(m, x_1) = y_1$ and $f_T(m, x_2) = y_2$. Then for $m > 0$,

$$\begin{bmatrix} a(m) \\ b(m) \end{bmatrix} = \frac{1}{(T - x_2)^m - (T - x_1)^m} \begin{bmatrix} y_1(T - x_2)^m - y_2(T - x_1)^m \\ -y_1 + y_2 \end{bmatrix}$$

Thus,

$$a(m) = \frac{y_2(T - x_1)^m - y_1(T - x_2)^m}{(T - x_1)^m - (T - x_2)^m} \tag{2.5}$$

and

$$b(m) = \frac{y_1 - y_2}{(T - x_1)^m - (T - x_2)^m} \tag{2.6}$$

Because the vector $\{\ln(T - i)\}_{i=1}^{L}$ is strictly monotonic and $L \geq 3$, there are unique coefficients $a(0)$ and $b(0)$ such that,

$$f_T(0, x_1) = y_1 \quad \text{and} \quad f_T(0, x_2) = y_2$$

The coefficients $b(0)$ and $a(0)$ are

$$b(0) = \frac{y_2 - y_1}{\ln(T - x_2) - \ln(T - x_1)} \quad \text{and} \quad a(0) = \frac{y_1 \ln(T - x_2) - y_2 \ln(T - x_1)}{\ln(T - x_2) - \ln(T - x_1)}$$

Consequently,

$$f_T(0, x) = \frac{y_2 \ln \frac{T - x_1}{T - x} - y_1 \ln \frac{T - x_2}{T - x}}{\ln \frac{T - x_1}{T - x_2}}$$

The next proposition shows that $f_T(0,x)$ is continuously consistent with the analogous functions $f_T(m,x)$ for $m > 0$ that also pass through the points (x_1, y_1) and (x_2, y_2).

Proposition 1. *Suppose we have real numbers $x_1 < x_2$ in $[1, L]$ and have specified $a(m)$ and $b(m)$ so that $f_T(m, x_i) = y_i$. Then for all real $x \in [1, L]$,*

$$\lim_{m\to 0^+} f_T(m,x) = f_T(0,x) := \frac{y_2 \ln \frac{T-x_1}{T-x} - y_1 \ln \frac{T-x_2}{T-x}}{\ln \frac{T-x_1}{T-x_2}}$$

Proof. Let $\tau_i = T - x_i$ for $i = 1, 2$. Then, using Equations 2.5 and 2.6,

$$\begin{aligned}\lim_{m\to 0^+} f_T(m,x) &= \lim_{m\to 0^+} \frac{y_2\tau_1^m - y_1\tau_2^m + (y_1-y_2)(T-x)^m}{\tau_1^m - \tau_2^m} \\ &\overset{\ell\cdot H}{=} \lim_{m\to 0^+} \frac{y_2\tau_1^m \ln\tau_1 - y_1\tau_2^m \ln\tau_2 + (y_1-y_2)(T-x)^m \ln(T-x)}{\tau_1^m \ln\tau_1 - \tau_2^m \ln\tau_2} \\ &= \frac{y_2 \ln \frac{\tau_1}{T-x} - y_1 \ln \frac{\tau_2}{T-x}}{\ln \frac{\tau_1}{\tau_2}}\end{aligned}$$

□

Lemma 2. *Any two non-constant general first level JLS models with distinct non-linear parameters $m \in [0, 1]$ cross each other at most two times.*

Proof. We assume $T \geq L+1$. All general first level JLS models, h, have one of two forms: there are constants a and b such that, (1) for all $x \in [1, L]$, $h(x) = a + b(T-x)^m$ for some $m > 0$; or (2) for all $x \in [1, L]$, $h(x) = a + b\ln(T-x)$. For non-constant models, $b \neq 0$ and the models are strictly increasing on $[1, L]$ or strictly decreasing on $[1, L]$. So if two non-constant models cross, say at (x_1, y_1) and (x_2, y_2) with $x_1 \neq x_2$, we know that $y_1 \neq y_2$. Suppose we have the assumptions outlined at the beginning of Section 2.1 and WOLG, assume $y_1 < y_2$. Given $m_1 < m_2 \in [0, 1]$, we claim x_1 and x_2 are the only cross points between $f_T(m_1, x)$ and $f_T(m_2, x)$. Suppose $\exists x_3 \notin \{x_1, x_2\}$ such that $f_T(m_1, x_3) = f_T(m_2, x_3)$. Let $h(x) = f_T(m_1, x) - f_T(m_2, x)$. Then $h(x)$ has at least three distinct zeros. By Rolle's Theorem, $\frac{\partial}{\partial x}h(x)$ has at least two distinct zeros. Thus, $\exists x_4 \neq x_5$ such that $\frac{\partial}{\partial x}f_T(m_1, x_i) = \frac{\partial}{\partial x}f_T(m_2, x_i)$ for $i = 4, 5$. If $m_1 \neq 0$, this implies that

$$\frac{m_1 b(m_1)}{m_2 b(m_2)} = (T - x_i)^{m_2 - m_1}, \quad i = 4, 5 \tag{2.7}$$

If $m_1 = 0$, then

$$\frac{\frac{y_1 - y_2}{\ln \frac{T-x_1}{T-x_2}}}{m_2 b(m_2)} = (T - x_i)^{m_2}, \quad i = 4, 5 \tag{2.8}$$

By Equation 2.6, the LHS of Equations 2.7 and 2.8 are positive. Furthermore, both are completely independent of the choice of x so they are the same at x_4 and x_5. However, for $x < T$, the mapping, $x \mapsto (T-x)^{m_2-m_1}$ is one-to-one on $(-\infty, T)$, and in particular, $[1, L]$. Thus, the RHS of Equations 2.7 and 2.8 are both different at x_4 and x_5 $\Rightarrow\Leftarrow$ Therefore, $f_T(m_1, x)$ and $f_T(m_2, x)$ cross exactly two times. □

Proposition 3. *Suppose we have real numbers $x_1 < x_2$ in $[1, L]$ and have specified $a(m)$ and $b(m)$ so that $f_T(m, x_i) = y_i$. Then,*

$$\frac{\partial}{\partial x} f_T(0, x) = \frac{y_2 - y_1}{(T - x)\ln \frac{T - x_1}{T - x_2}}$$

Furthermore, $\frac{\partial}{\partial x} f_T(m, x)$ is right-continuous at $m = 0$ for every $x < T$.

Proof. Since $f_T(m, x)$ is right-continuous at $m = 0$ by Proposition 1, then

$$\begin{aligned}\frac{\partial}{\partial x} f_T(0, x) &= \frac{\partial}{\partial x} \frac{y_2 \ln \frac{T - x_1}{T - x} - y_1 \ln \frac{T - x_2}{T - x}}{\ln \frac{T - x_1}{T - x_2}} \\ &= \frac{y_2 - y_1}{(T - x)\ln \frac{T - x_1}{T - x_2}}\end{aligned}$$

For $x < T$, we have right continuity since,

$$\begin{aligned}\lim_{m \to 0^+} \frac{\partial}{\partial x} f_T(m, x) &= \lim_{m \to 0^+} -mb(m)(T - x)^{m-1} \\ &= \frac{(y_2 - y_1)}{T - x} \lim_{m \to 0^+} \frac{m}{(T - x_1)^m - (T - x_2)^m} \\ &\overset{\ell\cdot H}{=} \frac{(y_2 - y_1)}{T - x} \lim_{m \to 0^+} \frac{1}{(T - x_1)^m \ln(T - x_1) - (T - x_2)^m \ln(T - x_2)} \\ &= \frac{y_2 - y_1}{(T - x)\ln \frac{T - x_1}{T - x_2}}\end{aligned}$$

□

Lemma 4. *Let $k \geq 2$. Then for $x, y \in \mathbb{R}$,*

$$x^k - kxy^{k-1} + (k - 1)y^k = (x - y)^2 \sum_{p=0}^{k-2} (p + 1)x^{k-2-p} y^p \tag{2.9}$$

Proof. Since $(x - y)^2 = x^2 - 2xy + y^2$, we can write the RHS of Equation 2.9 as,

$$\sum_{p=0}^{k-2} (p + 1)x^{k-p} y^p - 2 \sum_{p=0}^{k-2} (p + 1)x^{k-1-p} y^{p+1} + \sum_{p=0}^{k-2} (p + 1)x^{k-2-p} y^{p+2} \tag{2.10}$$

We would like to write Equation 2.10 using a single summation. Each of the three summations can be expressed as,

$$\sum_{p=0}^{k-2} (p + 1)x^{k-p} y^p = x^k + 2x^{k-1}y + \sum_{p=2}^{k-2} (p + 1)x^{k-p} y^p \tag{2.11}$$

$$-2 \sum_{p=0}^{k-2} (p + 1)x^{k-1-p} y^{p+1} = -2x^{k-1}y - 2 \sum_{p=2}^{k-2} px^{k-p} y^p - 2(k - 1)xy^{k-1} \tag{2.12}$$

$$\sum_{p=0}^{k-2}(p+1)x^{k-2-p}y^{p+2} = \sum_{p=2}^{k-2}(p-1)x^{k-p}y^{p} + (k-2)xy^{k-1} + (k-1)y^{k} \tag{2.13}$$

Adding up the RHS of Equations 2.11, 2.12, 2.13, we have

$$\begin{aligned}
&x^k + (-2k+2+k-2)xy^{k-1} + (k-1)y^k + \sum_{p=2}^{k-2}(p+1-2p+p-1)x^{k-p}y^p \\
&= x^k - kxy^{k-1} + (k-1)y^k
\end{aligned}$$

□

Lemma 5. *Suppose we have real numbers $x_1 < x_2$ in $[1, L]$ and have specified $a(m)$ and $b(m)$ so that $f_T(m, x_i) = y_i$ with $y_1 < y_2$. Then*

$$f_T(m_1, x) > f_T(m_2, x), \quad x \in [1, x_1) \cup (x_2, L]$$

and

$$f_T(m_1, x) < f_T(m_2, x), \quad x \in (x_1, x_2)$$

Proof. By Lemma 2, $f_T(m_1, x)$ and $f_T(m_2, x)$ cross each other exactly twice, at x_1 and x_2. For every $x_0 \in [1, L]$, let $g_{x_0}(m) = \frac{\partial}{\partial x} f_T(m, x_0)$. Let $\tau_i = (T - x_i)$, $i = 1, 2$. For $m > 0$, $\frac{\partial}{\partial m} b(m) = b(m) \frac{\tau_2^m \ln \tau_2 - \tau_1^m \ln \tau_1}{\tau_1^m - \tau_2^m}$. So for $m > 0$,

$$\begin{aligned}
\frac{\partial}{\partial m} g_{x_0}(m) &= \frac{\partial}{\partial m}\left[-mb(m)(T-x_0)^{m-1}\right] \\
&= -b(m)(T-x_0)^{m-1} - m\left(\frac{\partial}{\partial m}b(m)\right)(T-x_0)^{m-1} \\
&\quad - mb(m)(T-x_0)^{m-1}\ln(T-x_0) \\
&= -b(m)(T-x_0)^{m-1}\left[1 + m\frac{\tau_2^m \ln \tau_2 - \tau_1^m \ln \tau_1}{\tau_1^m - \tau_2^m} + m\ln(T-x_0)\right] \\
&= -b(m)(T-x_0)^{m-1}\left[\frac{\tau_1^m - \tau_2^m + m\tau_1^m \ln\frac{T-x_0}{\tau_1} - m\tau_2^m \ln\frac{T-x_0}{\tau_2}}{\tau_1^m - \tau_2^m}\right] \\
&= -b(m)(T-x_0)^{m-1}\left[\frac{\tau_1^m\left[1 + m\ln\frac{T-x_0}{\tau_1}\right] - \tau_2^m\left[1 + m\ln\frac{T-x_0}{\tau_2}\right]}{\tau_1^m - \tau_2^m}\right]
\end{aligned} \tag{2.14}$$

Let $N(m, x) := \tau_1^m\left[1 + m\ln\frac{T-x}{\tau_1}\right] - \tau_2^m\left[1 + m\ln\frac{T-x}{\tau_2}\right]$. We now make the comparable calculations for the special case, $m = 0$. By Proposition 3 $g_{x_0}(m)$ is right continuous at $m = 0$. Therefore, $\frac{\partial}{\partial m} g_{x_0}(0) = \lim_{m\to 0^+} \frac{\partial}{\partial m} g_{x_0}(m)$, where

$$
\begin{aligned}
&\lim_{m\to 0^+} \frac{T-x_0}{y_2-y_1}\cdot\frac{\partial}{\partial m}g_{x_0}(m)\\
&= \lim_{m\to 0^+} \frac{\tau_1^m\left[1+m\ln\frac{T-x_0}{\tau_1}\right]-\tau_2^m\left[1+m\ln\frac{T-x_0}{\tau_2}\right]}{(\tau_1^m-\tau_2^m)^2}\\
&\overset{\ell\cdot H}{=} \lim_{m\to 0^+} \frac{\tau_1^m\ln\tau_1\left[1+m\ln\frac{T-x_0}{\tau_1}\right]+\tau_1^m\ln\frac{T-x_0}{\tau_1}-\tau_2^m\ln\tau_2\left[1+m\ln\frac{T-x_0}{\tau_2}\right]-\tau_2^m\ln\frac{T-x_0}{\tau_2}}{2(\tau_1^m-\tau_2^m)(\tau_1^m\ln\tau_1-\tau_2^m\ln\tau_2)}\\
&\overset{\ell\cdot H}{=} \lim_{m\to 0^+} \frac{\tau_1^m\ln^2\tau_1\left[1+m\ln\frac{T-x_0}{\tau_1}\right]+2\tau_1^m\ln\tau_1\ln\frac{T-x_0}{\tau_1}}{2[(\tau_1^m\ln\tau_1-\tau_2^m\ln\tau_2)^2+(\tau_1^m-\tau_2^m)(\tau_1^m\ln^2\tau_1-\tau_2^m\ln^2\tau_2)]}\\
&\qquad + \frac{-\tau_2^m\ln^2\tau_2\left[1+m\ln\frac{T-x_0}{\tau_2}\right]-2\tau_2^m\ln\tau_2\ln\frac{T-x_0}{\tau_2}}{2[(\tau_1^m\ln\tau_1-\tau_2^m\ln\tau_2)^2+(\tau_1^m-\tau_2^m)(\tau_1^m\ln^2\tau_1-\tau_2^m\ln^2\tau_2)]}\\
&= \frac{\ln^2\tau_1+2\ln\tau_1\ln\frac{T-x_0}{\tau_1}-\ln^2\tau_2-2\ln\tau_2\ln\frac{T-x_0}{\tau_2}}{2\ln^2\frac{\tau_1}{\tau_2}}
\end{aligned}
$$

For $m>0$, since $x_0<T$ and $b(m)<0$, we have $-b(m)(T-x_0)^{m-1}>0$. Consider the bracketed factor in the last line of Equation 2.14. Since $T-x_1>T-x_2$, the denominator is positive for $m>0$ and we can just consider $N(m,x_0)$ since the sign of $\frac{\partial}{\partial m}g_{x_0}(m)$ and $N(m,x_0)$ are the same. Using the Taylor series representations for $\tau_i^m=(T-x_i)^m$, $i=1,2$ as functions of m centered at 0 we have $\tau_i^m=\sum_{k=0}^{\infty}\frac{\ln^k\tau_i}{k!}m^k$, $i=1,2$. Then,

$$
\begin{aligned}
&\tau_1^m\left[1+m\ln\frac{T-x_0}{\tau_1}\right]-\tau_2^m\left[1+m\ln\frac{T-x_0}{\tau_2}\right]\\
&= \left[1+m\ln\frac{T-x_0}{\tau_1}\right]\sum_{k=0}^{\infty}\frac{\ln^k\tau_1}{k!}m^k-\left[1+m\ln\frac{T-x_0}{\tau_2}\right]\sum_{k=0}^{\infty}\frac{\ln^k\tau_2}{k!}m^k\\
&= \sum_{k=0}^{\infty}\frac{\ln^k\tau_1}{k!}m^k+\sum_{k=0}^{\infty}\frac{\ln\frac{T-x_0}{\tau_1}\ln^k\tau_1}{k!}m^{k+1}-\sum_{k=0}^{\infty}\frac{\ln^k\tau_2}{k!}m^k-\sum_{k=0}^{\infty}\frac{\ln\frac{T-x_0}{\tau_2}\ln^k\tau_2}{k!}m^{k+1}\\
&= \sum_{k=2}^{\infty}\frac{\ln^k\tau_1+k\ln\frac{T-x_0}{\tau_1}\ln^{k-1}\tau_1-\ln^k\tau_2-k\ln\frac{T-x_0}{\tau_2}\ln^{k-1}\tau_2}{k!}m^k\\
&= \sum_{k=2}^{\infty}\frac{-(k-1)\ln^k\tau_1+k\ln(T-x_0)\ln^{k-1}\tau_1+(k-1)\ln^k\tau_2-k\ln(T-x_0)\ln^{k-1}\tau_2}{k!}m^k
\end{aligned}
$$

Since $m>0$, this series will be positive (negative), and thus, $\frac{\partial}{\partial m}g_{x_0}(m)$ will be positive (negative), if for all $k\geq 2$,

$$
\nu_{x_0}(k) := -(k-1)\ln^k\tau_1+k\ln(T-x_0)\ln^{k-1}\tau_1+(k-1)\ln^k\tau_2-k\ln(T-x_0)\ln^{k-1}\tau_2>0\,(<0)
$$

Let $x_0 = x_1$. For $k \geq 2$, we have

$$\begin{aligned}\nu_{x_1}(k) &= -(k-1)\ln^k \tau_1 + k \ln \tau_1 \ln^{k-1} \tau_1 + (k-1)\ln^k \tau_2 - k \ln \tau_1 \ln^{k-1} \tau_2 \\ &= \ln^k \tau_1 - k \ln \tau_1 \ln^{k-1} \tau_2 + (k-1)\ln^k \tau_2\end{aligned}$$

Since $\ln \tau_1 > 0$ and $\ln \tau_2 \geq 0$, by Lemma 4 (with $x = \ln \tau_1$, $y = \ln \tau_2$) we have $\nu_{x_1}(k) > 0$, $\forall k \geq 2$. Thus, $\frac{\partial}{\partial m} g_{x_1}(m) > 0$, for $m > 0$.

Now, let $x_0 = x_2$. For $k \geq 2$, we have

$$\begin{aligned}\nu_{x_2}(k) &= -(k-1)\ln^k \tau_1 + k \ln \tau_2 \ln^{k-1} \tau_1 + (k-1)\ln^k \tau_2 - k \ln \tau_2 \ln^{k-1} \tau_2 \\ &= -(k-1)\ln^k \tau_1 + k \ln \tau_2 \ln^{k-1} \tau_1 - \ln^k \tau_2\end{aligned}$$

Since $\ln \tau_1 > 0$ and $\ln \tau_2 \geq 0$, then by Lemma 4 (with $x = \ln \tau_2$, $y = \ln \tau_1$) we have $\nu_{x_2}(k) < 0$, $\forall k \geq 2$. Thus, $\frac{\partial}{\partial m} g_{x_2}(m) < 0$, for $m > 0$.

By Proposition 3, $\frac{\partial}{\partial x} f_T(m, x)$ is right-continuous at $m = 0$. Because, for $m > 0$, we have $\frac{\partial}{\partial m} g_{x_1}(m) > 0$ and $\frac{\partial}{\partial m} g_{x_2}(m) < 0$, then for all $m_1 < m_2 \in [0, 1]$,

$$\frac{\partial}{\partial x} f_T(m_1, x_1) < \frac{\partial}{\partial x} f_T(m_2, x_1) \quad \& \quad \frac{\partial}{\partial x} f_T(m_1, x_2) > \frac{\partial}{\partial x} f_T(m_2, x_2)$$

Finally, since there are exactly two cross points (at x_1 and x_2) then,

$$f_T(m_1, x) > f_T(m_2, x), \quad x \in [1, x_1) \cup (x_2, L]$$

and

$$f_T(m_1, x) < f_T(m_2, x), \quad x \in (x_1, x_2)$$

□

Proposition 6. *Suppose we have real numbers $x_1 < x_2$ in $[1, L]$ and have specified $a(m)$ and $b(m)$ so that $f_T(m, x_i) = y_i$. Then,*

$$\frac{\partial^2}{\partial x^2} f_T(0, x) = \frac{y_2 - y_1}{(T - x)^2 \ln \frac{T - x_1}{T - x_2}}$$

Furthermore, $\frac{\partial^2}{\partial x^2} f_T(0, x)$ is right-continuous at $m = 0$ for every $x < T$.

Proof. By Proposition 3,

$$\frac{\partial^2}{\partial x^2} f_T(0,x) = \frac{\partial}{\partial x} \frac{y_2 - y_1}{(T-x)\ln\frac{T-x_1}{T-x_2}}$$
$$= \frac{y_2 - y_1}{(T-x)^2\ln\frac{T-x_1}{T-x_2}}$$

Note $\frac{\partial^2}{\partial x^2} f_T(m,x) = (1-m)(T-x)^{-1}\frac{\partial}{\partial x} f_T(m,x)$. For $x < T$, we have right continuity since,

$$\lim_{m\to 0^+} \frac{\partial^2}{\partial x^2} f_T(m,x) = \lim_{m\to 0^+} (1-m)(T-x)^{-1}\frac{\partial}{\partial x} f_T(m,x)$$
$$= \lim_{m\to 0^+} (1-m)(T-x)^{-1} \cdot \lim_{m\to 0^+} \frac{\partial}{\partial x} f_T(m,x)$$
$$= (T-x)^{-1} \cdot \frac{\partial}{\partial x} f_T(0,x)$$
$$= \frac{y_2 - y_1}{(T-x)^2\ln\frac{T-x_1}{T-x_2}}$$

□

Proposition 7. *Suppose we have real numbers $x_1 < x_2$ in $[1, L]$ and have specified $a(m)$ and $b(m)$ so that $f_T(m, x_i) = y_i$ with $y_1 < y_2$. Then*

$$\frac{\partial}{\partial x} f_T(m,x) > 0, \quad \forall x < T, \quad \forall m \in [0,1]$$

Proof. Let $m \in (0,1]$. Since $y_1 < y_2$, then $b(m) < 0$ by Equation 2.6. Since $T - x > 0$, $\forall x$, then $\frac{\partial}{\partial x} f_T(m,x) = -mb(m)(T-x)^{m-1} > 0$, $\forall x < T$. Let $m = 0$. By Proposition 3, $\frac{\partial}{\partial x} f_T(0,x) := \frac{y_2-y_1}{(T-x)\ln\frac{T-x_1}{T-x_2}}$. Because $T - x_1 > T - x_2$, $\ln\frac{T-x_1}{T-x_2} > 0$. Consequently, $\frac{\partial}{\partial x} f_T(0,x) > 0$, $\forall x < T$. □

Proposition 8. *Suppose we have real numbers $x_1 < x_2$ in $[1, L]$ and have specified $a(m)$ and $b(m)$ so that $f_T(m, x_i) = y_i$ with $y_1 < y_2$. Then*

$$\frac{\partial^2}{\partial x^2} f_T(m,x) \geq 0, \quad \forall x < T, \quad \forall m \in [0,1]$$

with equality if and only if $m = 1$.

Proof. Let $m \in (0,1)$. Since $y_1 < y_2$, then $b(m) < 0$ by Equation 2.6. Since $T - x > 0$, $\forall x < T$, then $\frac{\partial^2}{\partial x^2} f_T(m,x) = -(1-m)mb(m)(T-x)^{m-2} > 0$, $\forall x < T$. Let $m = 0$. By Proposition 6, $\frac{\partial^2}{\partial x^2} f_T(0,x) := \frac{y_2-y_1}{(T-x)^2\ln\frac{T-x_1}{T-x_2}}$. Because $T - x_1 > T - x_2$ the factor $\ln\frac{T-x_1}{T-x_2} > 0$. Because $T - x > 0$, $\forall x < T$ and $y_1 < y_2$, $\frac{\partial^2}{\partial x^2} f_T(0,x) > 0$. Let $m = 1$. Then $\frac{\partial^2}{\partial x^2} f_T(1,x) = 0$. Conversely, suppose $\frac{\partial^2}{\partial x^2} f_T(m,x) = -(1-m)mb(m)(T-x)^{m-2} = 0$, $m \in (0,1]$. Because $b(m) < 0$ and $(T-x)^m > 0$, $1 - m = 0 \iff m = 1$. □

2.2 Extreme Vector Decomposition

We are interested in the extreme vectors that are specified by Equations 2.3 and 2.4 because any non-decreasing, concave up vector in $\mathbb{R}^L$ can be written as a linear combination of these extreme vectors, say $Y = \sum_{k=1}^{L} \delta_k Y_k$ with $\delta_k \geq 0$ for $k \geq 2$.

Proposition 9. *Let Y be an increasing, concave up vector of length L. Then Y can be uniquely expressed as a linear combination of the extreme vectors, $\{Y_k\}_{k=1}^{L}$, say $Y = \sum_{k=1}^{L} \delta_k Y_k$, with $\delta_k \geq 0$ for $k \geq 2$.*

Proof. Let $Y = \{y_j\}_{j=1}^{L}$ be an increasing, concave up vector. By hypothesis, we assume $y_{j-1} \leq y_j$ for $j \in \{2, \cdots, L\}$ and $y_{j-1} - y_{j-2} \leq y_j - y_{j-1}$ for $j \in \{3, \cdots, L\}$. Set $\delta_1 = y_1$, $\delta_2 = y_2 - y_1$, and $\delta_k = y_k - 2y_{k-1} + y_{k-2}$ for integers $k \in [3, L]$. We shall prove that $Y = \sum_{k=1}^{L} \delta_k Y_k$ by induction on the index number for each coordinate. Let $y_{k,j}$ refers to the j-th component of extreme vector Y_k. We check first $j = 1$. Then

$$y_1 = \delta_1 Y_{1,1} = y_1 \cdot 1 = y_1$$

For an induction hypothesis, assume $y_j = \sum_{k=1}^{L} \delta_k Y_{k,j}$ for some $j < L$. Consider y_{j+1}. Then,

$$\begin{aligned} \sum_{k=1}^{L} \delta_k y_{k,j+1} &= \delta_1 \cdot y_{1,j+1} + \sum_{k=2}^{j+1} \delta_k (j+2-k) + \sum_{k=j+2}^{L} \delta_k \cdot 0 \\ &= y_1 + \sum_{k=2}^{j} \delta_k (j+1-k) + \sum_{k=2}^{j+1} \delta_k \\ &= y_1 + \sum_{k=2}^{j} \delta_k y_{k,j} + \left[y_2 - y_1 + \sum_{k=3}^{j+1} (y_k - 2y_{k-1} + y_{k-2}) \right] \end{aligned} \tag{2.15}$$

By the induction hypothesis, the left two terms in the last line of Equation 2.15 add to y_j. For $j = 1$, the last term equals $y_2 - y_1$; when that is added to the first two terms, we obtain y_2 as desired. For $j = 2$, the last term equals

$$y_2 - y_1 + y_3 - 2y_2 + y_1 = y_3 - y_2$$

Thus, when all three terms are summed, we obtain y_3 as desired. For $j \geq 3$, the last term telescopes a bit:

$$y_2 - y_1 + \sum_{k=3}^{j+1} (y_k - 2y_{k-1} + y_{k-2}) = y_{j+1} - y_j$$

Consequently, when all three terms are summed, we obtain y_{j+1} as desired. $\square$

2.3 Limiting the Intervals of m

Proposition 10. *For any vector $Y \in \mathbb{R}^L$, we can define*

$$F_T^*(0,x) := \lim_{m\to 0^+} F_T^*(m,x) = \frac{1}{L}\sum_{i=1}^{L} y_i + \frac{Y\cdot W_1}{AW_1\cdot W_1}\left(\ln(T-x) - \frac{1}{L}\sum_{j=1}^{L}\ln(T-j)\right)$$

Proof. Let $\gamma_i = T - i$ for $i \in \{1, \cdots, L\}$. Let $G(m,x) = (T-x)^m - \frac{1}{L}\sum_{j=1}^{L}(T-j)^m$. Note that $\lim_{m\to 0^+} G(m,x) = 0$, $\lim_{m\to 0^+} \frac{\partial}{\partial m}G(m,x) = \ln(T-x) - \frac{1}{L}\sum_{i=j}^{L}\ln(T-j)$, and $\lim_{m\to 0^+}\frac{\partial^2}{\partial m^2}G(m,x) = \ln^2(T-x) - \frac{1}{L}\sum_{j=1}^{L}\ln^2(T-j)$. Then

$$\begin{aligned}
&\lim_{m\to 0^+} F_T^*(m,x)\\
&= \lim_{m\to 0^+}\left[\frac{1}{L}\sum_{i=1}^{L} y_i + \frac{Y\cdot AX}{AX\cdot X}G(m,x)\right]\\
&\overset{\ell\cdot H}{=} \frac{1}{L}\sum_{i=1}^{L} y_i + \lim_{m\to 0^+}\frac{(Y\cdot AX')G(m,x) + (Y\cdot AX)G'(m,x)}{2(AX\cdot X')}\\
&\overset{\ell\cdot H}{=} \frac{1}{L}\sum_{i=1}^{L} y_i + \lim_{m\to 0^+}\frac{(Y\cdot AX'')G(m,x) + 2(Y\cdot AX')G'(m,x) + (Y\cdot AX)G''(m,x)}{2[(AX'\cdot X') + (AX\cdot AX'')]}\\
&= \frac{1}{L}\sum_{i=1}^{L} y_i + \frac{Y\cdot AW_1}{AW_1\cdot W_1}\left(\ln(T-x) - \frac{1}{L}\sum_{j=1}^{L}\ln(T-j)\right)
\end{aligned}$$

□

Remark 11. *By the Gram-Schmidt orthogonalization process, the closest vector to Y, say $\widehat{Y}$, in the linear span of $\mathbf{1}$ and W_1 is given by*

$$\widehat{Y} = \frac{1}{L}(Y\cdot\mathbf{1})\mathbf{1} + (Y\cdot Z_2(0))Z_2(0)$$

where $Z_2(0) = \frac{1}{\|AW_1\|}AW_1$ as specified at the beginning of Chapter 2. So for each integer, $i \in [1, L]$, we have

$$\widehat{Y}_i = \frac{1}{L}\sum_{j=1}^{L} y_j + \frac{Y\cdot AW_1}{AW_1\cdot W_1}\left(\ln(T-i) - \sum_{j=1}^{L}\ln(T-j)\right)$$

So $F_T^*(0,x)$ as defined in Proposition 10 interpolates nicely $\widehat{Y}$. This supports the definition of $Z_2(0)$ that has been chosen; one can also show that $Z_2(0) = \lim_{m\to 0^+} Z_2(m)$.

Proposition 12. *For any vector $Y \in \mathbb{R}^L$, we have*

$$\frac{\partial}{\partial x}F_T^*(0,x) = \frac{-(Y\cdot AW_1)}{(T-x)(AW_1\cdot W_1)}$$

Furthermore, $\frac{\partial}{\partial x}F_T^*(m,x)$ is right-continuous at $m = 0$ for every $x < T$.

Proof. By Proposition 10,

$$\frac{\partial}{\partial x}F_T^*(0,x) = \frac{\partial}{\partial x}\left(\frac{1}{L}\sum_{i=1}^{L} y_i + \frac{Y\cdot W_1}{AW_1\cdot W_1}\left(\ln(T-x) - \frac{1}{L}\sum_{j=1}^{L}\ln(T-j)\right)\right)$$
$$= \frac{-(Y\cdot AW_1)}{(T-x)(AW_1\cdot W_1)}$$

Let $\tau_x = T - x$ for $x < T$. We have right-continuity since,

$$\lim_{m\to 0^+}\frac{\partial}{\partial x}F_T^*(m,x)$$
$$= \lim_{m\to 0^+} -mb^*(m)(T-x)^{m-1}$$
$$= \lim_{m\to 0^+}\frac{-m(Y\cdot AX)}{\tau_x^{1-m}(AX\cdot X)}$$
$$\overset{\ell\cdot H}{=} \lim_{m\to 0^+}\frac{-(Y\cdot AX) - m(Y\cdot AX')}{-\tau_x^{1-m}\ln\tau_x(AX\cdot X) + 2\tau_x^{1-m}(AX\cdot X')}$$
$$\overset{\ell\cdot H}{=} \frac{-2(Y\cdot AX') - m(Y\cdot AX'')}{\tau_x^{1-m}\ln^2\tau_x(AX\cdot X) - 4\tau_x^{1-m}\ln\tau_x(AX\cdot X') + 2\tau_x^{1-m}[(AX'\cdot X') + (AX\cdot X'')]}$$
$$= \frac{-(Y\cdot AW_1)}{\tau_x(AW_1\cdot W_1)}$$

□

Lemma 13. *Let $S_n = \{s_k\}_{k=1}^n$ be an increasing sequence and $R_n = \{r_k\}_{k=1}^n$ be an decreasing (increasing) sequence each with at least two distinct terms. Then $(AS\cdot AR) < 0\,(>0)$.*

Proof. For a base case, let $n = 2$. Suppose first that $AS = S$ and $AR = R$. Then $s_1 < 0 < s_2$, $r_1 > 0 > r_2$ $(r_1 < 0 < r_2)$ and

$$(AS\cdot AR) = \sum_{k=1}^{2} s_k r_k = s_1 r_1 + s_2 r_2 < 0\,(>0)$$

Now consider general R and S of length 2 meeting the hypothesis of the lemma. Because $A^2R = AR$ and $A^2S = AS$, we know that $(A^2R)\cdot(A^2S) < 0(>0)$ and hence $(AR)\cdot(AS) < 0(>0)$. We shall proceed with mathematical induction. Assume Lemma 13 holds for vectors of length $n-1 \geq 2$ and let $R, S \in \mathbb{R}^n$ satisfy the hypothesis of Lemma 13. Given $V \in \mathbb{R}^n$, we have $AV = A(V + c\mathbf{1})$ for all constant vectors $c\mathbf{1}$ in $\mathbb{R}^n$. We may therefore assume that $\sum_{j=1}^{n-1} r_j = \sum_{j=1}^{n-1} s_j = 0$. For the inductive step, let $n > 2$. Then

$s_n > 0$, $\widehat{s} := \frac{1}{n}\sum_{k=1}^{n} s_k = \frac{s_n}{n}$, $r_n < 0\,(> 0)$, $\widehat{r} := \frac{1}{n}\sum_{k=1}^{n} r_k = \frac{r_n}{n}$. Then

$$\begin{aligned}
(AR \cdot AS) &= \sum_{k=1}^{n}(s_k - \widehat{s})(r_k - \widehat{r}) \\
&= \sum_{k=1}^{n} s_k r_k - \widehat{r}\sum_{k=1}^{n} s_k - \widehat{s}\sum_{k=1}^{n} r_k + n\widehat{s}\widehat{r} \\
&= \sum_{k=1}^{n} s_k r_k - n\widehat{s}\widehat{r} \\
&= \sum_{k=1}^{n-1} s_k r_k + \left(1 - \frac{1}{n}\right) s_n r_n
\end{aligned} \tag{2.16}$$

In the last line of Equation 2.16, the first term is negative (positive) by the inductive hypothesis, or zero. Since $1 - \frac{1}{n} > 0$ then the second term is negative (positive). Thus, we have the desired result. ☐

Corollary 14. *Let Y be non-constant and increasing. Then $b^*(m) < 0$, $\forall m \in [0, \infty)$.*

Proof. Solving for the linear parameter $b^*(m)$ as in Equation 2.1, we have

$$b^*(m) = \begin{cases} \frac{Y \cdot AX}{AX \cdot X} & : m > 0 \\ \\ \frac{Y \cdot AW_1}{AW_1 \cdot W_1} & : m = 0 \end{cases}$$

Since Y is non-constant and increasing, for $m > 0$, $Y \cdot AX < 0$ and $AX \cdot X > 0$ and for $m = 0$, $Y \cdot AW_1 < 0$ and $AW_1 \cdot W_1 > 0$ all by Lemma 13. Thus, $b^*(m) < 0$, $\forall m \in [0, \infty)$. ☐

Corollary 15. *Suppose $T \geq L+1 \geq 4$. Let $Y = \{y_i\}_{i=1}^{L}$ be non-constant and increasing with $\frac{1}{L}\sum_{i=1}^{L} y_i \geq 0$, then $a^*(m) > 0$, $\forall m \in [0, 1]$.*

Proof. Solving for the linear parameter $a^*(m)$ as in Equation 2.2, we have

$$a^*(m) = \begin{cases} \frac{1}{L}\left(\sum_{i=1}^{L} y_i - b^*(m)\sum_{i=1}^{L}(T-i)^m\right) & : m > 0 \\ \\ \frac{1}{L}\left(\sum_{i=1}^{L} y_i - b^*(0)\sum_{i=1}^{L}\ln(T-i)\right) & : m = 0 \end{cases}$$

Since Y is non-constant and increasing, then $b^*(m) < 0$, $\forall m \in [0, \infty)$, by Corollary 14. Because $T \geq L+1$ and $L \geq 3$, we know that $T - i \geq 1$ and thus $\ln(T-i) \geq 0$ for all integers $i \in [1, L]$. Clearly, $\sum_{i=1}^{L}(T-i)^m > 0$. For the case of $m = 0$, we have $T - 1 \geq L \geq 3$ and thus,

$$\sum_{i=1}^{L}\ln(T-i) \geq \ln(T-1) \geq \ln(3) > 0$$

Consequently, $a^*(m) > 0$ for $m \in [0, \infty)$. $\square$

Lemma 16. *Let $V = \{v_i\}_{i=1}^L$ be strictly monotonic, $L \geq 3$. Let $H = span(\{1\}_{i=1}^L, V)$. For any $Y = \{y_i\}_{i=1}^L \notin H$ and its corresponding unique closest $\widehat{Y} = \{\widehat{y}_i\}_{i=1}^L \in H$, $\exists i < j < k \in \{1, \cdots, L\}$ such that $(y_i - \widehat{y}_i)(y_j - \widehat{y}_j) < 0$ and $(y_j - \widehat{y}_j)(y_k - \widehat{y}_k) < 0$.*

Proof. Let $d(Y, H)$ be the distance from Y to the linear subspace, H. Because H is a closed, convex set in $\mathbb{R}^L$, then $\exists! \widehat{Y} \in H$ such that $d(Y, H) = d(Y, \widehat{Y})$ and $Y - \widehat{Y} \perp H$, leading to the normal equations:

$$Y \cdot \{1\}_{i=1}^L = \widehat{Y} \cdot \{1\}_{i=1}^L \tag{2.17}$$

$$Y \cdot V = \widehat{Y} \cdot V \tag{2.18}$$

Suppose $y_i \leq \widehat{y}_i$, $\forall i \in \{1, L\}$. Then

$$\sum_{i=1}^L y_i \leq \sum_{i=1}^L \widehat{y}_i$$

However, by Equation 2.17, $y_i = \widehat{y}_i$, $\forall i$ and $Y \in H \Rightarrow\Leftarrow$. Thus, $\exists r \in \{1, L\}$ such that $y_r > \widehat{y}_r$. A symmetric argument proves that $\exists s \in \{1, L\}$ such that $y_s < \widehat{y}_s$. Let $i_1 = \min(s, r)$ and $i_2 = \max(s, r)$. Then $(y_{i_1} - \widehat{y}_{i_1})(y_{i_2} - \widehat{y}_{i_2}) < 0$. Let i_1' be the largest integer $< i_2$ such that $(y_{i'} - \widehat{y}_{i'})(y_{i_2} - \widehat{y}_{i_2}) < 0$. Note, $i_1 \leq i_1'$. Let i_2' be the smallest integer $> i_1'$ such that $(y_{i_1'} - \widehat{y}_{i_1'})(y_{i_2'} - \widehat{y}_{i_2'}) < 0$. Note, $i_2' \leq i_2$. Suppose $(y_i - \widehat{y}_i)(y_{i_1'} - \widehat{y}_{i_1'}) \geq 0$, $\forall i < i_1'$ and $(y_i - \widehat{y}_i)(y_{i_2'} - \widehat{y}_{i_2'}) \geq 0$, $\forall i > i_2'$. Let $x_1 = i_1' + \frac{1}{2}$.

Since V is strictly monotonic, there is a function v_x for real $x \in [1, L]$ that is strictly monotonic and interpolates V at the integers of $[1, L]$. Let v_x be any such function. Then, either

$$v_i - v_{x_1} < 0, \quad i < x_1, \qquad \& \qquad v_i - v_{x_1} > 0, \quad x_1 < i \tag{2.19}$$

or

$$v_i - v_{x_1} > 0, \quad i < x_1 \qquad \& \qquad v_i - v_{x_1} < 0, \quad x_1 < i \tag{2.20}$$

By Equation 2.18:

$$\begin{aligned}0 &= \sum_{i=1}^{L} y_i v_i - \sum_{i=1}^{L} \widehat{y}_i v_i \\ &= \sum_{i=1}^{L} (y_i - \widehat{y}_i)\, v_i \\ &= \sum_{i=1}^{L} (y_i - \widehat{y}_i)\, v_i - \sum_{i=1}^{L} (y_i - \widehat{y}_i)\, v_{x_1} + \sum_{i=1}^{L} (y_i - \widehat{y}_i)\, v_{x_1} \\ &= \sum_{i=1}^{L} (y_i - \widehat{y}_i)\,(v_i - v_{x_1}) + 0 \\ &= \sum_{1 \le i < x_1} (y_i - \widehat{y}_i)\,(v_i - v_{x_1}) + \sum_{x_1 < i \le L} (y_i - \widehat{y}_i)\,(v_i - v_{x_1})\end{aligned} \tag{2.21}$$

By the supposition, either

$$y_i - \widehat{y}_i \le 0, \quad i < x_1 \qquad \& \qquad y_i - \widehat{y}_i \ge 0, \quad x_1 < i \tag{2.22}$$

or

$$y_i - \widehat{y}_i \ge 0, \quad i < x_1 \qquad \& \qquad y_i - \widehat{y}_i \le 0, \quad x_1 < i \tag{2.23}$$

On the RHS of the last line of Equation 2.21, since $i_1 < x_1$ and $i_2 > x_1$, then there exists at least one non-zero term in each summation. Furthermore, any of the four combinations of one of Equations 2.19 and 2.20 with one of Equations 2.22 and 2.23 force every non-zero term to have the same sign. $\Rightarrow\Leftarrow$. Thus, $\exists i < i_1'$ such that $(y_i - \widehat{y}_i)(y_{i_1'} - \widehat{y}_{i_1'}) < 0$ or $\exists i > i_2'$ such that $(y_i - \widehat{y}_i)(y_{i_2'} - \widehat{y}_{i_2'}) < 0$. □

Theorem 17. *Given a non-constant increasing concave up vector, Y, the optimal m is in $[0, 1]$.*

Proof. Let $m > 1$ and Y be a non-constant increasing concave up vector. The vector $X(m)$ is both strictly decreasing and strictly concave up. By Corollary 14, the coefficient, $b^*(m)$, of $X(m)$ for the optimized first level JLS model fit is negative. By Lemma 16, Y and $F_T^*(m, x)$ cross each other at least two times. Let two of these points be labeled (s_1, r_1) and (s_2, r_2) with $s_1 < s_2$ and $r_1 < r_2$. Note $r_1 < r_2$ since $b^*(m) < 0$ and thus $x \mapsto a^*(m) + b^*(m)(T - x)^m$ is strictly increasing and strictly concave down on $(-\infty, T]$. These two points define a linear function $Q(x)$. By standard properties of concavity, we have

$$F_T^*(m, x) < Q(x) \le Y(x), \quad x < s_1$$

$$F_T^*(m, x) > Q(x) \ge Y(x), \quad s_1 < x < s_2$$

$$F_T^*(m, x) < Q(x) \le Y(x), \quad s_2 < x$$

Thus, $\phi(1) > \phi(m)$. Moreover, this prohibits $F_T^*(m, x)$ and Y from crossing more than two times. □

2.4 Optimizing m

Proposition 18. *Let Y be a non-constant, increasing vector. Then $\frac{\partial}{\partial x}F_T^*(m,x) > 0$, $\forall x < T$, $\forall m \in [0,1]$.*

Proof. Let $x < T$. Let $m \in (0,1]$. By Corollary 14, $b^*(m) < 0$. Therefore, $\frac{\partial}{\partial x}F_T^*(m,x) = -mb^*(m)(T-x)^{m-1} > 0$. Let $m = 0$. By Proposition 12, $\frac{\partial}{\partial x}F_T^*(0,x) = \frac{-(Y\cdot AW_1)}{(T-x)(AW_1\cdot W_1)}$. By Lemma 13, $(Y\cdot AW_1) < 0$ and $(AW_1\cdot W_1) > 0$. Therefore, $\frac{\partial}{\partial x}F_T^*(0,x) > 0$. □

Proposition 19. *For any vector $Y \in \mathbb{R}^L$, we have*

$$\frac{\partial^2}{\partial x^2}F_T^*(0,x) = \frac{-(Y\cdot AW_1)}{(T-x)^2(AW_1\cdot W_1)}$$

Furthermore, $\frac{\partial^2}{\partial x^2}F_T^*(0,x)$ is right-continuous at $m = 0$ for every $x < T$.

Proof. By Proposition 12, we have

$$\begin{aligned}\frac{\partial^2}{\partial x^2}F_T^*(0,x) &= \frac{\partial}{\partial x}\frac{-(Y\cdot AW_1)}{(T-x)(AW_1\cdot W_1)}\\ &= \frac{-(Y\cdot AW_1)}{(T-x)^2(AW_1\cdot W_1)}\end{aligned}$$

Note $\frac{\partial^2}{\partial x^2}F_T^*(m,x) = (1-m)(T-x)^{-1}\frac{\partial}{\partial x}F_T^*(m,x)$, for $x < T$. We have right-continuity since,

$$\begin{aligned}\lim_{m\to 0^+}\frac{\partial^2}{\partial x^2}F_T^*(m,x) &= \lim_{m\to 0^+}(1-m)(T-x)^{-1}\frac{\partial}{\partial x}F_T^*(m,x)\\ &= \lim_{m\to 0^+}(1-m)(T-x)^{-1}\cdot\lim_{m\to 0^+}\frac{\partial}{\partial x}F_T^*(m,x)\\ &= (T-x)^{-1}\cdot\frac{\partial}{\partial x}F_T^*(0,x)\\ &= \frac{1}{(T-x)}\cdot\frac{-(Y\cdot AW_1)}{(T-x)(AW_1\cdot W_1)}\\ &= \frac{-(Y\cdot AW_1)}{(T-x)^2(AW_1\cdot W_1)}\end{aligned}$$

□

Proposition 20. *Let Y be a non-constant, increasing vector. For $m \in [0,1)$ and $x < T$ we have $\frac{\partial^2}{\partial x^2}F_T^*(m,x) > 0$ and thus $F_T^*(m,x)$ is strictly concave up on $(-\infty,T)$. However, for $m = 1$, $F_T^*(1,x)$ is a linear function whose second derivative is zero everywhere.*

Proof. Let $x < T$. Let $m \in (0,1)$. By Corollary 14, $b^*(m) < 0$. Therefore, $\frac{\partial^2}{\partial x^2}F_T^*(m,x) = -(1-m)mb^*(m)(T-x)^{m-2} > 0$. Let $m = 0$. By Proposition 19, $\frac{\partial^2}{\partial x^2}F_T^*(0,x) = \frac{-(Y\cdot AW_1)}{(T-x)^2(AW_1\cdot W_1)}$. By Lemma 13, $(Y\cdot AW_1) < 0$ and $(AW_1\cdot W_1) > 0$. Therefore, $\frac{\partial^2}{\partial x^2}F_T^*(0,x) > 0$. □

Lemma 21. *Let Y_k, $3 \le k \le L$, be an extreme vector. Then the linearly interpolated $Y_k(x)$ and any optimized first level JLS model must cross for at least two distinct $x \in (1, L)$ and no more than three distinct $x \in (1, L]$ or three distinct $x \in [1, L)$. Furthermore, they cross for at most one $x \in [1, k-1]$ and at most two distinct $x \in [k-1, L]$.*

Proof. In the context of the optimized first level JLS model, each extreme vector, Y_k, $3 \le k \le L$ is not in the span of the constant vector, $\mathbf{1}$, and $X(m)$. Let $Y_k(x)$ be the linear interpolation of Y_k. Since $X(m)$ is strictly decreasing, then by Lemma 16, $\exists i < j < k \in \{1, \cdots L\}$ such that the difference function of $F_T^*(m, x)$ and $Y_k(x)$ has strict sign changes between i and j and between j and k. So there are crossings, say $x_1 < x_2$, in the interval $(1, L)$ by the continuity of the functions $Y_k(x)$ and $F_T^*(m, x)$ on $[1, L]$. Suppose $x_1 < x_2 \in [1, k-1]$ at which $F_T^*(m, x_i) = Y_k(x_i)$, $i = 1, 2$. This would force $\frac{\partial}{\partial x} F_T^*(m, x) = 0$ at some $x \in (x_1, x_2)$, contradicting Proposition 18. So there is at most one crossing in $[1, k-1]$. For $m \in [0, 1)$, by Proposition 20 we know that $F_T^*(m, x)$ is strictly concave up on $[1, L]$ and thus, can match $Y_k(x)$ for at most two values of x in $[k-1, L]$ since $Y_k(x)$ is a linear function on $[k-1, L]$.

Suppose $m = 1$. The functions $F_T^*(1, x)$ is linear; by Proposition 18 it has positive slope. As proved in the first paragraph, $F_T^*(1, x)$ has at most one cross point in $[1, k-1]$. We will argue by contradiction that the same is true on $[k-1, L]$. Suppose there are $x_3 \neq x_4$ in $[k-1, L]$ such that $F_T^*(1, x_i) = Y_k(x_i)$, $i = 3, 4$. Then the slope of $F_T^*(1, x)$ would equal 1 and $F_T^*(1, x) = Y_k(x)$ on $[k-1, L]$. Then for all $x \in [1, k-1)$, we would have $F_T^*(1, x) < Y_k(x)$. This contradicts Lemma 16.

So $F_T^*(m, x) = Y_k(x)$ for at most three values of x on $[1, L]$, say $x_1 < x_2 < x_3$. If $x_1 = 1$ and $x_3 = L$ then Lemma 16 is contradicted. Thus, if there are three crossings, they must be in $[1, L)$ or in $(1, L]$. ☐

Theorem 22. *Let Y_k be an extreme vector, $3 \le k \le L$. If $\exists m_0 \in [0, 1)$ such that $F_T^*(m_0, 1) > 0$, then $\exists \delta > 0$ such that for any $m \in (m_0, m_0 + \delta)$, $\phi(m) > \phi(m_0)$.*

Proof. Let $F_T^*(m, x)$ be the interpolation $x \mapsto a^*(m) + b^*(m)(T - x)^m$ of $F_T^*(m, i)$. Since Y_k is piecewise linear, we can write the linear interpolation of Y_k as

$$Y_k(x) = \begin{cases} 0 & \text{if } 1 \le x < k-1 \\ \\ x - k + 1 & \text{if } k-1 \le x \le L \end{cases}$$

Suppose $F_T^*(m_0, 1) > 0$. Since $\frac{\partial}{\partial x} F_T^*(m_0, x) > 0$ by Proposition 18, we have $F_T^*(m_0, x) > Y_k(x)$ on $[1, k-1]$. Therefore, Lemma 21 gives the existence of exactly two cross points between $F_T^*(m_0, x)$ and $Y_k(x)$, say $x_1 < x_2 \in (1, L)$, with $k - 1 < x_1 < x_2$. We may choose x_1 between i and j and x_2 between j and l, where the integers i, j, and l satisfy the conclusion of Lemma 16. So some components of the vector $Y_k - \{F_T^*(m_0, i)\}_{i=1}^L$ alternate in sign from $R_1 = [1, x_1)$ to $R_2 = (x_1, x_2)$ and from R_2 to $R_3 = (x_2, L]$. By the continuity of $F_T^*(m_0, x)$ and $Y_k(x)$, their difference has one (alternating) sign on each of the intervals,

R_1, R_2, and R_3. Since $F_T^*(m_0, 1) > Y_k(1) = 0$ by hypothesis, then

$$F_T^*(m_0, x) > Y_k(x), \quad x \in R_1 \cup R_3$$

$$F_T^*(m_0, x) < Y_k(x), \quad x \in R_2$$

Let $f_T(m, x)$ be the general first level JLS model described in Section 2.1 such that for any $m \in [0, 1]$, $f_T(m, x_1) = F_T^*(m_0, x_1)$ and $f_T(m, x_2) = F_T^*(m_0, x_2)$. If $m = 0$, use the definitions given in Proposition 1 and 10, respectively. Note that $F_T^*(m, x)$ is, given m, the closest first level model to Y, where as $f_T(m, x)$ is a possibly different model from the same linear subspace as $F_T^*(m, x)$ and chosen by the criterion of passing through two given points, namely $(x_1, F_T^*(m_0, x_1))$ and $(x_2, F_T^*(m_0, x_2))$. Also note that $f_T(m_0, x) = F_T^*(m_0, x)$, $\forall x < T$. By Lemma 5, for $m > m_0$,

$$F_T^*(m_0, x) > f_T(m, x), \quad x \in R_1 \cup R_3 \tag{2.24}$$

$$F_T^*(m_0, x) < f_T(m, x), \quad x \in R_2 \tag{2.25}$$

Consider R_1. Since $Y_k(x)$ is a linear function on $[k-1, L]$ and by Proposition 8, $\frac{\partial^2}{\partial x^2} f_T(m, x) > 0$, $\forall x < T$, $\forall m \in [0, 1)$, then for any $m \in (m_0, 1)$, $f_T(m, x) > Y_k(x)$ on $[k-1, x_1)$. Since $f_T(m, x)$ is continuous at every fixed x (and right-continuous at $m = 0$ by Proposition 1), then $\exists \delta > 0$ (which we choose less than $1 - m_0$) such that for any $m \in (m_0, m_0 + \delta)$, $f_T(m, 1) \geq Y_k(1) = 0$. Since $\frac{\partial}{\partial x} f_T(m, x) > 0$, $\forall x < T$, $\forall m \in [0, 1]$, by Proposition 7, then $f_T(m, x) \geq Y_k(x) = 0$ on $[1, k-1]$. Since $(m_0, m_0 + \delta) \subset (m_0, 1)$, then for any $m \in (m_0, m_0 + \delta)$,

$$f_T(m, x) \geq Y_k(x), \quad x \in R_1 \tag{2.26}$$

Consider R_2. Let $m \in [0, 1)$. Since $\frac{\partial^2}{\partial x^2} f_T(m, x) > 0$, $Y_k(x)$ is linear on R_2, and $f_T(m, x_1) = Y_k(x_1)$, $f_T(m, x_2) = Y_k(x_2)$, then by the geometry of strictly concave up functions in relation to linear functions,

$$f_T(m, x) < Y_k(x), \quad x \in R_2 \tag{2.27}$$

Consider R_3. Let $m \in [0, 1)$. Since $f_T(1, x) = Y_k(x)$ is a linear function on $[k-1, L]$, $\frac{\partial^2}{\partial x^2} f_T(m, x) > 0$, and $f_T(m, x_2) = Y_k(x_2)$, then

$$f_T(m, x) > Y_k(x), \quad x \in R_3 \tag{2.28}$$

Let $m \in (m_0, m_0 + \delta)$. By Equations 2.24, 2.25, 2.26, 2.27, and 2.28,

$$\sum_{i=1}^{L} \left(Y_k(i) - F_T^*(m_0, i)\right)^2 > \sum_{i=1}^{L} \left(Y_k(i) - f_T(m, i)\right)^2$$

Since the error in the JLS model is measured by sums of squares, then a lower sum of squares equates to a better fit and, equivalently, a higher ϕ value. Since $f_T(m,i)$ is not necessarily the optimized fit, then

$$\sum_{i=1}^{L}(Y_k(i)-f_T(m,i))^2 \geq \sum_{i=1}^{L}(Y_k(i)-F_T^*(m,i))^2$$

Thus, $\phi(m) > \phi(m_0)$.

□

Theorem 23. *Let Y_k be an extreme vector, $3 \leq k \leq L$. If $\exists m_0 \in (0,1]$ such that $F_T^*(m_0,L) < Y_k(L)$, then $\exists \delta > 0$ such that for any $m \in (m_0-\delta, m_0)$, $\phi(m) > \phi(m_0)$.*

Proof. Let $F_T^*(m,x)$ be the interpolation $x \mapsto a^*(m)+b^*(m)(T-x)^m$ of $F_T^*(m,i)$ and let $Y_k(x)$ be the linear interpolation of Y_k. Suppose $F_T^*(m_0,L) < Y_k(L)$. Since $\frac{\partial^2}{\partial x^2}F_T^*(m) \geq 0$, $\forall m \in (0,1]$ by Proposition 20, then $F_T^*(m_0,x)$ can match $Y_k(x)$ at most once on $[k-1,L)$. Then, by Lemma 21, there exists exactly two cross points between $F_T^*(m_0,x)$ and $Y_k(x)$, say $x_1 < x_2 \in (1,L)$, with $x_1 < k-1 < x_2$. Let $R_1 = [1,x_1)$, $R_2 = (x_1,x_2)$, and $R_3 = (x_2,L]$. By the continuity of $F_T^*(m_0,x)$ and $Y_k(x)$, their difference has one sign on each of the intervals, R_1, R_2, and R_3. These signs alternate by Lemma 16. Since $F_T^*(m_0,L) < Y_k(L)$ by hypothesis, then

$$F_T^*(m_0,x) < Y_k(x), \quad x \in R_1 \cup R_3$$

$$F_T^*(m_0,x) > Y_k(x), \quad x \in R_2$$

As in the proof of Theorem 22, let $f_T(m,x)$ be the general first level JLS model described in Section 2.1 such that for any $m \in [0,1]$, $F_T^*(m_0,x_1) = f_T(m,x_1)$ and $F_T^*(m_0,x_2) = f_T(m,x_2)$. If $m=0$, use the definitions given in Propositions 1 and 10, respectively. By Lemma 5, for $m < m_0$,

$$F_T^*(m_0,x) < f_T(m,x), \quad x \in R_1 \cup R_3 \tag{2.29}$$

$$F_T^*(m_0,x) > f_T(m,x), \quad x \in R_2 \tag{2.30}$$

Consider R_1. Let $m \in [0,1]$. Since $Y_k(x_1) = 0 < Y_k(x_2)$, then $\frac{\partial}{\partial x}f_T(m,x) > 0$, $\forall x < T$, by Proposition 7. Thus, for any $m \in [0,m_0)$,

$$f_T(m,x) < Y_k(x) = 0, \quad x \in R_1 \tag{2.31}$$

Consider R_2. A similar argument to the previous one shows for $m \in [0,1]$, $f_T(m,x) > Y_k(x) = 0$ on $(x_1,k-1]$. Because the coefficients $a(m)$ and $b(m)$ that define $f_T(m,x) = a(m)+b(m)(T-x)^m$ are continuous in m for $m > 0$ (see Equations 2.5 and 2.6), the function value $f_T(m,L)$ is continuous for $m > 0$. Choose $\delta > 0$ with $\delta < m_0$ so that, for $m \in (m_0-\delta, m_0]$ we have $f_T(m,x) < Y_k(L)$. Consier any $m \in (m_0-\delta, m_0)$. By Proposition 8 (note that $m < 1$), we know that $f_T(m,x)$ is strictly concave up. From the theory of concave up functions, the slope, s_1 from $(x_2, f_T(m,x_2))$ to $(L, f_T(m,L))$ is greater than or equal to $\frac{\partial}{\partial x} f_T(m,x_2)$. By hypothesis, s_1 is strictly less than the slope, s_2 from $(x_2, Y_k(x_2))$ to $(L, Y_k(L))$. So $1 = s_2 > s_1 \geq \frac{\partial}{\partial x} f_T(m,x_2)$. Because $f_T(m,x)$ is strictly concave up, we have $\frac{\partial}{\partial x} f_T(m,x) < \frac{\partial}{\partial x} f_T(m,x_2) < 1 = \frac{\partial}{\partial x} Y_k(x)$. So $\frac{\partial}{\partial x}\left(Y_k(x) - f_T(m,x)\right) > 0$ on $[k-1, x_2)$ with $Y_k(x_2) - f_T(m,x_2) = 0$. Thus, $Y_k(x) - f_T(m,x) < 0$ on $[k-1, x_2)$ and hence, $f_T(m,x) > Y_k(x)$ on $[k-1, x_2)$. Thus, for $m \in (m_0 - \delta, m_0)$

$$f_T(m,x) > Y_k(x), \quad x \in R_2 \tag{2.32}$$

Consider R_3. Suppose $\exists x_3 \in (x_2, L]$ such that $f_T(m,x_3) = Y_k(x_3)$. Because $Y_k(x)$ is linear on $(x_2, L]$ and $\frac{\partial^2}{\partial x^2} f_T(m,x) > 0$ on $[1, L]$ by Proposition 8, then $f_T(m,x) > Y_k(x)$ on $(x_3, L]$ and hence at $x = L$. This contradicts our choice of $m \in (m_0 - \delta, m_0)$. Thus, for any $m \in (m_0 - \delta, m_0)$,

$$f_T(m,x) < Y_k(x), \quad x \in R_3 \tag{2.33}$$

For $m \in (m_0 - \delta, m_0)$, Equations 2.29, 2.30, 2.31, 2.32, and 2.33 imply that,

$$\sum_{i=1}^{L} \left(Y_k(i) - F_T^*(m_0, i)\right)^2 > \sum_{i=1}^{L} \left(Y_k(i) - f_T(m,i)\right)^2$$

Since the error in the JLS model is measured by sums of squares, then a lower sum of squares equates to a better fit and, equivalently, a higher ϕ value. Since $f_T(m,i)$ is not necessarily the optimized fit, then

$$\sum_{i=1}^{L} \left(Y_k(i) - f_T(m,i)\right)^2 \geq \sum_{i=1}^{L} \left(Y_k(i) - F_T^*(m,i)\right)^2$$

Thus, $\phi(m) > \phi(m_0)$.

□

2.5 Unique Optimal m Values for Extreme Vectors

Let $L \geq 3$. Then

$$Y_2 = \{i-1\}_{i=1}^{L}, \quad \& \quad Y_{L,j} = \begin{cases} 0 & \text{if } j < L \\ \\ 1 & \text{if } j = L \end{cases}$$

Theorem 24. *$F_T^*(1,t)$ is the unique best optimized first level JLS model fit to the extreme vector, Y_2.*

Proof. By the Cauchy-Schwarz Theorem,

$$|Y \cdot Z_2(m)| = |AY \cdot Z_2(m)| \overset{CS}{\leq} \|AY\| \|Z_2(m)\| = \|AY\|$$

Since $Y_2 = \{i-1\}_i^L$, then $AX(1) = -AY_2$. Then

$$|Y_2 \cdot Z_2(1)| = \left\| \frac{Y_2 \cdot AX(1)}{AX(1)} \right\| = \left\| \frac{AY_2 \cdot AX(1)}{AX(1)} \right\| = \left\| \frac{AY_2 \cdot -AY_2}{AY_2} \right\| = \|AY_2\|$$

Therefore, $\forall m$, $|Y_2 \cdot Z_2(1)| \geq |Y_2 \cdot Z_2(m)|$ and $m = 1$ maximizes $|Y_2 \cdot Z_2(m)|^2$. By the Cauchy-Schwarz Theorem,

$$|V_1 \cdot V_2| = \|V_1\|\|V_2\| \iff (\exists r)(V_1 = rV_2 \,\text{or}\, V_2 = rV_1)$$

Suppose $m \neq 1$, is also a global maximum location. Then $|AY_2 \cdot Z_2(m)| = |AY_2 \cdot Z_2(1)| = \|AY_2\|$. Since $\|Z_2(m)\| = 1$, $\forall m$, then $AY_2 = rZ_2(m)$ or $Z_2(m) = rAY_2$ for some $r \in \mathbb{R}$. Neither vector is 0, so $r \neq 0$ and WOLG,

$$AY_2 = rZ_2(m) = r\frac{AX(m)}{\|AX(m)\|} \implies -AX(1) = \frac{rAX(m)}{\|AX(m)\|}$$

Since $-AX(1)$ is a vector whose linear interpolation is a linear function on $[1, L]$, the same holds for $AX(m)$. This only happens when $m = 1$. □

Proposition 25. *For any vector Y, we can define*

$$\phi(0) := \lim_{m \to 0^+} \phi(m) = \frac{(Y \cdot AW_1)^2}{AW_1 \cdot W_1}$$

Proof.

$$\begin{aligned}
\lim_{m \to 0^+} \phi(m) &= \lim_{m \to 0^+} \frac{(Y \cdot AX)^2}{AX \cdot X} \\
&\overset{\ell \cdot H}{=} \lim_{m \to 0^+} \frac{2(Y \cdot AX)(Y \cdot AX')}{2(AX' \cdot X)} \\
&\overset{\ell \cdot H}{=} \lim_{m \to 0^+} \frac{(Y \cdot AX)(Y \cdot AX'') + (Y \cdot AX')^2}{(AX'' \cdot X) + (AX' \cdot X')} \\
&= \frac{(Y \cdot AW_1)^2}{AW_1 \cdot W_1}
\end{aligned}$$

□

Proposition 26. *For any vector Y, the right derivative of $\phi(m)$ exists at $m = 0$ and*

$$\frac{\partial}{\partial m}\phi(0+) = \lim_{m \to 0^+} \frac{\partial}{\partial m}\phi(m) = \frac{(Y \cdot AW_1)[(Y \cdot AW_2)(AW_1 \cdot W_1) - (Y \cdot AW_1)(AW_1 \cdot W_2)]}{(AW_1 \cdot W_1)^2}$$

Proof. Let

$$C_1(m) = (Y \cdot AX')(AX \cdot X) - (Y \cdot AX)(AX \cdot X') \quad \& \quad D_1 = (AX \cdot X)^2$$

and for $j \in \{2, 3, 4, 5\}$,

$$C_j(m) = \frac{\partial^j}{\partial m^j} C_1(m) \quad \& \quad D_j = \frac{\partial^j}{\partial m^j} D_1$$

Suppressing the variable m, we have,

$$
\begin{aligned}
&\lim_{m\to 0^+} \frac{\partial}{\partial m}\phi(m) \\
&= 2 \lim_{m\to 0^+} \frac{(Y\cdot AX)C_1}{D_1} \\
&\overset{\ell\cdot H}{=} 2 \lim_{m\to 0^+} \frac{(Y\cdot AX')C_1 + (Y\cdot AX)C_2}{D_2} \\
&\overset{\ell\cdot H}{=} 2 \lim_{m\to 0^+} \frac{(Y\cdot AX'')C_1 + 2(Y\cdot AX')C_2 + (Y\cdot AX)C_3}{D_3} \\
&\overset{\ell\cdot H}{=} 2 \lim_{m\to 0^+} \frac{(Y\cdot AX''')C_1 + 3(Y\cdot AX'')C_2 + 3(Y\cdot AX')C_3 + (Y\cdot AX)C_4}{D_4} \\
&\overset{\ell\cdot H}{=} 2 \lim_{m\to 0^+} \frac{(Y\cdot AX'''')C_1 + 4(Y\cdot AX''')C_2 + 6(Y\cdot AX'')C_3}{D_5} \\
&\qquad\qquad + \frac{4(Y\cdot AX')C_4 + (Y\cdot AX)C_5}{D_5}
\end{aligned}
$$

As $m \to 0^+$,

$$
\begin{aligned}
&C_j \to 0, \quad j \in \{1,2,3\} \\
&C_4 \to 3[(Y\cdot AW_2)(AW_1\cdot W_1) - (Y\cdot AW_1)(AW_1\cdot W_2)] \\
&C_5 \to 8(Y\cdot AW_3)(AW_1\cdot W_1) + 6(Y\cdot AW_2)(AW_1\cdot W_2) \\
&\qquad\quad - 2(Y\cdot AW_1)[3(AW_2\cdot W_2) + 4(AW_1\cdot W_3)] \\
&D_j \to 0, \quad j \in \{1,2,3,4\} \\
&D_5 \to 24(AW_1\cdot W_1)^2
\end{aligned}
$$

Thus, $\lim_{m\to 0^+} \frac{\partial}{\partial m}\phi(m) = \frac{(Y\cdot AW_1)[(Y\cdot AW_2)(AW_1\cdot W_1) - (Y\cdot AW_1)(AW_1\cdot W_2)]}{(AW_1\cdot W_1)^2}$. ☐

Lemma 27. *Let $L \geq 3$, $T \geq L+1$. Let $\gamma_j = (T-j)$, $j \in \{1,\cdots,L\}$. Let $1 \leq p < q < r \leq L$. Then*

$$
(\gamma_p\gamma_q)^m \ln\left(\frac{\gamma_p}{\gamma_q}\right) + (\gamma_p\gamma_r)^m \ln\left(\frac{\gamma_r}{\gamma_p}\right) + (\gamma_q\gamma_r)^m \ln\left(\frac{\gamma_q}{\gamma_r}\right) > 0, \quad \forall m \in (0,\infty)
$$

Proof. Note that the second term is negative since the quotient inside the logarithm is less than 1. Therefore, we would like to show the following inequality

$$
(\gamma_p\gamma_r)^m \ln\left(\frac{\gamma_p}{\gamma_r}\right) < (\gamma_p\gamma_q)^m \ln\left(\frac{\gamma_p}{\gamma_q}\right) + (\gamma_q\gamma_r)^m \ln\left(\frac{\gamma_q}{\gamma_r}\right)
$$

We can simplify the problem by multiplying the first term on the right hand side by $\left(\frac{\gamma_r}{\gamma_r}\right)^m$ and the second term on the right hand side by $\left(\frac{\gamma_p}{\gamma_p}\right)^m$. This yields,

$$
(\gamma_p\gamma_r)^m \ln\left(\frac{\gamma_p}{\gamma_r}\right) < (\gamma_p\gamma_r)^m \left(\frac{\gamma_q}{\gamma_r}\right)^m \ln\left(\frac{\gamma_p}{\gamma_q}\right) + (\gamma_p\gamma_r)^m \left(\frac{\gamma_q}{\gamma_p}\right)^m \ln\left(\frac{\gamma_q}{\gamma_r}\right)
$$

Factoring out and canceling $(\gamma_p \gamma_r)^m$, we have

$$\ln\left(\frac{\gamma_p}{\gamma_r}\right) < \left(\frac{\gamma_q}{\gamma_r}\right)^m \ln\left(\frac{\gamma_p}{\gamma_q}\right) + \left(\frac{\gamma_q}{\gamma_p}\right)^m \ln\left(\frac{\gamma_q}{\gamma_r}\right)$$

The left hand side is equivalent to $\ln\left(\frac{\gamma_p}{\gamma_q}\right) + \ln\left(\frac{\gamma_q}{\gamma_r}\right)$. Combining each respective logarithm with its corresponding logarithm on the right hand side, we have

$$\left[1 - \left(\frac{\gamma_q}{\gamma_p}\right)^m\right] \ln\left(\frac{\gamma_q}{\gamma_r}\right) < \left[\left(\frac{\gamma_q}{\gamma_r}\right)^m - 1\right] \ln\left(\frac{\gamma_p}{\gamma_q}\right)$$

Adding $m \ln\left(\frac{\gamma_q}{\gamma_r}\right) \ln\left(\frac{\gamma_q}{\gamma_p}\right)$ to both sides yields

$$\left[1 + m\ln\left(\frac{\gamma_q}{\gamma_p}\right) - \left(\frac{\gamma_q}{\gamma_p}\right)^m\right] \ln\left(\frac{\gamma_q}{\gamma_r}\right) < \left[\left(\frac{\gamma_q}{\gamma_r}\right)^m - 1 - m\ln\left(\frac{\gamma_q}{\gamma_r}\right)\right] \ln\left(\frac{\gamma_p}{\gamma_q}\right)$$

Rewriting the general exponential using e,

$$\left[1 + m\ln\left(\frac{\gamma_q}{\gamma_p}\right) - e^{m\ln\left(\frac{\gamma_q}{\gamma_p}\right)}\right] \ln\left(\frac{\gamma_q}{\gamma_r}\right) < \left[e^{m\ln\left(\frac{\gamma_q}{\gamma_r}\right)} - 1 - m\ln\left(\frac{\gamma_q}{\gamma_r}\right)\right] \ln\left(\frac{\gamma_p}{\gamma_q}\right)$$

On the LHS, we have $1 + x_1 - e^{x_1}$ with $x_1 = m\ln\left(\frac{\gamma_q}{\gamma_p}\right)$. Note x_1 is negative since $\ln\left(\frac{\gamma_q}{\gamma_p}\right) < 0$. and $m > 0$. Since $1 + x - e^x < 0$ for $x < 0$, then the LHS is negative. On the RHS, we have $e^{x_2} - 1 - x_2$ with $x_2 = m\ln\left(\frac{\gamma_q}{\gamma_r}\right)$. Note, $x_2 > 0$ since $\ln\left(\frac{\gamma_q}{\gamma_r}\right) > 0$ and $m > 0$. Since $e^x - 1 - x > 0$ for $x > 0$, then the RHS is positive. The inequality is true and equivalent to the desired result. □

Theorem 28. *$F_T^*(0,t)$ is the unique best optimized first level JLS model fit to the extreme vector, Y_L.*

Proof. To prove $m = 0$ is the best m parameter to fit Y_L, we show:

$$\frac{\partial}{\partial m}\phi(m) = \frac{2(Y_L \cdot AX)[(Y_L \cdot AX')(AX \cdot X) - (Y_L \cdot AX)(AX \cdot X')]}{(AX \cdot X)^2} < 0, \quad \forall m \in (0,1]$$

Let $m \in (0,1]$. By Lemma 13, $(Y_L \cdot AX) < 0$ and $(AX \cdot X) > 0$. Thus, it is sufficient to show $(Y_L \cdot AX')(AX \cdot X) - (Y_L \cdot AX)(AX \cdot X') > 0$, $m \in (0,1]$. Note, for any two vectors U and V, $(AU \cdot AV) = (AU \cdot V) = (U \cdot AV)$. Applying the matrix operator A to Y_L, we have

$$AY_{L,i} = \begin{cases} -\frac{1}{L} & \text{if } i < L \\ \\ \frac{L-1}{L} & \text{if } i = L \end{cases}$$

If we let $\gamma_j = T_j$, $\overline{\gamma}^m = \frac{1}{L}\sum_{j=1}^{L} \gamma_j^m$, and $\widetilde{\gamma}_j^m = \gamma_j^m - \overline{\gamma}^m$, $j \in \{1, 2, \cdots, L\}$, we have

$$(AY_L \cdot X')(AX \cdot X) = -\frac{1}{L}\sum_{i=1}^{L-1}\sum_{j=1}^{L}\gamma_i^m\gamma_j^m\widetilde{\gamma}_j^m \ln \gamma_i + \frac{L-1}{L}\sum_{j=1}^{L}\gamma_j^m\gamma_L^m\widetilde{\gamma}_j^m \ln \gamma_L$$

and

$$(AY_L \cdot X)(AX \cdot X') = -\frac{1}{L}\sum_{i=1}^{L-1}\sum_{j=1}^{L}\gamma_i^m\gamma_j^m\widetilde{\gamma}_j^m \ln \gamma_j + \frac{L-1}{L}\sum_{j=1}^{L}\gamma_j^m\gamma_L^m\widetilde{\gamma}_j^m \ln \gamma_j$$

When we subtract, all the terms where $i = j$ will cancel out. We can also throw away the factor of $1/L$ in the computation since it will not affect the sign of the derivative. The subtraction (omitting the $1/L$ factor) yields,

$$\sum_{i=1}^{L-1}\sum_{j\neq i}\gamma_i^m\gamma_j^m\widetilde{\gamma}_j^m \ln \frac{\gamma_j}{\gamma_i} + (L-1)\sum_{j=1}^{L-1}\gamma_j^m\gamma_L^m\widetilde{\gamma}_j^m \ln \frac{\gamma_L}{\gamma_j} \tag{2.34}$$

Now, the left-hand double sum in Equation 2.34 contains all the terms where $j = L$. For each such term (there will be exactly $L-1$ of them), expand it using the fact that $\widetilde{\gamma_L}^m = -\sum_{p=1}^{L-1}\widetilde{\gamma}_p^m$. So for each $i \in \{1, 2, \cdots, L-1\}$,

$$\gamma_i^m\gamma_L^m\widetilde{\gamma}_L^m \ln \frac{\gamma_L}{\gamma_i} = -\sum_{p=1}^{L-1}\gamma_i^m\gamma_L^m\widetilde{\gamma}_p^m \ln \frac{\gamma_L}{\gamma_i} = \sum_{p=1}^{L-1}\gamma_i^m\gamma_L^m\widetilde{\gamma}_p^m \ln \frac{\gamma_i}{\gamma_L}$$

For each i, there is exactly one time $p = i$. Note each of these terms is the negative of one of the terms in the right-hand sum in Equation 2.34. Equation 2.34 becomes:

$$\sum_{i=1}^{L-1}\sum_{j\neq i<L}\gamma_i^m\gamma_j^m\widetilde{\gamma}_j^m \ln \frac{\gamma_j}{\gamma_i} + \sum_{i=1}^{L-1}\sum_{p\neq i<L}\gamma_i^m\gamma_L^m\widetilde{\gamma}_p^m \ln \frac{\gamma_i}{\gamma_L} + (L-2)\sum_{j=1}^{L-1}\gamma_j^m\gamma_L^m\widetilde{\gamma}_j^m \ln \frac{\gamma_L}{\gamma_j} \tag{2.35}$$

The left-hand double sum in Equation 2.35 can be rewritten as a single summation:

$$\sum_{i=1}^{L-1}\sum_{j\neq i, j<L}\gamma_i^m\gamma_j^m\widetilde{\gamma}_j^m \ln \frac{\gamma_j}{\gamma_i} = \sum_{i<j<L}[\widetilde{\gamma}_i^m - \widetilde{\gamma}_j^m]\gamma_i^m\gamma_j^m \ln \frac{\gamma_i}{\gamma_j} \tag{2.36}$$

Now, consider the middle double sum in Equation 2.35. Notice that for each $i \in \{1, 2, \cdots, L-1\}$, the inner sum has exactly $L-2$ addends. As such, rewriting the right hand sum using the index i, we can combine the middle double sum and right-hand sum in Equation 2.35:

$$\begin{aligned}
&\sum_{i=1}^{L-1} \sum_{p\neq i, p<L} \gamma_i^m \gamma_L^m \widetilde{\gamma}_p^m \ln \frac{\gamma_i}{\gamma_L} + (L-2) \sum_{i=1}^{L-1} \gamma_i^m \gamma_L^m \widetilde{\gamma}_i^m \ln \frac{\gamma_L}{\gamma_i} \\
&\quad = \sum_{i=1}^{L-1} \sum_{p\neq i, p<L} \gamma_i^m \gamma_L^m \widetilde{\gamma}_p^m \ln \frac{\gamma_i}{\gamma_L} + \sum_{i=1}^{L-1} \sum_{p\neq i, p<L} \gamma_i^m \gamma_L^m \widetilde{\gamma}_i^m \ln \frac{\gamma_L}{\gamma_i} \\
&\quad = \sum_{i=1}^{L-1} \sum_{p\neq i, p<L} \gamma_i^m \gamma_L^m [\widetilde{\gamma}_i^m - \widetilde{\gamma}_p^m] \ln \frac{\gamma_L}{\gamma_i} \\
&\quad = \sum_{i=1}^{L-1} \left(\sum_{i<p<L} \gamma_i^m \gamma_L^m [\widetilde{\gamma}_i^m - \widetilde{\gamma}_p^m] \ln \frac{\gamma_L}{\gamma_i} + \sum_{p<i} \gamma_i^m \gamma_L^m [\widetilde{\gamma}_p^m - \widetilde{\gamma}_i^m] \ln \frac{\gamma_i}{\gamma_L} \right) \\
&\quad = \sum_{i=1}^{L-1} \left(\sum_{i<p<L} \gamma_i^m \gamma_L^m [\widetilde{\gamma}_i^m - \widetilde{\gamma}_p^m] \ln \frac{\gamma_L}{\gamma_i} + \sum_{i<p} \gamma_p^m \gamma_L^m [\widetilde{\gamma}_i^m - \widetilde{\gamma}_p^m] \ln \frac{\gamma_p}{\gamma_L} \right) \\
&\quad = \sum_{i<p} \gamma_i^m \gamma_L^m [\widetilde{\gamma}_i^m - \widetilde{\gamma}_p^m] \ln \frac{\gamma_L}{\gamma_i} + \gamma_p^m \gamma_L^m [\widetilde{\gamma}_i^m - \widetilde{\gamma}_p^m] \ln \frac{\gamma_p}{\gamma_L} \\
&\quad = \sum_{i<p} [\widetilde{\gamma}_i^m - \widetilde{\gamma}_p^m] \Big[\gamma_i^m \gamma_L^m \ln \frac{\gamma_L}{\gamma_i} + \gamma_p^m \gamma_L^m \ln \frac{\gamma_p}{\gamma_L} \Big]
\end{aligned} \tag{2.37}$$

We can replace p with j to match indices in the last line in Equation 2.37. Adding Equations 2.36 and 2.37, we have:

$$\sum_{1\leq i<j\leq L-1} [\widetilde{\gamma}_i^m - \widetilde{\gamma}_j^m] \left[\gamma_i^m \gamma_j^m \ln \frac{\gamma_i}{\gamma_j} + \gamma_i^m \gamma_L^m \ln \frac{\gamma_L}{\gamma_i} + \gamma_j^m \gamma_L^m \ln \frac{\gamma_j}{\gamma_L} \right]$$

For $m > 0$, AX is decreasing. So $\widetilde{\gamma}_i^m - \widetilde{\gamma}_j^m > 0$ for $i < j$. By Lemma 27, for $1 \leq i < j \leq L-1$,

$$\gamma_i^m \gamma_j^m \ln \frac{\gamma_i}{\gamma_j} + \gamma_i^m \gamma_L^m \ln \frac{\gamma_L}{\gamma_i} + \gamma_j^m \gamma_L^m \ln \frac{\gamma_j}{\gamma_L} > 0, \quad m \in (0,1]$$

Thus, $\phi(m)$ is decreasing for $m \in (0,1]$. Since we can define $\phi(0)$ to be right-continuous with a corresponding $F_T^*(0,x)$ by Propositions 25 and 10, respectively, then we have the desired result. □

CHAPTER 3
NUMERICAL ANALYSIS

As mentioned in the previous chapter, one of main criticisms of the JLS model is the uncertainty that a random sampling of 20 sets of parameters together with a minimizing method is sufficient enough to find a suitable model fit [20]. The purpose of this chapter is to use numerical analysis to provide a means to guarantee a sufficiently fine grid search through the search space provided in [4].

Recall, there are two versions of the second level JLS model: 1) the 4 non-linear parameter model and 2) the 3 non-linear parameter model. While it is more useful in application to use the 3 non-linear parameter model [4], for the purpose of this chapter, we use the 4 non-linear parameter model to obtain sufficiently fine grid searches for the 3 non-linear parameters (T, m, and ω).

We provide such a fineness using two methods. Firstly, we use the standard Euclidean norm to find an upper bound for the modulus of continuity of the error function of the JLS model. Secondly, we use an upper bound for the second partial derivative of the error function with respect to T, m, and ω together with an application of the Mean Value Theorem to obtain an upper bound for the absolute value of the partial derivative of the error function of the JLS model. By finding these upper bounds, we can control the maximum amount of change of the error function by decreasing the step size (increasing the fineness) of the grid search.

3.1 Modulus of Continuity Approach

Because the method in this section is applied to each of the 3 non-linear parameters T, m, and ω in a similiar manner, we provide only the full calculations for the modulus of continuity method for the non-linear parameter ω. Following that, we will provide the upper bounds for the modulus of continuities for the non-linear parameters T and m.

3.1.1 Bounds for the Non-Linear Parameter ω

Recall, the 4 non-linear parameter JLS model, for the logarithm of a price series $y(t)$, $1 \leq t \leq L$, is

$$JLS(t) := A + (T-t)^m \cdot \{B + C \cdot \cos(\theta + \omega \cdot \ln(T-t))\}$$

with $T \geq L+1$, $m \in [.1, .9]$, $\omega \in [6, 13]$, and θ a phase angle.

It is useful that the parameters A, B, and C are involved linearly and can be slaved to T, m, θ, and ω. For every choice of the parameters T, m, θ, and ω, let $\widehat{A}(T,m,\theta,\omega)$, $\widehat{B}(T,m,\theta,\omega)$, and $\widehat{C}(T,m,\theta,\omega)$ denote values of A, B, and C that minimize

$$\sigma(T,m,\theta,\omega,A,B,C) = \sum_{t=1}^{L}(y(t) - JLS(t))^2$$

Let $\tau(T,m,\theta,\omega) = \sigma(T,m,\theta,\omega,\widehat{A},\widehat{B},\widehat{C})$. Let Y be the column vector where $Y_t = \ln P_t$ for a price vector P. Let X be this $L \times 3$ matrix:

$$X_{t,j} = \begin{cases} 1 & \text{if } j = 1 \\ (T-t)^m & \text{if } j = 2 \\ (T-t)^m \cos(\theta + \omega \cdot \ln(T-t)) & \text{if } j = 3 \end{cases}$$

Then, using the $\|\cdot\|$ for the Euclidean vector norm on $\mathbb{R}^D$,

$$\sigma^{1/2}(T,m,\theta,\omega,A,B,C) = \left\| Y - X \cdot \begin{pmatrix} A \\ B \\ C \end{pmatrix} \right\|$$

Note that $\sigma^{1/2}$ is the distance from Y to a linear subspace S of $\mathbb{R}^L$. If X has full rank of 3, then this distance is achieved for exactly one choice of A, B, and C. We will now select an orthogonal generating set $\{X'_j\}_{i=1}^3$ for S; it will be a basis for S if S has dimension 3. M_j denotes the j-th column a matrix M.

Begin with $X'_1 = X_1$. Then, inductively, let

$$X'_2 = X_2 - \frac{X_2 \cdot X_1}{X_1 \cdot X_1} X_1$$
$$X'_3 = X_3 - \frac{X_3 \cdot X_1}{X_1 \cdot X_1} X_1 - \frac{X_3 \cdot X'_2}{X'_2 \cdot X'_2} X'_2$$

Then the closest vector to Y in S is

$$\widehat{Y} = \sum_{j=1}^{3} \frac{Y \cdot X'_j}{X'_j \cdot X'_j} X'_j := \sum_{j=1}^{3} \alpha_j X'_j$$

and

$$\tau(T,m,\theta,\omega) = \|Y - \widehat{Y}\|^2 = \|Y\|^2 - \sum_{j=1}^{3} \frac{(Y \cdot X'_j)^2}{X'_j \cdot X'_j}$$

Let Z, Z', and α'_j correspond to X, X', and α_j for parameters T, m, θ, and ω^*. We have $X'_1 = Z'_1$ and $X'_2 = Z'_2$. Therefore

$$\tau(T,m,\theta,\omega^*) - \tau(T,m,\theta,\omega) = \frac{(Y \cdot Z'_3)^2}{Z'_3 \cdot Z'_3} - \frac{(Y \cdot X'_3)^2}{X'_3 \cdot X'_3} \tag{3.1}$$

Note that, due to orthogonality relations, in $Y \cdot X'_3$, $Y \cdot Z'_3$, etc., we may replace Y in Equation 3.1 by

$U = Y - \alpha_1 X_1' - \alpha_2 X_2'$:

$$\tau(T, m, \theta, \omega^*) - \tau(T, m, \theta, \omega) = \frac{(U \cdot Z_3')^2}{Z_3' \cdot Z_3'} - \frac{(U \cdot X_3')^2}{X_3' \cdot X_3'} \tag{3.2}$$

Let $j = 3$. By the Mean Value Theorem we have, for some ω_t'' between ω and ω^*,

$$\begin{aligned} Z_{t,3} - X_{t,3} &= (T-t)^m \left[\cos(\theta + \omega^* \cdot \ln(T-t)) - \cos(\theta + \omega \cdot \ln(T-t))\right] \\ &= (T-t)^m \ln(T-t)(\omega^* - \omega) \sin(\theta + \omega_t'' \cdot \ln(T-t)) \end{aligned}$$

Thus

$$|Z_{t,3} - X_{t,3}| \leq |\omega^* - \omega|(T-t)^m \ln(T-t)$$

Consequently,

$$\|Z_3 - X_3\| \leq \Gamma|\omega^* - \omega| \quad \text{with} \quad \Gamma := \left(\sum_{t=1}^{L} (T-t)^{2m} \ln^2(T-t)\right)^{1/2}$$

Next we explore $Z_3' - X_3'$.

$$Z_3' - X_3' = Z_3 - X_3 - \sum_{j=1}^{2} \frac{(Z_3 \cdot X_j') - (X_3 \cdot X_j')}{X_j' \cdot X_j'} X_j'$$

Consequently, because $Z_3' - X_3'$ is orthogonal to X_1' and X_2',

$$\|Z_3' - X_3'\| \leq \|Z_3 - X_3\|$$

Next we bound $(Z_3' \cdot Z_3') - (X_3' \cdot X_3')$. We have

$$(Z_3' \cdot Z_3') - (X_3' \cdot X_3') = Z_3' \cdot (Z_3' - X_3') + X_3' \cdot (Z_3' - X_3')$$

Consequently, because $\|Z_3'\| \leq \|X_3'\| + \Gamma|\omega^* - \omega|$,

$$\begin{aligned} |(Z_3' \cdot Z_3') - (X_3' \cdot X_3')| &\leq (\|Z_3'\| + \|X_3'\|)\Gamma|\omega^* - \omega| \\ &\leq \Gamma|\omega^* - \omega| \cdot (2\|X_3'\| + \Gamma|\omega^* - \omega|) \end{aligned} \tag{3.3}$$

It follows that

$$Z_3' \cdot Z_3' \geq (X_3' \cdot X_3') - \Gamma|\omega^* - \omega| \cdot (2\|X_3'\| + \Gamma|\omega^* - \omega|) \tag{3.4}$$

Let $f(\omega^*) = (U \cdot Z_3'(\omega^*))^2$, $f(\omega) = (U \cdot X_3'(\omega))^2$, $g(\omega^*) = Z_3'(\omega^*) \cdot Z_3'(\omega^*)$, and $g(\omega) = X_3'(\omega) \cdot X_3'(\omega)$.

The RHS Equation 3.2 has the form:

$$\frac{f(\omega^*)}{g(\omega^*)} - \frac{f(\omega)}{g(\omega)} = \frac{f(\omega^*)g(\omega) - g(\omega^*)f(\omega)}{g(\omega)g(\omega^*)} \tag{3.5}$$

By Equation 3.4, we can provide a lower bound on the size of the denominator on the RHS of Equation 3.5:

$$|g(\omega)g(\omega^*)| \geq |X_3' \cdot X_3'| \cdot (|X_3' \cdot X_3'| - \Gamma|\omega^* - \omega| \cdot (2\|X_3'\| + \Gamma|\omega^* - \omega|))$$

We will impose a constraint on $\delta = |\omega^* - \omega|$ so that

$$\Gamma|\omega^* - \omega| \cdot (2\|X_3'\| + \Gamma|\omega^* - \omega|) \leq \frac{1}{2} X_3' \cdot X_3'$$

as follows:

$$\delta(2\|X_3'\| + \Gamma\delta) \leq \frac{X_3' \cdot X_3'}{2\Gamma} \iff \delta(\delta + A) \leq \frac{X_3' \cdot X_3'}{2\Gamma^2} = \frac{A^2}{8}$$

with $A = \frac{2\|X_3'\|}{\Gamma}$. Since $\delta \geq 0$ this is achieved if

$$\delta \leq \sqrt{A^2/8 + A^2/4} - \frac{A}{2} = \frac{(\sqrt{3} - \sqrt{2})A}{2\sqrt{2}} = \frac{(\sqrt{3} - \sqrt{2})\|X_3'\|}{\sqrt{2}\Gamma}$$

This restriction allows us to give this lower bound:

$$|g(\omega)g(\omega^*)| \geq \frac{1}{2}(X_3' \cdot X_3')^2 \tag{3.6}$$

Next, we factor the numerator on the RHS of Equation 3.5:

$$f(\omega^*)g(\omega) - g(\omega^*)f(\omega) = g(\omega)(f(\omega^*) - f(\omega)) + f(\omega)(g(\omega) - g(\omega^*))$$

Note that,

$$\begin{aligned} f(\omega^*) - f(\omega) &= (U \cdot Z_3')^2 - (U \cdot X_3')^2 \\ &= [U \cdot (Z_3' - X_3')] \cdot [U \cdot (Z_3' + X_3')] \end{aligned}$$

Consequently, by Equation 3.3 and the discussion leading to it,

$$|f(\omega^*) - f(\omega)| \leq \|U\|^2 \cdot \Gamma|\omega^* - \omega| \cdot (2\|X_3'\| + \Gamma|\omega^* - \omega|) \tag{3.7}$$

Likewise, by Equation 3.3,

$$|g(\omega) - g(\omega^*)| \leq \Gamma|\omega^* - \omega|(2\|X_3'\| + \Gamma|\omega^* - \omega|) \tag{3.8}$$

Consequently, by Equations 3.7 and 3.8, the numerator on the RHS of Equation 3.5 is bounded by

$$\begin{aligned}
|g(\omega^*)f(\omega) - f(\omega^*)g(\omega)| &= |g(\omega)(f(\omega^*) - f(\omega)) + f(\omega)(g(\omega) - g(\omega^*))| \\
&\le g(\omega) \cdot \|U\|^2 \cdot \Gamma|\omega^* - \omega|(2\|X_3'\| + \Gamma|\omega^* - \omega|) \\
&\quad + f(\omega) \cdot \Gamma|\omega^* - \omega|(2\|X_3'\| + \Gamma|\omega^* - \omega|) \\
&= |\omega^* - \omega|\Gamma(2\|X_3'\| + \Gamma|\omega^* - \omega|)(f(\omega) + g(\omega)\|U\|^2) \\
&= |\omega^* - \omega| \cdot K
\end{aligned} \tag{3.9}$$

where

$$K := \Gamma[(U \cdot X_3')^2 + (U \cdot U)(X_3' \cdot X_3')](2\|X_3'\| + \Gamma|\omega^* - \omega|)$$

Consequently, by Equation 3.6 and 3.9, for $|\omega^* - \omega| \le \frac{(\sqrt{3}-\sqrt{2})\|X_3'\|}{\sqrt{2}\Gamma}$,

$$|\tau(T, m, \theta, \omega^*) - \tau(T, m, \theta, \omega)| \le |\omega^* - \omega| \cdot \frac{2K}{(X_3' \cdot X_3')^2} \tag{3.10}$$

A higher upper bound is given by $\widetilde{K}$ in place of K where

$$\widetilde{K} = 2\Gamma \cdot \|U\|^2 \cdot \|X_3'\|^2 \left[2\|X_3'\| + \Gamma|\omega^* - \omega|\right]$$

and thus,

$$|\tau(T, m, \theta, \omega^*) - \tau(T, m, \omega)| \le |\omega^* - \omega| \cdot (4\Gamma\|U\|^2) \cdot \left[\frac{2}{\|X_3'\|} + \frac{\Gamma|\omega^* - \omega|}{\|X_3'\|^2}\right]$$

Of course, $\|U\| \le \|Y\|$, so relative to $\|Y\|^2$ (for $Y \neq 0$), we can estimate numerically some effective moduli of continuity (to control relative error independently of the price vector).

3.1.2 Other Bounds

We have similarly achieved bounds corresponding to the non-linear parameters T and m. Let $\widehat{U} = Y - \alpha X_1'$. Consider the non-linear parameter T. For $|T^* - T| \le \frac{(\sqrt{3}-\sqrt{2})}{\sqrt{2}} \min\left\{\frac{\|X_2'\|}{\Psi_1}, \frac{\|X_3'\|}{\Psi_1 + \Psi_2}\right\}$, where $\Psi_1 := \left(\sum_{t=1}^{L} m^2(\overline{T} - t)^{2m-2}\right)^{1/2}$, $\Psi_2 := \left(\sum_{t=1}^{L}(m^2 + \omega^2)(\overline{T} - t)^{2m-2}\right)^{1/2}$, $\overline{T} = \max(T^*, T)$, we have

$$|\tau(T^*, m, \theta, \omega) - \tau(T, m, \theta, \omega)| \le |T^* - T| \cdot \left(\frac{2I_2}{(X_2' \cdot X_2')^2} + \frac{2I_3}{(X_3' \cdot X_3')^2}\right) \tag{3.11}$$

where

$$I_2 := \Psi_1[(\widehat{U} \cdot X_2')^2 + (\widehat{U} \cdot \widehat{U})(X_2' \cdot X_2')](2\|X_2'\| + \Psi_1|T^* - T|)$$

$$I_3 := (\Psi_1 + \Psi_2)[(\widehat{U} \cdot X_3')^2 + (\widehat{U} \cdot \widehat{U})(X_3' \cdot X_3')](2\|X_3'\| + (\Psi_1 + \Psi_2)|T^* - T|)$$

Now, consider the non-linear parameter m. For $|m^* - m| \le \frac{(\sqrt{3}-\sqrt{2})}{\sqrt{2}\Delta} \min\left\{\|X_2'\|, \frac{\|X_3'\|}{2}\right\}$, where $\Delta := \left(\sum_{t=1}^{L}(T-t)^{2\overline{m}} \ln^2(T-t)\right)^{1/2}$, $\overline{m} = \max(m^*, m)$, we have

$$|\tau(T, m^*, \theta, \omega) - \tau(T, m, \theta, \omega)| \le |m^* - m| \cdot \left(\frac{2J_2}{(X_2' \cdot X_2')^2} + \frac{2J_3}{(X_3' \cdot X_3')^2}\right) \tag{3.12}$$

where

$$J_2 := \Delta[(\widehat{U} \cdot X_2')^2 + (\widehat{U} \cdot \widehat{U})(X_2' \cdot X_2')](2\|X_2'\| + \Delta|m^* - m|)$$

$$J_3 := 4\Delta[(\widehat{U} \cdot X_3')^2 + (\widehat{U} \cdot \widehat{U})(X_3' \cdot X_3')](\|X_3'\| + \Delta|m^* - m|)$$

3.2 Taylor Theory Approach

Similarly to the previous section, we provide the full computations for bounding the second partial derivative of $\tau(T, m, \theta, \omega)$ for one of 3 non-linear parameters in the 3 non-linear parameter JLS model. We then present the corresponding bounds for the other two non-linear parameters. In this section we will present the computations for T in full.

3.2.1 Bounds for the Non-Linear Parameter T

Using the same notation as Section 3.1, we seek apriori bounds for the absolute value of $\frac{\partial \tau}{\partial T}$. We select an orthonormal generating set $\{Q_j\}_{j=1}^{3}$ for the linear subspace S of $\mathbb{R}^L$ of dimension 3. Begin with $Q_1 = \frac{X_1}{\|X_1\|}$. That is,

$$Q_{1,t} = L^{-1/2}, \quad \text{for } 1 \le t \le L$$

Then let

$$Q_2 = \frac{1}{\|X_2'\|} X_2', \quad \text{where} \quad X_2' = X_2 - (X_2 \cdot Q_1)Q_1$$

Finally, set

$$Q_3 = \frac{1}{\|X_3'\|} X_3' \quad \text{where} \quad X_3' = X_3 - (X_3 \cdot Q_1)Q_1 - (X_3 \cdot Q_2)Q_2$$

Then the closest vector to Y in L is

$$\widehat{Y} = \sum_{j=1}^{3}(Y \cdot Q_j)Q_j$$

with $Y - \widehat{Y}$ perpendicular to $\widehat{Y}$. Hence

$$\begin{aligned}\tau(T, m, \theta, \omega) &= \|Y - \widehat{Y}\|^2 = (Y - \widehat{Y}) \cdot (Y - \widehat{Y}) \\ &= (Y - \widehat{Y}) \cdot Y - (Y - \widehat{Y}) \cdot \widehat{Y} \\ &= Y \cdot Y - \sum_{j=1}^{3} (Y \cdot Q_j)(Y \cdot Q_j) - 0 \\ &= Y \cdot Y - \sum_{j=1}^{3} (Y \cdot Q_j)(Y \cdot Q_j)\end{aligned}$$

Note that

$$\frac{\partial \tau(T, m, \theta, \omega)}{\partial T} = -\frac{\partial}{\partial T}\left[(Y \cdot Q_2)^2 + (Y \cdot Q_3)^2\right]$$

By the orthogonality of Q_1, Q_2 and Q_3 we have $Y \cdot Q_2 = U \cdot Q_2$ and $Y \cdot Q_3 = U \cdot Q_3$, where $U = Y - (Y \cdot Q_1)Q_1$. Therefore

$$\frac{\partial \tau(T, m, \theta, \omega)}{\partial T} = -\frac{\partial}{\partial T}\left[(U \cdot Q_2)^2 + (U \cdot Q_3)^2\right] \tag{3.13}$$

Consider the first term on the RHS of Equation 3.13. Because U does not depend on T we have

$$-\frac{\partial}{\partial T}\left[(U \cdot Q_2)^2\right] = -2(U \cdot Q_2)\left(U \cdot \frac{\partial}{\partial T} Q_2\right)$$

Above, by applying the Cauchy-Schwarz inequality to both inner products on the right, we have

$$\left|-\frac{\partial}{\partial T}\left[(U \cdot Q_2)^2\right]\right| \leq 2\|U\|^2 \cdot \left\|\frac{\partial}{\partial T} Q_2\right\|$$

Since U and $(Y \cdot Q_1)Q_1$ are an orthogonal decomposition of Y [1], we have

$$\|U\| = \|Y\|^2 - (Y \cdot Q_1)^2$$

Consider next the second term on the RHS of Equation 3.13:

$$-\frac{\partial}{\partial T}\left[(U \cdot Q_3)^2\right] = -2(U \cdot Q_3)\left(U \cdot \frac{\partial}{\partial T} Q_3\right)$$

As with the first term on the RHS of Equation 3.13,

$$\left|-\frac{\partial}{\partial T}\left[(U \cdot Q_3)^2\right]\right| \leq 2\|U\|^2 \cdot \left\|\frac{\partial}{\partial T} Q_3\right\|$$

[1] Geometrically, $\|U\|$ is the distance from Y to the linear subspace generated by X_1

From 3.13 we also have

$$\frac{\partial^2\tau(T,m,\theta,\omega)}{\partial T^2} = -2\left[\left(U\cdot\frac{\partial Q_2}{\partial T}\right)^2 + (U\cdot Q_2)\left(U\cdot\frac{\partial^2 Q_2}{\partial T^2}\right) + \left(U\cdot\frac{\partial Q_3}{\partial T}\right)^2 + (U\cdot Q_3)\left(U\cdot\frac{\partial^2 Q_3}{\partial T^2}\right)\right]$$

It follows from the Cauchy-Schwarz Inequality,

$$\left\|\frac{\partial^2\tau(T,m,\theta,\omega)}{\partial T^2}\right\| \le 2\|U\|^2\left[\left\|\frac{\partial Q_2}{\partial T}\right\|^2 + \left\|\frac{\partial^2 Q_2}{\partial T^2}\right\| + \left\|\frac{\partial Q_3}{\partial T}\right\|^2 + \left\|\frac{\partial^2 Q_3}{\partial T^2}\right\|\right] \tag{3.14}$$

We proceed to bound each of the four terms on inside the brackets on the RHS of Equation 3.14.

Bounding $\left\|\frac{\partial Q_2}{\partial T}\right\|$

We have

$$\begin{aligned}\frac{\partial}{\partial T}Q_2 &= \frac{\partial}{\partial T}\left(\|X_2'\|^{-1}X_2'\right)\\ &= \left[\frac{\partial}{\partial T}\left(\|X_2'\|^{-1}\right)\right]X_2' + \|X_2'\|^{-1}\cdot\frac{\partial}{\partial T}X_2'\end{aligned} \tag{3.15}$$

We have

$$\frac{\partial}{\partial T}X_2 = \left\{m(T-t)^{m-1}\right\}_{t=1}^{L}$$

Since U and Q_1 do not depend on T we have

$$\begin{aligned}\frac{\partial}{\partial T}X_2' &= \frac{\partial}{\partial T}\left[X_2 - (X_2\cdot Q_1)Q_1\right]\\ &= \frac{\partial}{\partial T}X_2 - \left[\left(\frac{\partial}{\partial T}X_2\right)\cdot Q_1\right]Q_1\end{aligned}$$

Consequently,

$$\left\|\frac{\partial}{\partial T}X_2'\right\|^2 = \left\|\frac{\partial}{\partial T}X_2\right\|^2 - \left(\left(\frac{\partial}{\partial T}X_2\right)\cdot Q_1\right)^2 \le \left\|\frac{\partial}{\partial T}X_2\right\|^2 \tag{3.16}$$

Then we have

$$\begin{aligned}\left\|\frac{\partial}{\partial T}X_2'\right\| \le \left\|\frac{\partial}{\partial T}X_2\right\| &= \|\{m(T-t)^{m-1}\}_{t=1}^{L}\|\\ &= \left(\sum_{t=1}^{L} m^2(T-t)^{2m-2}\right)^{1/2}\\ &=: \Lambda_1(T)\end{aligned} \tag{3.17}$$

It follows that the length of $\frac{\partial}{\partial T}X_2'$ is equal to the distance of vector $\frac{\partial}{\partial T}X_2$ to the linear subspace generated by X_1. Note that $\frac{\partial}{\partial T}X_2'$ is perpendicular to Q_1.

Starting with $\|X_2'\|^2 = X_2'\cdot X_2'$ we have

$$2\|X_2'\|\cdot\frac{\partial}{\partial T}\|X_2'\| = 2X_2'\cdot\left(\frac{\partial}{\partial T}X_2'\right) \tag{3.18}$$

Consequently,

$$\frac{\partial}{\partial T}\|X_2'\| = \frac{X_2' \cdot \left(\frac{\partial}{\partial T}X_2'\right)}{\|X_2'\|} \tag{3.19}$$

and thus,

$$\left|\frac{\partial}{\partial T}\|X_2'\|\right| \leq \left\|\frac{\partial}{\partial T}X_2'\right\| \tag{3.20}$$

By Equations 3.15 and 3.19,

$$\begin{aligned}\frac{\partial}{\partial T}Q_2 &= \left[\frac{\partial}{\partial T}\left(\|X_2'\|^{-1}\right)\right]X_2' + \|X_2'\|^{-1}\cdot\frac{\partial}{\partial T}X_2' \\ &= \frac{-1}{\|X_2'\|^2}\cdot\frac{X_2'\cdot\left(\frac{\partial}{\partial T}X_2'\right)}{\|X_2'\|}X_2' + \|X_2'\|^{-1}\cdot\frac{\partial}{\partial T}X_2'\end{aligned}$$

Therefore, by Equations 3.17 and 3.20,

$$\left\|\frac{\partial}{\partial T}Q_2\right\| \leq \frac{2\left\|\frac{\partial}{\partial T}X_2'\right\|}{\|X_2'\|} = \frac{2\Lambda_1(T)}{\|X_2'\|} \tag{3.21}$$

Bounding $\left\|\frac{\partial Q_3}{\partial T}\right\|$

We have

$$\begin{aligned}\frac{\partial}{\partial T}Q_3 &= \frac{\partial}{\partial T}\left(\|X_3'\|^{-1}X_3'\right) \\ &= \left[\frac{\partial}{\partial T}\left(\|X_3'\|^{-1}\right)\right]X_3' + \|X_3'\|^{-1}\cdot\frac{\partial}{\partial T}X_3'\end{aligned} \tag{3.22}$$

Note that Q_1 does not depend on T but Q_2 does:

$$\begin{aligned}\frac{\partial}{\partial T}X_3' &= \frac{\partial}{\partial T}X_3 - \sum_{j=1}^{2}\frac{\partial}{\partial T}\left[(X_3\cdot Q_j)Q_j\right] \\ &= \frac{\partial}{\partial T}X_3 - \left[\left(\frac{\partial}{\partial T}X_3\right)\cdot Q_1\right]Q_1 \\ &\quad - \left[\left(\frac{\partial}{\partial T}X_3\right)\cdot Q_2\right]Q_2 - \left[X_3^T\frac{\partial}{\partial T}Q_2\right]Q_2 - (X_3\cdot Q_2)\frac{\partial}{\partial T}Q_2\end{aligned}$$

Consequently, since $\left\|\frac{\partial}{\partial T}X_3 - (\frac{\partial}{\partial T}X_3\cdot Q_1)Q_1 - (\frac{\partial}{\partial T}X_3\cdot Q_2)Q_2\right\| \leq \left\|\frac{\partial}{\partial T}X_3\right\|$ by the usual orthogonal decomposition argument,

$$\left\|\frac{\partial}{\partial T}X_3'\right\| \leq \left\|\frac{\partial}{\partial T}X_3\right\| + 2\|X_3\|\left\|\frac{\partial}{\partial T}Q_2\right\| \tag{3.23}$$

Then we have

$$\begin{aligned}\left\|\frac{\partial}{\partial T}X_3\right\| &= \left\|\left\{(T-t)^{m-1}\sqrt{m^2+\omega^2}\sin(\theta+\omega\ln(T-t)+\arctan\left(-\frac{m}{\omega}\right)+\pi)\right\}_{t=1}^{L}\right\| \\ &= \left(\sum_{t=1}^{L}(T-t)^{2m-2}(m^2+\omega^2)\sin^2(\theta+\omega\ln(T-t)+\arctan\left(-\frac{m}{\omega}\right)+\pi)\right)^{1/2} \\ &=: \Lambda_2(T)\end{aligned} \tag{3.24}$$

Starting with $\|X_3'\|^2 = X_3' \cdot X_3'$ we have

$$2\|X_3'\| \cdot \frac{\partial}{\partial T}\|X_3'\| = 2X_3' \cdot \left(\frac{\partial}{\partial T}X_3'\right)$$

Consequently,

$$\frac{\partial}{\partial T}\|X_3'\| = \frac{X_3' \cdot \left(\frac{\partial}{\partial T}X_3'\right)}{\|X_3'\|} \tag{3.25}$$

and thus,

$$\left|\frac{\partial}{\partial T}\|X_3'\|\right| \leq \left\|\frac{\partial}{\partial T}X_3'\right\| \tag{3.26}$$

By Equations 3.22 and 3.25,

$$\frac{\partial}{\partial T}Q_3 = \frac{-1}{\|X_3'\|^2} \cdot \frac{X_3' \cdot \left(\frac{\partial}{\partial T}X_3'\right)}{\|X_3'\|} \cdot X_3' + \|X_3'\|^{-1}\frac{\partial}{\partial T}X_3'$$

Therefore, by Equations 3.21, 3.23, 3.24, and 3.26,

$$\left\|\frac{\partial}{\partial T}Q_3\right\| \leq \frac{2\left\|\frac{\partial}{\partial T}X_3'\right\|}{\|X_3'\|} \leq \frac{2\left[\Lambda_2(T) + \frac{4\|X_3\|\Lambda_1(T)}{\|X_2'\|}\right]}{\|X_3'\|} = \frac{2\|X_2'\|\Lambda_2(T) + 8\|X_3\|\Lambda_1(T)}{\|X_2'\|\|X_3'\|} \tag{3.27}$$

Bounding $\left\|\frac{\partial^2 Q_2}{\partial T^2}\right\|$

Set $\widehat{X}_2$ equal to the $\frac{\partial^2}{\partial T^2}X_2$. We have

$$\frac{\partial^2}{\partial T^2}Q_2 = 2\left(\frac{\partial}{\partial T}\|X_2'\|^{-1}\right)\left(\frac{\partial}{\partial T}X_2'\right) + X_2'\left(\frac{\partial^2}{\partial T^2}\|X_2'\|^{-1}\right) + \|X_2'\|^{-1}\left(\frac{\partial^2}{\partial T^2}X_2'\right) \tag{3.28}$$

Note that

$$\frac{\partial^2}{\partial T^2}X_2' = \widehat{X}_2 - \left(\widehat{X}_2 \cdot Q_1\right)Q_1 \tag{3.29}$$

The RHS in Equation 3.29 is the difference between $\widehat{X}_2$ and its orthogonal projection onto the subspace generated by X_1; hence this difference is perpendicular to that subspace. Consequently,

$$\left\|\frac{\partial^2}{\partial T^2}X_2'\right\|^2 = \|\widehat{X}_2\|^2 - (\widehat{X}_2 \cdot Q_1)^2 \leq \|\widehat{X}_2\|^2$$

Note that

$$\widehat{X}_{2,t} = m(m-1)(T-t)^{m-2}$$

Consequently, we have

$$\begin{aligned}\left\|\frac{\partial^2}{\partial T^2}X_2'\right\| &\le \|\widehat{X}_2\| \\ &= \|\{m(m-1)(T-t)^{m-2}\}_{t=1}^{L}\| \\ &= \left(\sum_{t=1}^{L} m^2(m-1)^2(T-t)^{2m-4}\right)^{1/2} \\ &=: \Lambda_3(T)\end{aligned} \tag{3.30}$$

By Equation 3.18,

$$2\|X_2'\|\cdot\frac{\partial}{\partial T}\|X_2'\| = 2X_2'\cdot\frac{\partial}{\partial T}X_2'$$

Taking the partial derivative of both sides, we have,

$$2\|X_2'\|\cdot\frac{\partial^2}{\partial^2 T}\|X_2'\| + 2\left(\frac{\partial}{\partial T}\|X_2'\|\right)^2 = 2\left(\frac{\partial}{\partial T}X_2'\cdot\frac{\partial}{\partial T}X_2'\right) + 2X_2'\cdot\frac{\partial^2}{\partial^2 T}X_2'$$

Hence,

$$\frac{\partial^2}{\partial^2 T}\|X_2'\| = \frac{\left(\frac{\partial}{\partial T}X_2'\cdot\frac{\partial}{\partial T}X_2'\right) + X_2'\cdot\frac{\partial^2}{\partial^2 T}X_2' - \left(\frac{\partial}{\partial T}\|X_2'\|\right)^2}{\|X_2'\|}$$

Note by Equation 3.20, $\left\|\frac{\partial}{\partial T}X_2'\right\|^2 - \left(\frac{\partial}{\partial T}\|X_2'\|\right)^2 \ge 0$. Thus, by Equations 3.17, 3.20, and 3.30,

$$\left|\frac{\partial^2}{\partial^2 T}\|X_2'\|\right| \le \frac{\left\|\frac{\partial}{\partial T}X_2'\right\|^2}{\|X_2'\|} + \left\|\frac{\partial^2}{\partial^2 T}X_2'\right\| \le \frac{\Lambda_1^2(T)}{\|X_2'\|} + \Lambda_3(T) \tag{3.31}$$

By Equation 3.28,

$$\frac{\partial^2}{\partial^2 T}Q_2 = -\frac{2}{\|X_2'\|^2}\frac{\partial}{\partial T}\|X_2'\|\frac{\partial}{\partial T}X_2' + \frac{2X_2'}{\|X_2'\|^3}\left(\frac{\partial}{\partial T}\|X_2'\|\right)^2 - \frac{X_2'}{\|X_2'\|^2}\frac{\partial^2}{\partial^2 T}\|X_2'\| + \frac{1}{\|X_2'\|}\frac{\partial^2}{\partial^2 T}X_2'$$

and thus, by Equations 3.17, 3.20, 3.30, and 3.31,

$$\begin{aligned}\left\|\frac{\partial^2}{\partial T^2}Q_2\right\| &\le \frac{2\Lambda_1^2(T)}{\|X_2'\|^2} + \frac{2\Lambda_1^2(T)}{\|X_2'\|^2} + \frac{1}{\|X_2'\|}\left(\frac{\Lambda_1^2(T)}{\|X_2'\|} + \Lambda_3(T)\right) + \frac{1}{\|X_2'\|}\Lambda_3(T) \\ &= \frac{5\Lambda_1^2(T) + 2\|X_2'\|\Lambda_3(T)}{\|X_2'\|^2}\end{aligned} \tag{3.32}$$

Bounding $\left\|\frac{\partial^2 Q_3}{\partial T^2}\right\|$

Set $\widetilde{\cdot}$ equal to $\frac{\partial}{\partial T}\cdot$ and $\widehat{\cdot}$ equal to $\frac{\partial^2}{\partial T^2}\cdot$. We have

$$\frac{\partial^2}{\partial T^2}Q_3 = 2\left(\frac{\partial}{\partial T}\|X_3'\|^{-1}\right)\left(\frac{\partial}{\partial T}X_3'\right) + X_3'\left(\frac{\partial^2}{\partial T^2}\|X_3'\|^{-1}\right) + \|X_3'\|^{-1}\left(\frac{\partial^2}{\partial T^2}X_3'\right) \tag{3.33}$$

Note that

$$\begin{aligned}\frac{\partial^2}{\partial T^2}X_3' = \widehat{X}_3 - \left(\widehat{X}_3\cdot Q_1\right)Q_1 - \left(\widehat{X}_3\cdot Q_2\right)Q_2 - \left(X_3\cdot\widehat{Q}_2\right)Q_2 - \left(X_3\cdot Q_2\right)\widehat{Q}_2 \\ - 2\left(\widetilde{X}_3\cdot\widetilde{Q}_2\right)Q_2 - 2\left(\widetilde{X}_3\cdot Q_2\right)\widetilde{Q}_2 - 2\left(X_3\cdot\widetilde{Q}_2\right)\widetilde{Q}_2\end{aligned}$$

Then

$$\left\|\frac{\partial^2}{\partial T^2}X_3'\right\| \leq \left\|\frac{\partial^2}{\partial T^2}X_3\right\| + 2\left\|X_3\right\|\left\|\frac{\partial^2}{\partial T^2}Q_2\right\| + 4\left\|\frac{\partial}{\partial T}X_3\right\|\left\|\frac{\partial}{\partial T}Q_2\right\| + 2\left\|X_3\right\|\left\|\frac{\partial}{\partial T}Q_2\right\|^2$$

Note that

$$\begin{aligned}\left\|\widehat{X}_3\right\| &= \Bigg\|\Big\{(T-t)^{m-2}\sqrt{(m(m-1)-\omega^2)^2+((-2m+1)\omega)^2} \\ &\qquad\cdot\sin\left(\theta+\omega\ln(T-t)+\operatorname{atan2}(m(m-1)-\omega^2,(-2m+1)\omega)\right)\Big\}_{t=1}^{L}\Bigg\| \\ &= \Bigg(\sum_{t=1}^{L}(T-t)^{2m-4}[(m(m-1)-\omega^2)^2+((-2m+1)\omega)^2] \\ &\qquad\cdot\sin^2\left(\theta+\omega\ln(T-t)+\operatorname{atan2}(m(m-1)-\omega^2,(-2m+1)\omega)\right)\Bigg)^{1/2} \\ &=: \Lambda_4(T)\end{aligned} \tag{3.34}$$

Consequently, by Equations 3.21, 3.32, 3.24, and 3.34, we have,

$$\begin{aligned}\left\|\frac{\partial^2}{\partial T^2}X_3'\right\| &\leq \Lambda_4(T) + 2\|X_3\|\left(\frac{5\Lambda_1^2(T)+2\|X_2'\|\Lambda_3(T)}{\|X_2'\|^2}\right) \\ &\qquad + 4\Lambda_2(T)\left(\frac{2\Lambda_1(T)}{\|X_2'\|}\right) + 2\|X_3\|\left(\frac{2\Lambda_1(T)}{\|X_2'\|}\right)^2 \\ &\leq \frac{\|X_2'\|^2\Lambda_4(T)+8\|X_2'\|\Lambda_1(T)\Lambda_2(T)+18\|X_3\|\Lambda_1^2(T)+4\|X_2'\|\|X_3\|\Lambda_3(T)}{\|X_2'\|^2}\end{aligned} \tag{3.35}$$

Differentiating $\|X_3\|^2 = X_3\cdot X_3$ give us,

$$2\|X_3'\|\cdot\frac{\partial}{\partial T}\|X_3'\| = 2X_3'\cdot\frac{\partial}{\partial T}X_3'$$

Taking the partial derivative of both sides, we have,

$$2\|X_3'\| \cdot \frac{\partial^2}{\partial T^2}\|X_3'\| + 2\left(\frac{\partial}{\partial T}\|X_3'\|\right)^2 = 2\left(\frac{\partial}{\partial T}X_3' \cdot \frac{\partial}{\partial T}X_3'\right) + 2X_3' \cdot \frac{\partial^2}{\partial T^2}X_3'$$

Hence

$$\frac{\partial^2}{\partial T^2}\|X_3'\| = \frac{\left(\frac{\partial}{\partial T}X_3' \cdot \frac{\partial}{\partial T}X_3'\right) + X_3' \cdot \frac{\partial^2}{\partial T^2}X_3' - \left(\frac{\partial}{\partial T}\|X_3'\|\right)^2}{\|X_3'\|}$$

Note by Equation 3.26, $\left\|\frac{\partial}{\partial T}X_3'\right\|^2 - \left(\frac{\partial}{\partial T}\|X_3'\|\right)^2 \geq 0$. Thus, by Equations 3.21, 3.23, 3.24, 3.26, and 3.35,

$$\begin{aligned}\left|\frac{\partial^2}{\partial T^2}\|X_3'\|\right| &\leq \frac{\left\|\frac{\partial}{\partial T}X_3'\right\|^2}{\|X_3'\|} + \left\|\frac{\partial^2}{\partial T^2}X_3'\right\| \\ &\leq \frac{\left(\|X_2'\|\Lambda_2(T) + 4\|X_3\|\Lambda_1(T)\right)^2}{\|X_2'\|^2\|X_3'\|} \\ &\quad + \frac{\|X_2'\|^2\Lambda_4(T) + 8\|X_2'\|\Lambda_1(T)\Lambda_2(T) + 18\|X_3\|\Lambda_1^2(T) + 4\|X_2'\|\|X_3\|\Lambda_3(T)}{\|X_2'\|^2}\end{aligned} \tag{3.36}$$

By Equation 3.33,

$$\frac{\partial^2}{\partial T^2}Q_3 = -\frac{2}{\|X_3'\|^2}\frac{\partial}{\partial T}\|X_3'\|\frac{\partial}{\partial T}X_3' + \frac{2X_3'}{\|X_3'\|^3}\left(\frac{\partial}{\partial T}\|X_3'\|\right)^2 - \frac{X_3'}{\|X_3'\|^2}\frac{\partial^2}{\partial T^2}\|X_3'\| + \frac{1}{\|X_3'\|}\frac{\partial^2}{\partial T^2}X_3'$$

and thus, by Equations 3.21, 3.23, 3.24, 3.26, 3.35, and 3.36,

$$\begin{aligned}\left\|\frac{\partial^2}{\partial T^2}Q_3\right\| &\leq \frac{5\left(\|X_2'\|\Lambda_2(T) + 4\|X_3\|\Lambda_1(T)\right)^2}{\|X_2'\|^2\|X_3'\|^2} \\ &\quad + \frac{2\left[\|X_2'\|^2\Lambda_4(T) + 8\|X_2'\|\Lambda_1(T)\Lambda_2(T) + 18\|X_3\|\Lambda_1^2(T) + 4\|X_2'\|\|X_3\|\Lambda_3(T)\right]}{\|X_2'\|^2\|X_3'\|}\end{aligned} \tag{3.37}$$

3.2.2 Other Bounds

We have similarly achieved bounds corresponding to the non-linear parameters m and ω. For the non-linear parameter m, we have,

$$\left\|\frac{\partial^2\tau(T,m,\theta,\omega)}{\partial m^2}\right\| \leq 2\|U\|^2\left[\left\|\frac{\partial Q_2}{\partial m}\right\|^2 + \left\|\frac{\partial^2 Q_2}{\partial m^2}\right\| + \left\|\frac{\partial Q_3}{\partial m}\right\|^2 + \left\|\frac{\partial^2 Q_3}{\partial m^2}\right\|\right] \tag{3.38}$$

with

$$\left\|\frac{\partial}{\partial m}Q_2\right\| \leq \frac{2\Pi_1(m)}{\|X_2'\|} \tag{3.39}$$

$$\left\|\frac{\partial}{\partial m}Q_3\right\| \leq \frac{2\|X_2'\|\Pi_2(m) + 8\|X_3\|\Pi_1(m)}{\|X_2'\|\|X_3'\|} \tag{3.40}$$

$$\left\|\frac{\partial^2}{\partial m^2}Q_2\right\| \leq \frac{5\Pi_1^2(m) + 2\|X_2'\|\Pi_3(m)}{\|X_2'\|^2} \tag{3.41}$$

$$\left\|\frac{\partial^2}{\partial m^2}Q_3\right\| \le \frac{5\left(\|X_2'\|\Pi_2(m)+4\|X_3\|\Pi_1(m)\right)^2}{\|X_2'\|^2\|X_3'\|^2} + \frac{2\left[\|X_2'\|^2\Pi_4(m)+8\|X_2'\|\Pi_1(m)\Pi_2(m)+18\|X_3\|\Pi_1^2(m)+4\|X_2'\|\|X_3\|\Pi_3(m)\right]}{\|X_2'\|^2\|X_3'\|} \tag{3.42}$$

where

$$\Pi_1(m) := \left\|\frac{\partial}{\partial m}X_2\right\| = \left(\sum_{t=1}^{L}(T-t)^{2m}\ln^2(T-t)\right)^{1/2}$$

$$\Pi_2(m) := \left\|\frac{\partial}{\partial m}X_3\right\| = \left(\sum_{t=1}^{L}(T-t)^{2m}\ln^2(T-t)\cos^2(\theta+\omega\ln(T-t))\right)^{1/2}$$

$$\Pi_3(m) := \left\|\frac{\partial^2}{\partial m^2}X_2\right\| = \left(\sum_{t=1}^{L}(T-t)^{2m}\ln^4(T-t)\right)^{1/2}$$

$$\Pi_4(m) := \left\|\frac{\partial^2}{\partial m^2}X_3\right\| = \left(\sum_{t=1}^{L}(T-t)^{2m}\ln^4(T-t)\cos^2(\theta+\omega\ln(T-t))\right)^{1/2}$$

For the non-linear parameter ω, let $\overline{U} = Y - (Y\cdot Q_1)Q_1 - (Y\cdot Q_2)Q_2$. Then we have

$$\left\|\frac{\partial^2\tau(T,m,\theta,\omega)}{\partial\omega^2}\right\| \le 2\|\overline{U}\|^2\left[\left\|\frac{\partial Q_3}{\partial\omega}\right\|^2 + \left\|\frac{\partial^2 Q_3}{\partial\omega^2}\right\|\right] \tag{3.43}$$

with

$$\left\|\frac{\partial}{\partial\omega}Q_3\right\| \le \frac{2\Xi_1(\omega)}{\|X_3'\|} \tag{3.44}$$

$$\left\|\frac{\partial^2}{\partial\omega^2}Q_3\right\| \le \frac{5\Xi_1^2(\omega)+2\|X_3'\|\Xi_2(\omega)}{\|X_3'\|^2} \tag{3.45}$$

where

$$\Xi_1(\omega) := \left\|\frac{\partial}{\partial\omega}X_3\right\| = \left(\sum_{t=1}^{L}(T-t)^{2m}\ln^2(T-t)\sin^2(\theta+\omega\ln(T-t))\right)^{1/2}$$

$$\Xi_2(\omega) := \left\|\frac{\partial^2}{\partial\omega^2}X_3\right\| = \left(\sum_{t=1}^{L}(T-t)^{2m}\ln^4(T-t)\cos^2(\theta+\omega\ln(T-t))\right)^{1/2}$$

3.3 Applications Using the JLS Model

3.3.1 Using Moduli of Continuity to Guide a Grid Search

Consider any log price vector $Y = \{y_i\}_i$ such that the mapping $i \mapsto y_i$ is not linear and thus $\mathrm{SS}_{gmb} > 0$ where SS_{gmb} is the least sum of squared errors from fitting a linear function to $i \mapsto y_i$. Let $\widehat{\tau}$ be the least sum of squared errors from fitting a JLS model to Y. In Section 4.4 we will define Efficiency by

$$\text{Efficiency} = 1 - \frac{\widehat{\tau}}{\text{SS}_{gmb}}$$

In this section we describe a method for finding an estimate, E^*, of efficiency with an absolute error of at most ϵ, for any given $\epsilon > 0$. This goal is equivalent to finding an estimate τ^* of $\widehat{\tau}$ such that

$$|\tau^* - \widehat{\tau}| \leq \epsilon \cdot \text{SS}_{gmb} := \epsilon_\tau$$

For this section, we assume that L has been fixed as $L = 65 + 60i$ for some positive integer $i \leq 11$, with T restricted to the real interval $[L+1, L+60]$. We also assume that $m \in [.1, .9]$ and $\omega \in [6, 13]$. Using the techniques of the previous subsections of Chapter 3, there are bounds B_i and radii $\delta_i > 0$ such that, given parameters (T, m, ω), we have, for T^*, m^*, and ω^* within the search space,

$$|T^* - T| \leq \delta_1 \quad \Rightarrow \quad |\tau(T^*, m, \omega) - \tau(T, m, \omega)| \leq B_1 |T^* - T|$$

$$|m^* - m| \leq \delta_2 \quad \Rightarrow \quad |\tau(T, m^*, \omega) - \tau(T, m, \omega)| \leq B_2 |m^* - m|$$

$$|\omega^* - \omega| \leq \delta_3 \quad \Rightarrow \quad |\tau(T, m, \omega^*) - \tau(T, m, \omega)| \leq B_3 |\omega^* - \omega|$$

The numbers δ_i and B_i may be locally valid, or chosen to be globally valid—for simplicity we assume the latter. The simplest grid search for τ^* is by brute force. To keep the explanation as short as possible, let

$$\lambda = \min \left\{ \delta_1,\ \delta_2,\ \delta_3,\ \frac{\epsilon_\tau}{3B_1},\ \frac{\epsilon_\tau}{3B_2},\ \frac{\epsilon_\tau}{3B_3} \right\}$$

Note that $\lambda > 0$ and, if any of T, m or ω is varied by at most λ while keeping the other two variables fixed, the change in $\tau(T, m, \omega)$ will be at most $\frac{\epsilon_\tau}{3}$.

Set $T_i = L + i\lambda$ for integers i such that $1 \leq i \leq \left\lfloor \frac{60}{\lambda} \right\rfloor$; set $m_j = 0.1 + j\lambda$ for integers j such that $1 \leq j \leq \left\lfloor \frac{0.8}{\lambda} \right\rfloor$; and set $\omega_k = 6 + k\lambda$ for integers k such that $1 \leq k \leq \left\lfloor \frac{7}{\lambda} \right\rfloor$. It follows that for every (T, m, ω) in the search space, there is at least one (i, j, k) such that $|T - T_i| \leq \lambda$, $|m - m_j| \leq \lambda$ and $|\omega - \omega_k| \leq \lambda$. With that closeness we have

$$\begin{aligned} |\tau(T, m, \omega) - \tau(T_i, m_j, \omega_k)| &\leq |\tau(T, m, \omega) - \tau(T_i, m, \omega)| \\ &\quad + |\tau(T_i, m, \omega) - \tau(T_i, m_j, \omega)| \\ &\quad + |\tau(T_i, m_j, \omega) - \tau(T_i, m_j, \omega_k)| \\ &\leq \frac{\epsilon_\tau}{3} + \frac{\epsilon_\tau}{3} + \frac{\epsilon_\tau}{3} = \epsilon_\tau \end{aligned}$$

Thus, wherever $\widehat{\tau}$ is realized, say at $(\widehat{T}, \widehat{m}, \widehat{\omega})$, there is some (i, j, k) such that

$$|\widehat{\tau} - \tau(T_i, m_j, \omega_k)| \leq \epsilon_\tau \quad \text{and thus} \quad \widehat{\tau} \geq \tau(T_i, m_j, \omega_k) - \epsilon_\tau$$

Set

$$\tau^* = \min\{\, \tau(T_i, m_j, \omega_k) \,:\, \text{all } (i, j, k) \text{ of the grid} \,\}$$

By the the previous paragraph $\widehat{\tau} \geq \tau^* - \epsilon_\tau$. However, because $\widehat{\tau}$ is the global minimum of τ over the entire search space, and the grid is a subset of the search space, we have $\widehat{\tau} \leq \tau^*$. Therefore $|\widehat{\tau} - \tau^*| \leq \epsilon_\tau$ as desired.

3.3.2 Searching for Interior Minima Points of τ

A more sophisticated search, with more logical overhead, involves carving up the search space in a binary tree fashion until one reaches rectangular cells in which the minimum value of τ can be ignored. The tree search can be more efficient in that large cells have the potential to be rejected early, if τ at some point in the cell is especially large relative to (a progressively updated) τ^*. To be more specific about this last aspect, let (T, m, ω) be the lower, left, back corner of a rectangular cell S of lengths $\rho_1 \leq \delta_1$, $\rho_2 \leq \delta_2$ and $\rho_3 \leq \delta_3$ in the directions of the variables T, m and ω respectively. Then, for (T^*, m^*, ω^*) in S we have

$$|\tau(T, m, \omega) - \tau(T^*, m^*, \omega^*)| \leq B_1\rho_1 + B_2\rho_2 + B_3\rho_3$$

and thus

$$(T^*, m^*, \omega^*) \in S \quad \Rightarrow \quad \tau(T^*, m^*, \omega^*) \geq \tau(T, m, \omega) - B_1\rho_1 - B_2\rho_2 - B_3\rho_3$$

We can "ignore" S (except for $\tau(T, m, \omega)$) if, given the most current estimate τ^* for $\widehat{\tau}$, we have

$$\tau(T, m, \omega) - B_1\rho_1 - B_2\rho_2 - B_3\rho_3 \geq \tau^* - \epsilon_\tau$$

This last is equivalent to

$$\tau(T, m, \omega) \geq \tau^* - \epsilon_\tau + B_1\rho_1 + B_2\rho_2 + B_3\rho_3$$

That is the sense in which $\tau(T, m, \omega)$ is large enough to ignore the rest of S. The main challenge of this search method is the logical overhead of organizing it.

3.3.3 Refinement of the Search Over ω.

Suppose we have a bound $C > 0$ for $\left|\frac{\partial^2 \tau}{\partial \omega^2}\right|$ that is valid for all (T, m, ω) in the search space. Given fixed values for T and m, this permits use of a Taylor formula for τ of degree 2 in the change in ω:

$$\tau(T, m, \omega^*) = \tau(T, m, \omega) + \frac{\partial \tau}{\partial \omega}(T, m, \omega) \cdot (\omega^* - \omega) + E(\omega^*)$$

where

$$|E(\omega^*)| \leq \frac{1}{2} C |\omega^* - \omega|^2$$

Let τ^* be the current (continuously updated) estimate of $\widehat{\tau}$. If $\tau^* \leq \epsilon_\tau$ we are done because $0 \leq \widehat{\tau} \leq \tau^*$ and thus $|\tau^* - \widehat{\tau}| \leq \epsilon_\tau$. So assume $\tau^* > \epsilon_\tau$. We seek $\lambda > 0$ so that $\omega^* \in [\omega, \omega + \lambda]$ satisfy

$$\tau(T, m, \omega^*) \geq \tau^* - \epsilon_\tau$$

Clearly it suffices that, for $\omega^* \in [\omega, \omega + \lambda]$,

$$\tau(T, m, \omega) + \frac{\partial \tau}{\partial \omega}(T, m, \omega) \cdot (\omega^* - \omega) - \frac{1}{2} C (\omega^* - \omega)^2 \geq \tau^* - \epsilon_\tau$$

With $x = \omega^* - \omega$, this is equivalent $F + Gx - \frac{1}{2}Cx^2 \geq 0$, where F and G are the corresponding coefficients of the powers of x. Note that $F \geq \epsilon_\tau > 0$ because τ^* is the minimum value observed so far and $\tau(T, m, \omega)$ is one of the observations. We have $G^2 + CF > G^2$ and

$$x \in \left[\frac{G - \sqrt{G^2 + 2CF}}{C}, \frac{G + \sqrt{G^2 + 2CF}}{C}\right]$$

Note that $\left[0, \frac{G + \sqrt{G^2 + 2CF}}{C}\right]$ is always a subset of this interval. We can set $\lambda = \frac{G + \sqrt{G^2 + 2CF}}{C}$. Then $\lambda > 0$ and that we can "ignore" $\omega^* \in [\omega, \omega + \lambda]$.

There is a universal lower bound on λ. Note that, because the search space is compact and the function τ is is continuously differentiable, there is some universal bound $D \geq 0$ for the first derivative of τ with respect to ω. We use this to assert a universal lower bound on λ. Clearly, if $G \geq 0$, then $\lambda \geq \frac{\sqrt{2CF}}{C} \geq \sqrt{\frac{2\epsilon_\tau}{C}}$. Suppose $G < 0$ and $|G| \leq \sqrt{CF}$. Then,

$$G + \sqrt{G^2 + 2CF} \geq -\sqrt{CF} + \sqrt{2CF} \geq (\sqrt{2} - 1)\sqrt{CF}$$

and thus

$$\lambda \geq \frac{(\sqrt{2} - 1)\sqrt{CF}}{C} \geq \frac{(\sqrt{2} - 1)\sqrt{\epsilon_\tau}}{\sqrt{C}}$$

Finally, suppose $G < 0$ and $|G| > \sqrt{CF}$. Then,

$$\begin{aligned} G + \sqrt{G^2 + 2CF} &= \sqrt{G^2 + 2CF} - |G| \\ &= |G| \left[\sqrt{1 + 2CF/G^2} - 1 \right] \\ &> \sqrt{CF} \left[\sqrt{1 + 2CF/D^2} - 1 \right] \end{aligned}$$

and thus,

$$\lambda \geq \frac{\sqrt{CF} \left[\sqrt{1 + 2CF/D^2} - 1 \right]}{C} \geq \frac{\sqrt{\epsilon_\tau} \left[\sqrt{1 + 2C\epsilon_\tau/D^2} - 1 \right]}{\sqrt{C}}$$

By the previous paragraph, there is global bound in the number of steps searching through $\omega \in [6, 13]$, a bound that is valid at every T and m in the search space.

Clearly, and in practice, for very small step sizes this use of the second derivative speeds up the search. One aspect of the speed up is that, by using the first derivative of τ with respect to ω, there is a decided increase in step size when this derivative is positive.

3.3.4 Searching for Interior Critical Points of τ.

By interior critical points we mean (T, m, ω) in the interior of the search space where the gradient of τ is the zero vector. If one has a priori bounds on all second derivatives (including the mixed partial derivatives), one can use a binary tree search of rectangular cells to identify (many) cells of small volume where the norm of the gradient might be close to 0. This is probably far too much information for practical use, as the candidate cells might include local maxima or saddle points of τ. To reject a cell as having no critical points, one needs only one of the three partial derivatives to be significantly away from 0 at some point in the cell; clearly that helps to winnow the cells.

CHAPTER 4
EXPERIMENT USING THE JLS MODEL

4.1 Trading Strategy Experiment

The purpose of this experiment is to test the feasibility of the JLS model on real time stock data. The experiment began on March 1, 2016 and was concluded at the close of June 15, 2016. The pool of stock data used is the NYSE. Note that in the theory and derivation of the model, we can reformulate the JLS model to predict anti-bubbles [22]. In this experiment, we track the predictions of both bubbles (crashes) and anti-bubbles (rebounds), simultaneously. We note that by [18], crashes and rebounds do not necessarily mean opposite price behavior, but rather, the ceasing of current price behavior. We obtain stock data via the FinancialData command in Mathematica which receives its data from Yahoo! Finance. This experiment runs multiple data sets, each with a pre-determined window length, for each stock through a filter program to determine which stocks are exhibiting behavior that is unusual enough to be judged concave up and increasing or concave down and decreasing. A subset of these tested data sets become tracking stocks, which are monitored by the JLS model to produce prediction dates. Note, throughout this experiment, by days we mean trading days, not calendar days.

4.2 Fitting Method

For each stock in the NYSE, we consider 11 distinct windows, $L_i = [t_i, t_0]$, $i \in \{1, 2, \cdots, 11\}$ where t_0 is the date of the most recent price data and $t_i = t_0 - 64 - 60i$, $i \in \{1, 2 \cdots 11\}$. This gives us the 11 data sets per stock ranging from the last 6 months to the last 3 years. Each data set for each stock is then put through a filter program to determine whether that particular data set is exhibiting unusual behavior in either the positive bubble direction or negative bubble direction. Unusual in this context means by comparison with hypothetical price paths generated under a null hypothesis of Geometric Brownian motion for price paths (a common starting assumption in finance as in [16]).

4.3 Filter Program

Due to the previously described characteristics of bubble behavior, it seems to be a necessary trait of bubbles (both positive and negative) to be, in general, matching in direction and curvature. By this we mean we would expect a data set exhibiting a positive bubble to have log-prices that are generally increasing and concave up. Likewise, we would expect a data set exhibiting a negative bubble to have log-prices that are generally decreasing and concave down. The filter program we use has two steps. First, it screens each of the 11 data sets for each stock in the NYSE for this general behavior. Second, it assesses how unusual the observed statistic is among those computed from Geometric Brownian motion paths (under which the logarithms of prices are assumed to be Brownian motion with constant drift and volatility).

Given a data set for a given stock, the filter program first calculates the average slope of the entire data set. Next, it partitions the data set in two equal subsets. The average slope is calculated over each of the two subsets. Either the second slope is larger than the first or the second slope is smaller than the first. In the former case, the process is repeated for equal partitions of three subsets, four subsets, and so on until there is a partition of $p+1$ subsets where there exists an i-th slope smaller than the $i-1$-st slope, for at least one $i \in \{2, 3, \cdots, p+1\}$. We call the previous partition P and the test statistic for this data set is the integer p. In the latter case, the process is similarly repeated until there is a partition $q+1$ where there exists an i-th slope larger than the $i-1$-st slope, for at least one $i \in \{2, 3, \cdots, q+1\}$. We call the previous partition Q and the test statistic is the integer q. Finally, in the partition P for increasing prices or Q for decreasing prices, we check the first slope in that partition. If the first slope of partition P is positive or the first slope in partition Q is negative, then that data set will be compared to randomly generated Geometric Brownian motion price paths. Otherwise, it will be discarded.

Assuming a particular data set passes through the first process, it either has some integer valued test statistic $p \geq 2$ or $q \geq 2$. The next step in the filter program is to compare the observed statistic to those in a (more or less) permanent table of corresponding values for a set of 1000 Geometric Brownian motion paths. We count how many $p_k \geq p$ and how many $q_k \geq q$, $k \in \{1, 2 \cdots, 1000\}$ and use these counts to assign a p-value and q-value to the observed statistics. If either the p-value or q-value is less than a significance level of 0.05, then that data set is considered a filtered stock.

We randomly generated 1000 Geometric Brownian motion price paths with mean 0 and variance 1. We ran each of these 1000 paths through the partitioning process described in the previous paragraphs and generated 1000 test statistics for increasing slopes, $\{p_i\}_{i=1}^{100}$, and 1000 test statistics for decreasing slopes, $\{q_i\}_{i=1}^{100}$. We ordered these 1000 slopes in ascending order and formed a table of control test statistics. We created 11 such tables, one for each of the 11 window lengths, and used the same table for any data set with the appropriate window lengths. By Proposition 29 (below) we can use a single table for each window length regardless of the stock symbol, because the probability distribution of the statistics under the null hypothesis is independent of the drift and volatility assumed in the null hypothesis (as long as the volatility parameter is non-zero).

Proposition 29. *Using the filter program as described in Section 4.3, the significance level of a data set of length L is independent of the mean and variance for the randomly generated Geometric Brownian motion paths of length L.*

Proof. Let $W(t)$ be a Geometric Brownian motion process with mean μ and standard deviation $\sigma > 0$. As in [16, p.148], a sample path, $S(t)$, with constant drift and volatility can be described by

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$$

If we consider $\ln(S(t))$, by the Ito-Doeblin formula [16, p.147], we have

$$\begin{aligned}
d\ln(S(t)) &= \frac{1}{S(t)}dS(t) - \frac{1}{2S^2(t)}dS(t)dS(t) \\
&= \frac{1}{S(t)}\left[\mu S(t)dt + \sigma S(t)dW(t)\right] \\
&\qquad - \frac{1}{2S^2(t)}\left[\mu^2 S^2(t)dtdt + 2\mu\sigma S^2(t)dtdW(t) + \sigma^2 S^2(t)dW(t)dW(t)\right] \\
&= \mu dt + \sigma dW(t) - \frac{1}{2}\sigma^2 dt \\
&= \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dW(t)
\end{aligned}$$

It follows by integration that for $t_1 < t_2$,

$$\ln(S(t_2)) - \ln(S(t_1)) = \left(\mu - \frac{\sigma^2}{2}\right)(t_2 - t_1) + \sigma\left[W(t_2) - W(t_1)\right] \tag{4.1}$$

We use Equation 4.1 to generate simple Geometric Brownian motion paths in the filter program, where $W(t_2) - W(t_1)$ is a normal random variable with mean 0 and variance $t_2 - t_1$. Computing the slopes of these log-prices from t_1 to t_2, we have

$$\frac{\ln(S(t_2)) - \ln(S(t_1))}{t_2 - t_1} = \left(\mu - \frac{\sigma^2}{2}\right) + \frac{\sigma[W(t_2) - W(t_1)]}{t_2 - t_1} \tag{4.2}$$

Let $\varphi(S(t)) = \left(S(t)e^{-\left(\mu - \frac{\sigma^2}{2}\right)t}\right)^{1/\sigma}$. Since $S(t)$ is a stochastic process, then $\varphi(S(t))$ is a stochastic process. Using Equation 4.1 for $t_1 < t_2$, we have

$$\begin{aligned}
\ln(\varphi(S(t_2))) - \ln(\varphi(S(t_1))) &= \frac{1}{\sigma}\left[\ln(S(t_2)e^{-\left(\mu - \frac{\sigma^2}{2}\right)t_2}\right] - \frac{1}{\sigma}\left[\ln(S(t_1)e^{-\left(\mu - \frac{\sigma^2}{2}\right)t_1}\right] \\
&= \frac{1}{\sigma}\left[\ln(S(t_2)) - \ln(S(t_1))\right] + \left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right)(t_2 - t_1) \\
&= W(t_2) - W(t_1)
\end{aligned}$$

Computing the slopes of these log-prices from t_1 to t_2, we have

$$\frac{\ln(\varphi(S(t_2))) - \ln(\varphi(S(t_1)))}{t_2 - t_1} = \frac{W(t_2) - W(t_1)}{t_2 - t_1} \tag{4.3}$$

Comparing Equations 4.2 and 4.3 on two adjacent intervals, $t_1 < t_2 < t_3$, for $\sigma > 0$, we have

$$\begin{aligned}
&\left(\mu - \frac{\sigma^2}{2}\right) + \frac{\sigma[W(t_2) - W(t_1)]}{t_2 - t_1} < \left(\mu - \frac{\sigma^2}{2}\right) + \frac{\sigma[W(t_3) - W(t_2)]}{t_3 - t_2} \\
\iff &\frac{W(t_2) - W(t_1)}{t_2 - t_1} < \frac{W(t_3) - W(t_2)}{t_3 - t_2}
\end{aligned} \tag{4.4}$$

The if and only if applies if we substitute for the inequality in Equation 4.4 the symbols $\leq$, $>$, or $\geq$ as well. Consequently, in the filter program, $S(t)$ and $\varphi(S(t))$ will have the same test statistic. Let $\lambda(\cdot)$ be the name of this statistic; thus, $\lambda(S(t)) = \lambda(\varphi(S(t)))$. Finally, we show that $\omega \to \lambda(S(\cdot, \omega))$ is measurable with respect to the probability space (Ω, P, Σ) underlying the Brownian motion, so that we can apply probabilities to certain sets defined by the values of λ. The subset of Ω defined by $\lambda(S(t, \omega)) \leq r$ is equal to

$$\bigcup_{p=2}^{r+1} \left[\bigcup_{1<i\leq p} \left(\frac{W(t_{p,i}) - W(t_{p,i-1})}{t_{p,i} - t_{p,i-1}} > \frac{W(t_{p,i+1}) - W(t_{p,i})}{t_{p,i+1} - t_{p,i}} \right) \right], \text{ for positive bubbles} \tag{4.5}$$

and

$$\bigcup_{p=2}^{r+1} \left[\bigcup_{1<i\leq p} \left(\frac{W(t_{p,i}) - W(t_{p,i-1})}{t_{p,i} - t_{p,i-1}} < \frac{W(t_{p,i+1}) - W(t_{p,i})}{t_{p,i+1} - t_{p,i}} \right) \right], \text{ for negative bubbles} \tag{4.6}$$

where $\{t_{p,i}\}_{i=1}^{p+1}$ are the boundary index integers for dividing the integers $[1, L]$ into nearly equal consecutive subintervals with $t_{p,1} = 1$ and $t_{p,p+1} = L$. Given integers p and i with $p \in [2, r+1]$ and $i \in [2, p]$, we can write each component set in Equations 4.5 and 4.6 as

$$W(t_{p,i+1}) < (t_{p,i+1} - t_{p,i}) \left[\frac{W(t_{p,i}) - W(t_{p,i-1})}{t_{p,i} - t_{p,i-1}} \right] + W(t_{p,i}) \tag{4.7}$$

and

$$W(t_{p,i+1}) > (t_{p,i+1} - t_{p,i}) \left[\frac{W(t_{p,i}) - W(t_{p,i-1})}{t_{p,i} - t_{p,i-1}} \right] + W(t_{p,i}) \tag{4.8}$$

respectively. Note at time $t_{p,i+1}$ for $i \in \{2, \cdots p\}$, the values of $W(t_{p,i})$ and $W(t_{p,i-1})$ are known. Because $W(t, \omega)$ is a stochastic process, each set in Equations 4.7 and 4.8 is measurable for every integer p and i with $p \in [2, r+1]$ and $i \in [2, p]$. Consequently, by Equations 4.5 and 4.6, the set $\lambda(S(t, \omega)) \leq r$ is a measurable subset of Ω. □

4.4 Filtered Stocks

The filtered stocks represent the subset of tested stocks who exhibit positive, increasing slopes or negative, decreasing slopes. Each filtered stock is then run through the JLS model. We fit the model to the log-prices of the data set as in [4]. We first slave the linear parameters A, B, C, and D to the non-linear parameters T, m, and ω. For a data set of length $w_i = t_0 - t_i + 1 = 65 + 60i$, $i \in \{1, 2, \cdots, 11\}$, we use a three-dimensional grid search using the following bounds for the search space:

$$T \in [w_i + 1, w_i + 60], \qquad m \in [.1, .9], \qquad \omega \in [6, 13]$$

The bound selection for m and ω are as in [4]. The bound selection for T was chosen as we are only concerned with data sets for which we can apply a trading strategy that ends at the close of June 15, 2016. Searching for T values beyond three months is not useful for the purpose of this experiment. In addition to fitting the JLS model to the log-prices of the data set, we also fit a straight line to the log-prices, representing a Geometric Brownian motion fit to the data set prices. Once both models are fit to the log-prices, we calculate the efficiency of the JLS model to Geometric Brownian motion where

$$\text{Efficiency} = 1 - \frac{\text{least sum of squares of JLS model}}{\text{least sum of squares of Geometric Brownian motion}}$$

If Efficiency ≥ 0.75 and $T \in (w_i + 1, w_i + 60)$, then the data set is considered a tracking stock. From a mathematical point of view, we do not consider a data set to be a tracking stock if $T = w_i + 1$ or $T = w_i + 60$ since it is unclear whether such T values provide true minima for the error or all truly error-minimizing values of T satisfy $T < w_i + 1$ or $T > w_i + 60$. From a practical point of view, such T values may point toward the crash or rebound already happening or the crash or rebound occurring well off in the future. Any filtered stock data set fulfilling these two requirements is considered a tracking stock and is monitored on a daily basis for a prediction. We should note here that our definition for a tracking stock does not prohibit a single stock from having multiple data sets pass the filter program. While such instances have not been common, we have decided to keep data sets with the same ticker symbol and different window lengths. However, as seen in Section 4.7, we do not consider them distinct when applying a trading strategy.

4.5 Tracking Stocks

Each tracking stock has four traits:

1. ticker symbol
2. window length
3. date it passed the filter program
4. positive or negative bubble behavior

Since this experiment is conducted in real time, the date a data set passes the filter program and exhibits an efficiency ≥ 0.75 is, reasonably, the first instance it does so. At the beginning of the experiment or when conducting back-tests, it can be unclear which date to use for this trait. For this experiment, a reasonable first instance means that the data set has not passed the filter program prior to January 4, 2016 (the first trading day of 2016) but did pass the filter program at a date from January 4, 2016 or later. When monitoring a tracking stock, the window length provides the size of the data set with which to apply the JLS and Geometric Brownian motion models. Since the models are applied to the tracking stock each day, the data set changes slightly to remove the previous day's first data point (with respect to time) and adds the current day as the new last data point. For example, suppose a tracking stock has a window length of 245

Figure 4.1: Example of a CLIP: FBR 125 3/11/16 (-)

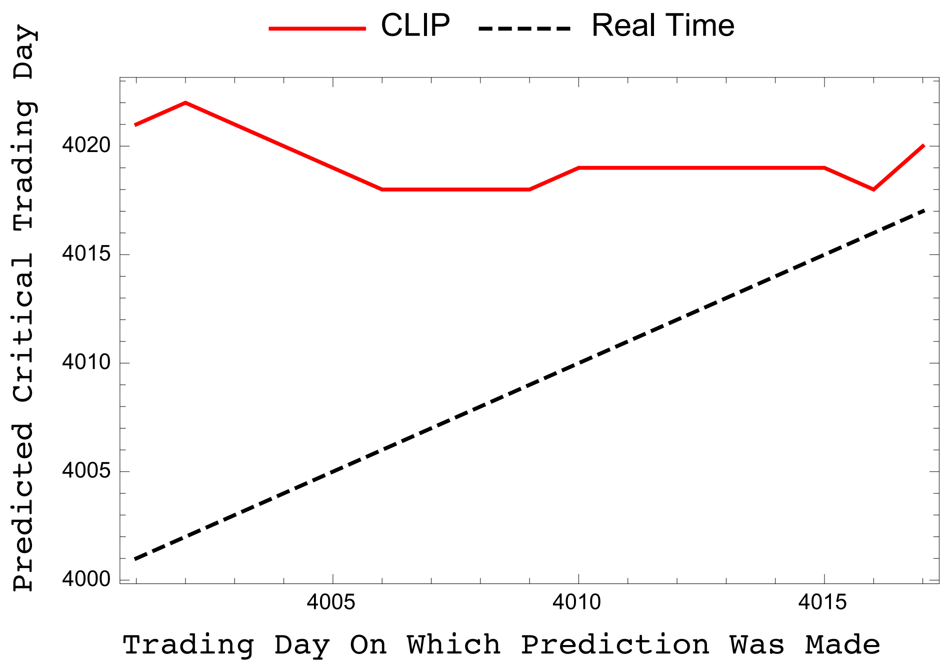
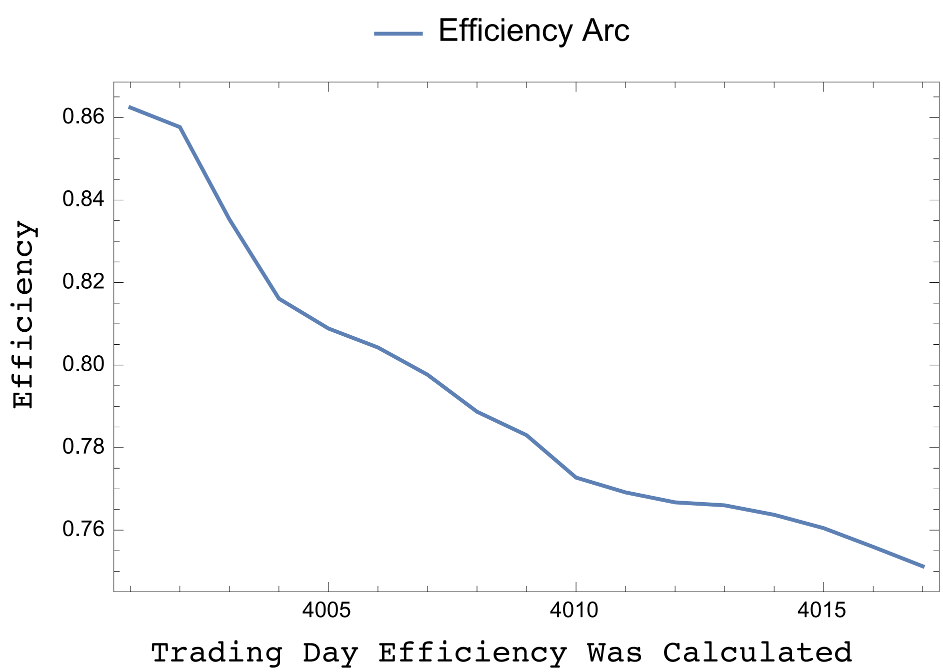


Figure 4.2: Example of an efficiency arc: FBR 125 3/11/16 (-)



days. On day n, the data set being tracked is $\{y_i\}_{i=n-244}^{n}$. On day $n+1$, the data set being tracked changes to $\{y_i\}_{i=n-243}^{n+1}$. Each day the JLS and Geometric Brownian motion models provide new values for the critical time, T, and efficiency. We can plot these two values over time to create crash lock-in plots (CLIPs) [6] and efficiency arcs, respectively. On the x-axis of a CLIP are the stocks' trading days (trading day 1 corresponds to the first day the stock was traded on the NYSE, trading day 2 corresponds to the second day the stock was traded on the NYSE, etc.) beginning with the day the stock first became a tracking stock. In Figure 4.1, the CLIP for FBR with window length 125 days is plotted from March 11, 2016 (trading day 4001) to April 5, 2016 (trading day 4017). On the y-axis of a CLIP is the adjusted critical date. It is adjusted to make consistent critical dates display as a horizontal line (see Figure 4.1). The dashed diagonal line represents real time. Ideally, CLIPs will stabilize over time as the JLS model consistently provides a fixed prediction date in the future for the impending crash or rebound. Efficiency arcs are plots with the stocks' trading days, beginning with the day corresponding to the date it became a tracking stock, on the x-axis and efficiencies on the y-axis (see Figure 4.2). We would hope that efficiency arcs behave nicely, in the sense that as a positive or negative bubble develops, it should become increasingly unlikely to be explained by Geometric Brownian motion. In this experiment, we assume this "other" dynamic to be explained by the log periodic power law model. In the next section, we develop three rules to identify a prediction date for each tracking stock.

4.6 Prediction Dates

We present a set of rules with which to make predictions for tracking stocks followed by several examples to illustrate how predictions are made. On the day the filtered stock becomes a tracking stock, we record the critical date and efficiency provided by fitting the JLS model. On each subsequent day, we refit the JLS model to the most recent data and record new values for the critical date and efficiency associated with that day. As time passes, we choose one of these critical dates to be the predicted critical date. We do so by the following rules:

1. Imminency: if the JLS model outputs a critical date of two days in the future or less (but strictly greater than 1), then that critical date becomes the predicted critical date.

2. CLIP Stability: in the absence of imminency, if the JLS model outputs at least 5 consecutive critical dates that stabilize to the same future date, then that date becomes the predicted critical date.

3. Max Efficiency: in the absence of both imminency and CLIP stability, the critical date with the highest corresponding efficiency becomes the critical date.

Note that these rules are ordered by precedence. As long as the critical date has yet to arrive, the rule governing the predicted critical date may change, but only in the upward direction (i.e. rule 3 to rule 2, etc).

Figure 4.3: CLIP for AUY 125 5/6/16 (+)

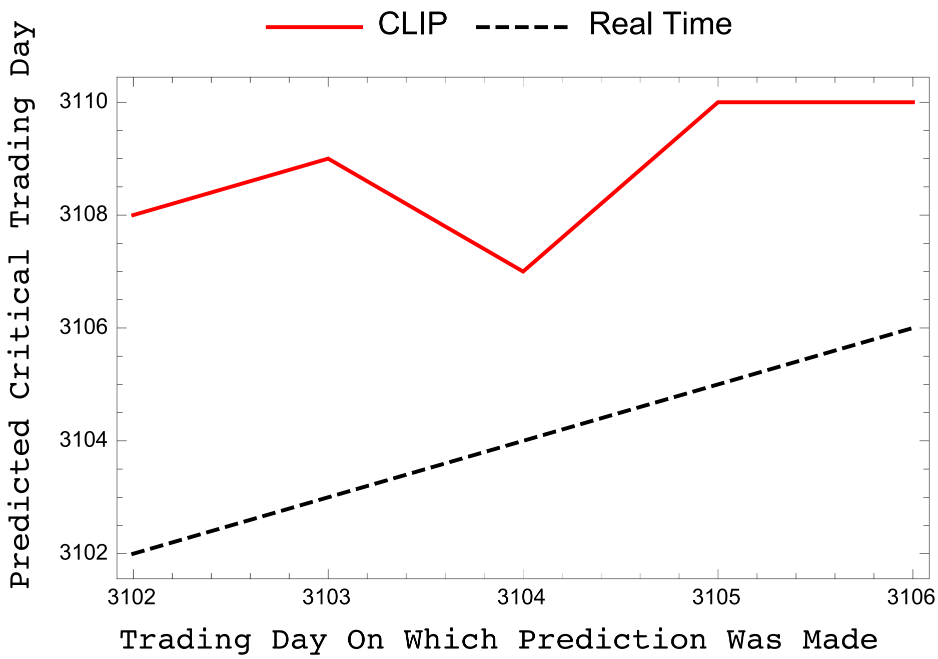
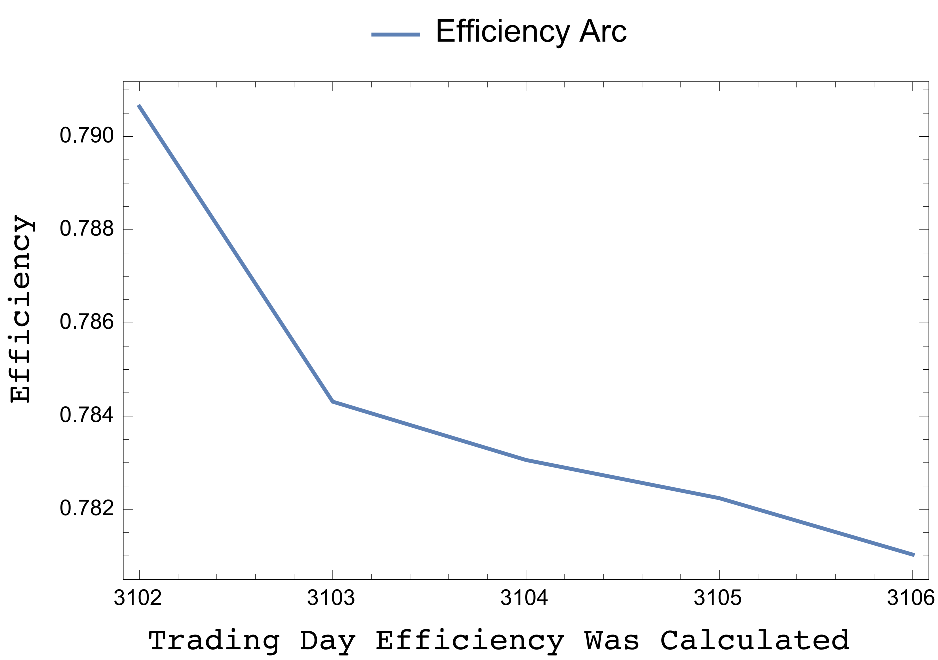


Figure 4.4: Efficiency arc for AUY 125 5/6/16 (+)



In Figures 4.3 and 4.4, notice the stock AUY had a maximum efficiency on the day it became a tracking stock on May 6, 2016 (day 3102). However, because on May 10, 2016 (day 3104) the JLS model predicted an imminent critical date of just two trading days later, we forewent the critical date associated with the maximum efficiency (May 16, 2016) in favor of the date associated with the prediction of two trading days in the future (May 13, 2016).

Figures 4.5 and 4.6 display an example where the initial predicted critical date was chosen by maximum efficiency. However, after the initial prediction was made, but before the initial predicted critical date arrived, the CLIP stabilized to a critical date different from the initial predicted critical date. By CLIP stability, we forewent the initial predicted critical date in favor of the new CLIP stabilized critical date. Note that even though the efficiencies began to increase after the CLIP stabilized, we still held the prediction date to the date determined by the CLIP stability. Here, the initial predicted critical date (February 19, 2016) was chosen based off of the JLS model fit on January 27, 2016 (day 6512). However, from February 11, 2016 (day 6523) to February 18, 2016 (day 6527), for 5 consecutive days, the CLIP stabilized to the critical date March 1, 2016. By order of precedence, we changed the predicted critical date from February 19, 2016 to March 1, 2016.

Figures 4.7 and 4.8 display an example where we defaulted to maximum efficiency in the absence of imminency and CLIP stability. For the stock CBD, we used the critical date given on January 28, 2016 (day 5007). Since the critical date is 24 trading days in the future, we used a predicted critical date on March 4, 2016.

4.7 Trading Strategy

With a set of rules with which to make predictions for tracking stocks, we can turn toward developing a trading strategy. It should be noted that not every predicted tracking stock will be traded. As described in Section 4.4, we consider multiple data sets from the same stock with different window lengths to be distinct for tracking purposes. However, to avoid a trading strategy that commits to investing in the same stock twice in the same time period (an impossibility) and to avoid skewing our tabulated results, we trade only on the predicted tracking stock exhibiting the earliest predicted critical date. Furthermore, we consider the earliest predicted critical date to be the predicted critical date for that particular bubble and we do not consider any other tracking stocks with the same stock symbol until the number of days passed beyond the previously traded-on predicted critical date is greater than or equal to the window length of the new tracking stock. Tracking stocks to be traded will be considered trade stocks.

For this experiment, there are two trading strategies we employ depending on whether a trade stock has positive or negative bubble behavior. For each trade stock, its prediction is in one of three states: success, pending, or failure. For positive (negative) bubbles, we take a short (long) position in the stock at the opening of the critical date. Call this the initial price. Going forward, the prediction state is pending if

Figure 4.5: CLIP for BRS 545 1/22/16 (-)

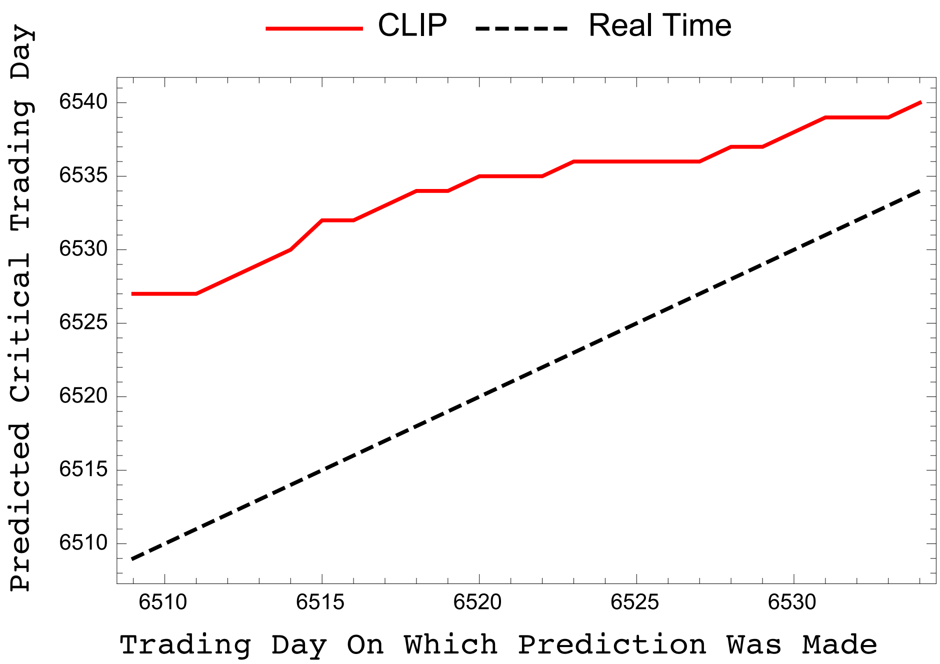


Figure 4.6: Efficiency arc for BRS 545 1/22/16 (-)

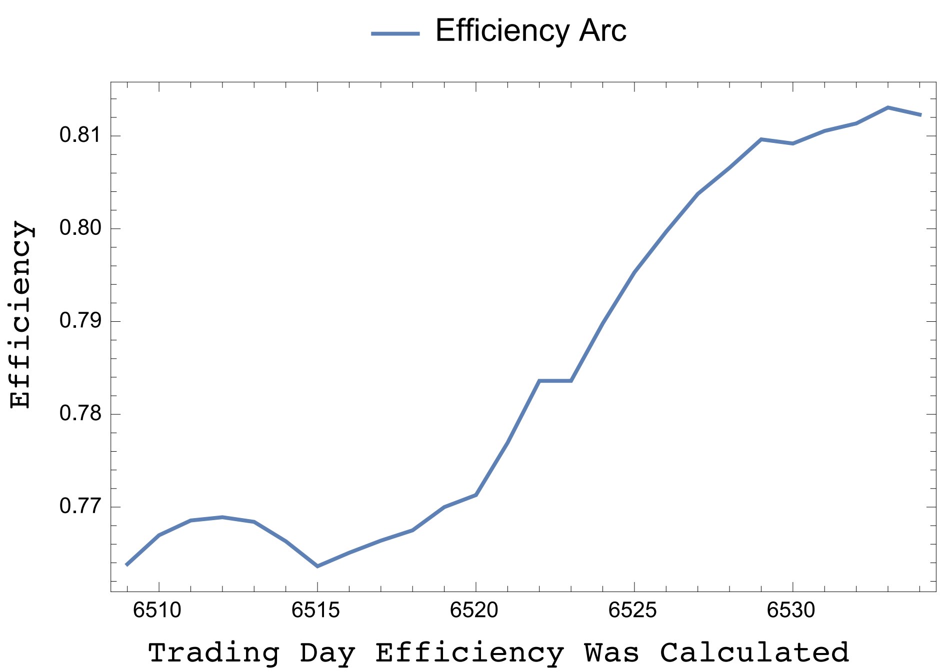


Figure 4.7: CLIP for CBD 605 1/7/16 (-)

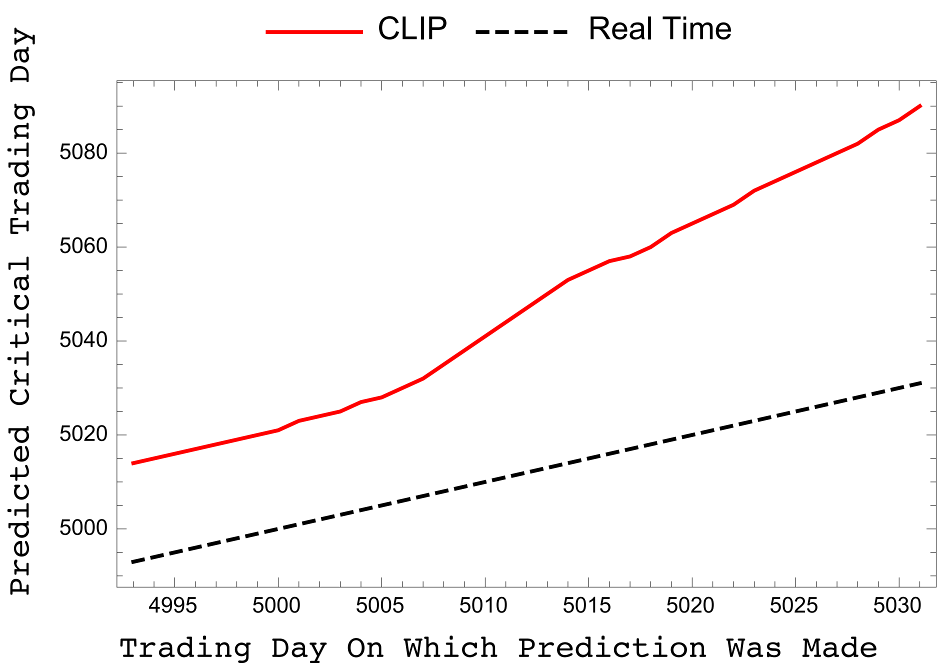
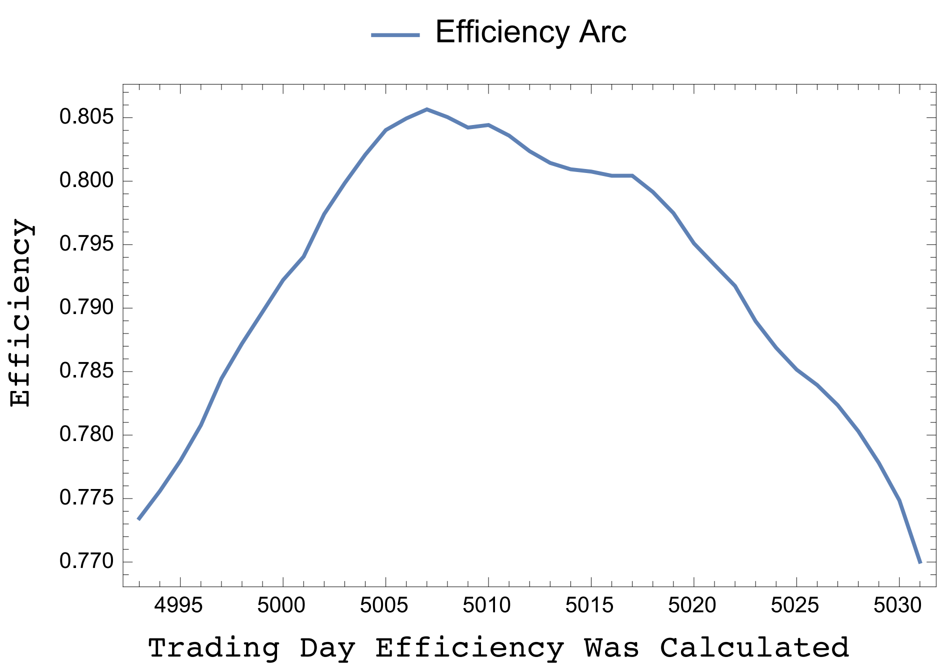


Figure 4.8: Efficiency arc for CBD 605 1/7/16 (-)



the price is within ±10% of the initial price. The prediction state is success if the stock price decreases (increases) at least 10% below (above) the initial price before the stock price increases (decreases) at least 10% above (below) the initial price. The prediction state is failure if the stock price increases (decreases) at least 10% above (below) the initial price before the stock price decreases (increases) at least 10% below (above) the initial price. We report our results in the next section.

4.8 Tabulated Results

We organize the following results table with each row consisting of the stock ticker symbol, window length, critical time (days from the date the prediction was made to the critical date), predicted critical date, opening price, target price (for success), bust price (for failure), current prediction state, and wait time (how many days it took for a success, given a success occurred). We print our results in two tables: Table 4.1 for positive bubbles and Table 4.2 for negative bubbles. All stocks with a pending status at the close of June 15, 2016 had their positions cleared using the lowest price on June 15, 2016.

Table 4.1: Positive bubble prediction results.

Stock	Window	Critical Time	Critical Date	Opening Price	Target Price	Bust Price	Prediction Status	Wait Time
EPR	185	1	03/08/16	$61.12	$55.01	$67.23	FAILURE	-
TSN	365	2	03/08/16	$65.35	$58.82	$71.89	CLEARED	-
FN	245	18	03/30/16	$33.04	$29.74	$36.34	FAILURE	-
AWK	305	2	04/01/16	$68.49	$61.64	$75.34	FAILURE	-
CPB	545	2	04/06/16	$64.15	$57.74	$70.57	CLEARED	-
ATO	365	7	04/20/16	$73.68	$66.31	$81.05	CLEARED	-
DFT	245	2	05/03/16	$40.41	$36.37	$44.45	FAILURE	-
COR	305	2	05/04/16	$76.98	$69.28	$84.68	CLEARED	-
HRL	245	49	05/06/16	$38.20	$34.38	$42.02	SUCCESS	13
AUY	125	2	05/13/16	$4.73	$4.26	$5.20	SUCCESS	7
KGC	125	7	05/18/16	$5.20	$4.68	$5.72	SUCCESS	1
MUX	185	2	05/31/16	$2.11	$1.90	$2.32	FAILURE	-
NSP	125	3	06/01/16	$72.12	$64.91	$79.33	CLEARED	-
HL	125	19	06/09/16	$4.65	$4.19	$5.12	SUCCESS	3
XYL	305	1	06/09/16	$46.15	$41.54	$50.77	CLEARED	-

With regard to Table 4.1, there are 15 traded stocks: 5 failures, 4 successes, and 6 cleared. With regard to Table 4.2, there are 43 traded stocks: 17 failures, 26 successes, and 0 cleared. The ultimate question this experiment tries to answer is whether it seems feasible that an investor could make money from utilizing the JLS model. Let us suppose an investor were to invest a fixed amount of money, X, on the opening of each critical date. Assuming all of X could be moved in and out of the market, a success would reward the investor with a positive gain of $0.1X$ and a failure would only leave the investor with a net loss of $0.1X$. From these two tables, it seems that investing in positive bubble stocks would be a relatively rare event compared to investing in negative bubble stocks. Clearing the positions of all positive bubble stocks at the close of June 15, 2016 using the lowest prices would net the investor $0.038X$ (rounded conservatively). Of completed predictions, the investor would have seen a net total loss of $-0.1X$ from a total investment of $15X$. This calculation ignores transaction costs and assumes that X can be fully invested for each stock. Investing in negative bubbles might be the more accelerated option. If an investor invested a fixed amount of money, X, on the opening of each critical date, the investor would have zero from the cleared positions on June 15, 2016 using the lowest prices since there were no pending stocks at the close of June 15, 2016. From the completed predictions, the investor would have seen a net total gain of $0.9X$ from a total investment of $43X$, again, ignoring transaction costs and imperfections. Note that because this experiment had a termination date and thus potentially some accounting due clearing positions in stocks with a pending status, in this experiment none of the positions in "rebound" stocks had a pending status. While there are many different ways an investor might go about moving money in and out of the stock market using the JLS model, from this experiment, it seems three and a half months is not a long enough time period to determine the feasibility of using the JLS model to make money off of positive bubble stocks. However, it may be feasible that an investor could use the JLS model to see positive gains in their portfolios for negative bubble stocks.

Table 4.2: Negative bubble prediction results.

Stock	Window	Critical Time	Critical Date	Opening Price	Target Price	Bust Price	Prediction Status	Wait Time
NMM	245	4	01/28/16	\$2.50	\$2.75	\$2.25	FAILURE	-
ATW	485	2	01/29/16	\$5.68	\$6.25	\$5.11	FAILURE	-
BXC	305	6	02/04/16	\$0.41	\$0.45	\$0.37	FAILURE	-
TGP	365	6	02/16/16	\$10.77	\$11.85	\$9.69	FAILURE	-
IPI	605	4	02/18/16	\$2.36	\$2.60	\$2.12	FAILURE	-
CHK	425	1	02/22/16	\$2.09	\$2.30	\$1.88	SUCCESS	0
PACD	665	5	02/23/16	\$3.70	\$4.07	\$3.33	SUCCESS	0
TDW	665	13	02/23/16	\$5.85	\$6.44	\$5.27	FAILURE	-
SDRL	665	1	02/26/16	\$1.76	\$1.94	\$1.58	SUCCESS	0
BRS	545	8	03/01/16	\$15.42	\$16.96	\$13.88	SUCCESS	2
CS	125	4	03/01/16	\$12.97	\$14.27	\$11.67	SUCCESS	2

Stock	Window	Critical Time	Critical Date	Opening Price	Target Price	Bust Price	Prediction Status	Wait Time
TOO	665	1	03/01/16	$2.95	$3.25	$2.66	SUCCESS	0
FCX	425	20	03/03/16	$8.66	$9.53	$7.79	SUCCESS	1
CBD	605	24	03/04/16	$12.38	$13.62	$11.14	SUCCESS	11
BHP	725	7	03/09/16	$26.58	$29.24	$23.92	FAILURE	-
MRO	305	8	03/09/16	$10.51	$11.56	$9.46	SUCCESS	5
SEMG	245	12	03/14/16	$20.60	$22.66	$18.54	FAILURE	-
GZT	365	8	03/15/16	$8.47	$9.32	$7.62	SUCCESS	23
DVN	305	12	03/17/16	$27.03	$29.73	$24.33	SUCCESS	17
GNW	305	11	03/17/16	$2.94	$3.23	$2.65	SUCCESS	2
BG	305	1	03/18/16	$55.71	$61.28	$50.14	SUCCESS	28
NR	605	15	03/18/16	$4.61	$5.07	$4.15	FAILURE	-
CJES	305	10	03/21/16	$1.88	$2.07	$1.69	FAILURE	-
PPP	725	1	03/22/16	$2.01	$2.21	$1.81	FAILURE	-
HSC	605	25	03/30/16	$5.64	$6.20	$5.08	FAILURE	-
FBR	125	1	04/06/16	$8.20	$9.02	$7.38	SUCCESS	11
EEP	365	25	04/08/16	$17.28	$19.01	$15.55	SUCCESS	4
MUR	605	26	04/08/16	$25.14	$27.65	$22.63	SUCCESS	2
NRP	605	41	04/13/16	$8.95	$9.85	$8.06	SUCCESS	1
KEY	125	32	04/18/16	$11.52	$12.67	$10.37	SUCCESS	26
WPX	665	24	04/18/16	$6.99	$7.69	$6.29	SUCCESS	0
EEQ	365	31	04/19/16	$19.70	$21.67	$17.73	SUCCESS	6
NAV	605	35	04/20/16	$13.64	$15.00	$12.28	SUCCESS	5
CDI	485	33	04/22/16	$7.50	$8.25	$6.75	FAILURE	-
DAC	185	9	04/26/16	$3.98	$4.38	$3.58	FAILURE	-
WDR	365	34	05/03/16	$19.99	$21.99	$17.99	SUCCESS	23
BCEI	605	42	05/04/16	$3.14	$3.45	$2.83	SUCCESS	0
LM	125	50	05/09/16	$31.06	$34.17	$27.95	SUCCESS	12
TAL	605	36	05/09/16	$15.15	$16.67	$13.64	FAILURE	-
WTI	665	48	05/10/16	$2.27	$2.50	$2.04	FAILURE	-
BTE	665	49	05/11/16	$4.55	$5.01	$4.10	SUCCESS	1
MOS	245	57	05/16/16	$24.76	$27.24	$22.28	SUCCESS	14
TK	245	57	06/02/16	$9.98	$10.98	$8.98	FAILURE	-

CHAPTER 5
CONCLUSION

Researchers like JLS and others are convinced the JLS model sufficiently explains numerous historical cases of both bursting bubbles and rebounding crashes. Skeptics have replied with several valid criticisms of the model. JLS responded to these criticisms with varying degrees of detail. This dissertation was motivated by two of these criticisms: (1) the uniqueness of the best fit and (2) the adequacy of the search method to locate a best fit. In the previous chapters, we have explored some of the more mathematical aspects of the JLS model as it attempts to diagnose, time, and predict terminations of both positive bubbles and negative bubbles.

In Chapter 2, we found we could extend the search space for m to $[0, 0.9]$ to expand the applicability of the JLS model and, in doing so, have also found increasing and concave up log-price vectors for which the first level JLS model is not a good fit. In addition, we also provided the beginnings of a new search method to locate the optimal m value of the first order JLS model to an increasing "extreme" vector in the following sense. By Theorems 22 and 23, we can potentially decrease the length of the interval of viable m values when searching for the optimal m that minimizes the objective error function for the first order JLS model on extreme vectors. Given some $L \geq 3$ and some extreme vector Y_k, if hypothesis of Theorem 22 is satisfied then there exists some $m_0 \in [0, 1)$ such that $F_T^*(m_0, 1) > 0$. By the continuity of $F_T^*(m, 1)$, there exists

$$m_1^* := \sup\{m \in (m_0, 1) \mid F_T^*(m, 1) \geq 0 \text{ and } F_T^*(\widehat{m}, 1) > 0, \forall \widehat{m} < m\}$$

Furthermore, by Proposition 21 there exists

$$m_2^* := \sup\{m \in (m_1^*, 1) \mid F_T^*(m, L) \geq L - k + 1 \text{ and } F_T^*(\widehat{m}, L) > L - k + 1, \forall \widehat{m} < m\}$$

Thus we can limit the search for m to $[m_1^*, m_2^*]$. Likewise, if the hypothesis of Theorem 23 is satisfied then there exists some $m_0 \in (0, 1]$ such that $F_T^*(m_0, L) < Y_k(L)$. By the continuity of $F_T^*(m, L)$, there exists

$$m_2' := \inf\{m \in (0, m_0) \mid F_T^*(m, L) \leq L - k + 1 \text{ and } F_T^*(\widehat{m}, L) < L - k + 1, \forall \widehat{m} > m\}$$

By Proposition 21 there exists

$$m_1' := \inf\{m \in (0, m_2') \mid F_T^*(m, 1) \leq 0 \text{ and } F_T^*(\widehat{m}, 1) < 0, \forall \widehat{m} > m\}$$

Thus, we can limit the search for m to $[m_1', m_2']$. Note, if $m_2' = 0$, then $m_1' = 0$ and the optimal m value is $m = 0$.

In Chapter 3, we found various upper bounds on varying moduli of continuity as well as first and second order derivatives for the parameters T, m, and ω using two methods. The first method is fairly direct, using the Cauchy-Schawrz Theorem to find upper bounds for the varying moduli of continuity for each of the 3 non-linear parameters. The second method used upper bounds on the first and second derivatives for each non-linear parameter and applied Taylor's Theorem to get an approximation for the objective function. We concluded with a description of an exhaustive method for searching through fine grids to confidently locate optimal fits within specified error bars.

In Chapter 4, we outlined an experiment aimed at determining the feasibility of utilizing the second level JLS model to make money via a specific trading strategy on the NYSE. The process included filtering stocks to determine which stocks are exhibiting unusual behavior relative to the Geometric Brownian Motion null hypothesis. These filtered stocks were then run through the JLS model to determine whether they were to be tracked for a critical date. Tracking stocks were then monitored daily and, using an explicit set of rules, were diagnosed with a predicted date. According to a specified trading strategy, predicted stocks were followed until either a success or failure presented itself or a clearing of positions was necessary. We reported multiple tables of results for both positive bubbles and negative bubbles.

While much experimentation has been performed on historical market data using the JLS model, a more rigorous mathematical exploration of its validity is still in its infancy. This dissertation was motivated by mathematically rooted questions regarding the fitting of the JLS model. After exploring the theoretical proofs regarding the model to increasing and concave up log-price curves, the numerical analysis of the robustness of the grid search method, and an experiment to determine the feasibility of the application of the JLS model to produce positive monetary results in the stock market, the JLS model's place as a mathematically viable financial model is still unclear. We hope this dissertation helps pave the way for more questions, more research, and more answers aiming to find a sliver of certainty in what has always been an uncertain world.

BIBLIOGRAPHY

[1] D. Brée, D. Challet, and P. Peirano. "Prediction Accuracy and Sloppiness of Log-Periodic Functions". In: *Quantitative Finance* 13.2 (2013), pp. 275–280. DOI: 10.1080/14697688.2011.607467. arXiv: 1006.2010 [q-fin.ST] (cit. on p. 2).

[2] D. Brée and N. Joseph. "Testing for Financial Crashes Using the Log Periodic Power Law Model". In: *International Review of Financial Analysis* 30.C (2013), pp. 287–297. DOI: 10.1016/j.irfa.2013.05.005. arXiv: 1002.1010 [q-fin.ST] (cit. on p. 2).

[3] B. Derrida, L. De Seze, and C. Itzykson. "Fractal structures of zeros in hierarchical models". In: *Journal of Statistical Physics* 33.3 (1983), pp. 559–569. DOI: 10.1007/BF01018834 (cit. on p. 6).

[4] V. Filimonov and D. Sornette. "A Stable and Robust Calibration Scheme of the Log-Periodic Power Law Model". In: *Physica A* 392.17 (2013), pp. 3698–3707. DOI: 10.1016/j.physa.2013.04.012. eprint: 1108.0099 (q-fin.GN) (cit. on pp. 1, 2, 33, 54).

[5] V. Filimonov and D. Sornette. "Power Law Scaling and "Dragon-Kings" in Distributions of Intraday Financial Drawdowns". In: *Chaos, Solitons & Fractals* 74.5 (2015), pp. 27–45. DOI: 10.1016/j.chaos.2014.12.002. arXiv: 1407.5037 [q-fin.ST] (cit. on p. 3).

[6] P. Geraskin and D. Fantazzini. "Everything You Always Wanted to Know about Log Periodic Power Laws for Bubble Modelling but Were Afraid to Ask". In: *European Journal of Finance* 19.5 (2011). DOI: 10.1080/1351847X.2011.601657 (cit. on pp. 2–6, 57).

[7] N. Goldenfeld. *Lectures on Phase Transitions and the Renormalization Group*. Westview Press, 1992. ISBN: 978-0-201-55409-0. DOI: 10.1088/0022-3719/6/7/010 (cit. on p. 4).

[8] J. Grantham. "Looking for Bubbles Part One: A Statistical Approach". URL: https://www.gmo.com/docs/default-source/research-and-commentary/strategies/gmo-quarterly-letters/looking-for-bubbles-and-in-defense-of-risk-aversion.pdf (cit. on p. 7).

[9] John Hussman. *Investment Research & Insight*. Hussman Funds. URL: http://www.hussmanfunds.com/researchInsight.html (cit. on p. 7).

[10] Z. Jiang et al. "Bubble Diagnosis and Prediction of the 2005-2007 and 2008-2009 Chinese Stock Market Bubbles". In: *Journal of Economic Behavior & Organization* 74.3 (2010), pp. 149–162. DOI: 10.1016/j.jebo.2010.02.007. arXiv: 0909.1007 [q-fin.ST] (cit. on p. 2).

[11] A. Johansen, O. Ledoit, and D. Sornette. "Crashes as Critical Points". In: *International Journal of Theoretical and Applied Finance* 3.2 (2000), pp. 219–255. DOI: 10.1142/S0219024900000115. arXiv: cond-mat/9810071 [cond-mat] (cit. on pp. 3–7).

[12] T. M. Liggett. *Interacting Particle Systems*. Springer-Verlag, 2005. ISBN: 978-3-540-26962-5. DOI: 10.1007/b138374. URL: http://www.springer.com/us/book/9783540226178 (cit. on p. 5).

[13] T. M. Liggett. “Stochastic Models of Interacting Systems”. In: *The Annals of Probability* 25.1 (1997), pp. 1–29. DOI: 10.1214/aop/1024404276 (cit. on p. 5).

[14] Y. Malevergne and D. Sornette. *Extreme Financial Risks: From Dependence to Risk Management*. Springer-Verlag, 2006. ISBN: 978-3-540-27266-3. DOI: 10.1007/b138841. URL: http://www.springer.com/us/book/9783540272649 (cit. on pp. 1, 7).

[15] R. Shiller. *Shiller PE Ratio*. Multpl. URL: http://www.multpl.com/shiller-pe/ (cit. on p. 7).

[16] S. Shreve. *Stochastic Calculus for Finance II: Continuous Time Models*. Springer-Verlag, 2004. ISBN: 978-0-387-40101-0. URL: http://www.springer.com/us/book/9780387401010 (cit. on pp. 1, 51–53).

[17] A. Smithers. *US CAPE and Q Chart*. Ed. by Andrew Smithers. URL: http://www.smithers.co.uk/page.php?id=34 (cit. on p. 7).

[18] D. Sornette. “Critical Market Crashes”. In: *Physics Reports* 378.1 (2003), pp. 1–98. DOI: 10.1016/S0370-1573(02)00634-8. arXiv: cond-mat/0301543 [cond-mat.stat-mech] (cit. on pp. 7, 51).

[19] D. Sornette. *Financial Crisis Observatory*. ETH Zürich: Department of Management, Technology and Economics. URL: http://www.er.ethz.ch/financial-crisis-observatory.html (cit. on p. 2).

[20] D. Sornette et al. “Clarifications to Questions and Criticisms on the Johansen-Ledoit-Sornette Bubble Model”. In: *Physica A* 392.19 (2013), pp. 4417–4428. DOI: 10.1016/j.physa.2013.05.011. arXiv: 1107.3171 [q-fin.GN] (cit. on pp. 2, 8, 33).

[21] H. E. Stanley. *Introduction to Phase Transitions and Critical Phenomena*. University Oxford Press, 1989. ISBN: 978-0-195-05316-6. DOI: 10.1002/qua.560350412. URL: http://www.oupcanada.com/catalog/9780195053166.html (cit. on p. 4).

[22] Y. Wanfeng et al. “Detection of Crashes and Rebounds in Major Equity Markets”. In: *International Journal of Portfolio Analysis and Management* 1.1 (2012), pp. 59–79. DOI: 10.1504/IJPAM.2012.046909. arXiv: 1108.0077 [q-fin.GN] (cit. on pp. 2, 51).