Basic Stochastic Framework for Designing and Understanding Real Estate Models

David Park[1]

[1] Campbell Hall Episcopal School, CA 91607, USA

Correspondence: David Park, Campbell Hall Episcopal School, CA 91607, USA.

Received: January 8, 2025 Accepted: January 31, 2025 Online Published: February 10, 2025

doi:10.5430/ijfr.v16n1p35 URL: https://doi.org/10.5430/ijfr.v16n1p35

Abstract

The real estate market offers significant opportunities for financial success, yet navigating its complexities—such as tax benefits and regulatory influences—can be challenging. Stochastic processes play a critical role in minimizing risk and optimizing returns in this domain. Despite its historically stable growth, real estate remains underutilized, particularly among lower-income households, partly due to the limitations of traditional financial models like Black-Scholes. These models often struggle to account for the volatility, interest rate fluctuations, and government policies that influence real estate markets. This paper demonstrates the efficacy of integrating advanced stochastic models, such as the Heston model and jump diffusion models, to address these limitations. Additionally, we explore diverse methods for validating these models and highlight key considerations in their design to ensure greater accuracy and practical relevance. These insights aim to enhance the robustness of real estate modeling, providing a pathway for improved financial decision-making.

Keywords: stochastic calculus, probability, martingales, expected values, differential equations, simulations, market volatility, real estate, stochastic asset price modeling, random variables, financial risk management, tax manipulation, existence and uniqueness of solutions, randomness, processes, distributions

1. Introduction

48% of Americans are not holding any investment assets (Almazora, a2024). Why is nearly half of one of the most powerful economic countries not putting their money in assets? Investments provide ways to generate more income and support any financial situation. Ben Rizzuto, a wealth strategist working at Janus Henderson, remarks that it is due to a "lack of understanding around how to invest" *(Almazora, b2024) The question isn't why, but how. Real median household income charts, according to the Federal Reserve BAnk of St. Louis, have statistics indicating that the median income is around 80,000-81,000 dollars, which in states like California, is barely enough to survive and not enough to own a home (Real).

The real estate market offers a way towards financial freedom. But it is full of unexpected surprises, in terms of how the market fluctuates, and it can be extremely risky. Just a few mistakes in investments could damage your entire financial portfolio, and thus is a reason why many Americans do not attempt to become large renters. However, advances in the fields of Stochastic Calculus and Measure Theory have shown that even volatile markets like stock markets can be modeled to reduce risk. While the Black-Scholes isn't as widely accepted today due to its flaws, it revolutionized the industry by showcasing the power of stochastic concepts and predicting randomness. And, in a more stable market, such as the real estate market, it can be inferred that such models utilized in these fields can be even stronger.

The pair of questions that are now remaining to be answered, are what even are we modeling, and how will we model such a thing? Simply put, there are many different things to price in real estate. This can come in the forms of mortgage house pricing, predicting the median house cost, interest, etc. But the bigger question is, how will we model something that is relatively and mathematically practical?

This paper will provide a framework that will help form an understanding of how future stochastic models should be designed to help real estate investors manage their investments with minimal risk and high returns. It will give an introduction to several important concepts, demonstrate how they are significant or can be significant to the real estate market given that they usually model asset prices, and give particular areas of innovation to improve the accuracy of simulated stochastic models.

2. Why Stochastic Processes Are Required

2.1 Failures of the Riemann Integral

Fundamentally, it is important to understand the basics and all the contents of calculus taught during high school and the first years of undergraduate mathematics. Calculus deals with problems of change. Particularly, differential equations can model things like population capacities, and several other important every-day scenarios. These are all under the forms of Reimann calculus, where integration by partitioning rectangles and taking a limit of the sum of the amount of partitions, will result in the area of a given curve y=f(x).

Now, take a function f(x), where f(x) = 0 for all rationals, and 1 for all irrationals. How would you integrate this, say from the interval [0,1]? This function is discontinuous everywhere, and thus, is not differentiable, but more importantly, Riemann integration fails here. We need to now define an alternative solution of integration to deal with functions like these.

We begin to see how the Reimann integration method of rectangular x-value width partitions are not very effective for more complicated functions which have more unique analytical problems.

2.2 The Lebesgue Integral

To put it simply, we need to now focus on a new method of integration, using non-rectangle partitions on the x-axis. In fact, we are going to use a set of points in the range of the function. We will take the “measure” of each range value given that it is determined. Now, for those of you who haven’t taken measure theory, think of a “measure” as a “length” of some set in a very generalized way. With this, you may now think of the Lebesgue Integral as a “horizontal” integral, and now, the function f(x) which is 1 for irrationals and 0 for rationals can be integrated because the discontinuities are now irrelevant when applying the Lebesgue Integral. It is notated with the integral from 0 to 1 of f(x) with respect to μ, or dμ more specifically.

This integral has very nice properties, as it can have a wider class of functions, including those with complex discontinuities, and superior theoretical properties. In fact, it can integrate much better towards higher dimension functions. For example, integrating a volume using riemann techniques results in cubes as partitions, which makes integrating very complicated. There are also significant properties that come as a result of the method of integration, such as it being possible to exchange limits and integration under conditions, while doing this for a Riemann integral is entirely baroque.

2.3 Time

I will mention time very early in this paper. Time will be constant, and measured on the horizontal axis. Almost all models will have discrete time increments, such as taking measures of time from 0 to 1 second.

2.4 Randomness and Probability Space

Stochastic processes deal with random variables, as modeling asset prices is not consistent because stock prices are not deterministic. Probability is really at the core of stochastic calculus, as it is all about chances of events. In any model, it is important to recognize the space at which the model works. For modeling Real Estate pricing models, or general models that predict some price of a factor in this market, should be using a probability space noted with (Ω,F,P). The omega here is key, representing the sample space of all possible outcomes. A common and great example is a sample space of heads or tails. All possible outcomes of flipping a coin 10 times and getting a certain result (like 5 heads) can be denoted by this omega. The F (which in models will be the greek form), is a sigma-algebra, and this is very common to know for students who have taken measure theory. For the rest of us, think of it as the collection of subsets including omega and is closed under countable unions. P here is just a probability measure with a property of it being countably additive.

2.5 Probability Foundations

To effectively navigate stochastic processes, a solid understanding of probability theory is essential. Key concepts include random variables, which describe outcomes of random phenomena; probability distributions, which assign likelihoods to different outcomes; and expectation and variance, which provide measures of central tendency and variability. Additionally, familiarity with conditional probability and independence is critical for understanding how events relate to one another. Advanced topics such as martingales, Brownian motion, and the law of large numbers are foundational for stochastic calculus. These concepts equip practitioners to model uncertainty and randomness, as seen in real-world systems like financial markets and real estate. A deep grasp of probability enables the transition from deterministic models to stochastic frameworks, which better reflect the complexities of dynamic environments. A strong foundation in this field will guarantee a good foundation of stochastic calculus.

In fact, all stochastic processes are based on probability. What path a stochastic process chooses to take, the random variables in SDEs, are all based on "chance," and thus fit into probability theory concepts. The ideas of variance, or really standard-deviation in many models, and the central limit theorem, will all be fundamental probability concepts that will be taught in stochastic finance courses.

2.6 Measure Theory

A strong foundation in this field will guarantee success in developing extraordinary models. The reason for this is that measure theory is in almost every aspect of stochastic processes. Measure Theory contains concepts of spaces, measures of sets, the famous Lebesgue integral, and a large subset of unique properties of "measures" that allow stochastic processes to be created and proved. Standard integration techniques, theorems of elementary and ordinary differential equations, and more, will fail for SDEs. Thus, when taking an approach of an SDE with measure theory backgrounds, that's where the magic happens.

So, how does this play into stochastic calculus? The idea begins with probability spaces. All models must define a space to work with, each having their own unique properties, particularly with norms. Then, random variables can be defined as measurable functions. For example, a random variable X is a measurable function X: Ω-> R, meaning it maps outcomes to real numbers in a way that respects the structure of the sigma algebra.

A filtration is a family of sigma-algebras representing the information available up to time t, and thus can provide a framework for analyzing filtrations, which are central to modeling the evolution of a random process over time. Martingales also come from measure-theory concepts. We will get to this in a different section.

2.7 Combining Concepts and Stochastic Paths

Other concepts are, of course, always necessary. Understanding concepts of economics, abstract algebra, and even topology, can extremely enhance the accuracies of models if employed correctly. However, in general the real estate market can be modeled with simply stochastic knowledge itself. Arguably, it is clear from statistical data that real estate markets are much less volatile (NerdWallet).

Due to the nature of random elements incorporated into stochastic models, real estate pricing can be predicted with high accuracies that are previously undeterminable using ordinary differential equations which cannot model the unpredictabilities within asset prices.

Later, we will discuss simulations and errors that may be encountered. It is important to mention what "paths" are in stochastic calculus to end this section. Suppose you have a coin, and you can obviously only flip and receive one of heads or tails as a result. Say you do 2 flips. Your possibilities are {H,H}, {H,T}, {T,T}, and {T,H}. Think of the various different routes of outcomes as stochastic paths, and that a simulation through code can produce different possibilities. That is what a path is. You can imagine then, that as a larger discrete time interval for possibilities is given, the more variations and paths form to exist, and thus making predictions more difficult in the long run. We will return to this idea in the penultimate section discussing simulations.

2.8 Aims

This paper aims to create a basic framework and introductory guide to create stochastic models for the market of real estate, and of course, can be extrapolated to other financial markets, such as the commonly modeled stock market. The paper will demonstrate the fundamental requirements for designing a model, potential and inevitable errors of results, and applications of stochastic processes.

3. Real Estate Potential

3.1 Real Estate Market Trends

Recent historical data gathers that Real Estate markets have had an average annual return of 10.6% when including rental income. Commercial Real Estate returns are slightly lower. While it may be fair to argue that stock markets have had a 12.25% return (S&P 500) annually, it is important to note the extreme volatility of stock markets. Stock markets are very volatile, unlike Real Estate investments, which increase in returns as the value of the house increases statistically around 4% per year. It is clear that such investments are a practical and rational approach when choosing between the stock market (Real).

In fact, volatility in the stock market is so severe that many stochastic models fail. We will return to the failures of past models, such as the Black Scholes and Heston Models, and see where the innovations faltered, after we discuss the ideas of stochastic calculus.

Real Estate moreover, being in the rental market, offers a significant method of financial freedom over other forms of investments. Successful realtors and agents have made large fortunes off this market, and oftentimes, most wealthy people will have some connection to being successful in this market. It is even tied to the idea of the American Dream, and despite the political and social opinions surrounding this topic, it is important to note that the idea of ownership of property is rooted in the nation.

3.2 Tax Manipulation in Real Estate

A fantastic example of a way to “not pay taxes legally” is through the real estate market, particularly through something called a 1031 exchange. In short, this method allows a person to postpone capital gains taxes on the sale of an investment property by using the proceeds to buy a similar property. By using the equity of selling one investment property to purchase another property, realtors and people in this market can postpone their taxes (not make them go away). Yes, there are lots of restrictions and rules, but this is just a starting point. Refinancing or borrowing against home equity may also allow for manipulations of tax codes. Depreciation Deduction is another method where you can recover the “cost income-producing rental property through annual tax deductions called depreciation”. The IRS allows property owners to deduct the cost of wear and tear on income-producing properties over their useful life. This non-cash expense reduces taxable income and can often offset significant portions of rental income, lowering the overall tax liability. These are just a few of the most common methods to “avoid” taxes. Taking advantage of such manipulations, which are legal, can benefit one’s financial portfolio when executed safely.

These should show you just how amazing the real estate market can be for someone who seeks a better financial situation, and thus, making modeling with stochastic processes for such a market extremely attractive.

3.3 Mortgage and Interest

The main two things that stochastic models in this market should focus on are mortgage rates and interest rates. These two are by far the biggest factors that affect real estate deals. Accurately predicting mortgage rates in particular demand stochastic processes that evolve accurately over time, followed by some non-normal distribution. Mortgage rates fluctuate between states, with California’s being uniquely high due to population, standard of living, and other factors, while other states have lower mortgages. Interest rates are also very unpredictable with standard linear ODEs. Most importantly, interest rates will be a parameter of most models, because predicting the interest rate a year from now using stochastic processes is almost unintuitive. It is mostly the government that affects interest rates. Macroeconomics tells us this very clearly, when discussing loanable funds topics. The government’s interest rate charged to banks who borrow money will fluctuate and cause severe changes for the demand of loanable funds for the housing market. Now, it is generally somewhat predictable depending on how much political knowledge you have of the president and congress’s ideals of interest rates, but for the most part, interest will either be a fixed assumption or an element of your SDE model.

4. Stochastic Concepts, Model Failures, Potential Innovations

4.1 Stochastic Calculus and Brownian Motion

Top firms like J.P. Morgan, big investors, quants, and more, all utilize the extremely powerful mechanisms behind Stochastic Calculus, and particularly the Stochastic Differential Equations that will be prevalent throughout this section. Tagging on many of the ideas of ordinary differential equations, stochastic differential equations (noted as SDE’s from here on out), include the addition of a random element. SDE’s take in a deterministic element and a random element, and it is fitting for finance, because the price of an asset is not even close to being deterministic. If you knew the price of APPL tomorrow, and so did everyone in the world, then the act of simply knowing the price tomorrow will influence the price. The nature of the market is random to the core. Prices fluctuate, and they can be extremely volatile, a huge factor of consideration for many models that want to make more accurate predictions. However, it is possible to make models to predict prices, using stochastic processes.

Brownian Motion is a large factor when modeling SDE’s. The concept originates in the physics fields of mathematics, though unsurprisingly, as the movements of particles are random. This mechanism of randomness is then substituted into the world of finance to handle the extreme randomness of asset pricing. This is often written as B(t) ot W(t) in SDE’s. You could see how physicists would actually be brilliant collaborators in the financial mathematics world of stochastic calculus, as the factors of randomness are similar and often exchangeable.

Those studying this field should have strong foundations of measure theory and probability theory, as this field of mathematics is incredibly demanding and deals with significantly complicated situations.

4.2 Ito's Lemma

Ito's Lemma is one of the foundational results in stochastic calculus, playing a role analogous to the chain rule in classical calculus but adapted for functions of stochastic processes. It allows us to compute the differential of a function of a stochastic process, such as a stock price modeled by a Stochastic Differential Equation (SDE). This tool is crucial in quantitative finance, particularly in the derivation of option pricing models and risk management frameworks.

The derivation of Ito's lemma can be found through numerous sources, and will not be thoroughly derived, as the purpose of this paper is to help with navigating real estate risk and investments with introductions in the field of stochastic calculus. The derivation will require the use of Taylor's Theorem and expand a function f(X(t)) , t) around a small time increment/differential dt.

This lemma is highly relevant for modeling real estate markets when prices are treated as stochastic processes. For example, consider a real estate price P(t) modeled as:

$$dP(t) = \mu P(t)dt + \sigma P(t)dW(t)$$

If we are interested in a function of the price, such as the logarithm ln(P(t)), we can use Ito's Lemma to derive its dynamics.

The result of this will be of the following:

$$d(\ln(P(t))) = (\mu - 0.5\sigma^2)dt + \sigma dW(t)$$

It is important to note that σ^2 represents variance from statistics with σ simply by itself being the volatility term. From statistics, this is also referred to as the standard deviation, and it is of a random component in this field of mathematics. In real estate, it will often be represented as the variability or volatility of property prices due to factors like market shocks, policy changes, or economic conditions. μ is the drift term (or more specifically, μ(X(t), t)) meaning that it represents the deterministic rate of change of the process over time. It's often referred to as the expected return or growth rate of the process. In the context of real estate, it may represent the average annual appreciation rate of property prices.

An example in a Real Estate market may be something like this SDE (sample, not a legit model for use, and thus will require no proofs or simulations).

For a real estate price P(t):

$$dP(t) = \mu P(t)dt + \sigma P(t)dW(t)$$

The first term on the right hand side represents the average rate of price appreciation, while the second term represents the randomness in price changes due to market volatility,

4.3 Why Is the Black Scholes a Weak Model? What About the Heston?

The Black Scholes model fails particularly because of its underlying assumptions. It not only has too many assumptions, but too many incorrect ones. For starters, the model assumes that stocks have log-normal distributed returns. Real stock prices do not have such a distribution, and in fact, have skewed distributions (Chen). So why were these assumptions made? Well, one might notice that the gaussian distribution is very easy to integrate. The integral of e^-x^2, a simplified form of the gaussian, on the infinite interval, can be integrated by either Feynmann's Technique (Note 1) of differentiating under the integral sign with respect to some parameter, or by using double integrals and polar conversions, or even by some other method that results in root(pi) as a part of the result. Either way, this distribution is easy to work with, and so it was assumed. Other distributions have properties that make it very difficult to integrate elements of models with, or simply even to work with calculus, as they have various conditions and restrictions that are unique to each distribution. As an investor, it is important to avoid this model, and as an innovator, it is important to design models that have accurate distributions of asset prices, otherwise the model will fail. The model assumes that markets have no jumps or discontinuities, which may be okay for periods of time where the market is stable, but any sudden shock of stock prices can make the model ineffective. The model also assumes constant volatility, which does not hold in real markets where volatility often fluctuates. Future workers must ensure that new models incorporate volatility of volatility, and also incorporate potential randomness of jumps within the model, to make pricing as accurate as possible for large amounts of cases.

In the real estate context, real estate markets would not be able to adopt such models for predicting housing prices, as such markets exhibit changing volatility. In addition, real estate is an illiquid asset, and prices can experience sharp jumps due to sudden changes in policy or economic conditions.

The Heston model is an extension of the Black-Scholes that introduces stochastic volatility, addressing the previous issue of constant volatility. However, this model still lacks accounts for jumps and large discontinuities, and also makes several key assumptions that are not entirely true for markets. For example, volatility is assumed to revert to a long-term average denoted by the greek letter "theta", which may not be always true in markets subject to persistent shocks.

4.4 Poisson Distribution

It is important to understand why this distribution is recommended in the various forms of models that have been mentioned. We've already mentioned this once, but it is important to emphasize the weakness of the Gaussian Distribution. First of all, the Gaussian Distribution does not describe the changes in asset prices, and definitely not the real estate market. In fact, it isn't even close. Those who have created stochastic models that use such a distribution understand this. This is not saying that Fisher Black and Myron Scholes were ignorant, but they simply knew the limits of their mathematics. The normal distributions allow for easier solutions to be made, considering the smooth and easily-integrable nature of the normal distribution. However, innovations for real estate price modeling need to avoid such an inaccurate distribution. That's where poisson comes in.

The poisson process is useful in several different areas of financial investing. This distribution is a tool used in probability theory based statistics that predicts the amount of variation from a known average rate of occurrence in a given time frame (Corporate).

4.5 The Significance of Logarithms in Financial Calculus

Statistics and math majors love logarithms. Logarithmic properties open up doors to finding many unique solutions. For example, the Bernoulli Integral (Note 2) can be solved using the properties of logs to bring down the exponent "x" and using the gamma function and a Taylor expansion of e^x. Beyond this, logarithms allow us to do many things in calculus, such as using its properties to differentiate implicitly for functions like y= x^x. But the significance isn't about the practical applications, but more in the graphical applications of the function. For example, we can enhance the correlation of a statistics graph by taking the natural logarithm of both sides. If you take a graph of f(x) = x^2, a basic parabola, applying the logarithm to the right hand side will make it a logarithmic function times a scalar 2, which significantly flattens the curve. Now imagine a data set that is very chaotic, with points everywhere making the regression and correlation unclear. Applying a logarithmic term will significantly flatten such chaotic points to a clearer result. In statistics, we formally describe this process as making skewed or heteroscedastic data into symmetric and homoscedastic data for better analysis. We can apply this same technique to stochastic calculus. Regression models, especially those involving exponential growth or decay, can be linearized (De la Calle). As we will discuss later, this idea of linearization (or linear growth) will be extremely important for showing the existence of solutions. The logarithim's "flattening" nature is extremely useful for toning down complex and chaotic data that can be misleading or misread if without the log term. In statistics, it can transform skewed data, linearize relationships between variables, and most importantly, reduce the impact of outliers, as the data's range is compressed and thus the extreme values are brought closer to the expected value (De La Calle).

4.6 Martingales

To put it simply, martingales in stochastic processes is a sequence of random variables where the next value's expected value is equivalent to the present value. In more formal mathematical notation, this is more clearly defined as: $E[X_{n+1} \mid X_1, X_2 \ldots X_n] = X_n$. You can intuitively understand what this is really saying, which is that, in the case where Xn represents the price of a stock at some time, the knowledge you have about the price up to today should resemble the price tomorrow. In other words, the price shouldn't really change at all from the present day to the next day. For real-world stock prices, this is relatively true. Unless there are some extreme market factors, which yes, are almost always impossible to predict with any stochastic model, the price of tomorrow's stock shouldn't be unexpectedly different from today. Martingales are also great tools to utilize for betting. Take a fair game, where a martingale betting strategy will be implemented. You win the game (and the bet) if you land on heads.

As mentioned in section 1.6, martingales are generally very useful for finding conditional expectations. An example of what such a martingale might look like is this: $E[M_t \mid F_s]=M_s$, for $s \leq t$. The "E" represents taking an expected value, which is really just mean. Martingales are also central to the no-arbitrage principle in financial modeling. In real estate, this can be applied to ensure that the pricing of real estate derivatives, such as futures or options, reflects a fair market value. Furthermore, martingales are used to construct hedging strategies for real estate investments. For example, real estate-backed securities or mortgage-backed securities can be analyzed using martingale approaches to minimize risk.

5. Sub-Models and Other Areas for Innovation

5.1 Levy Noise

Take a continuous function f, with it being a smooth curve. Suppose you tried to model asset prices with such a curve. Of course, this is very inaccurate to the very spiky graphs you often see regarding financial asset prices. In order to make a model adapted to such a graph using continuous functions and algebraic/transcendental functions, we need to input a new element called “noise”. Noise is what makes a graph have that “spiky” nature, and it is an important part of all stochastic differential equation models. A levy process is a fantastic element to include in new innovations, and deals with a poisson distribution with an intensity defined with λ. Models in the real estate market must include some sort of noise factor, otherwise the entire point of the stochastic process is entirely meaningless.

5.2 Jump Diffusions

Previous models, such as the Black-Scholes, fail to incorporate jump diffusions, making any irregular shocks in the market entirely unpredictable for the investor using the model. Thus, new models should incorporate such a model. A sample would take form of something like this:

$$dS_t = \mu S_t dt + \sigma S_t dW_t + S_t dJ_t$$

dJ_t will depend on the jump process and is modeled, as stated earlier, by a poisson process of intensity λ.

Real Estate pricing models should follow such jump diffusions to handle market irregularities. Without jump diffusions, any external factors that make significant impacts to the real estate market, often coming from government regulations on interest rates, are essentially ignored by a given model.

5.3 Fractional Brownian Motion

Fraction Brownian Motion, or often referred to as Fractal Brownian Motion, is a generalization of standard Brownian motion that incorporates memory effects and long-range dependencies, making it more suitable for certain real estate dynamics. Unlike standard B.M., which assumes independent increments, F.B.M. allows for positively or negatively correlated increments, as determined by its Hurst parameter (H). This feature aligns well with real estate markets, where price movements and trends often exhibit persistence (e.g., prolonged growth in housing markets) or mean-reverting behaviors (e.g., correction after a bubble). Additionally, fBM captures the heteroscedasticity and irregular volatility patterns observed in real estate prices, which are influenced by factors like economic cycles, policy changes, and demographic shifts. By incorporating these realistic market characteristics, fractional Brownian motion provides a more nuanced framework for analyzing and forecasting property prices, risk assessment, and investment strategy development compared to traditional models based on standard Brownian. It may be wise to take on such a random element and incorporate it into pricing models.

5.4 Markov Chains

Lastly, it is important to mention Markov chains. These are a type of stochastic process that describes a sequence of events where the probability of transitioning to a future state depends only on the present state and not on the sequence of past states. In other words, a Markov chain is a stochastic model that uses mathematics to predict the probability of a sequence of events occurring based on some recent or present event. It is how search engines can predict what you may be typing based on your previous search history, and is part of how they promote content that fits your interests (Markov). This property is called the Markov Property. Formally, for a Markov chain with states S_1, S_2,...S_n, the transition satisfies $P(X_{t+1}=S_j \mid X_t = S_i, X_{t-1},...X_0) = P(X_{t+1}=S_j \mid X_t=S_i)$.

6. Hybrid Systems

6.1 What Are Hybrid Stochastic Systems?

Innovations in this area of stochastic processes can be the most influential if done correctly. Hybrid stochastic systems are mathematical frameworks that combine continuous dynamics with discrete events, making them highly suitable for modeling real-world systems characterized by a mix of deterministic, stochastic, and event-driven behaviors. These systems are increasingly relevant in fields such as finance, real estate, and engineering, where interactions between continuous processes and abrupt changes play a critical role. These incorporate two key components: Continuous stochastic processes, and discrete event dynamics.

Continuous stochastic processes include models like Geometric Brownian Motion or SDE's, which describe gradual changes driven by randomness (e.g., price fluctuations, interest rate movements).

Discrete event dynamics capture sudden shifts or transitions at specific times, such as market shocks, regulatory changes, or structural alterations in a real estate portfolio.

6.2 Applications of Hybrid Systems to Real Estate Markets

Any system of hybrid elements containing continuous and discrete equations may come up within new innovations. These are particularly well-suited for real estate markets due to their ability to model both continuous trends (e.g., steady price appreciation) and discrete events (e.g., zoning changes, economic recessions). Take, for example, this hypothetical stochastic model: $dP(t) = \mu P(t)dt + \sigma P(t)dW(t) + J(t)dq(t)$

Here, $J(t)$ is the jump magnitude, and $q(t)$ is a counting process (like a poisson process mentioned earlier) that triggers jumps at random intervals. It takes on both continuous and discrete events. Particularly, the discrete events include economic shocks, policy changes, and other elements that may impact the model's design.

Hybrid Stochastic systems offer several advantages. By incorporating both continuous and discrete dynamics, these systems provide a more accurate representation of real estate markets, with outside influences that are discrete now being factored into the randomness of the simulated paths. Of course, as a consequence, these models will be more effective for minimizing risk for the investor, and thus optimizing the return on investments.

7. Errors and Difficulties of Modeling

7.1 Simulation Errors

Simulations are a great starting point for showing a model's applicability. As a start, simulating predicted asset prices against real-time or recently historical asset prices of the real estate market is great, particularly if the simulation is running against a fairly non-volatile market. Think of this as a "test run" for future simulations, or some sort of indicator of what future simulations may turn out to be.

Often when simulating stochastic results, say 10 paths against a real live asset, say AAPL, the simulation will have an obvious trade off. It is simple. The less accurate the model, such as less incorporations of different models and elements of randomness, the faster and more likely that a result appears. However, making a model more accurate, will result in a slower computation speed of the simulations, and in many cases, not even super-computers would be able to handle such simulations and stochastic paths.

Now, this problem of course, can simply be solved with computer advancements, but as of right now, unless one has access to computers such as those in NASA or any prototypes of quantum computers, one must make a decision between these trade-offs. Typically, simulations are shown to the most optimized result, where the simulation isn't too accurate, but isn't also inaccurate, and thus being able to produce at least some simulated path. In addition, over larger time intervals, the predicted outcomes will become less and less accurate, which will be an inevitable consequence to these models. As the time intervals sparse out, the approximations and randomness become less and less meaningful to real-time events, and inputting historical data into simulations will still result in far out investment predictions to become less accurate, unless an asset is very non-volatile.

Despite the real estate market being much more consistent than the stock market, it is an unfortunate truth that even this market will be susceptible to inaccurate approximations as the years go farther out. A great example of what this really means is to think of linear approximations. If you use a tangent line to approximate values of a function $f(x)$, it is clear that as you get further and further away from the point of tangency, the approximation is less and less accurate (although this isn't entirely true for some functions whose behavior may be more resistant to such methods of approximations).

7.1.1 Coded Simulations

This subsection will include simulations of previous models to further emphasize the motivations behind needing improved stochastic models. We will begin with a simulation of the BSM Model, the foundation for all stochastic innovations.

```
import numpy as np
import yfinance as yf
import matplotlib.pyplot as plt
import pandas as pd
```

```
# Fetch historical AAPL data
ticker = 'AAPL'
start_date = '2024-01-01'
end_date = '2025-01-01'

# Download the data
aapl_data = yf.download(ticker, start=start_date, end=end_date)

# Print columns to check the correct column name for closing price
print(aapl_data.columns)

# If "Adj Close" is not available, use "Close"
aapl_data['Returns'] = aapl_data['Close'].pct_change()    # Use "Close" if "Adj Close" isn't available

# Estimate drift (mu) and volatility (sigma) from historical returns
mu = aapl_data['Returns'].mean() * 252    # Annualized drift (mean return * 252 trading days)
sigma = aapl_data['Returns'].std() * np.sqrt(252)    # Annualized volatility

# Simulation parameters
S0 = aapl_data['Close'].iloc[0]    # Initial stock price
T = 1    # Time horizon in years (1 year for simplicity)
dt = 1/252    # Daily time step, assuming 252 trading days in a year
N = len(aapl_data)    # Number of time steps (same as the number of trading days in the period)
n_simulations = 10    # Number of simulations to run

# Generate random Brownian motion (Wiener process)
np.random.seed(42)    # For reproducibility
dW = np.random.normal(0, np.sqrt(dt), size=(n_simulations, N))    # Wiener process increments

# Simulate stock price paths using the Black-Scholes model (GBM)
S = np.zeros((n_simulations, N))
S[:, 0] = S0    # Set initial stock price

for i in range(n_simulations):
    for j in range(1, N):
        S[i, j] = S[i, j-1] * np.exp((mu - 0.5 * sigma**2) * dt + sigma * dW[i, j-1])

# Plot historical AAPL data
plt.figure(figsize=(14, 7))
plt.plot(aapl_data.index, aapl_data['Close'], label='Historical AAPL Price', color='blue')

# Plot the simulated stock price paths
t = aapl_data.index    # Use the same dates as the historical data
for i in range(n_simulations):
    plt.plot(t, S[i], lw=1, alpha=0.6)

# Add labels and title
plt.title('Historical vs. Simulated AAPL Stock Prices')
plt.xlabel('Date')
plt.ylabel('Stock Price ($)')
plt.legend()
plt.grid()
plt.show()
```

The following simulation results in various different stochastic paths as shown in Figure 1.

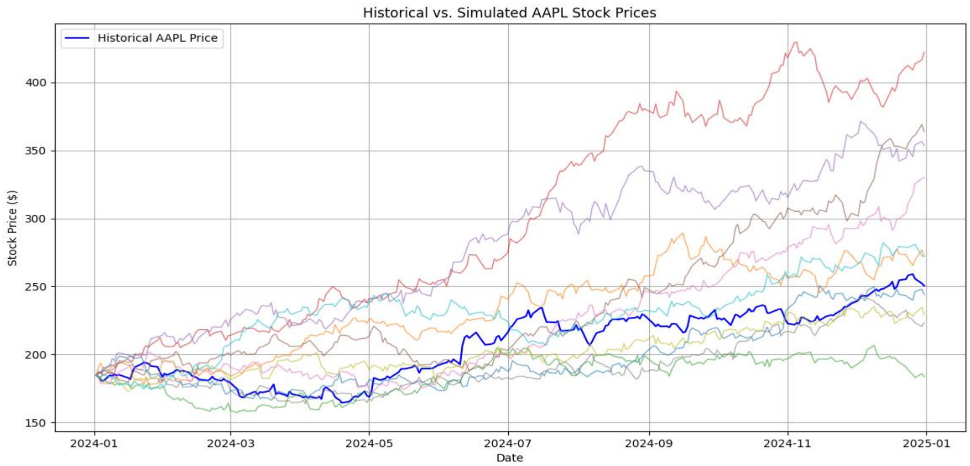

Figure 1. Black Scholes Simulation against AAPL Prices within 2024-2025

We see the Black-Scholes Model is not entirely inaccurate for certain paths, but even so, the gaps are sufficient to motivate any mathematician to innovate.

Next, we will simulate the Bates model. The Bates Model is an extension and an improvement of the Heston Model, though often critiqued for its difficulties with parameters and extremely theoretical nature. It captures continuous time Bates-Stochastic volatility, and uses MS-GARCH dynamics. While it captures the leverage effect and handles jumps, we see through the simulation of this model against Microsoft's prices, that there are still gaps present. Here is the code for the simulation:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import yfinance as yf
from scipy.stats import norm, poisson

# Fetch Microsoft stock data
ticker = "MSFT"
data = yf.download(ticker, start="2020-01-01", end="2025-01-01")

# Ensure data was downloaded
if data.empty:
    raise ValueError("Failed to fetch data for MSFT. Check ticker or internet connection.")

# Print column names for debugging
print("Available columns:", data.columns)

# Use 'Close' instead of 'Adj Close' to avoid KeyError
log_returns = np.log(data['Close'] / data['Close'].shift(1)).dropna()

# Estimate initial parameters from data
S0 = data['Close'].iloc[0]    # Initial price
mu = log_returns.mean() * 252      # Annualized drift
```

```
sigma_0 = log_returns.std() * np.sqrt(252)    # Initial volatility estimate

# Bates Model Parameters (Heston + Jumps)
v0 = sigma_0 ** 2      # Initial variance
kappa = 2.0             # Mean reversion speed
theta = v0              # Long-run variance
sigma_v = 0.2           # Volatility of variance
rho = -0.5              # Correlation between asset and volatility
lambda_j = 0.3          # Jump intensity (Poisson mean)
mu_j = -0.02            # Mean jump size (negative for downward jumps)
sigma_j = 0.1           # Jump volatility
T = len(log_returns)    # Number of days
dt = 1 / 252    # Daily time step
n_sims = 1    # Number of simulations (single path for comparison)

# Simulating the Bates Model
np.random.seed(42)
S_paths = np.zeros((T, n_sims))
V_paths = np.zeros((T, n_sims))
S_paths[0] = S0
V_paths[0] = v0

for t in range(1, T):
    Z1 = norm.rvs()
    Z2 = rho * Z1 + np.sqrt(1 - rho ** 2) * norm.rvs()

    # Stochastic variance (Heston)
    V_paths[t] = np.maximum(V_paths[t-1] + kappa * (theta - V_paths[t-1]) * dt + sigma_v * np.sqrt(V_paths[t-1] * dt) * Z2, 0)

    # Jump process
    Jumps = poisson.rvs(lambda_j * dt)
    jump_size = np.exp(mu_j + sigma_j * norm.rvs(size=Jumps)).sum() if Jumps > 0 else 0

    # Asset price dynamics
    S_paths[t] = S_paths[t-1] * np.exp((mu - 0.5 * V_paths[t-1]) * dt + np.sqrt(V_paths[t-1] * dt) * Z1) * (1 + jump_size)

# Plot Real vs Simulated Prices
plt.figure(figsize=(12,6))
plt.plot(data.index[1:], data['Close'][1:], label="MSFT Real Price", color="blue")
plt.plot(data.index[1:], S_paths[:, 0], label="Bates Model Simulated", color="red", linestyle="dashed")
plt.title("Bates Model Simulation vs. Microsoft Stock Price")
plt.xlabel("Date")
plt.ylabel("Stock Price")
plt.legend()
plt.show()
```

Figure 2 will showcase a sample simulation with a single path.

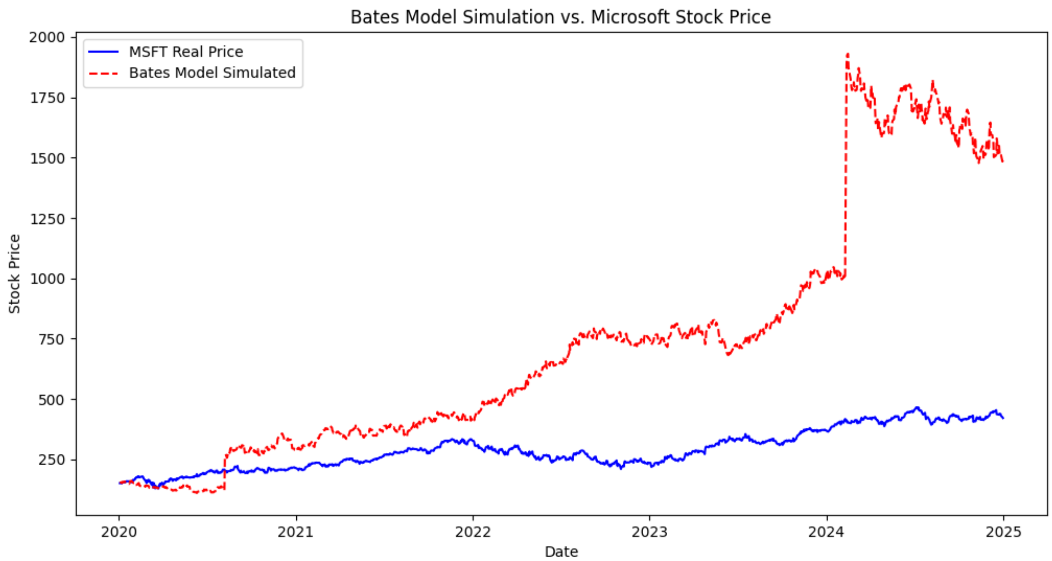

Figure 2. Bates Model against Real-Time Microsoft Prices

It is important to note that simulations also only showcase stochastic paths under favorable conditions, with constraints that often ignore outside real-world factors that can increase volatility and thus the inaccuracy of simulations. Historical data is also not enough to showcase a model's full capacity, because historical data is still insufficient to external factors. Thus, while successful simulations are a great indicator of a successful model, more evidence and derivations are required if such a model is to be deemed effective. It is also clear, from the simulations, that indeed the further out we go into the future, the less accurate the model's predictions become. In fact, the Bates Model is extremely inaccurate past the 2022 time stamp. Such gaps are great motivations for new innovation.

7.2 Uniqueness and Existence of Solutions Using Lipschitz Conditions

Once any model has been designed for this market, it must be shown to have a unique solution, and even before that, an existence of solutions. A common approach to showing this is through the Lipschitz continuity condition, and through showing that the model passes the linear growth test. Yes, these were used originally for ordinary differential equations, but it can be extended to stochastic differential equation solutions.

When studying ordinary differential equations, showing a uniqueness of solutions was done through the Picard-Lindelof Theorem. Suppose you have an ordinary differential equation $dy(t) / dt = f(t, y(t))$, with $y(t_0) = y_0$, and $y(t)$ is the unknown solution, t is the independent variable, and y_0 is the initial condition. The theorem states that this O.D.E. has a unique solution if it follows two conditions: 1. $f(t,y)$ is continuous with respect to t and y; 2. $f(t,y)$ satisfies a Lipschitz condition with respect to y on some domain. The first part can be shown easily. However, the second condition is fairly new for many students.

Now, before moving on, it is important to showcase what this second condition is. A function that maps the reals is said to be Lipschitz continuous on a domain "D" of real numbers if there exists a constant $L \geq 0$, where L is a Lipschitz constant such that for all $x,y \in D$.

While this theorem does not directly apply to SDEs in its original form because it is a result for a deterministic ordinary differential equation, there is an analogous result in the stochastic setting that provides existence and uniqueness of solutions under similar conditions. Suppose you have the following SDE of the form: $dX_t = f(t, X)dt + g(t, X)dW_t$ with $X_0 = x_0$, and where: X_t is the unknown stochastic process, $f(t,X_t)$ is the drift term (deterministic), $g(t,X_t)$ is the diffusion term (random), W_t is the Wiener process or Brownian motion element, and x_0 is the initial condition. For existence and uniqueness of solutions, they can be guaranteed if both $f(t,x)$ and $g(t,x)$ are Lipschitz continuous in x, uniformly in t. In other words, there exists a constant $L \geq 0$ such that

$\| f(t,x_1)-f(t,x_2) \| + \| g(t,x_1)-g(t,x_2) \| \leq L \| x_1-x_2 \|$ for all x_1 and x_2 and t. But moreover, the process must also follow a linear growth condition to ensure that the solution does not blue up in finite time.

The intuition behind the Lipschitz condition for f(drift) and g(diffusion) ensures that small changes in the state do not lead to disproportionate changes in the output. This prevents the stochastic process from branching into multiple paths, analogous to the deterministic case. Similarly, the growth condition ensures the process remains well-behaved and finite over time.

Overall, there may be alternative methods to the Lipschitz condition, especially for models that showcase uniqueness of solutions through simulations, but cannot be easily done through these conditions. However, it remains clear that a model cannot be determined to be persistently accurate if it cannot be shown that it has a unique solution and an existence of such a solution.

7.3 Proofs Behind the SDE's

Finally, any mathematical innovation requires a proof. Despite simulations showing uniqueness of solutions or even existence of solutions, it cannot be extrapolated to show that a model is legitimately effective. Simulations can be argued with, but proofs cannot. Proofs are definite, absolutes, and nobody can argue against it. If you can prove that your model works by any means of derivation, then it is sufficient to show that the model is applicable to general markets, and in our case, real estate markets.

A great starting point for proofs is to showcase the existence and uniqueness of solutions, otherwise the point of attempting a proof is impractical. While numerical methods such as this can be used as "proofs," they are much weaker than other forms of proofs. For stochastic processes, similar to derivations of the Black-Scholes Model, once a solution can be shown to be in existence and unique, the proof can now move on to other methods. A common approach is to use Ito's Lemma, the chain rule counterpart for stochastic processes. Other methods may include derivations from first principles, showing solutions of PDEs like the Black-Scholes match the outputted data taking in historical trends, and more. But one cannot get away from the fact that proofs are necessary.

Before moving on to the conclusion, it is important to note that SDE proofs will become much more difficult as more accurate elements are incorporated. For example, it is much easier to show a proof for a SDE without the Levy Process element, which yes, makes the model less accurate, but also, can be proven.

7.4 Assumptions

Without a doubt, the assumptions for any model must be made. Often trade-offs will be made, whether it is losing accuracy for faster simulations, or increasing the accuracy of the model, but making the task of simulating any results or coming up with concrete proofs much more difficult. Regardless, when choosing a stochastic optimization process with an objective function with certain constraints, one must note these constraints and any key assumptions behind stochastic systems. For example, you must note the distribution behind the SDE systems. Is it normal? You must be very clear about the assumptions, otherwise, it is very easy to point out obvious mistakes and be discredited for designs of models.

8. Conclusion

Stochastic mathematics offers a wide variety of opportunities to enhance investment portfolios in the real estate market, as well as minimizing risk. Due to the real estate market's consistent growth, averaging above a 4% annual rate of return for investors, stochastic models can be even more effective in this market than a stock market, which is much more volatile.

This paper has described the several basic and introductory methods, and requirements, into designing a stochastic model. In addition, I have described the necessary requirements for showcasing a stochastic model's validity. Simulations can be shown to be effective, but further proofs such as demonstrations of uniqueness of solutions and derivations using techniques such as Ito's Lemma can lead to concrete validations of stochastic models for real estate pricing. This review has showcased various important concepts that any model designer should incorporate into their studies and eventually into their models in order to optimize the success and accuracy of the stochastic model. Moreover, it can be concluded that the future of stochastic asset pricing revolves around combining various models that can cover each other's flawed assumptions, and also potentially branching more into the world of levy processes and hybrid systems, where both discrete and continuous equations can join into a single model.

Common flaws will occur most frequently in the area of proving such stochastic models. Incorporating the Levy Process, without significant graduate level mathematical knowledge of measure theory and financial mathematics, will come as a great obstacle for demonstrating proofs. Often, when using hybrid stochastic systems, proofs using certain inequalities such as the Gronwall Inequality or applications of the Variance Gamma Process, as examples, are difficult to reason about. It may be unclear if a certain inequality or formula is able to be utilized under certain parameters and conditions.

Current limitations of the study mainly stem from the technological issues with basic computers today. While the future of quantum computing looks very promising, due to the fact that we are unable to run simulations of great accuracy for longer periods of time (thus models that incorporate lots of elements of jumps, variability, and hybrid systems), this study's descriptions are mainly theoretical. Combinations and proofs of the models indicated possible for development may turn out with not even any existence of solutions, but until simulations are run to numerically solve complicated SDE's, this study is limited to theory itself. Nevertheless, many of the insights are still applicable, as the math is clear and consistent.

If this framework is followed carefully, given one takes time to learn what the individual elements are and how to incorporate them into new models, then more accurate models that reflect realistic circumstances of the real estate market will significantly improve the risk management of current investments using quantitative financial models. Too often are models such as the Black-Scholes evade the consequences of creating false assumptions about the market for ease of use. This paper has provided many of the missteps, troubles, and errors of past innovations, and through careful and thorough applications of these ideas, mathematicians and financial quants will develop more accurate models.

Acknowledgments

None

Authors' contributions

David Park was the sole contributor to this paper, and has provided all the contents.

Funding

None

Competing interests

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

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal and publisher adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data is available.

References

Almazora, L. (2024). Why Don't Nearly Half of Americans Have Any Investments?. *InvestmentNews*, 9 Oct. 2024. Retrieved 14 Jan. 2025, from https://www.investmentnews.com/industry-news/why-dont-nearly-half-of-americans-have-any-investments/257620

Brilliant.org. Differentiate Through the Integral. *Brilliant.* Retrieved 16 Jan. 2025, from https://brilliant.org/wiki/differentiate-through-the-integral/

Chen, J. (2015). How to Circumvent the Limitations of the Black-Scholes Model. *Investopedia*, Dotdash Meredith, 10 Apr. 2015. Retrieved 9 Jan. 2025, from www.investopedia.com/articles/active-trading/041015/how-circumvent-limitations-blackscholes-model.asp

Corporate Finance Institute. Poisson Distribution. *Corporate Finance Institute*. Retrieved from corporatefinanceinstitute.com

De la Calle, J. (2023). Best Tips and Tricks: When and Why to Use Logarithmic Transformations in Statistical Analysis. *Medium,* 9 May. 2023. Retrieved 13 Jan. 2025, from juandelacalle.medium.com/best-tips-and-tricks-when-and-why-to-use-logarithmic-transformations-in-statistical-analysis-9f1d72e83cfc

Kumaresan, M. Bernoulli Trials and the Binomial Distribution. *Scarsdale Public Schools.* Retrieved 18 Jan. 2025, from https://www.scarsdaleschools.k12.ny.us/cms/lib/NY01001205/Centricity/Domain/1382/bernoulli.pdf

Markov Chain. *Built In.* Retrieved 18 Jan. 2025, from https://builtin.com/machine-learning/markov-chain#:~:text=A%20Markov%20chain%20is%20a%20stochastic%20model%20that%20uses%20mathematics,your%20previous%20entry%20within%20Gmail

Real Estate Vs Stock Investing: Historical Performance | The Luxury Playbook. 9 July 2024. Retrieved from https://theluxuryplaybook.com/real-estate-vs-stock-investing-historical-performance/

Notes

Note 1. A Technique of Integration often used for improper integrals that cannot be evaluated with elementary integration techniques. It was made popular by Richard Feynman, but originally written by Gottfried Wilhelm Leibniz. Brilliant.org, "Differentiate Through the Integral," *Brilliant*, https://brilliant.org/wiki/differentiate-through-the-integral/, accessed January 16, 2025.

Note 2. The integral of x^x on [0,1]. This famous integral cannot be directly solved, and the common solution results in a summation. Michael Kumaresan, *Bernoulli Trials and the Binomial Distribution*, Scarsdale Public Schools, https://www.scarsdaleschools.k12.ny.us/cms/lib/NY01001205/Centricity/Domain/1382/bernoulli.pdf, accessed January 18, 2025.