# Beyond Black-Scholes Option Theory

Dennis Yang
*dennis.yang@atmif.com*

April 3, 2007

**Abstract**

This article presents a new option pricing principle that is more useful than the no-arbitrage principle, especially for incomplete markets. The focus here is on ideas behind mathematics—why the new theory is warranted, and how common sense dictates its construction.

## I. Complete and Incomplete Markets

The real financial world is too complicated for anyone to understand it completely. Thus a simplified model world is needed for two purposes: (i) To help us understand the real world, in the sense that a real-life phenomenon is explained if it is expected in a model world. (ii) To help us make decisions in the real world, *i.e.*, applying rules proven to be correct in the model world to the real world.

There are two types of model world in option theory; one is the so-called complete markets and the other, incomplete markets. The difference is that in complete markets, as in the famous Black-Scholes framework, derivatives carry no risks because they can be completely eliminated by the delta hedging scheme; whereas in incomplete markets, derivatives have inherent risks because perfect hedges are impossible. An example of incomplete markets is where the underlying is not a tradable instrument, *e.g.*, real options and weather derivatives. Even if the underlying is continuously tradable, the presence of transaction costs, stochastic volatilities or jumps etc. renders the model market incomplete. In short, any imperfection of a complete market model makes it incomplete, so there are a lot more incomplete markets than complete ones.

Complete markets are total fiction; reality is better modeled by incomplete markets because of unavoidable risks. For example, to take on a larger position in the real world, you demand a better unit price to compensate for the extra risks. But this natural risk averse behavior contradicts predictions of any complete market model where there is a unique price regardless trading sizes.

The Black-Scholes-Merton complete market option theory is *complete*, *i.e.*, well understood. Pricing options by replication is widely written and taught. However, despite its importance, the current option theory for incomplete markets is very *incomplete*. I will explain in this article why this is the case, and how it can be *completed*.

## II. Position and Risk

When you are making a risk related decision, it is *your* risk that really matters. Your risk comes from your position, without a position you have no risk. I view this statement as common sense.

Should you disagree with me on this point, then there is no point of reading the rest of this article, as other conclusions all rely on this observation.

Unlike complete markets, position and risk are inseparable in incomplete markets. Because different people have different positions, the concept risk, and anything related to it, should be viewed from an *individual* perspective. By the same logic, risk from the market aggregate perspective is not meaningful, as the net derivative position for the whole market is always zero.

Without further ado, I now present the central theme of this article: Incomplete markets are about risks that cannot be hedged away, risks come from having a position, therefore any consistent option theory for incomplete markets *must* contain derivative position information.

# III. Current Status

I now review some attempts people have made in extending the complete market option theory to incomplete markets.

(i) No-arbitrage principle: In incomplete markets, applying the no-arbitrage principle gives a range of possible option prices, in contrast to that of the complete market case where it produces a unique price for an option. In most incomplete market models, this range is so wide that it becomes useless as a "price". As an example, for a defaultable zero coupon bond with unit face value, the arbitrage-free price range is between zero and one [5].

(ii) Delta hedging methodology: In complete markets, risks associated with any derivative position can be eliminated by delta hedging. Does this methodology have natural extensions in incomplete markets? It is obvious that delta hedging is not possible for those incomplete markets where the underlying is not tradable. As it turns out, even when the underlying is continuously tradable in incomplete markets, the delta hedging scheme is not optimal (see Section 7.6 of [4]) in a sense to be made clear in the next section.

(iii) Risk neutral valuation: As far as I know, the phrase "risk neutral" has a two-fold meaning. One is that instead of being risk averse or risk seeking, one's preference towards risk is neutral. The other meaning is not having any risk, which is what "risk neutral" means in complete markets. So what does risk neutral mean in incomplete markets? Clearly not in the preference sense; how about not having risks? That is impossible, because by definition an option's payoff function in incomplete markets cannot be replicated. Thus "risk neutral" is an oxymoron in incomplete markets. If the phrase "risk neutral measure" is meant for the pricing measure in which the drift of the continuously tradable underlying is replaced by the risk-free interest rate, then such a pricing measure is suboptimal in incomplete markets [4, 6].

Because risks can be completely eliminated in complete markets, the position information is irrelevant. Since the aforementioned approaches all originate from the complete market option theory, none contains position information. As argued before, position is the source of risks in incomplete markets. Without recognizing this point, any attempt to generalize the complete market option theory to incomplete markets is doomed to fail.

In practice, people use an ad hoc fitting scheme (market calibration) to finesse the difficulties of incomplete markets. To be concrete, I use stochastic volatility models in equity markets as an example. The option pricing equation derived from the no-arbitrage principle in this case contains an unspecified risk related term (market-price-of-risk), which is supposed to be a *deterministic* function of state variables. To use the option pricing equation, the market-price-of-risk term must be identified. The common practice is to deduce this term by fitting model prices of vanilla options to those of the observed market prices, then use the equation to price exotic options. The fitting framework is unsatisfactory because it is unable to price the basic instruments—vanilla options. In addition, there is a more serious problem. Due to unpredictable short term supply and demand, the market price of a vanilla option, or rather its implied volatility fluctuates in a stochastic manner. Thus instead of being deterministic, the fitted market-price-of-risk actually changes stochastically with each calibration. This creates a logical inconsistency, *i.e.*, the fitting procedure introduces a source of randomness that is not modeled. Market calibration procedures for other incomplete markets suffer similar inconsistency problems.

Inconsistency in a model world is unacceptable, because the very purpose of using a model is to setup a logically consistent framework to prove or disprove certain assertions. Saying the common practice is logically inconsistent is akin to shouting "The emperor has no clothes." I am neither the first nor the only one to make such an observation. This view has even made into some textbooks such as [1, 3]. However, the difference this time is that I have made a royal robe for the emperor. The calibration procedure will be re-examined later to shed new light on it.

I hope that I have convinced you that the current option theory for incomplete markets is indeed insufficient. I will explain two concepts next before presenting the remedy that fixes the problem.

# IV. Optimality

In incomplete markets, the final wealth, or profit and loss (P&L), of applying certain trading strategy is a random variable. For example, suppose the outcomes of using hedging strategies A and B are Gaussian random variables with means 1.0 and 0.6, and standard deviations 1.0 and 0.5, respectively, which hedging strategy do you prefer to use? Note that this is a preference question, so there is no right or wrong answer. The key point is that you must answer the question of how to make decisions under uncertainty when markets are incomplete.

It is customary in economics and game theory to use the expected utility framework to resolve the preference issue. By taking expectation of a utility function, each probability distribution of the final wealth is now mapped into a real number that can be ranked. The basics of the expected utility theory are not difficult to grasp; introductory literature is available on the Internet in case you are unfamiliar with the subject and want to learn.

In complete markets, after applying the delta hedging strategy, the probability distribution of the final wealth is a Dirac delta function, *i.e.*, a certainty, hence no risks. Given two delta functions representing two wealth levels, there is no need to use a utility function to make a choice. Partly because of this, there is a common misconception that utility theory is futility theory for pricing derivatives. The misconception may also be due to the fact that current academic research on utility theory and option pricing has yielded little tangible results for practitioners. This misconception is unfortunate and unwarranted. The bottom line is that utility theory is superfluous in complete markets, but necessary in incomplete markets.

Once a utility function is chosen, the goal is to adopt a strategy that maximizes its expected value. The word optimal in this article is to be understood in the sense of utility maximization.

# V. Liquidity

There is a school of academic work on incomplete markets, which I call "completing the market" approach. The idea is to eliminate some risk factors through dynamic hedging by assuming that certain derivatives are continuously tradable. If all risk factors are eliminated, then the market becomes complete. The wonderful machinery of complete markets can then be applied to price other derivatives (Section VII of [6] offers a good example). There is nothing wrong with the mathematical logic of this approach, the issue here is about modeling.

When can a security be regarded as continuously tradable? The answer is that if its bid-ask spread is almost always very narrow. It is not unusual for bid-ask spreads on the most liquid derivatives (*e.g.*, options on IWM and QQQQ) to be on the order of five to ten percent, which is far from being small. Thus the continuously tradable assumption for derivatives is rarely, if ever, justified in reality. Just because it is mathematically convenient does not mean it is the right modeling choice. Continuous tradability is an aggressive assumption on liquidity; ignoring illiquidity risks underestimates the true risk associated with a derivative position.

I will take the opposite, more conservative approach by assuming derivatives are completely illiquid, which means that you are prepared for the worst-case scenario of holding your position to maturity. The key insight is that you expect your derivative position to remain constant after the initial trade. So illiquidity implies that you only consider the constant position strategy when trading derivatives. This strategy will be used later in portfolio optimization.

# VI. Local Equilibrium Principle

Recognizing the problem is one thing, fixing it is another. The challenge is to find a theoretically consistent framework that includes the position information. The foundation of the new theory is the local equilibrium principle, which I now describe.

Suppose you use a model to price a derivative, let us call the model's output price *your* fair value. The local equilibrium principle says that you act as follows: if the market price, or trading price, of the derivative is lower than your fair value, then you are a buyer, conversely you are a seller if it is higher. Naturally you stay put when the market price and your fair value agree. The name local equilibrium comes from the fact that you are in equilibrium with the market when your fair value agrees with the market price. But this equilibrium is temporary (local); as soon as the two prices diverge, you need to adjust your position.

The local equilibrium principle is so intuitive such that it is hard to imagine using a model in any other way. How can this common sense approach solve the problem? The next observation is crucial: When the market price and your fair value agree, you *choose* to stay put. This can mean only one thing—you think your current position is optimal. Aha! *There is a link between derivative pricing and portfolio optimization.*

After realizing that derivative pricing should be embedded in a portfolio optimization problem, the rest is straightforward. Optimality means certain conditions must be satisfied. For example, in calculus a maximum point of a smooth function requires its first order derivative be zero. Here the necessary condition of the current option portfolio being optimal (a constant position strategy) will give rise to the option pricing equations from which fair values can be computed. The new set of option pricing equations is the royal robe for the naked emperor. The position dependency is built into the equations, as they are derived under the assumption of the current position being optimal. Details of derivations for specific models can be found in [4, 5, 6].

It is widely regarded that the most important concept in option theory is the no-arbitrage principle. Of course, the no-arbitrage principle needs to be satisfied everywhere. Nevertheless, the local equilibrium principle is even more useful for pricing derivatives, both in complete and incomplete markets. I would not make such an extraordinary statement without having evidence to back it up.

Let us examine the complete market case first, no arbitrage versus local equilibrium. It is in all textbooks that applying the delta hedging argument and the no-arbitrage principle leads to the

Black-Scholes equation. As it turns out, applying the local equilibrium principle can also lead to the Black-Scholes equation (see Section 3.2 of [4]). Note the position information drops out in complete markets. Hence by inspecting position dependency of the resulting option pricing equation, the new derivation automatically diagnoses whether a model belongs to complete or incomplete markets. Furthermore, instead of having the delta hedging argument as an input, the new approach has it as an output, which implies that the local equilibrium principle is more productive.

In incomplete markets, there is no contest. The no-arbitrage principle cannot pin down an option's price, whereas the local equilibrium principle produces a set of equations that uniquely price an option. The application of the new set of option pricing equations and their consequences will be examined in the next few sections.

# VII. Automatic Trading

Real-life traders use models to help them make trading decisions. Can derivatives trading be automated? To be specific, is it possible to find a system that can be proven to make optimal trading decisions in a model world? In this man versus machine contest, I come down on the machine side, as I now explain.

It is clear that complete market models cannot be used to make rational automatic trading decisions. Because when the market and model prices of a derivative disagree, the complete market based system thinks it is an arbitrage opportunity, hence to trade infinite size, which is obviously nonsensical. The point is that without the position information in the theoretical framework, there will be no systematic way of determining a finite trading size that can be proven to be optimal. Intuition says that no risk averse trader will ever put on an infinite size position when risks are present. Thus the combination of market incompleteness and risk aversion makes the optimal trading size problem well defined. The tricky part is to obtain a quantitative answer.

The new position dependent option theory provides a natural solution. Since your fair value is position dependent, it is possible for you to adjust your position to make your fair value of a derivative equal to its market price, *i.e.*, the optimal trading size can be found by solving the local equilibrium equation *fair value = market price*. Moreover, the new theory guarantees that the local equilibrium equation has a unique solution.

Inventory control, which is used by all practitioners, means that you lower your price after buying and raise it after selling. The position dependent fair value has this feature automatically built in. In addition, the theory predicts that you demand a better unit price when trading bigger sizes, which again fits intuition. In summary, qualitative behaviors of real-life traders can be reproduced by a logically consistent quantitative trading system.

# VIII. Mutually Beneficial

With derivatives trading so wide spread nowadays, it may seem silly to ask the question of why people trade derivatives. The asymmetric information explanation of one party taking advantage of another just does not ring true. It is more likely that options are used in the context of portfolio optimization, and that derivative transactions are mutually beneficial. As I said at the beginning of this article that to constitute an understanding of a phenomenon, it must be quantitatively explained in a model world.

First, complete market models can never offer a rational explanation. Derivatives are redundant in complete markets, so there is no economical reason for their existence. Thus any possible explanation must come from incomplete market models. It is widely believed that derivatives' function in the economy is to transfer risk, which is a viewpoint I share. To demonstrate endogenously the benefits of risk transfer, the position information must be included in the theoretical framework, as position and risk are intimately related.

Within the new framework of the position dependent option theory, it is now possible to make quantitative computations. For each derivative transaction, the new theory can compute the expected utility change for each party involved. In general people have different fair values for a derivative instrument due to various differences: models, parameters, positions, and constraints, among others. The new theory says that whenever two people disagree on a derivative's fair value, they can always eliminate the valuation disparities, *i.e.*, reach a local equilibrium, by trading with each other (see Section 5.7 of [4]). Since the equilibrium state is optimal, both parties increase expected utilities through trading, thus trading is mutually beneficial, *i.e.*, it is a win-win game utility wise (see [4, 5] for specific examples). Finally we have a theory that offers quantitative explanations on why people engage in derivative transactions in the real world.

# IX. Revisiting Market Calibration

Again, I use stochastic volatility models in equity markets as an example in this section, but general conclusions are valid for other incomplete markets. Recall that the old option pricing equation derived from the no-arbitrage principle contains an unspecified market-price-of-risk term, but the new option pricing equation derived from the local equilibrium principle has this risk related term identified (see Section 7.4 of [4]), which, unsurprisingly, is position dependent. Since you know your position exactly, there is no longer any missing information in the new option pricing equation. Thus, just as the Black-Scholes equation in complete markets, the new equation for incomplete markets can price vanilla options as well as exotics ones.

Market calibration means pricing exotic options when vanilla options' prices are given. This is how things work in the new paradigm. First you estimate your model parameters based on historical stock data. When your fair values of vanilla options disagree with market prices, you put on a position by trading with the market, which implies that you are betting on history repeating itself (statistically speaking). The new theory guarantees that by adjusting your position, you can match your fair values with the corresponding market prices. After establishing a local equilibrium with the market on vanilla options, you then use the option pricing equation to compute your fair value of the exotic option. Note since the fair value is position dependent, you would have gotten a different valuation for the exotic option had you not established a local equilibrium with the market on vanilla options.

At this point, we have two linear option pricing equations (the old and the new) that differ only by one risk related term. Through fitting and trading, respectively, these two equations produce almost the same answers for a wide variety of payoff functions—vanilla options of all strikes and maturities. It is not difficult to imagine that these two linear equations will give similar results for an exotic payoff function as well. So the new systematic theory confirms the usefulness of the logically inconsistent fitting framework, which is adopted by practitioners as an interpolation device [2]. There is a saying that the market will find its way. Indeed it did find a good approximation in the case of pricing exotic options in incomplete markets.

The new theory offers much more than pricing: (i) It tells what size to trade given the trading price of the exotic option. (ii) It provides a consistent way to do static hedging, which works as follows: After trading the exotic option, your fair values for vanilla options will change, as your position has changed. So you need to trade vanilla options again with the market to re-establish a local equilibrium. The position difference of vanilla options before and after trading the exotic option is precisely what is needed to statically hedge the exotic option. If subsequently the implied volatility surface of the vanilla options changes, you follow the same procedure to do static hedging, *i.e.*, adjust your position until all your fair values of vanilla options are within their market bid-ask spreads. (iii) It gives the optimal dynamic hedging strategy if the underlying is continuously tradable. The optimal dynamic hedging takes into account the portfolio effect, *i.e.*, it is not equivalent to a sum of hedging each instrument separately. Thus it differs from the usual delta hedging scheme.

The final P&L of following the new paradigm consists of three parts, one from the exotic option, one from trading vanilla options, and the last from dynamic hedging with the underlying. As long as the underlying is modeled correctly, the new theory guarantees that the final P&L is positive *on average*, no matter how the implied volatility surface of the vanilla options fluctuates.

# X. Conclusion

The existing mainstream option theories for incomplete markets do not contain derivative position information. Without the source of risks being taken into consideration, risks are not properly accounted for. Such theories often exhibit logical inconsistency. The new personal approach based on the local equilibrium principle results in a position dependent option theory, which produces quantitative results that make common sense.

# Acknowledgment

The author is grateful for the helpful comments from Peter Leopold, Larry Lu, Marek Musiela, Jeremy Staum, and Qiang Zhang on earlier drafts of this article.

# References

[1] Andrew J. G. Cairns. *Interest Rate Models: An Introduction*. Princeton University Press, New Jersey, 2004.

[2] Emanual Derman and Nassim Taleb. The illusions of dynamic replication. *Quantitative Finance*, 4:323–326, 2005.

[3] Paul Wilmott. *Quantitative Finance*. John Wiley & Sons, New York, 2nd edition, 2006.

[4] Dennis Yang. *Quantitative Strategies for Derivatives Trading*, ATMIF, New Jersey, 2006. Excerpt available at http://www.atmif.com/qsdt

[5] Dennis Yang. Derivatives Pricing and Trading in Incomplete Markets: A Tutorial on Concepts. Available at www.atmif.com/papers

[6] Dennis Yang. A Simple Jump to Default Model. Available at http://www.atmif.com/papers