THE EXPLICIT FINITE DIFFERENCE METHOD: OPTION PRICING UNDER STOCHASTIC VOLATILITY

April 28, 2013

Abstract

This paper provides an overview of the finite difference method and its application to approximating financial partial differential equations (PDEs) in incomplete markets. In particular, we study German's [6] stochastic volatility PDE derived from indifference pricing. In [6], it is shown that the first order-correction to derivatives valued by indifference pricing can be computed as a function of the stochastic volatility PDE. We present three explicit finite difference models to approximate the stochastic volatility PDE and compare the resulting valuations to an Euler-Muruyama Monte Carlo pricing algorithm. We also discuss the significance of boundary condition choice for explicit finite difference models.

Acknowledgements

I would like to thank Professor Henry Schellhorn for his patience, flexibility, and clear explanations. I am also extremely grateful to Dr. David German for suggesting and building my interest in the research topic, for many helpful discussions and comments, and for his continued advising. Without the help of Nathan Lenssen, I would not have been able to compile this project, and I am thankful to him for sharing his knowledge of programming at odd hours. In addition, I appreciate the guidance I have received from Kees de Graaf and Dr. Domingo Tavella, as well as Professor Mark Huber. Lastly, I thank my loving family and my friends for their enduring support.

Contents

1 Introduction

1.1 Background to Research Problem

By nature, the structure of financial markets gives rise to derivative trading. Markets depend on the exchange of one asset for another, and the process of trading is composed of two distinct exchange processes [12]. First, the agents of a transaction must agree on "fair prices" for the assets involved in the trade. This is the issue of settling exactly what that will be exchanged. Second, terms for the delivery of the trade are necessary to specify how and when the assets will be exchanged. The matter of delivery best motivates the development of derivatives.

Depending on the terms outlined in a transaction, payment and delivery are not constrained to take place at the same time. This flexibility "fills out" the market and allows for participants to enter trading positions today that depend on future events. These transactions derive their value from future events, and they help bring liquidity to typically illiquid markets. For example, Poitras documents use of derivative securities as early as the 1600s in Antewerp where a "to arrive" contract market emerged for grain and commodities still in transport, supporting the expansion of trade by increasing the liquidity of the commodities market as a whole [12]. Nonetheless, a more robust marketplace introduces additional complexity, a consequence of the task of determining a "fair" price for a more complex security.

The basics of option pricing are founded on the principle of *complete markets*, financial markets where the payoffs of every security in the market at every possible state of the world can be constructed as a combination of the other, existing securities. In complete markets, a replication strategy forms a portfolio of securities such that the payoffs of the constructed portfolio equal the payoffs of the original security in every possible state of the world. The price of the initial security then can be inferred from the known prices of the securities included in the replicating portfolio. Since any market participant can reconstruct the security in similar fashion, by extension of the argument, the price for the security must be unique. Otherwise, an economic agent could exploit the price difference between the security and the replicating portfolio to "buy-low and sell-high" and secure a riskless profit. Explicitly, consider a replication strategy V for a security X on a finite probability space Ω. V is then a function that takes as input a state (denoted by ω), begins with some initial capital X_0, and generates exactly the same payoff at every time t as the security X over all possible states of the world ($\omega \in \Omega$) such that $X_t(\omega) = V_t(\omega)$, $\forall t > 0$ and $\forall \omega \in \Omega$.

This unique price condition precludes arbitrage, and it allows for valuation based on the assumption that all economic agents in the market are risk-neutral, an assumption that will be relaxed when we move to markets that better approximate reality. Most famously, Merton Black and Myron Scholes utilized this assumption to show that a European call option could be replicated with only the underlying asset and a bank account. Given the right proportions of the stock and the bond, the payoffs of a European call can be replicated regardless of the underlying asset price movements over the life of the option. As a result, the value of the security is independent of

the probability the underlying moves up or down in reality and is instead dependent of a *risk-neutral probability measure*. Under this probability measure, all assets are "fair" investments, and the price of any asset can be computed as the mathematical expectation under the risk-neutral probability measure of $V_T(\omega)$ (adjusted for the time value of money) where T denotes the terminal time period and ω the outcome at time T.

For example, without loss of generality, assume a zero risk-free interest rate and suppose that our state space Ω consists only of two states, {Heads, Tails} for a coin flip with "reality" probability measure $\mathbb{P}$ such that p and $1-p$ are the probabilities of heads and tails, respectively. Shreve [15] shows how we can find a risk-neutral probability measure $\tilde{\mathbb{P}}$ where $\tilde{p}$ and $1-\tilde{p}$ are the probabilities of heads and tails in a risk-neutral world. With a risk-neutral probability measure, we can then determine the derivative security's initial value X_0 by solving

$$X_0 = \tilde{\mathbb{E}}_{\forall\omega\in\Omega}\left[V_T(\omega)\right] \tag{1}$$

for any termination time T, replication strategies $V(\omega)$ as before, and mathematical expectation $\tilde{\mathbb{E}}$ under the $\tilde{\mathbb{P}}$ probability measure. If $T = 1$, then we solve equation (1) for X_0 as

$$X_0 = (\tilde{p})(V_T(Heads)) + (1-\tilde{p})(V_T(Tails)). \tag{2}$$

In real markets, however, determining risk-neutral probability measures becomes much more complicated. Relative to the simplified financial market model above, real markets are increasingly robust (in terms of securities available to meet specific needs), but they are frequently *incomplete*. This means that generally it is not possible to replicate every security in a real market by a combination of the other existing, traded securities. For example, if a security has payoffs that are dependent on exogenous factors beyond market prices, both buyer and seller are exposed to non-hedgable risk when trading securities [6]. For example, if the asset that underlies our contingent claim exhibits stochastic volatility rather than constant volatility, it is not possible to hedge the random volatility factor unless there is a traded asset in the associated volatility (either artificially constructed or already existing in the marketplace). Even volatility indices such as the VIX may present systematic errors when calculating a the "true" volatility [9]. Hence, barring artificial completion of incomplete markets, the classic no-arbitrage argument is not used to value derivative securities in incomplete markets. We must then consider alternative valuation approaches that account for market incompleteness.

1.2 Incomplete Markets: Indifference Pricing

One such method used to value contingent claims in incomplete markets is *indifference pricing* [2]. Under incomplete markets, since exact replication is not possible, there exists some non-hedgable risk associated with buying or selling a derivative. This risk, and an investor's appetite for holding it, can be quantified with a utility function.

The introduction of a investor preferences with a utility function can be thought of in the framework of hedging and investing. In complete markets, construction of a replicating portfolio allows an investor to *hedge*, or mitigate the risk of holding a derivative by shorting the replicating portfolio. Conversely, in markets barring replication, a position in an option without a perfect hedge represents an *investment*, an attempt to capitalize on a particular view of the market by taking a position. In incomplete markets without replication possibilities, buying or selling a derivative security involves non-hedgable risk, and the decision to go-through with buying or selling a derivative can should be assessed in the context of furthering an *investment* objective [3].

For example, consider an economic agent has initial wealth x_0. The agent is faced with the decision to invest this initial endowment immediately in a number of different trading strategies all requiring x_0 initial capital. These trading strategies are random variables on a certain probability space Ω with filtration $\mathcal{F} = \mathcal{F}_T, (\mathcal{F}_t)_{0 \leq t \leq T}$ and probability measure $\mathbb{P}$ such that $\forall \pi \in \theta, X^{\pi}_{x_0}(t)$ represents the value of the portfolio at time t when the investor deploys x_0 initial wealth in trading strategy $\pi \in \theta$, where θ denotes the entire set of possible trading strategies. For trading strategies 1 and 2 both requiring initial capital x_0 in θ, the agent has no reason to believe that the time one values $X^1_{x_0}(1)$ and $X^2_{x_0}(1)$ differ from each other. If the time one values of the strategies did differ, then the two strategies would require different amounts of initial capital as determined by the laws of supply and demand. In greater detail, beginning with initial wealth x_0, if one strategy were known to yield a higher payoff for every $\omega \in \Omega$, demand dictates that it would cost more to enter this strategy, thereby diminishing the amount of the trading strategy x_0 can buy at time zero and equalizing the return. Still, the agent is faced with the issue of deciding the optimal investment strategy.

Following Davis [2], suppose the investor's preferences are characterized by a utility function

$$U : \mathbb{R}_+ \to \mathbb{R}_+$$

such that for all portfolios of value x,

$$\begin{aligned} U'(0) &= \lim_{x \to 0} U'(x) = \infty \\ U'(\infty) &= \lim_{x \to \infty} U'(x) = 0. \end{aligned} \tag{3}$$

Additionally, U is assumed to be non-decreasing and continuous. The introduction of a utility function then allows for indirect comparison between two strategies' values at time one, $U(X^1_{x_0}(1))$ and $U(X^2_{x_0}(1))$, and it addresses the initial no-arbitrage problem. However, these quantities are themselves random variables, so we cannot compare them directly. Rather, we must consider the mathematical expectation of each for direct comparison.

Thus, to compare strategies $X^1_{x_0}$ and $X^2_{x_0}$ and the portfolios $X^1_{x_0}(1)$ and $X^2_{x_0}(1)$ at time one, we must solve the optimization problem to maximize the investor's expected utility over all possible portfolios resulting from trading strategies $X^{\pi}_{x_0}$ for $\pi \in \theta$ at

terminal time one

$$V(x) = \sup_{\pi \in \theta} \mathbb{E}[U(X^{\pi}_{x_0}(1))] \tag{4}$$

The above equation is the expected utility of wealth at time one, and a rational agent seeks to maximize this terminal value. Given an optimal portfolio $X^{\pi^*}_{x_0}$, we can distinguish between portfolios by buying (or selling) an investment in a European option priced at time zero at p with nonnegative, random payoff B that is $\mathcal{F}_1$ measurable. With purchase price p at time zero, Davis [2] shows how we can use a marginal rate of substitution argument to determine how much of the investor's initial wealth x_0 can be diverted to purchase (or sell) the option security without affecting the investor's maximum utility. This quantity $U'(X^{\pi}_{x_0}(1))$ is often easier to calculate than the pure terminal wealth $U(X^{\pi}_{x_0}(1))$ [6]. Explicitly, this results in the function

$$W(\delta, x, p) = \sup_{\pi \in \theta,\ \delta \in \mathbb{R}} \left\{ \mathbb{E}\left[U\left(X^{\pi}_{x_0 - \delta} + \frac{\delta}{p} B \right) \right] \right\}. \tag{5}$$

where δ represents the nominal cash used to purchase $\frac{\delta}{p}$ options at price p. The function W then represents the conjugate function of the investor's utility function and describes how the investor's terminal utility is affected by the introduction of $\frac{\delta}{p}$ costing δ while the remaining cash $x_0 - \delta$ is invested in the optimal trading strategy $X^{\pi^*}_{x_0}(1)$. That is, we are attempting to capture how an investor's utility changes from the optimal solution where marginal utility equals zero by introducing a position in the derivative security.

Further mechanics of the solution are beyond the scope of this paper (interested readers are referred to [14], [2], [7]), but German [6] notes that in the general, incomplete financial model we study above (where an investor is endowed at initial time with cash and some amount of non-traded contingent claims), explicit solution is impractical, moreover it is often not possible. Instead, the prices generated by this method are linear approximations rather than "true" values. To calculate the option prices with more accuracy, it is first necessary to compute the approximation. It is the goal of this paper to approximate the stochastic volatility (SV) model that German [6] derives by indifference pricing. In doing so, the resulting approximation may be used in turn to compute the correction factor to the option prices that it generates.

The remaining paper is partitioned into four sections. Section 2 outlines the fundamentals of numerical PDE approximation theory necessary to solve German's [6] SV model. The third section advances the study of numerical option pricing by reviewing the Black-Scholes framework and analyzing it with the approximation techniques discussed in section two. Section 4 section extends Black-Scholes to German's [6] SV framework that better encompasses real-world complexities. It applies the explicit finite difference method to approximate German's SV PDE with three models. Finally, Section 5 investigates the resulting numerical approximations and compares them to prices generated by Monte Carlo simulation. Section 5 also discusses avenues for further research.

2 PDEs and the Finite Difference Method

In engineering, the natural sciences, and economics, mathematical models involving differential equations have helped describe systems of interest [10]. In 1973, financial markets saw a major breakthrough in modeling the fair price of derivative securities following Myron Scholes and Fischer Black's European option pricing formula. Of particular significance is the equation's capability to "relate spot returns to cross-sectional properties of the options [8]. These relationships are derived from studying the dependent variable (option value) as its independent parameters (time or underlying price) vary. The relationships thus lead to a formula involving partial derivatives. Such formulae are classified as partial differential equations (PDEs) and are highly non-trivial to solve analytically. Thus, alongside increases in computational capability, the numerical approximation of PDEs has taken on increased importance.

Computational advances have stimulated research into the theory of efficient numerical solution techniques. Most numerical treatment of PDEs fall into one of two broad categories:

- Solution
- Simulation

Solution techniques approximate the value of the PDE at a finite number of points. This paper employs an explicit technique where the PDE values at a certain time are based on approximations at previously calculated points (that ultimately trace their origins to an initial condition, i.e. option payoff at expiry). Simulation techniques, such as Monte Carlo methods, value options by sampling many different, possible underlying asset price paths from the governing PDE. The value of a specific option is then approximated from the randomly generated outcomes of the price path simulation.

Often, the level of complexity involved in implementing a particular solution technique mirrors the complexity of the PDE itself [1]. It is therefore convenient to have some means of classifying a PDE so that undue complexity in implementing a solution technique may be avoided. Note that the author limits the scope of classification to that which is relevant for understanding the stochastic volatility PDE under analysis in Sections 3, 4, and 5.

2.1 Classification of PDEs

Consider a generic PDE of the form

$$a(x,y)\frac{\partial u}{\partial^2 x} + b(x,y)\frac{\partial u}{\partial x \partial y} + c(x,y)\frac{\partial u}{\partial^2 y} + d(x,y)\frac{\partial u}{\partial x} + e(x,y)\frac{\partial u}{\partial y} + fu + g = 0 \quad (6)$$

Classification begins with defining a PDEs *order*. This is a measure of the PDEs highest order derivative, and in the example PDE (6), the second derivative is its highest order, so equation (6) a *second-order* PDE. Furthermore, it is a *linear* PDE

because the coefficients of the unknown function u depend only on the independent variables x and y, and not u itself. Solving linear PDEs requires significantly less computational effort than non-linear PDEs such as

$$a\left(x, y, \frac{\partial u}{\partial^2 y}\right)\frac{\partial u}{\partial^2 x} + b(x,y)\frac{\partial u}{\partial x \partial y} = 0 \tag{7}$$

where the first term depends on the independent variables x and y as before, but also the dependent variable u. Were this a formula describing the option value, it would imply that the option value (at any given specification of parameters) also depends on the option value's rate of change.

Lastly, depending on the value of $b(x,y) - 4a(x,y)c(x,y)$ in equation (6), the PDE can be classified as one of three *types*:

- Elliptic, if $b^2 4ac < 0$, e.g. $a = c = 1$ and $b = 0$ and $\dfrac{\partial^2 U}{\partial x^2} + \dfrac{\partial^2 U}{\partial y^2} = 0$,
- Parabolic, if $b^2 4ac = 0$, e.g. $b = c = 0$ and $\dfrac{\partial^2 U}{\partial t^2} = k\dfrac{\partial^2 U}{\partial x^2} = 0$, or
- Hyperbolic, if $b^2 4ac > 0$, e.g. $4\rho^2 > 0$ and $\dfrac{\partial^2 U}{\partial t^2} - \rho^2\dfrac{\partial^2 U}{\partial x^2} = 0$.

Depending on its type, a PDE may require two types of auxiliary conditions to specify, or "tie down, the particular solution of interest from the field of possible solutions. The first type of conditions are *initial conditions*, which are conditions that must be satisfied everywhere for a single, given value of an independent variable [10]. Intuitively, these provide a "starting point" for the PDE. The second set of conditions are *boundary conditions*, and these must be satisfied everywhere along the PDE's boundary region. They help describe the solution's behavior as the independent variables take on their minima and maxima.

Deriving initial conditions is often relatively straightforward. For option pricing, initial conditions describe the payoff of the security at maturity, and they compel backward induction techniques. Boundary conditions, on the other hand, come in three varieties

- Dirichlet, which explicitly describe the value of the function along the boundary,
- Von Neumann, which describe the derivative of the function along the boundary, and
- Robin, which specify a point-by-point combination of the preceding two types

and are derived by studying the theoretical behavior of the PDE along its edges.

In option pricing, the independent parameters often specify a deterministic function along the boundary of the PDE, resulting in Dirichlet conditions. For example, in the Black-Scholes framework, an almost certain-exercise vanilla European Call option has time as the only effect on option value, and thus the boundary condition reflects

the time-value-of money principle. For the PDEs throughout this paper, Dirichlet boundary conditions are primarily used.

Of the three types of second order PDEs, an elliptic equation results in a boundary value problem where boundary conditions alone are required; a parabolic equation gives rise to an initial boundary value problem where both boundary and initial conditions are required; and a hyperbolic equation results in an initial value problems where initial conditions alone are required to specify the particular solution [10].

2.1.1 Ex: Classification of the Black-Scholes Equation

Now consider the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{8}$$

where V represents the value of the option, S represents the stock or underlying asset, σ denotes the constant volatility, r is the per-period interest rate, and t represents time.

From our prior overview, we can describe the above Black-Scholes equation as a linear, second-order, parabolic PDE. As such, it requires an initial condition paired with multiple boundary conditions.

2.2 The Finite Difference Method (FDM)

2.2.1 Solution Mesh Discretization for FDM

The FDM approximates a PDE at a finite number of points [16]. The process of translating a continuous PDE to a discrete set of points is called *discretization*. Often, the first step in the discretization process is describing the solution domain. Each variable that appears in the PDE must be discretized to describe the set of points for which the method will approximate the PDE's value. For example, consider the standard heat equation (second order, linear, parabolic) in two variables, length (x) and time (t), with an arbitrary constant of diffusivity $k \in \mathbb{R}$ that describes the temperature of an infinitely thin and arbitrarily long rod over time

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2}. \tag{9}$$

Suppose the portion of the rod that is of interest has endpoints x_{min} and x_{max}. This continuous sequence can be translated to a discrete approximation with a total of m nodes forming the equally spaced sequence $[x_{min}, x_{max}]$. For $\{1, 2, ..., m\}]$, the sequence describing the nodes along the rod where temperature will be computed can be represented as

$$[1\Delta l, 2\Delta l, ..., m\Delta l] \tag{10}$$

where $m\Delta l = x_{max}$. Similarly, we must define a period of time where in which we are interested in the rod's temperature. Suppose time is a sequence of discrete

points defined on $[t_{min}, t_{max}]$ with n equally spaced (with increment Δt) nodes. These two discretizations form a lattice of points where each point ($time$, $length$) can be represented as

$$(t_t, x_l) = (n\Delta t, m\Delta l) \tag{11}$$

for $n = 0, 1, ..., n$ and $m = 0, 1, ..., m$ that describe the solution of the PDE at every combination of length and time. For more convenient notation moving forward, we define

$$u(t_t, x_l) \equiv u_l^t. \tag{12}$$

2.2.2 PDE Discretization for FDM

After describing the solution lattice, we now focus on determining the numerical approximation at at each node (mesh point) in the lattice by substituting discrete approximations for each partial derivative term in the PDE in much the same way the lattice was created. Ultimately, the FDM hinges on the discretization of each partial derivative in a given PDE by approximating its analytic form with a difference quotient derived from Taylor expansion [1].

For example, consider a Taylor approximation of a function $v : \mathbb{R} \to \mathbb{R}$ around $x + h$

$$v(x+h) = v(x) + hv'(x) + \frac{1}{2}h^2v''(x) + \frac{1}{6}h^3v''' + ... = \sum_{i=0}^{n} \frac{1}{i!}h^i v^{(i)}(x) \tag{13}$$

and similarly around $x - h$

$$v(x-h) = v(x) - hv'(x) + \frac{1}{2}h^2v''(x) - \frac{1}{6}h^3v''' + ... = \sum_{i=0}^{n} (-1)^i \frac{1}{i!}h^i v^{(i)}(x) \tag{14}$$

where the (i) notation on$v$ denotes a derivative of ith order. Neglecting the terms of second-order or higher, simple rearrangement yields an approximation for the first derivative

$$v'(x) \approx \frac{v(x+h) - v(x)}{h}, \tag{15}$$

which is a *forward* difference because of the forward-looking use of $x + h$. Similarly, the backward difference derived from $x - h$ is rearranged as

$$v'(x) \approx \frac{v(x) - v(x-h)}{h}. \tag{16}$$

As $h \to 0$, the approximations of the analytic derivative become better, and ultimately the true derivative is defined as the limit when h tends to 0. When computing these approximations numerically, it is not possible to compute the value when $h = 0$. Instead, it is useful to define a measure of "fit" for each approximation as h gets closer and closer to zero. Encapsulating the second-order and higher terms in equation (15) into a truncation error term ξ such that ξ is a bounded function of the approximating

equation whose range is in $\mathbb{R}$, we can rewrite the Taylor approximation around $x+h$ as an equality in the form

$$v(x+h) = v(x) + hv'(x) + \frac{h^2}{2}v''(\xi). \tag{17}$$

Rearranging as

$$v'(x) = \frac{v(x+h) - v(x)}{h} + \frac{h}{2}v''(\xi) \tag{18}$$

shows that the coefficient on the error term ξ is bounded above and below by h, i.e. $\Theta(h)$. Thus, we call this forward approximation a *first-order* approximation. Were the coefficient a multiple of h^2, it would be a *second-order* approximation, i.e. $\Theta(h^2)$. A recurring concept in numerical approximation is that approximations for higher order derivatives can be determined by combining approximations of lower order derivatives. The centered difference approximation for the first derivative (with truncation error proportional to h^2) can be computed by summing the first order Taylor approximations (15) and (16) and solving for the sought variable u' in the resulting equation. Taking the linear average of (15) and (16) by summing the halved first-order Taylor approximations for $x+h$ and $x-h$ yields the centered difference approximation with second-order truncation error (ξ proportional to h^2)

$$v'(x) = \frac{v(x+h) - u(x-h)}{h} + \frac{2h^2\xi}{3!} \tag{19}$$

Extending this concept, it is also possible to approximate higher order derivatives from the principles of Taylor expansion. For example, consider again summing the expansion for $x+h$ and $x-h$ in equations (13) and (14).

$$v(x+h) + v(x-h) = 2v(x) + h^2v''(x) + \frac{2h^4}{4!}v^{(iv)}(x) \tag{20}$$

where the numerals (iv) indicate the fourth derivative. Once again we can solve for the desired derivative, this case $v''(x)$ yielding

$$v''(x) = \frac{v(x+h) - 2v(x) + v(x-h)}{h} + 2\frac{h^2\xi}{4!} \tag{21}$$

Finally, discretization techniques can be applied in succession to compute approximations for mixed partial terms. Consider the approximation for the second-order mixed partial of v with respect to both x and y as

$$\frac{\partial^2 v}{\partial x \partial y} = \frac{\partial}{\partial y}\left[\frac{\partial v}{\partial x}\right] \tag{22}$$

by approximating $\frac{\partial v}{\partial x}$ first with a centered difference as above. In turn, differencing with respect to the y coordinate of the center-difference x terms yields the full discretization

$$\frac{\partial^2 v}{\partial x \partial y} \approx \frac{v_{x+1,y+1} - v_{x+1,y-1} - v_{x-1,y+1} + v_{x-1,y-1}}{4\Delta x \Delta y} \tag{23}$$

2.2.3 Ex: Heat Equation Approximation with FDM

Returning to the heat equation (9), the first step in approximating the PDE is to classify the PDE. It is straightforward to show that the heat equation is second-order, parabolic, and linear. Thus, it will require both initial and boundary conditions. Now consider substituting the Taylor difference approximations for each derivative term. Some care must be given when choosing which approximation to use for each term. Since the initial condition provides a starting point for the temperature at a given time, it is sensible that the scope of analysis will be to analyze the temperature of the rod as time moves forward. Thus, it is reasonable to choose a forward difference approximation for the temporal derivative (i.e. the derivative term describing how the temperature changes with time)

$$\frac{\partial u}{\partial t}(n\Delta t, m\Delta l) = \frac{u(n\Delta t, m\Delta l) - u((n+1)\Delta t, m\Delta l)}{\Delta t}, \tag{24}$$

or in condensed notation,

$$\frac{\partial u}{\partial t}(n\Delta t, m\Delta l) = \frac{u_l^{t+1} - u_l^t}{\Delta t}. \tag{25}$$

Note that if the scope of analysis moved the opposite direction, a backward difference would work, too.

For the spatial derivative (i.e. the derivative term describing how the temperature changes at different lengths), the centered difference

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{l+1}^t - 2u_l^t + u_{l-1}^t}{(\Delta l)^2} \tag{26}$$

is reasonable to use over either the forward or backward differences because it provides greater accuracy. Having discretized both the time and spatial components of the heat equation, we can now substitute equations (25) and (26) in for the LHS and RHS, respectively, in equation (9)

$$\frac{u_l^{t+1} - u_l^t}{\Delta t} = k\frac{u_{l+1}^t - 2u_l^t + u_{l-1}^t}{(\Delta l)^2}. \tag{27}$$

If we consider the system at the time where the initial condition specifies the temperature along the length of the rod, say $t = 0$, known quantities in equation (27) include the terms evaluated at time t. This leaves u_l^{t+1} as the sole unknown in the equation, and it describes the temperature at a specific length (l) along the rod at the very next time step ($t + 1$). Thus, it is easily solvable. Rearrangement yields

$$u_l^{t+1} = \lambda\left(u_{l+1}^t - 2u_l^t + u_{l-1}^t\right) - u_l^t \tag{28}$$

where λ is defined as the Courant number $k\frac{\Delta t}{\Delta x^2}$, which will be important to our subsequent analysis. This system is called the *explicit* finite difference method since

the unknown is calculated from explicitly known quantities. For a full example of the code implemented in the R language, please refer to the Online Appendix. Alternatively, *implicit* evaluation of the spatial derivative at $t+1$ gives a system of simultaneous equations (rather than a single, governing equation) that solves for the same unknowns as before. Separating the system of equations into a coefficients matrix, C

$$\begin{pmatrix} 1+2\lambda & -\lambda & & 0 \\ -\lambda & \ddots & \ddots & \\ & & \ddots & \ddots \\ 0 & & -\lambda & 1+2\lambda \end{pmatrix} \tag{29}$$

where λ is as before, unknown matrix U (that is, the vector we are solving for)

$$\begin{pmatrix} u_1^{t+1} \\ u_2^{t+1} \\ \vdots \\ u_{m-1}^{t+1} \\ u_m^{t+1} \end{pmatrix} \tag{30}$$

and known matrix, K (evaluated at time t)

$$\begin{pmatrix} u_1^t + \lambda u_0^t \\ u_2^t \\ \vdots \\ u_{m-2}^t \\ u_{m-1}^t + \lambda u_m^t \end{pmatrix} \tag{31}$$

the system can be written as $CU = K$. At each new time step, U can be found from $U = KC^{-1}$ solving the system at time $t+1$

$$\begin{pmatrix} \frac{1}{\Delta t}+\frac{2k}{\Delta l^2} & -\frac{k}{\Delta l^2} & 0 & \dots & 0 \\ -\frac{k}{\Delta l^2} & \frac{1}{\Delta t}+\frac{k}{\Delta l^2} & -\frac{k}{\Delta l^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & -\frac{k}{\Delta l^2} & \frac{1}{\Delta t}+\frac{k}{\Delta l^2} & -\frac{k}{\Delta l^2} \\ 0 & \dots & 0 & -\frac{k}{\Delta l^2} & \frac{1}{\Delta t}+\frac{k}{\Delta l^2} \end{pmatrix} \begin{pmatrix} u_1^{t+1} \\ u_2^{t+1} \\ \vdots \\ u_{m-2}^{t+1} \\ u_{m-1}^{t+1} \end{pmatrix} = \begin{pmatrix} u_1^t \\ u_2^t \\ \vdots \\ u_{m-2}^t \\ u_{m-1}^t \end{pmatrix} \tag{32}$$

2.3 Convergence Analysis for FDM

The explicit and implicit methods shown above represent two extremes on a continuum of FDM techniques; the explicit method is *conditionally* stable and the implicit method is *unconditionally* stable. Numerical analysis of each scheme is important in helping the practitioner understand the differences, and when it is appropriate to use one scheme over another. The first distinction that must be made among different methods is the distinction between schemes whose solutions approximate the PDE solution as the lattice discretization becomes finer (and the grid spacings tend to zero) and those whose solutions that do not. The first type of schemes are called *convergent*, and this property is the most basic property a in useful schemes. Checking for convergence is not a trivial matter, but checking for the *consistency* and *stability* of a scheme are more straightforward and provide insight into the solution mechanism. If a scheme is shown it be both consistent and stable, then by the Lax-Richtmyer Equivalence Theorem [16] the scheme is convergent (provided that the problem is well-posed).

According to Strikwerda's [16] definition, a *consistent* finite difference scheme is one where the difference between the PDE and the numerical approximation tends to zero (for a smooth function) as the grid spacing tends to zero. For example, again consider the heat equation (27). Written in operator notation where the operator P is $\frac{\partial}{\partial t} - k\frac{\partial^2}{\partial l^2}$ so that

$$P\phi = \phi_t - k\phi_{ll} \tag{33}$$

where subscripts represent partial derivatives.

Denoted by $P_{\Delta t,\Delta l}$, we apply the explicit FDM scheme to the PDE as before and get

$$P_{\Delta t,\Delta l} = \frac{\phi_l^{t+1} - \phi_l^t}{\Delta t} - k\frac{\phi_{l+1}^t - \phi_l^t + \phi_{l-1}^t}{\Delta l^2}, \tag{34}$$

where

$$\phi_l^t = \phi\left(n\Delta t, m\Delta l\right), \tag{35}$$

or equivalently

$$\phi_l^t = \phi\left(t, l\right) \tag{36}$$

where $t \in \{1, 2, , n\}$ and $l \in \{1, 2, , n\}$.

Expanding the function ϕ around (t_n, l_m) we have the approximations

$$\phi_l^{t+1} = \phi_l^t + \Delta t\phi_{(t)} + \frac{1}{2}\Delta t^2\phi_{(tt)} + O(\Delta t^3),$$

$$\phi_{l+1}^t = \phi_l^t + \Delta t\phi_{(l)} + \frac{1}{2}\Delta l^2\phi_{(ll)} + O(\Delta l^3),$$

and

$$\phi_{l-1}^t = \phi_l^t - \Delta l\phi_{(l)} + \frac{1}{2}\Delta l^2\phi_{(ll)} + O(\Delta l^3).$$

Substitution of the above equation into (34) yields

$$\phi_{(t)} + \frac{1}{2}\Delta t\phi_{(tt)} + O(\Delta t^2) = k(\phi_{(ll)} + O(\Delta l)). \tag{37}$$

where the $O(\cdot)$ notation refers to the truncation error. Thus

$$P\phi - P_{\Delta t,\Delta l} = \frac{1}{2}\phi_{(tt)} + O(\Delta t^2) + O(\Delta l) \to 0, \tag{38}$$

as $(\Delta t, \Delta l) \to (0, 0)$ and the scheme is consistent.

Next, to determine convergence, we must show the scheme is stable. Strikwerda [16] notes that for a scheme to be convergent, it must be bounded, and this boundedness is the underlying principle of stability. Intuitively, if the multiplier from step to step is too large, the scheme will diverge. Thus, stability analysis determines the bounds on the multiplier limit from step to step.

Stability can be determined from *von Neumann analysis*, which involves applying Fourier transforms to the errors of the approximation [4]. To isolate the errors, the numerical approximation can be thought of as the analytic solution plus the error terms where the errors are of the same form as the PDE. For a scheme to be stable, the approximation errors caused from truncating the Taylor expansions cannot be amplified from step to step. Hence, the ratio of the next steps error to the current error cant be larger than one; otherwise, the schemes error will diverge. To check for stability, we must then determine a bound on the coefficient term λ, where

$$\lambda = k^2 \frac{\Delta t}{(\Delta x)^2}. \tag{39}$$

Substituting λ into the discretized heat equation, the difference equation becomes

$$u_l^{t+1} = \lambda u_{l+1}^t + (1 - 2\lambda)u_l^t + \lambda u_{l-1}^t \tag{40}$$

Now, we can check for stability using Fourier transforms. Transforming the numerical difference method, we assume that the scheme provides a solution in the form

$$u_l^t = a^{(t)}(\omega)e^{(il\omega\Delta x)} \tag{41}$$

with $i = \sqrt{-1}$ and wave number (i.e. spatial frequency of the Fourier wave) ω. If we define the amplification error ratio as

$$G(\omega) = \frac{a^{(t+1)}(\omega)}{a^{(t)}(\omega)}, \tag{42}$$

the von Neumann stability condition requires $G(\omega) \leq 1$ for $0 \leq \omega\Delta x \leq \pi$.

Substituting our assumed form of the solution (41) into the difference equation (40) and solving for λ yields the restriction [3]

$$0 \leq \lambda \leq \frac{1}{2}. \tag{43}$$

Thus, the explicit method is only stable provided that

$$\Delta t \leq \frac{(\Delta x)^2}{2k^2}, \tag{44}$$

which means that if the spatial discretization is refined by a factor of 2, the temporal discretization must be refined by a factor of 4. This increases total computation by a factor of 8. It would be convenient, then, for a scheme to be unconditionally convergent (i.e. its convergence is independent of the fineness of the solution mesh).

The implicit method is one such technique that achieves unconditional convergence [16] but it is not discussed further in this paper.

3 FDM for Option Pricing

This section presents the full process of implementing an explicit FDM scheme to the Black-Scholes equation (8) and the key architectural decisions that affect the outcome. The explicit scheme has the benefit of easy implementation, even for more complicated PDEs (as we will see in the next section), but it comes far short of an ideal scheme in terms of convergence properties. To begin, recall the Black-Scholes equation (8)

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Here, we have two variables that require discretization: the underlying asset price S and time t. The first step is to build a mesh on which the FDM scheme can provide a numeric approximation to the Black-Scholes PDE solution. We begin by defining S and t in discrete terms and must consider the scope of the analysis. If we suppose a vanilla European Call option has a strike price of 75, want to make sure the grid is large enough to capture the range of possible underlying asset price movements. Suppose the maximum asset price is 150. This simplifies reality, but our approximation allows us to see the general trend in option value when the underlying price rises. On the other hand, we take 0 as the lower bound for the underlying price discretization. The final input discretization decision is that of time. We take three years from the point of valuation as the expiration date.

We now have the minimum and maximum values for our discretization. Subsequently, we must adhere to the results of stability analysis and choose grid spacing such that the scheme will be stable. We refine the asset discretization and the time discretization through trial and error until the scheme is stable. Before implementing the stepwise procedure of solving for each mesh point, we must specify initial and boundary conditions (since it is a parabolic PDE).

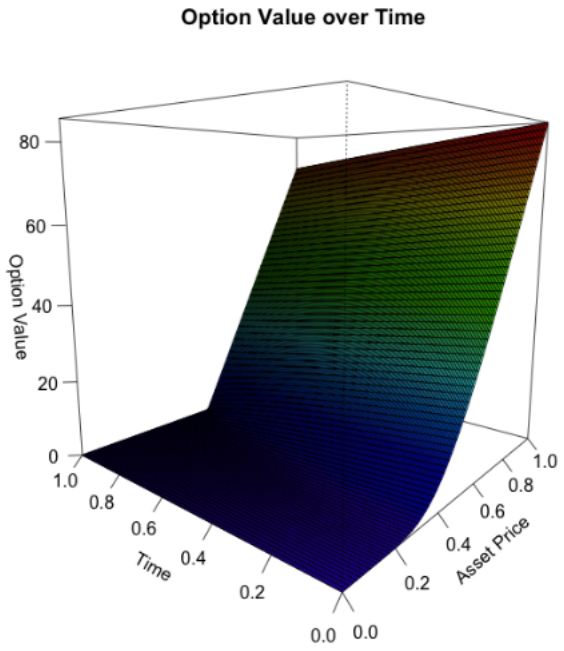
The initial condition is straightforward. At expiry, the payoff of the option is simply the positive value of the underlying minus the strike. Determining the boundary conditions requires slightly more work. As the underlying tends to infinity, the probability of exercise approaches one, and the option's boundary condition can be interpreted as a forward contract where only the time-value-of-money affects option value. As S approaches 150, our approximation for infinity, the underlying price will almost certainly be greater than the strike price, and hence, the option will almost certainly be exercised. Across time, the only difference then between the value with one year left to maturity and two years left to maturity is the present value of the

strike price. With a compounding schedule to match the discretization, it is possible to track the decrease in value moving from time 0 to time maturity by discounting the strike appropriately. Regardless of the time to expiration on the option, if the underlying is zero, the value of the option is also zero.

The prior steps have now provided a shell of the solution. The starting point is known, and the behavior of the solution over time is known. From here forward, we must compute the option value's approximation at each mesh point. This is the bulk of the FDM computation time. Making use of the previously calculated option values at prior time steps is essential, and for option pricing, it is often useful to move backwards in time.

Although less than ideal (for the aforementioned stability concerns), we solve for the sought unknown quantity at a specific grid node (S, t) in terms of known quantities, similar to our method solving for u_l^{t+1} from the heat equation (9), and then loop this equation through the interior discretizations of each variable.

Plotting the surface, we see the familiar results below.



For R code to explicit and implicit methods of solving the Black-Scholes equation, see the Online Appendix, which is in part based on Richardson's option pricing code [13].

4 Stochastic Volatility Extension

In this section, we present three models, M1, M2, and M3 for approximating German's [6] (eq. 28) SV PDE. The models differ in their selection of boundary conditions.

4.1 SV Models

We begin by tying our prior discussion of indifference pricing to the ideas underlying German's derivation of the SV PDE. If we suppose the payoff function for a contingent claim is a bounded function f, then $f(S_T)$ is the value of that security at time T. Assigning an arbitrary utility function U to the investor satisfying assumptions (3), German shows that a universal minimal risk-neutral probability measure $\tilde{\mathbb{Q}}$ exists and that the sensitivity of an investor's utility to the presence of non-traded derivative securities is a function of the derivative's vega, among other quantities [6]. An option's vega is the partial derivative of option value with respect to volatility. German denotes the value of a contingent claim $f(S_T)$ as

$$u(t, S_t, V_t) = \mathbb{E}_{\hat{\mathbb{Q}}}[f(S_T)|\mathcal{F}_t] \tag{45}$$

for time $t \leq T$, risk-neutral probability measure $\hat{\mathbb{Q}}$, and a bounded function $f = f(x)$. Differentiating this equation, German then presents an arbitrary stochastic volatility model with zero interest rate where

$$dS_t = \eta S_t \sqrt{V_t} dt + S_t \sqrt{V_t} dW_t \tag{46}$$

and

$$dV_t = \mu(t, V_t)dt + \sigma(t, V_t)d\widetilde{W}_t. \tag{47}$$

are Markov processes with $\widetilde{W}_t = \rho W_t + \sqrt{1-\rho^2}B_t$ and B is Brownian motion independent of W given on a filtered probability space $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$. By the Feller condition, $\mu(t, V_t)$ and $\sigma(t, V_t)$ are processes such that the volatility process (47) is non-negative with $\eta \in \mathbb{R}$ and $0 < \rho < 1$. This model then yields the PDE

$$\frac{\partial u}{\partial t} + \mu(t,\nu)\frac{\partial u}{\partial \nu} + \frac{1}{2}\nu S^2 \frac{\partial^2 u}{\partial S^2} + \frac{1}{2}\sigma(t,\nu)\frac{\partial^2 u}{\partial \nu^2} + \rho\sigma(t,\nu)\sqrt{\nu}S\frac{\partial^2 u}{\partial S \partial \nu} \tag{48}$$

where μ and σ represent parameters depending on time (t) and the volatility at time t denoted above as ν

For the first two models, we assume μ and σ are constant. In the third model, we choose μ such that

$$\mu(t, V_t) = a(b - V_t) \tag{49}$$

where $a \in \mathbb{R}$ and b is the running mean for volatility at time t, and σ such that

$$\sigma(t, V_t) = \sigma(t-1, V_{t-1})\sqrt{V_t} \tag{50}$$

and are ensured that the volatility process is nonnegative in the FDM since our discretization for V_t is always positive and $\sigma(t = T, V_T) \geq 0$ where T denotes the time at expiry.

The first model (M1) prices a vanilla European Call. Here we face the issue of observing potentially unbounded payoffs, which pose a problem for a grid limited by finite approximations.

The second model (M2) prices a European Bull-Call Spread to comply with the assumed boundedness of the derivative payoff function $f = f(x)$ in equation 45). In contrast to a vanilla European Call where the payoff can theoretically approach infinity, a Bull-Call Spread is a bounded derivative that pays a maximum value even when the asset price tends toward infinity. In this way, it is a useful derivative for studying the SV equation while reducing the complexity of the $S \to \infty$ boundary condition.

The third model (M3) also prices a vanilla Bull-Call Spread. Here, we propose an explicit method that calculates the value of the option at each node independently of boundary conditions. We select forward, backward, and centered approximations for the terms of equation (48) depending on the location of the node relative to the edge of the discretization.

We begin by describing the discretization of the SV PDE (48) and then turn to the construction of M1, M2, and M3.

4.2 Discretization

As with the Black-Scholes discretization, we substitute the backward difference quotient for the time term and the centered difference quotients for the remaining terms as outlined in Section 3, given the that the flow of analysis marches backwards in time from maturity. Rearranging to solve for $u_{x,y}^{t-1}$ (i.e. the option value at time $t-1$) yields the discretization

$$\begin{aligned} u_{x,y}^{t-1} = \Delta t \Bigg[& \mu \left(\frac{u_{x,y+1}^{t} - c_{x,y-1}^{t}}{\Delta \nu} \right) + \frac{1}{2} \Delta \nu y x^2 \left(u_{x+1,y}^{t} - 2u_{x,y}^{t} + u_{x-1,y}^{t} \right) \\ & + \frac{1}{2} \sigma^2 \left(\frac{u_{x,y+1}^{t} - u_{x,y}^{t} + u_{x,y-1}^{t}}{\Delta \nu^2} \right) \\ & + \rho \sigma \sqrt{\Delta \nu} \sqrt{y} x \left(\frac{u_{x+1,y+1}^{t} - u_{x+1,y-1}^{t} - u_{x-1,y+1}^{t} + u_{x-1,y-1}^{t}}{4 \Delta \nu} \right) \Bigg] + u_{x,y}^{t} \end{aligned} \tag{51}$$

where $u_{x,y}^{t} = u(x, y, t) = u(\text{asset price}, \text{volatility}, \text{time})$ with S and ν discretized as $x\Delta S$ and $y\Delta\nu$, respectively.

Gathering terms

$$\begin{aligned} \frac{u_{x,y}^{t-1}}{\Delta t} = & u_{x,y+1}^{t} \left(\frac{\mu(t,\nu)}{2\Delta\nu} + \frac{1}{2}\sigma(t,\nu) \right) + u_{x,y-1}^{t} \left(\frac{-\mu(t,\nu)}{2\Delta\nu} + \frac{1}{2}\sigma(t,\nu) \right) \\ & + u_{x+1,y}^{t} \left(\frac{1}{2} y \Delta\nu x^2 \right) + u_{x-1,y}^{t} \left(\frac{1}{2} y \Delta\nu x^2 \right) + u_{x,y}^{t} \left(\frac{1}{\Delta t} - y\Delta\nu x^2 - \frac{\sigma(t,\nu)}{(\Delta\nu)^2} \right) \\ & + \rho\sigma(t,\nu) \sqrt{(y)(\Delta\nu)} x \left(\frac{u_{x+1,y+1}^{t} - u_{x+1,y-1}^{t} - u_{x-1,y+1}^{t} + u_{x-1,y-1}^{t}}{4\Delta\nu\Delta S} \right) \end{aligned} \tag{52}$$

As before, this equation allows for the solution of the unknown quantity $\dfrac{u_{x,y}^{t-1}}{\Delta t}$ in terms of known quantities evaluated at time t.

Since the PDE (48) is three-dimensional (1 time variable, 2 space variables), we can ultimately store the numerical approximation in a three-dimensional array. To determine both the boundary and initial conditions, we take one of the three variables as fixed and consider how the approximation behaves in this newly restricted space. In moving from a two-dimensional equation like Black-Scholes to a three-dimensional equation like the SV equation, the dimensionality of the boundary and initial conditions also rises by one dimension. Thus, each boundary condition can be represented as a surface in three dimensions rather than a line in two dimensions.

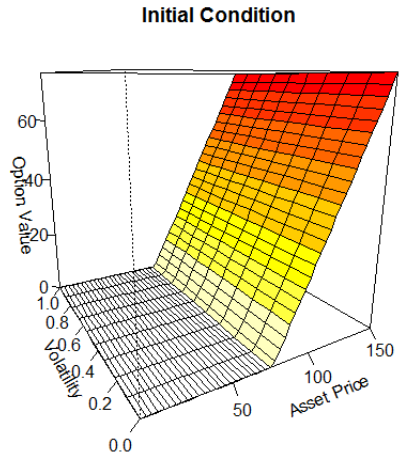
4.3 M1

4.3.1 M1 Initial Conditions

In this model, we consider a European call option, and thus we must update our initial (i.e. expiry time) payoff conditions. A call simply pays the maximum of the asset price S minus the strike price K, so at each node (S, ν), we solve for the initial condition as

$$u_{x,y}^{T} = \max\{(S - K, 0)\} \tag{53}$$

where T denotes the time at expiry. We notice that this boundary condition is independent of volatility.



4.3.2 M1 Boundary Conditions

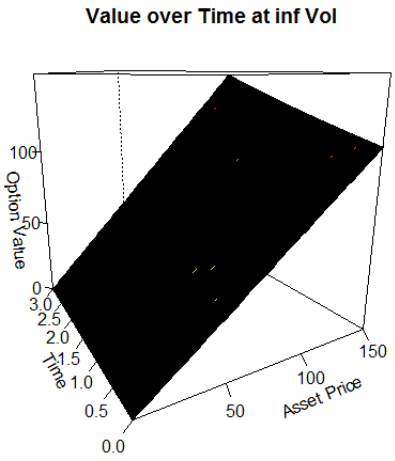
Here we utilize Shreve's suggested boundary conditions for a similar stochastic volatility PDE he presents [15]. Like M1, the $S = 0$ boundary condition is trivially 0 for every node (ν, t) when the asset price is zero. However, when we consider the $S \to \infty$ boundary condition, we notice that the slope of the option payoff is correlated one to one with the price of the underlying asset, i.e. the option's delta is one. Since the

derivative of the option value with respect to the underlying asset price approaches one, we replace the Neumann condition for the simpler Dirichlet condition and approximate the option value as the price of the underlying minus the strike.

$$u^t_{x=x_{max},y} = S - K. \tag{54}$$

Next consider the option value when volatility takes on its extremes of zero and the discrete approximation for infinity. When volatility is zero, we must be careful not to make the assumption that the stock price is a stopped process as this would contradict the nature of stochastic volatility. Instead, we solve for the option value at discrete nodes (S, t) as before, according to Shreve's suggested boundary conditions [15]. We believe this still allows for stochastic volatility since the random price process (46) determined by the stochastic volatility process (47) can be thought of as tracing a random path along the surface of the approximation solution [11]. Ideally a solution procedure would impose the PDE at each of the possible boundary conditions, for example, substituting $\nu = 0$ in the SV PDE to solve the zero volatility condition. As we show in M3 and Section 5, there is a tradeoff between imposing the PDE at the boundary condition for additional accuracy with increased computation time.

At the $\nu \to \infty$ and $\nu = 0$ boundary conditions, we require the value of the call to be $S - K$, following Shreve's [15] guidance. With (approximated) infinite volatility, the option price at each time approaches the price of the underlying payoff, in this case S.



Conversely, at zero volatility, we also solve for the boundary condition as $S - K$.

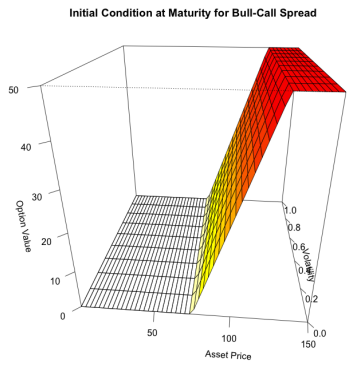
4.4 M2

4.4.1 M2 Initial Conditions

The initial condition (computed at maturity) depends on the specified type of option. We propose studying a Bull-Call Spread to reduce the complexity of the S_{max} bound-

ary condition and also comply with Germna's assumption that $f = f(x)$ is a bounded function. This type of option is created by a buying a call with a low strike, K_1, and selling another call with a high strike, K_2. Thus, the option pays the minimum of a vanilla call with strike K_1 and the difference between K_2 and K_1, as shown in the below diagram. When asset price tends to infinity, the holder receives the difference between the high and low strike, and thus the payoffs are bounded. Formally, the payoff at maturity is represented as

$$u(t_{max}, S, \nu) = \min\left[(S - K_1, 0)^+, (K_2 - K_1)\right] \forall S \geq 0, \nu \geq 0. \tag{55}$$



4.4.2 M2 Boundary Conditions

To solve the SV PDE (48), we need conditions that control the option's behavior as S and ν take on their minima and maxima. Consider the asset price boundary conditions. Like the $S = 0$ boundary condition from Black-Scholes, when the stock price is zero, the Bull-Call Spread will never be exercised, so the value of the option is zero.

$$u(S = 0, \nu, t) = 0 \,\forall\, 0 \leq t \leq t_{max},\ \nu \geq 0. \tag{56}$$

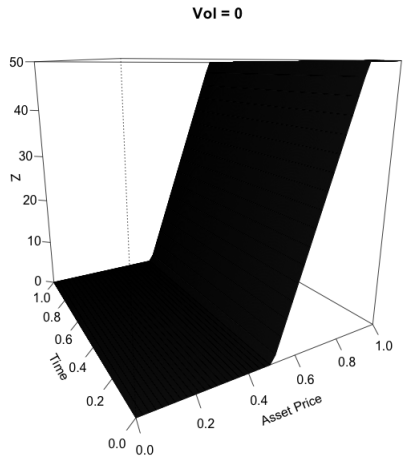
We solve this equation for each node (ν, t) except for the value at $t = t_{maturity}$.

Conversely, when the stock price tends toward infinity, we assume that the option will be exercised. Since we have assumed an interest rate of zero, we do not need to consider how the option value changes over time. Instead, even at an infinite asset price, the Bull-Call Spread remains bounded, and thus the derivative of option value with respect to asset price is zero. Accordingly, we are justified in using a Dirichlet condition since the Neumann boundary condition in determining the derivative of option value with respect to asset price is zero. The option value along this boundary is then similar to the initial condition, but we must adjust for time and replace the discretization for S with the approximation for $S \to \infty$. In other words, it is the minimum of the maximum of the approximation for S at infinity minus the low strike and zero, and the difference between the high and low strike prices

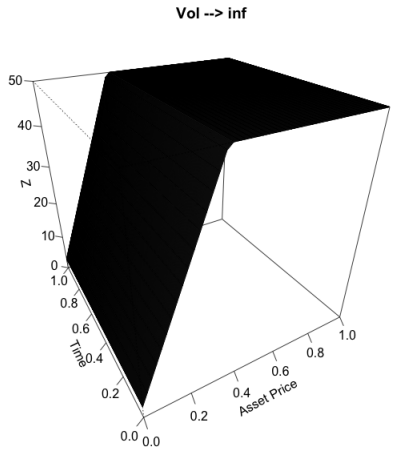
$$u(t, S_{max}, \nu) = \min\left[\max\left(S_{max} - K_1, 0\right), (K_2 - K_1)\right]\ \forall t \geq 0,\ \nu \geq 0. \tag{57}$$

In the model, we choose K_1 and K_2 to be 75 and 125, respectively, and thus since $S_{max} - K_1$ is always greater than $K_2 - K_1$, the surface is 50 for every node.

For $\nu \to \infty$ and $\nu = 0$ boundary conditions, we construct payoffs similar to M1. That is, we solve for the value of the option at each $(S, \nu = 0, t)$ node without imposing the PDE. We leave the stochastic price path as the determinant of randomness [11]. We assume the payoff of the option at each node depends on only on asset price (relative to the derivative's strike prices) and time. If $S \geq K_2$, then the payoff is the difference between the low and high strikes. This result is derived from the fact that the purchased call will pay off $S - K_1$ while the sold call costs $-(S - K_2)$ so that together for underlying prices above K_2, the Bull-Call Spread holder is guaranteed $(K_2 - K_1)$. If $S < K_2$, then the payoff is the positive part of the stock price minis the strike K_1, and we see



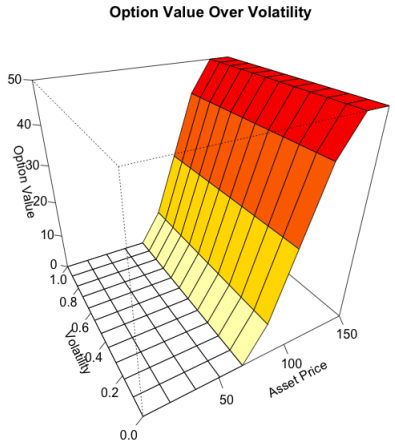
When volatility becomes unbounded, the option value approaches the value of the underlying payoff structure. A Bull-Call Spread provides an upper bound on the value of the option, simplifying the limit as $\nu \to \infty$ so that the maximum value an economic agent would ever pay for such an option would be $K_2 - K_1$. To determine the boundary condition across our stock price discretization, it is helpful to consider alternative methods of achieving the same payoff. Up to strike K_2, the same payoff could be achieved by simply buying the stock. With unbounded volatility, neither the price of the option nor the price of the underlying can fall below zero, but both securities stand to benefit equally (in the limit, the strike price becomes irrelevant in the worth of the option). Hence, for prices below K_2, the price of the option is simply S, the most an agent would be willing to pay for the chance to participate in the same risk-reward profile. With the bounded maximum condition from the bull-call spread, the analysis doesn't have to make the additional approximation that a plain vanilla call option would have to make for the value of the option as the underlying price rises above K_2. Thus, for asset prices above K_2, the option is worth $K_2 - K_1$.



4.5 M3

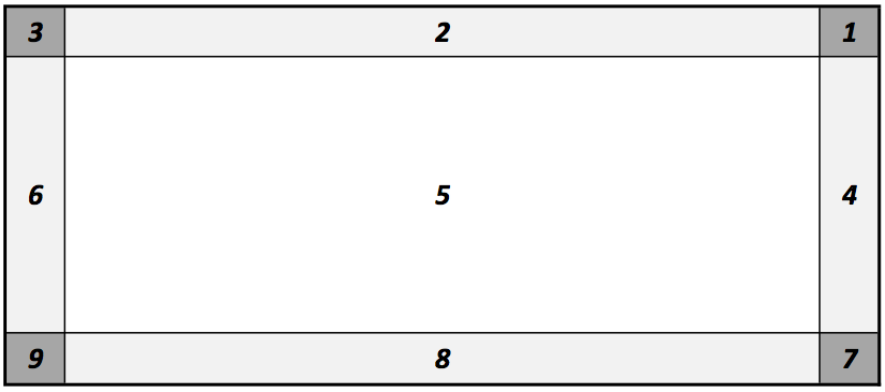
The explicit FDM can also be interpreted as a trinomial model. In this way, it is possible to construct a pricing surface without the use of boundary conditions. Instead, we must consider the terminal condition, and then at each subsequent time step, we must ensure that our point-by-point calculation of each node depends solely on previously calculated nodes. In both M1 and M2, we ensured that each node did not depend on "ghost" points outside of the discretization by solving for the boundary conditions based on the theoretical behavior of the option price as its factors (asset price, volatility) took on their extremes. In M3, by substituting forward, backward, and centered differences depending on where the node of calculation is relative to the edge of the discretized mesh, we can effectively eliminate the use of boundary conditions and approximate the PDE itself at every node with a combination of specifically ordered finite difference approximations.

We begin by defining the bounded Bull-Call Spread initial condition payoff as we have defined it in M2.



Notice that for asset values below 75, the surface yields a price of 0, indicated by the white coloring.

Next, we decompose the asset-volatility grid into nine sections, based on the asset and volatility index within our main for-loop.



We then use first order approximations for each term in equation (48), and construct these as follows. We denote the ith-order volatility derivative $\mu(t,\nu)\frac{\partial u}{\partial \nu}$ as $V_i(\cdot)$ and compute its forward, backward, and centered approximations as

$$V_1(Fwd) = \mu(t,\nu)\frac{u^t_{x,y+1} - u^t_{x,y}}{\Delta\nu} \tag{58}$$

$$V_1(Bwd) = \mu(t,\nu)\frac{u^t_{x,y} - u^t_{x,y-1}}{\Delta\nu} \tag{59}$$

$$V_1(Ctr) = \mu(t,\nu)\frac{u^t_{x,y+1} - u^t_{x,y-1}}{2\Delta\nu} \tag{60}$$

where x denotes the asset value index, y denotes the volatility index, and t denotes the time. We then continue with each term in equation (48) so that the second term is

$$S_2(Fwd) = \frac{1}{2}y\Delta\nu(x\Delta S)^2\frac{u^t_{x+2,y} - 2u^t_{x+1,y} + u^t_{x,y}}{(\Delta S)^2} \tag{61}$$

$$S_2(Bwd) = \frac{1}{2}y\Delta\nu(x\Delta S)^2\frac{u^t_{x,y} - 2u^t_{x-1,y} + u^t_{x-2,y}}{(\Delta S)^2} \tag{62}$$

$$S_2(Ctr) = \frac{1}{2}y\Delta\nu(x\Delta S)^2\frac{u^t_{x+1,y} - 2u^t_{x,y} + u^t_{x-1,y}}{(\Delta S)^2}. \tag{63}$$

The third term is then

$$V_2(Fwd) = \frac{1}{2}\sigma(t,\nu)\frac{u^t_{x,y+2} - 2u^t_{x,y+1} + u^t_{x,y}}{(\Delta\nu)^2} \tag{64}$$

$$V_2(Bwd) = \frac{1}{2}y\Delta\nu(x\Delta S)^2\frac{u^t_{x,y} - 2u^t_{x,y-1} + u^t_{x,y-2}}{(\Delta\nu)^2} \tag{65}$$

$$V_2(Ctr) = \frac{1}{2}y\Delta\nu(x\Delta S)^2\frac{u^t_{x,y+1} - 2u^t_{x,y} + u^t_{x,y-1}}{(\Delta\nu)^2}. \tag{66}$$

and finally the mixed term is composed of five discretizations

$$S(Fwd)V(Fwd) = \rho\sigma(t,\nu)\sqrt{(\nu)}S\frac{u^t_{x+1,y+1} - u^t_{x+1,y} + u^t_{x,y+1} + u^t_{x,y}}{(\Delta S)(\Delta\nu)} \tag{67}$$

$$S(Fwd)V(Bwd) = \rho\sigma(t,\nu)\sqrt{(\nu)}S\frac{u^t_{x+1,y} - u^t_{x+1,y-1} + u^t_{x,y} + u^t_{x,y-1}}{(\Delta S)(\Delta\nu)} \tag{68}$$

$$S(Bwd)V(Fwd) = \rho\sigma(t,\nu)\sqrt{(\nu)}S\frac{u^t_{x,y+1} - u^t_{x,y} + u^t_{x-1,y+1} + u^t_{x-1,y}}{(\Delta S)(\Delta\nu)} \tag{69}$$

$$S(Bwd)V(Bwd) = \rho\sigma(t,\nu)\sqrt{(\nu)}S\frac{u^t_{x-1,y-1} - u^t_{x-1,y} + u^t_{x,y-1} + u^t_{x,y}}{(\Delta S)(\Delta\nu)} \tag{70}$$

$$S(Ctr)V(Ctr) = \rho\sigma(t,\nu)\sqrt{(\nu)}S\frac{u^t_{x+1,y+1} - u^t_{x+1,y-1} + u^t_{x-1,y+1} + u^t_{x-1,y-1}}{4(\Delta S)(\Delta\nu)}. \tag{71}$$

Finally, we compute the option price moving backwards based on the above discretizations depending on the location of the node. For example, if the asset index exceeds the number of asset nodes minus 2 (to account for the second derivative approximations) and the volatility index exceeds the number of volatility nodes minus 2 (that is, we are in region 7), then we compute the node based on the equation

$$u^{t-1}_{x,y} = \Delta t(V_1(Bwd) + S_2(Bwd) + V_2(Bwd) + S(Bwd)V(Bwd)) + u^t_{x,y}. \tag{72}$$

We proceed in this fashion for every node, ultimately determining nine such equations to compute the option value accordingly so that our computation do not depend on nodes outside of the discretization, but only upon nodes known at prior times.

4.6 Implementing Explicit FDM for Stochastic Volatility Model

Here we present pseudocode for implementing an explicit FDM algorithm for approximating equation (48).

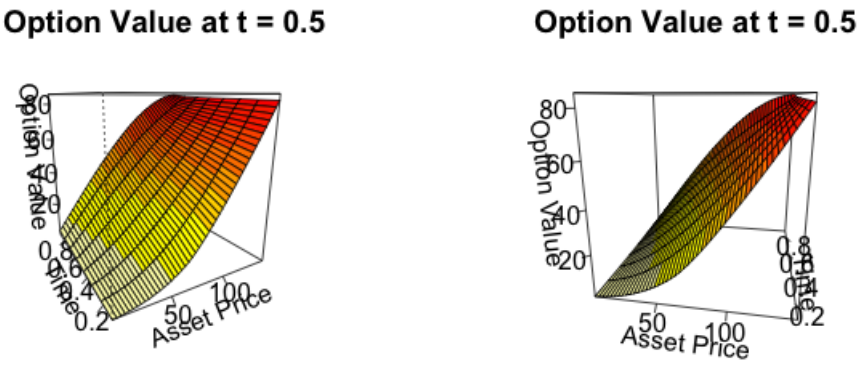
1. Define the solution mesh, a three-dimensional array of S, ν, and t
2. Discretize each variable with a uniformly spaced sequence
3. Impose boundary conditions according to the above discussion
4. Solve for the interior nodes using the discretized PDE
5. Refine grid spacing until scheme is stable

For a full example of the code to M1, M2, and M3, please refer to the Online Appendix.

5 Discussion and Analysis

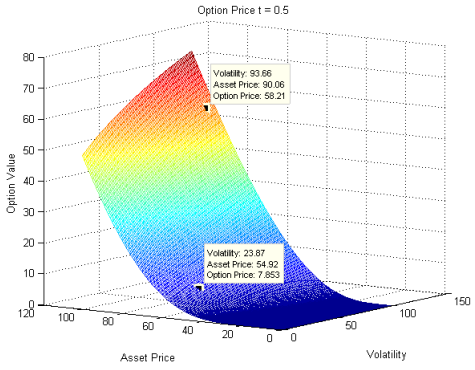
5.1 M1 Option Value Surface

Again taking a cross-section of the 3D solution array with time to expiry equal to one half the time period, from different perspectives we see



These surfaces exhibit the expected diffusion between the $\nu = 0$ and $\nu \to \infty$ boundary conditions. These show a marked inflection point for lower and lower asset prices as volatility increases. We attribute this change in concavity to the limits of our discretization mesh. Intuitively, the boundary condition "pulls" the option value down at large asset prices when the asset value approaches its maximum. This lowered approximation for S_{max} then diffuses throughout the future nodes that are calculated explicitly from it. Were the discretization to encompass larger asset price values, we hypothesize that the option value surface would see less "warping". If this is true, then it indicates that to study the value of an option more accurately, practitioners should expand the discretization of their grid to include asset values much higher than reasonably attainable and volatilities much higher than reasonably predicted.

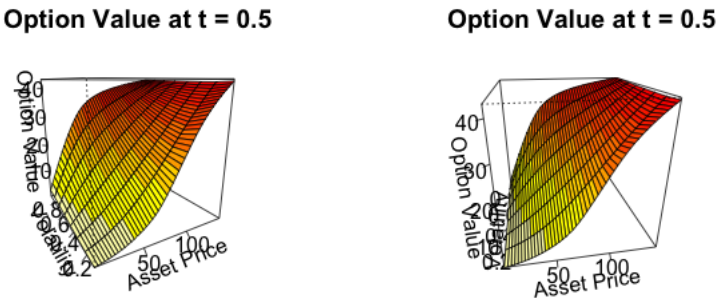
Compared to the Heston model below [5], [8]



M1 shows similar over low levels of volatility. The curvature grows more pronounced as volatility rises, and this phenomenon is likely attributable to a lack of discretization grid refinement. That is, the $\nu \to \infty$ boundary condition diffuses too quickly over the rest of the surface. Consequently, we must only evaluate the option prices M1 generates at lower volatilities.

5.2 M2 Option Value Surface

Taking a cross-section of the 3D approximation array at a single time describes how the option behaves over different volatility levels. Here we show two views of the option value with half its time left to expiry.



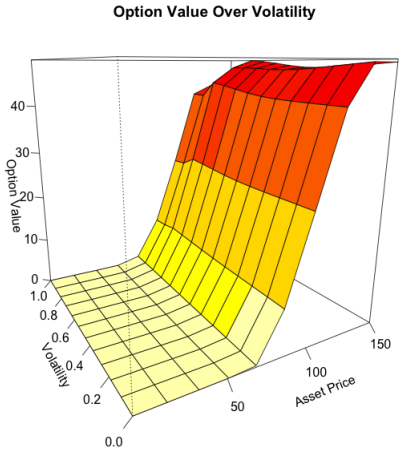
We see the explicit effects of volatility increasing option value in this figure. The stochastic volatility model describes how the option value "diffuses" between the boundary conditions.

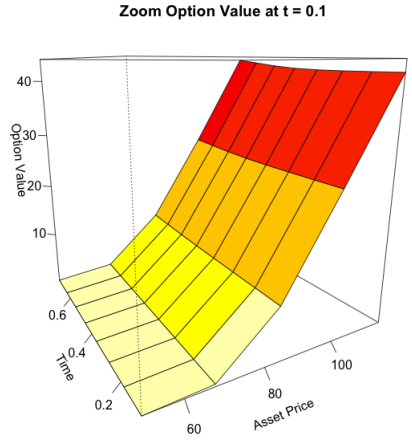
These plots show the general smoothing we expect. We see a more pronounced rise in option value for the lower underlying prices in this model than with a constant volatility model i.e. Black-Scholes. We attribute this phenomenon to the introduction of stochastic volatility which increases the chance the option will end up in-the-money for low initial underlying prices. Around the low strike K_1, we see an inflection point. Notice that above K_1 the option value rises at a constant or decreasing rate. This reflects the payoffs from selling the second call at a strike of K_2 we notice that the value approaches the upper bound value $K_2 - K_1$.

R code and data is posted on the Online Appendix. The run time for this algorithm is around five minutes.

5.3 M3 Option Surface

Plotting the surface for the Bull-Call Spread, we see





Since M3 calculates the nodes of the surface with nine different discretizations, it is not surprising that there exists warping near the edges in the first plot. Nonetheless, the first plot shows an increasing trend in option value as volatility rises. This is most pronounced for higher asset prices, indication that the volatility is correlated with the price of the underlying. As volatility approaches its upper bound, though, we see this trend fail. We attribute this inconsistency to our choice of exclusively backward difference approximations which continually lag behind the increasing forward trend beginning at the zero volatility boundary. Indicated by the light yellow coloring, when we compare the above surfaces to the surface at maturity, we see the rough beginnings of option value diffusion to lower strike prices. However, because of the high computational costs required to "set-up" the calculation at each node, the use of finer grid discretization (such as those used in M1 and M2) becomes impractical and we do not see the expected "smoothing". As we take a closer look at the option value near the strike price, we further see the beginnings of the familiar stochastic volatility smoothing.

5.4 Comparison to Monte Carlo

Here we compare the numeric approximations generated by the explicit FDM in models M1, M2, and M3 to Monte Carlo simulated option values. We refer to equations (46) and (47) to construct our price path.

We simulate the stock price path with an Euler-Maruyama algorithm, using the same parameters as in M1, M2, and M3. In [6], German defines $\mu(t, V_t)$ and $\sigma(t, V_t)$ as any functions such that the volatility process satisfies the Feller condition (i.e. remains non-negative). We define them as in Section 4 and use an Acceptance/Rejection algorithm to satisfy the Feller condition. After simulating the price path up to time t under the risk-neutral probability measure, we simply calculate the option payoff as $S_t - K$ and average the results over 10,000 trials.

We summarize the results of our Monte Carlo simulation and finite difference approximations below. We choose a low volatility so as not to capture the warping effects of the boundary conditions towards the upper edge of the grid discretization.

European Call Price

Stock Price (t = 0)	M1	Euler-Muryama MC (10,000)
50	11.39	0 +/- 0
75	25.85	1.91 +/- 0.01
83.33	42.74	8.39 +/- 0.05

Time to Expiry (Yrs): 3
Volatility (Annual): 30%

European Bull-Call Spread Price

Stock Price (t = 0)	M2	Euler-Muryama MC (10,000)
50	21.69	0 +/- 0
75	28.95	1.92 +/- 0.03
100	33.98	25.02 +/- 0.07

Time to Expiry (Yrs): 3
Volatility (Annual): 30%

European Bull-Call Spread Price

Stock Price (t = 0)	M3	Euler-Muryama MC (10,000)
58.33	0	0 +/- 0
75	0.98	0.73 +/- 0.01
83.33	9.13	8.02 +/- 0.02

Time to Expiry (Wks): 7
Volatility (Annual): 30%

Our results show massive FDM over-valuation relative to the MC simulation, particularly for M1. This lends support to the conclusion that the theoretical boundary conditions in M1 and M2 are at fault. Additionally this conclusion is supported by the M3 valuations. Though M3 has the sparsest grid discretization, the values it generates most closely follow the MC generated option values.

5.5 Future Research

The solution techniques presented in Section 4 are perhaps the most basic approaches to approximate the SV PDE (48) numerically. They calculate the value of option at time $t-1$ from the known values at time t. While straightforward to implement, their simplicity leaves much to be desired in terms of computational efficiency and accuracy. Future research may seek to apply semi-implicit finite difference schemes (such as the alternating direction implicit method) to solve German's SV PDE or linear, parabolic, 2nd-order, non-constant-coefficient PDEs similar to German's PDE.

Regardless of solution technique, approximating of the SV PDE (48) is vital to computing the first-order correction to the price of a contingent claim with stochastic volatility derived from indifference pricing. German [6] shows that this correction term is dependent on the vega of the option and the instantaneous volatility of the underlying. Future research may attempt to compute the first-order correction based on the M3 numeric approximations presented in this paper. Calculating the magnitude of correction to M1 and M2 prices relative to M3 prices would likely provide a good basis of comparison for determining the sensitivity of first order corrections to approximation errors.

References

[1] Paolo Brandimarte. *Numerical Methods in Finance and Economics: A MATLAB-based Introduction*. John Wiley Sons, Inc., 2nd edition edition, 2006.

[2] Mark H.A. Davis. Option pricing in incomplete markets. *Mathematics of Derivative Securities*, 1998.

[3] Mark H.A. Davis. Valuation, hedging, and investment in incomplete financial markets. In JM Hill and R. Moore, editors, *Applied Mathematics Entering the 21st Century*, pages 49–70. Society for Industrial and Applied Mathematics, 2004.

[4] C.S.L. de Graaf. Finite difference methods in derivatives pricing under stochastic volatility models. Master's thesis, Universiteit Leiden, September 2012.

[5] Vasillis Galliotos. Stochastic volatility and the volatility smile. Master's thesis, Uppsala University, August 2008.

[6] David German. Corrections to the prices of derivatives due to market incompleteness. *Applied Mathematical Finance*, pages 1–33, May 2010.

[7] David German. Overview of utility-based valuation. *Journal of Statiscal Theory and Practice*, 5(3):1–15, 2011.

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

[9] George J. Jiang and Yisong S. Tian. Extracting model-free volatility from option prices: An examination of the vix index. *Journal of Derivatives*, 14(3), 2007.

[10] Kenneth A Lindsay. Numerical solution of partial differential equations. University of Glasgow, Department of Mathematics, November 2005.

[11] Salih N. Neftci. *An Introduction to the Mathematics of Financial Derivatives*. Academic Press, p. 272, 1996.

[12] Geoffry Poitras. *The Early History of Financial Economics, 1478-1776. From Commercial Arithmetic to Life Annuities and Joint Stocks*. Edward Elgar Publishing, Inc., 2000.

[13] Mark Richardson. Numerical methods for option pricing. Oxford Department of Mathematics, 2009.

[14] Walter Schachermayer. Introduction to the mathematics of financial markets. Lecture Notes in Mathematics 1816 - Lectures on Probability Theory and Statistics, 2003.

[15] Steven E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance, 1st edition edition, 2004.

[16] John C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Society for Industrial and Applied Mathematicians, 2nd edition edition, 2007.

A Online Appendix

For source code, please see `https://github.com/jacob-roth/thesis-r-code`