

# An introduction to financial econometrics

Jianqing Fan

Department of Operation Research and Financial Engineering  
Princeton University  
Princeton, NJ 08544

November 14, 2004

## **What is the financial econometrics?**

This simple question does not have a simple answer. The boundary of such an interdisciplinary area is always moot and any attempt to give a formal definition is unlikely to be successful. Broadly speaking, financial econometrics is to study quantitative problems arising from finance. It uses statistical techniques and economic theory to address a variety of problems from finance. These include building financial models, estimation and inferences of financial models, volatility estimation, risk management, testing financial economics theory, capital asset pricing, derivative pricing, portfolio allocation, risk-adjusted returns, simulating financial systems, hedging strategies, among others.

Technological invention and trade globalization have brought us into a new era of financial markets. Over the last three decades, enormous number of new financial products have been created to meet customers' demands. For example, to reduce the impact of the fluctuations of currency exchange rates on a firm's finance, which makes its profit more predictable and competitive, a multinational corporation may decide to buy the options on the future of foreign exchanges; to reduce the risk of price fluctuations of a commodity (e.g. lumbers, corns, soybeans), a farmer may enter into the future contracts of the commodity; to reduce the risk of weather exposures, amuse parks (too hot or too cold reduces the number of visitors) and energy companies may decide to purchase the financial derivatives based on the temperature. An important milestone is that in the year 1973, the world's first options exchange opened in Chicago. At the very same year, Black and Scholes (1973) published their famous paper on option pricing and Merton (1973a) launched general equilibrium model for security pricing, two landmarks for modern asset pricing. Since then, the derivative markets have experienced extraordinary growth. Professionals in finance now routinely use sophisticated statistical techniques and modern computation power in portfolio management, securities regulation, proprietary trading, financial consulting and risk management.

Financial econometrics is an active field of integration of finance, economics, probability, statistics, and applied mathematics. Financial activities generate many new problems, economics provides useful theoretical foundation and guidance, and quantitative methods such as statistics, prob-

ability and applied mathematics are essential tools to solve quantitative problems in finance. To name a few, complex financial products pose new challenges on their valuations and risk management. Sophisticated stochastic models have been introduced to capture the salient features of underlying economic variables and used for security pricing. Statistical tools are employed to identify parameters of stochastic models, to simulate complex financial systems and to test economic theories via empirical financial data.

There are several books on financial econometrics and related areas. Campbell *et al.*(1997) is an excellent book on a comprehensive overview of financial econometrics. A distinguished feature of the book is that it includes many empirical studies. Gouriéroux and Jasiak (2001) give a concise account on financial econometrics, but some prerequisites are needed. Tsay (2002) is an excellent book on the analysis of time series. It emphasizes on the methodological power of time series techniques on the analysis of financial data. A very nice introductory book on finance econometrics is Ruppert (2004), which aims at undergraduate or master level.

I taught a financial econometrics class for the master students in finance at Princeton University. Last semester, I used the aforementioned books and the book by Fan and Yao (2003) as reference books in an attempt to gather the strengths from these books and to give students a comprehensive overview of the field. While the curriculum is expected to be revised from time to time, I listed the topics covered in my class to give readers an overview.

1. Overview of statistical methods
2. Predictability of asset returns
3. Discrete time volatility models
4. Efficient portfolio and CAPM
5. Multifactor pricing models
6. Intertemporal equilibrium and stochastic discount models
7. Expectation and present value relation
8. Simulation methods for financial derivatives
9. Econometrics of financial derivatives.
10. Forecast and management of market risks
11. Multivariate time series in finance\*
12. Nonparametric methods in financial econometrics\*

The field of financial econometrics is much wider than what I presented here. However, the above topics give students a sample of taste on the subject.

I am fully aware the challenge to give an overview of a fast growing field in such a limited space. Readers can easily imagine the undue task if you were asked to write an overview of statistics. Instead of covering every aspect of financial econometrics, I just briefly touch a few important topics where statistics plays an important role.

### **Econometrics of financial derivatives**

Financial derivatives are mainly introduced to reduce various exposures of market risks. They are also used in speculative trading to increase leverages. They include options on stocks, stock indices, bonds and bond indices, futures on commodities and currencies, and options on the futures and so on. An option gives the right to its holders to buy or sell certain asset at certain price, called strike price, at or before the expiration. An European option is the one that can only be exercised at the expiration date, while an American option can be exercised any time before its expiration. Options on the stock indices such as the standard and poor 500 index are usually European, while options on individual stocks are usually American. See Duffie (2001) and Hull (2003) for a comprehensive account on this subject.

Options are usually used to hedge against certain financial risks, weather conditions, and natural disasters such as catastrophe events. Consider, for example, a company which expects to receive 250 million euros in 3 months from its sales in Europe. The financial officers of the company do not wish to have unpredictable earnings due to the fluctuations of the exchange rates between the US dollars and euros, making the quarterly earning less predicable. One way to reduce the exchange rate risk is to buy the options, or more frequently the futures, on the currency exchange matured in 3 months with different strike prices. Indeed, the company does not have to buy the actual futures on the US dollars, but to buy the options on the futures of the US dollars. The idea is that if the value of the euro drops, the options on the future will rise to offset the lost in the currency exchanges. As one can see, combinations of different financial instruments can yield various payoffs. The aim is to reduce the market risks to a manageable level. In this sense, financial derivatives serve a similar purpose to that in insurance industry. One advantage of the financial derivatives is that it is very liquid, trading at a large volume every trading day.

Since the options give right for holders to buy or sell financial assets, they are valuable. The question is how to value them. The celebrated formula of the Black and Scholes (1973) was derived based on what so called “relative pricing”. The basic idea is to produce a trading strategy, which dynamically balances the holdings of the underlying asset and riskless bond, such that the portfolio is riskless and has the same payoff as the option. Through tracking the prices of such a dynamic portfolio, we obtain the value of its corresponding options. Under the Black-Scholes model, the price for option can be explicitly obtained. It depends on the stock price, the risk-free rate of interest, the strike price, the time to maturity, and the volatility of the asset.

The option writers incur substantial risk in selling underlying options without taking any risk hedging measures. Suppose that an option writer sells 100,000 shares of a call option on a stock

with strike price \$50 matured in 20 weeks. Suppose that the current stock price is \$49 and stock volatility is 20% per annum and the riskless interest rate is 5%. According to the Black-Scholes formula, the value of option is \$240,000. If the writer sells it for \$300,000, it is expected to make \$60,000 or 25% of the option's value in 20 weeks. If the option writer does nothing to manage his risk, i.e., takes a naked position, and if stock price goes up to \$70, he will lose \$2,000,000 in option, much more than the premium received. If he takes a fully covered position, namely, buys 100,000 shares at \$49 right after issuing the option, and if stock goes down to \$39, he will lose \$1,000,000 in stock. Clearly, some hedging strategy is needed.

Let me briefly introduce a hedging strategy which helps us understand the derivations of the Black-Scholes formula. The delta-hedging is to hold the shares of the underlying asset according to the sensitivity of the option value to the change of underlying stock price, namely the partial derivative, denoted by  $\Delta$ , of the option price formula with respect to the stock price. For the aforementioned example, at the 0<sup>th</sup> week (immediately after issuing the option), according to the Black-Scholes formula, it can be computed that  $\Delta = 0.522$  and the writer buys 52,200 shares or 52.2% of 100,000 shares of the option issued. Suppose that after the first week of issuance, the price goes down to \$48.1, it can similarly be computed that  $\Delta = .458$ . The writer needs to hold only 45,800 shares, and hence needs to sell  $52,200 - 45,800 = 6,400$  shares. This is a very intuitive decision, as the risk of upward price movement relaxes and the option writer needs to reduce the holding of shares to mitigate the downward risks. Rebalance is needed periodically to keep the risk in check. The cost for producing such a kind of riskless strategy can be computed (interest paid and loss or gain in rebalancing stock) and is approximately the same as the value of the option. The length of rebalance in this example is one week, but this can be chosen to optimize the hedging performance. As the interval of rebalance tends to zero, the cost of producing such a riskless strategy is the value of the option. See Hull (2003).

The art of the implementations of the Black-Scholes formula involves a lot of econometrics and statistics. First of all, one needs to estimate the volatility of an asset from a stock. The available data can be huge. In addition to the historical prices of an underlying stock, related stocks also provide the volatility information of the underlying stock. This results in a multivariate time series and picking relevant ones needs both statistics and fundamental and technical analysts. Further to the complication of the problems is that the stock prices are usually not stationary. Thus, one can not use the historical price volatility based on a large time window. Adaptive choice of time horizons requires statistical studies and empirical research. The precision of volatility estimation has direct implications on the value of the option and confidence intervals are needed for the valuation of options. Further, traders need to estimate the sensitivities of the option prices with respect to the changes of stock prices, interest rates, volatilities and the time to maturity (referred respectively to as delta, rho, vega and theta) and so on. They introduce additional statistical problems for investigation.

Options are actively traded every day. Another possibility is to infer the volatility from the

traded option price, which is the inverse of the Black-Scholes formula, called implied volatility. The implied volatility is generally less volatile than the historical one based on asset prices. It should be constant, if the Black-Scholes model holds. However, the implied volatility depends on the characteristics of options such as the strike price or ratio of stock price to the strike price, called moneyness and the time to maturity. Statistical techniques have been introduced to estimate the implied volatility function and the option value surface (as a function of moneyness and the time to maturity) and to understand the observed stylized features such as volatility smiles.

The Black-Scholes formula is derived under the assumption that the price dynamic follows the geometric Brownian motion. Various efforts have been made in extending its applicability to, for example, the yields of bond processes and catastrophe linked securities. The geometric Brownian model has also been extended to various other stochastic models, including stochastic volatility, nonparametric models, and jump diffusion models. These give rises to new financial econometrics problems. We will discuss some of these issues in the stochastic modeling section.

Large volume of option trading also enables one to infer the pricing formula directly from the traded options. This is done through the estimation of the state price density, which is the probability density of the value of an asset under the risk-neutral world or equivalent martingale measure. It does not depend on the payoff function. As such, it can be directly used to price financial derivatives with different payoff functions, which are simply the expected payoffs under the state price density (the integration with respect to the state price density). A nice feature of the state price density is that it can be used to evaluate the price of illiquid derivatives, once it is estimated from more liquid ones. Mathematically, the state price density is the second derivative of the price of the call options with respect to the strike price. Hence, the statistical problem becomes to build a regression model for the prices of options as a function of their option characteristics such as stock spot price, strike price, time to maturity, risk free interest rate and dividend yields. See Aït-Sahalia and Lo (1998) for further details on semiparametric and nonparametric approaches to this problem.

### **Asset pricing and CAPM**

Asset pricing tries to understand the prices of claims with uncertain payments. Stock holders, for example, are entitled to the dividend payments over the life time of the stock, which are uncertain. The derivative pricing in the last section is based on relative pricing, inferring the derivative value given the prices of some other assets such as the stocks and the riskfree bond. We do not ask where the price of other assets came. There is also a huge literature on “absolute pricing”, valuating each asset by reference to its exposure to fundamental sources of macroeconomic risk. An excellent book on this topic is Cochrane (2001).

The quantification of the tradeoff between risk and expected return is one of fundamental problems in financial econometrics. While common sense suggests that riskier investments will generally yield higher returns, it is not until the birth of the Capital Asset Pricing Model (CAPM)

that economists were able to quantify risk and the reward for bearing it. Markowitz (1959) laid the groundwork for the CAPM and postulated the investor's portfolio selection problem in terms of the expected return and variance of the return. He argued that investors would optimally hold a mean-variance efficient portfolio. Sharpe (1964) and Lintner (1965) developed further the Markowitz's work and showed that market portfolio (e.g. S&P 500 index, as a proxy) is a mean-variance efficient portfolio. As a consequence of this, they showed the following CAPM: The expected excessive return of any asset over a risk free bond is a multiple, called market  $\beta$ , of the excessive return of the market portfolio. The market  $\beta$  measures the risk of the asset relative to the market portfolio. The CAPM quantifies exactly how the expected return depends on the risk of the asset, measured by the market  $\beta$ .

Since its publication, various statistical techniques have been developed to verify the validity of the CAPM in empirical finance. The early evidence was largely positive. Yet, in the late 1970's, some evidence against the CAPM began to appear in the so-called anomalies literature in which firms can be grouped according to their characteristics to form a portfolio that can be more efficient than the market portfolio. While the evidences against the CAPM are still controversial, various extensions of the CAPM have been proposed to better capture the market risks. These include intertemporal CAPM (Merton, 1973b), multifactor pricing model such as the Arbitrage Pricing Theory (Ross, 1976), and consumptions based CAPM, among others. These models can be more generally represented by the stochastic discount factor model. A different CAPM amounts to choose a different stochastic discount factor (see, e.g., Cochrane, 2001). They form spectacularly beautiful theory on asset pricing.

Testing the validity of various versions of CAPMs attracts a lot attention in empirical finance. Various testing procedures and statistical methods have been proposed and studied. Statistical techniques have also been used to select risk factors that explain the expected returns of assets over a period of time. For example, Fama and French (1993) build a famous three-factor CAPM to explain the expected excessive returns of assets. They test CAPM using the following three factors: the CRSP value-weighted stock index (a proxy of the market portfolio), the difference of returns on a portfolio of low and high market value of equity firms, the difference of returns on a portfolio of high and low book-to-market value firms. Sophisticated statistical models have been introduced to model the behaviors of consumptions and habits and advanced statistical methods have been applied to test the consistency of these models with empirical financial data.

### **Stochastic modeling and statistical inferences**

The valuation of financial derivatives depends largely on the stochastic model assumptions on the price dynamics of underlying assets. A different asset class requires a different class of stochastic models. For example, the geometric Brownian motion, which has a constant rate of expected return and volatility, can not be used to model bond yields, which possess the heteroscedasticity and mean-reversion feature. As they rise, the interest rates tend to be more volatile and there is a positive

force pulling the rates down when they exceed a mean level, while as the interest rates go down, the volatility tends to be smaller and there is a positive force driving the rates up when they are below the mean reversion level. Managing and modeling the risks such as natural disaster and weather require a very different class of stochastic models. A stochastic model can only capture certain aspect of underlying stochastic dynamics. This is why there are many models being introduced for the price dynamics of various asset classes.

Options prices depend on the underlying parameters of stochastic processes. The question then arises how to efficiently estimate the parameters from a discretely observed stochastic diffusion process. For an overview, see Fan (2003). If the model has been parameterized, then the maximum likelihood method is a natural candidate. However, except for a few specific models, the likelihood function is difficult to derive analytically and hence hard to implement. One possible remedy is to use the generalized method of moments (Hansen, 1982) to derive some estimation equations and some other features such as local time of stochastic processes to derive a different set of equations. Other methods involve using approximate likelihood, resulting from the Euler approximation or higher-order approximations of stochastic processes when the sampling interval is small. This is more feasible nowadays, thanks to the availability of high frequency data. The biases of an approximated likelihood method can be reduced by using some calibration methods such as the indirect inference by Gouriéroux et al. (1993).

Nonparametric models arise naturally in financial modeling. They aim at reducing modeling biases of parametric models and validating their goodness of fit to financial data. Many of such parametric models are simple and convenient ones to facilitate mathematical derivations and statistical inferences. They are not derived from economics theory and can not be expected to fit all financial data. While the asset pricing theory gives nice pricing formulas when the underlying price dynamic is correctly specified, it offers little guidance in choosing or validating a model. Hence, there are genuine needs for flexible stochastic modeling, and nonparametric methods offer a unified and elegant treatment.

Nonparametric approaches have recently been introduced to estimate return, volatility, transition densities and state price densities of stock prices and bond yields. See Fan (2003). They are also useful for examining the extent to which the dynamics of stock prices and bond yields vary over time. They have immediate applications to the valuation of bond price and stock options and management of market risks. They can also be employed to test economic theory such as the CAPM and other stochastic discount models, and answer the questions such as if the geometric Brownian motion fits certain stock indices, whether the Cox-Ingersoll-Ross model fits yields of bonds, and if interest rates dynamics evolve with time. In these testing problems, nonparametric models serve as natural alternative models to the null hypotheses. Furthermore, based on empirical data, one can also fit directly the observed option prices with their associated characteristics and checks if the option prices are consistent with the theoretical pricing formula. They can also be used to testing whether an underlying asset follows a time-homogeneous Markovian process and can even

be used as estimation tools for parametric models.

### **Volatility, portfolio optimization and risk management**

Volatility pervades almost every facet of financial econometrics. It is used in pricing financial derivatives, portfolio allocation to control and manage risks, and computation of risk-adjusted returns for comparisons of relative performance of various financial investments (e.g. mutual funds). It measures the risk of a portfolio and is associated with capital requirement in banking regulations. The topic is prominently featured at the heart of the financial econometrics.

There are two popular classes of widely used models for volatility of discretely observed time series. One is the ARCH and GARCH models and their various extensions. For an overview of the subject, see Engle (1995). These models attempt to capture several important stylized features in financial markets. These include volatilities clustering, heavy tail and asymmetric distributions of returns, persistence of autocorrelation in absolute or squared returns, and leverage effect (stock price movement is negatively correlated with the change of volatility). Various statistical procedures, including quasi-maximum likelihood estimator and robust methods, have been introduced to fit the ARCH-GARCH models. Various efforts have been devoted to investigate the property of this class of models and the performance of statistical procedures. Modeling volatility matrices of multivariate time series poses many new statistical problems with new challenges, as the number of parameters grows quickly with the number of the assets. Recent arrival of high frequency data give rises to new interesting statistical problems.

Stochastic volatility models are another important class for capturing the stylized features of volatilities in financial data. The challenge is that the volatility is not directly observable. Instead, it is driven by a different unobservable random process. The problem here is similar to the error-in-variable regression in statistics (see, e.g., Carroll, Ruppert and Stefanski, 1995). There is a large literature on studying the statistical issues and probabilistic aspects of the models. See Shephard (2004) for an overview.

The portfolio allocation can be formed based on the mean and variance consideration in a similar way to the fundamental work of Markowitz (1959), Sharpe (1964) and Lintner (1965). Its basic idea is to maximize the expected returns while controlling the risks. It can be formulated as a constrained optimization problem. The expected return and volatility play a prominent role in the portfolio allocation, and statistical techniques are widely used for modeling the returns, volatilities as well as the risk management.

Risk management is to identify risk sources, to measure risks and to control and manage risks. There are large efforts in statistical community on defining and forecasting risk measures. They are directly related to the volatilities of asset returns. and have been widely used in security and bank regulations, proprietary trading, and risk managements. Statistical methods have been widely used in such an endeavor. For an overview, see Jorion (2000).

## References

- [1] Aït-Sahalia, Y. and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, **53**, 499-547
- [2] Black, F. and Scholes, M. (1973). The pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**, 637-659.
- [3] Campbell, J.Y., Lo, A. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press, Princeton, N.J.
- [4] Carroll, R.J., Ruppert, D. and Stefanski, L.A. (1995). *Measurement Error in Nonlinear Models*, Chapman and Hall, London.
- [5] Cochrane, J.H. (2001). *Asset Pricing*, Princeton University Press, Princeton.
- [6] Duffie, D. (2001). *Dynamic Asset Pricing Theory* (Third Edition). Princeton University Press, Princeton, New Jersey.
- [7] Engle, R. F. (1995). *ARCH, selected readings*, Oxford University Press, Oxford.
- [8] Fan, J. (2003). A selective review of nonparametric methods in financial econometrics. [www.princeton.edu/~jqfan](http://www.princeton.edu/~jqfan).
- [9] Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer-verlag, New York.
- [10] Gouriéroux, C. and Jasiak, J. (2001). *Financial Econometrics: Problems, Models, and Methods*. Princeton University Press, Princeton.
- [11] Gouriéroux, C., Monfort, A. and Renault, E. (1993). Indirect Inference . *Jour. Appl. Econometrics*, **8**, supplement, S85-S118.
- [12] Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica*, **50**, 1029-1054.
- [13] Hull, J. (2003). *Options, Futures, and Other Derivatives* (Fifth edition). Prentice Hall, Upper Saddle River, New Jersey.
- [14] Jorion, P. (2000). *Value at Risk: The new benchmark for managing financial risk* (2nd ed.). McGraw-Hill, New York.
- [15] Lintner, J. (1965). The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, **47**, 13-37.
- [16] Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, New York.
- [17] Merton, R. (1973a). Rational theory of option pricing. *Bell Journal of Economics and Management Science*, **4**, 141-183.
- [18] Merton, R.C. (1973b). An intertemporal Capital Asset Pricing Model. *Econometrica*, **41**, 867-887.
- [19] Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, **13**, 341-360.

- [20] Ruppert, D. (2004). *Statistics and Finance: An Introduction*. Springer-verlag, New York.
- [21] Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditons of risks. *Journal of Finance*, **19**, 425-442.
- [22] Shephard, N. (2004). *Stochastic volatility: selected readings*. Oxford University Press.
- [23] Tsay, R.S. (2002). *Analysis of Financial Time Series*. John-Wiley & Sons, Inc., New York.