The CPAM and the High Frequency Trading;

Will the CAPM hold good under the impact of high-frequency trading?

By

YoungHa Ki

Submitted to the graduate degree program in Economics
and the Graduate Faculty of the University of Kansas in partial fulfillment of the
requirements for the degree of Master of Arts.

Chairperson Prof. Dr. Shu Wu

Prof. Dr. Mohamed El-Hodiri

Prof. Dr. Ted Juhl

Date Defended: May 20, 2011

The Thesis Committee for YoungHa Ki
certifies that this is the approved version of the following thesis:

The CPAM and the High Frequency Trading;

Will the CAPM hold good under the impact of high-frequency trading?

Chairperson Prof. Dr. Shu Wu

Date approved: May 20, 2011

Abstract

The main purpose of this paper is to investigate the possible relationship between the Capital Asset Pricing Model – CAPM and the prevailing High Frequency Trading (HFT) method of stocks trading and to explain the relationship between them, if exist, with the references from research papers and advanced statistical method. This paper mainly follows Jagannathan and Wang's paper (The Conditional CAPM and the cross-section of expected return, 1996) to explain the capability of CAPM, especially with financial turmoil. However, instead of using the cross-sectional statistical method by following Jagannathan and Wang, the mixed model will be implemented. This paper draws the intermediate conclusion regarding the relationship and shows the existence of relationship, if exist, rather than introducing a new model.

Acknowledgement

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. I want to thank the Department of Economics for giving me permission to work on this thesis in the first instance.

I am deeply indebted to my supervisor Prof. Dr. Shu Wu from the University of Kansas whose generosity and time allowance helped me in all the time of research for and writing of this thesis. I furthermore want to give my greatest appreciation to my mentor Prof. Dr. El-Hodiri to encourage me all the time and to give me all the advice.

My deep appreciation goes to one of my best friends in colleagues, Dr. JaeHoon Lee, a research associate at the Bureau of Child Research in the University of Kansas for supporting me in the statistical method and all the related precious advice for my research and this thesis.

Especially, I would like to give my special thanks to my wife Jena whose patient love enabled me to complete this work.

Table of Contents

1. Introduction

1) Capital Asset Pricing Model

To find out the accurate predictability on the expected return has been a major issue in the modern financial field. Using a linear model like CAPM (Sharpe, 1964 and Litner, 1965) was a landmark and has been a proper way to predict the expected return on investment. CAPM sets up the relationship between the expected return on an asset in the financial market and the risk. And the risk is shown by a measurement called beta which is calculated with the variance of market and the covariance between the market and an asset. Theoretically speaking, with assumptions CAPM is designed to predict the expected return on an asset toward a considerably accurate projection. Based on the empirical research papers and statistical works, CAPM has been statistically significant and has shown an acceptable prediction on the return.

To have a better predictability and to take the unique characteristics of market and assets into the consideration, arbitrage pricing theory – APT and conditional CAPM models, derived from the unconditional model, with the consideration of the market factors have been developed and introduced. These models are based on the linear relationship between the market and an asset, like the unconditional CAPM, and added unique market factors as variables such as size of the company and dividends. Empirical tests and researches have shown that the conditional CAPM and the APT model seem to fit better than the unconditional CAPM. And especially when there are any unscheduled or sudden changes in the market, the projection of the expected future returns, with sudden changes in market, seem to depend less on the betas, the variance-covariance between the market and an asset, and more on the unsystematic risk which cannot be explained by the CAPM. So, the CAPM may not appear to work properly. [1]

Under the consideration of inability of the CAPM toward the accurate predictability, the High frequency trading (HFT), the stock trading method using internet such as day trading and intra-day trading and using algorithm software programs, which has become popular since the year of 2000 and the deregulation in 2005 could be one of market factors

[1] Febian and Herwany (2009). “The performance of asset pricing models before, during, and after financial crisis in emerging market: Evidence from Indonesia”

that the CAPM cannot explain. Currently, the daily trading volume of HFT consists of more than 50% of the total volume.[2]

2) High Frequency Trading (HFT)

HFT is a computer-determined trading and the algorithm in HFT decides the execution time and price without the human interaction. The main purpose of the HFT is to minimize risks and to post small deal sizes that enable to move in and out of trades extremely quickly, arbitraging between spreads available on different exchanges and platforms, and even between the speed of trading available on them.[3]There are several sub-categories of HFT methods but they are all unified by the identical purpose: getting a profit faster and smarter than everyone else. During the last five years from 2005, HFT methods had been getting popular and had shown a big leap in the amount and volume of the trading. In 2005, there was less than one quarter of the trading volume in U.S. stock market dealt by high frequency traders but approximately 2% of the stock trading firms in U.S. stock market generate 73% of the total trading volume and is estimated to have $15 billion to $25 billion in revenue.[4] Table 1 shows the proportion of HFT trading out of total equity trading. In table 1, there was a tendency of increasing trading volume from 2005 and sudden surge around the year 2007. In addition, the trading volume and the amount by HFT, especially in NYSE, have been increasing, according to Table 2, since 2005 when there was the deregulation. However, during the period from 2005 to 2009 every trading has been downsized more than half. Hence, smaller trading order size enables to have transactions approximately 10 times faster in 2009 than in 2005.

After computerized order system was initiated in the mid-1970 to the New York Stock Exchange and after NASDAQ recognized as the first electronic stock market, stock trading with computerized software and internet became to be a major tool to deal with orders in the financial market. Outstandingly, the deregulation in 2005 accelerated the development of HFT. Securities and Exchange Commission (SEC) enacted Regulation National Market System (Reg. NMS) for the purpose to modernize and to strengthen the regulatory structure

[2] Supra note 4, at 3606 (noting that HFT “is a dominant component of the current market structure and is likely to affect nearly all aspects of its performance”)
[3] Michael J. McGowan, The rise of computerized HFT: Use and controversy. page 4(2010), Duke Law & Technology Review
[4] Tyler Durden. Goldman’s $4 billion High Frequency Trading Wildcard (Jul. 17, 2009)

of the U.S. Equity markets for the efficient, competitive, and fair markets and to protect investors. Reg. NMS induced the inclusive use of advanced computer technologies so that the speed and the capacity of the trading functions from the market participants were much improved. Under Reg. NMS, market orders can be posted electronically and can immediately be executed at the best price quote and at the fastest speed. [5]

HFT method drew attention when there was stock market turmoil in June 2010. Active transactions by the HFT method caused, at first, the stock market to fall down drastically and to turn around later. This event cannot be explained by the CAPM because it is treated as the unsystematic risk. Due to the technology advance in stock market trading and competitiveness by stock market investors to have economic profit, this type of financial turmoil would be able to happen again and at each time CAPM would not be able to work as a tool for the return predictability. Then, does the CAPM still hold good for the stock market?

2. Related Literature Research

After Litner proposed the Capital Asset Pricing Model in 1961 and later improved the CAPM theory in 1964, many related academic literatures and empirical studies have been introduced and developed. The fundamentals of the theory is related to the covariance of an asset to the market, defined as beta, and academics and practitioners have empirically tested if this beta can explain the expected return of an asset or a portfolio properly. Among tests and researches, Roll (1977, 1978) criticized the unconditional CAPM regarding its incapability to predict the expected return and Merton (1973) introduced a revised CAPM, the inter-temporal CAPM, in which he mentioned that the expected returns defined by the unconditional CAPM can be generated only under special additional assumptions. Also, Ross (1976) initiated another asset pricing theory, Arbitrage Pricing Theory (APT) and the APT theory offered that the expected return of an asset can be formulized by a linear relationship among macroeconomic factors in the financial market, which can be shown as factor-specific betas. The CAPM cannot explain the size, book-to-market, past return continuity, and positive relationship between the average returns and illiquidity effects. Consequently, the

---

[5] Tara Bhupathi (2010). Recent development; Technology's latest market manipulator? High frequency trading: The strategies, tools, risks, and responses. N.C. Journal of Law & Technology.

incapability of the CAPM to explain these irregularities has brought the introduction of conditional and dynamic versions of the CAPM. Fama and French (1993) argue that value firms are known to perform well than growth firms and Carhart (1997) claims factors from firms and unique stock characteristics with the CAPM can generate well-established performance than the CAPM alone can do so.

Despite a long list of empirical studies on the conditional CAPM and its validity, Lewellen and Nagel (2006) claim that the conditional CAPM is not different to the unconditional CAPM in explaining the conventional irregularities. Yalcin and Ersahin (2010) test the asset pricing performance of the conditional CAPM and argue the conditional CAPM performs just as much as the unconditional CAPM can do in pricing portfolios they consider.

However, Soydemir (2001) shows evidence that the conditional model can price the market risk when the static CAPM cannot do so. Akdeniz et al (2001) use a conditional CAPM model similar to the two-factor model of Fama and French (1992) and conclude, using the Istanbul stock market data, that the size and Book-to-Market contribute to the stock returns whereas the market beat cannot explain any portion of returns. In 1996, Jagannathan and Wang argue that the incapability of the static CAPM caused to fail to capture the evidence of the CAPM validity and introduced one version of the conditional CAPM, called Premium-Labor model (P-L model) by using the concept from Fama-MacBeth technique of sampling and testing of cross-sectional models. In addition, Buss, Schlag, and Vilkov (2009) test the validity of a time-varying conditional CAPM, under the reason that both asset betas and the market risk premium are not constant over time but time-varying. By introducing the forward-looking factor into the time-varying conditional model, it is concluded that the revised model can produce the expected return more accurately. In addition, they find out the conditional model is sufficient to explain the expected returns, especially when the economy is in the good shape. However, when there is a fluctuation in the financial market or when the whole economy is in a poor shape, the conditional model may need additional market factors to explain the difference between the expected returns and the actual returns[6]. As a conclusion, it is argued that the expected return analysis seems to be majorly related to the economic states. Their research result is in favor and is followed by Febian and Herwany (2009). In a research on the performance of asset pricing models in Indonesian financial market, Febian

[6] Buss, Schlag, and Vilkov (2009). CAPM with option-implied betas: Another rescue attempt.

and Herwany, through the empirical study, find that the cross sectional regression test on the CAPM generates unknown variables which cannot be explained by the CAPM and which are positively correlated to the expected returns. Also, they conclude that the accuracy of the asset price models depends on the macroeconomic characteristics and CAPM is not capable to explain the expected return during the economic crisis.

In addition to the CAPM, Kirilenko et al (2010) investigate how the HFT trading affects the financial market, especially after the Flash Crash on May 6, 2010. They suggest that the HFT trading, even though it was not the main cause of the Flash Crash, aggravated the volatility to the financial market. But the increasing volatility generated by the HFT seems to counteract the predictability of asset pricing models to the financial market. Chakravarty et al (2010) argue that the Flash Crash in 2010 resulted from the liquidity problem by the Intermarket Sweep Orders (ISO) which is typically used by institutional algorithmic investors known as HFT investors. Smith (2010) claims that the HFT extends its impact on the microstructures of the trading dynamics. The more HFT becomes self-similar, the more investors in the financial market generate more volatility and the less predictable the financial market would be.

3. Data and Methodology

The main purpose of this research is to find out if the CAPM can interpret the volatility in the financial market, especially after the surge of HFT. The chosen model will be tested to see how much the chosen model can fit with data to detect if any other variable(s) is (are) necessary to explain the expected return which cannot be explained by the model. Even though three considerably important macro-economic variables are introduced and betas for all three variables are applied to the three-beta model, it is going to be the beginning point of period to think of new macro-economic variable if this chosen model is not sufficient for the expected return predictability. To implement the test, both unconditional and conditional CAPM are tested for the statistical significance. Regarding the models and method, I decide to follow the method of Jagannathan and Wang (1996), which is the P-L model. Reasons behind this model is, first of all, since it is aimed to see if the CAPM shows a good fit with the HFT, it is necessary to find a method to be suitable for both unconditional and conditional

CAPM, not taking into consideration of any derived models such as APT. Not only the result of the unconditional CAPM but also a couple of different results of conditional CAPM be generated to compare the statistical significances among models. Secondly, this research paper shows how to set each versions of CAPM to explain the cross-section of stock returns.

1) Data

First of all, the stock price data includes monthly data of closing prices of 27 different stocks in NYSE or NASDAQ, including S&P 500 index. To have a possibly good statistical interpretation and to represent the stock market well as a proxy, 26 individual stocks are selected by volume in the recent 20 years. In detail, 26 most active stocks per number of shares that changed volume for S&P 500, Dow, and NASDAQ are chosen. These 26 stocks account for more than 25% of all domestic volume in U.S. financial market.[7] Stock returns are derived from stock price as a growth rate in percentage. Originally, Jagannathan and Wang used the CRSP value-weighted index as proxy for the market. But their justification seemed weak and that index is for stock returns. Therefore, instead of CRSP value-weighted index, S&P 500 will replace it.

Secondly, the spread between BAA rating and AAA rating bond yields is used as a proxy for the risk premium as Jagannathan and Wang introduced. Thirdly, since human capital comprise of a large portion of the total capital in the economy, the human capital should be one of factors to be considered in the model.[8] This data is provided by National Income and Product Account section in Bureau of Economic Analysis (BEA). To distinguish the income effect from the stock market effect, the dividend income will be excluded from the per capita income.

Monthly data spans from June 1994 to December 2010 and consists of 199 individual values for each category. For the test of the best fit among several unconditional and conditional models, the whole period data will be employed. However, regarding the cross-sectional test which will result in if the selected model can explain the volatility well in the financial market, the whole period data will be divided into three sub-periods; from June

[7] Tyler Durden's article on Jul. 2010 on http://www.zerohedge.com, which listed the top 20 stocks in volume and also information on rateviewer.com regarding top 20 active stocks by volume. 26 stocks are sorted out since many stocks are overlapped from two information sources.

[8] Eka Katamadze (2008). Overview of conditional CAPM and the cross-section of expected returns.

1994 to December 2000, from January 2001 to May 2005, and June 2005 to December 2010. The reason behind this separation of data is to see if the CAPM can explain the market well before and after the major surge from June 2005.

2) Model

As mentioned previously, I will follow the methodology of Jagannathan and Wang (1996) for unconditional and conditional CAPM. Since necessary proofs are completed and provided by their work, here I will elaborate the necessary equations and procedures.

Let R_i denote the return on an asset i and R_m the return on the market. The unconditional CAPM is as below.

$$\mathrm{E}(R_i) = r_0 + r_1\beta_i, \tag{1}$$

Where β_i is defined as

$$\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)} \tag{2}$$

And even CRSP value-weighted index was used as the proxy for the market index in the original paper, the reason behind is not strong and since the S&P 500 can work as the good proxy for the market index, S&P 500 will, in this case, work as the proxy for the market index. Let R_t^{vw} denotes the return on the value-weighted stock index.

$$R_t^{vw} = \frac{price\ of\ S\&P\ 500_t - price\ of\ S\&P\ 500_{t-1}}{price\ of\ S\&P\ 500_{t-1}} \tag{3}$$

Then, there are some constants $\emptyset_0$ and $\emptyset_{vw}$ s.t.

$$R_{mt} = \emptyset_0 + \emptyset_{vw}R_t^{vw} \tag{4}$$

, where R_{mt} is the unconditional market return at time t.

Define β_i^{vw} as

$$\beta_i^{vw} = \frac{Cov(R_{it}, R_t^{vw})}{Var(R_t^{vw})} \quad (5)$$

By substituting (4) into (2) and using (5), unconditional CAPM, equation (1), can be re-defined as below to show the linear relationship between the unconditional expected return and the β_i^{vw}. So, there are some constants of c_0 and c_{vw}, s.t.

$$\mathrm{E}(R_{it}) = c_0 + c_{vw}\beta_i^{vw} \quad (6)$$

And, also as mentioned above at the data, the spread between BAA rating and AAA rating corporate bond yield will be used as the proxy for the market risk premium. Interest rate variable is one of the most powerful variables to predict the future economic situations. There are several representations for the interest rate variable such as the 3-month Treasury bill. However, according to Bernanke (1990), the spread between the T-bill rate and the commercial paper rate showed the best performance.[9] Hence, the spread between two corporate bond yield will be introduced and is denoted by R_{t-1}^{prem}. Under the assumption that the market risk premium is a linear function of R_{t-1}^{prem}, conditional market risk premium, r_{it-1}, can be defined as

$$r_{it-1} = k_0 + k_1 R_{t-1}^{prem} \quad (7)$$

, where k_0 and k_1 are constants.

For an asset *i*,

$$\beta_i^{prem} = \frac{Cov(R_{it}, R_{t-1}^{prem})}{Var(R_{t-1}^{prem})} \quad (8)$$

With using the definition of the premium beta that shows the beta-instability risk;

$$\beta_i^{r} = \frac{Cov(R_{it}, r_{1t-1})}{Var(r_{1t-1})} \quad (9)$$

[9] Jagannathan and Wang, (1996), The conditional CAPM and the cross-section of expected return, Federal Reserve Bank of Minneapolis, page 10

Suppose that β_i^r is not a linear function of β_i. Then, the unconditional expected returns can be re-formed as a linear function of the unconditional market beta and the premium beta. So, there are some constants a_0, a_1, and a_2 s.t.

$$\mathrm{E}(R_{it}) = a_0 + a_1\beta_i + a_2\beta_i^r \text{ for every asset } i \tag{10}$$

By substituting (7) into (9) and employing (8) and (10), it is shown that the expected return is linear in the premium beta and the market beta. So, there are constants c_0, c_m, and c_{prem} s.t.

$$\mathrm{E}(R_{it}) = c_0 + c_m\beta_i + c_{prem}\beta_i^{prem} \text{ for every asset } i. \tag{11}$$

It is assumed that the return on human capital is a linear function of the growth rate in per capita labor income. The realized capital gain part of the rate of return on human capital will be the realized growth rate in per capita labor income. So, the growth rate of per capita labor income at t is defined as below.

$$R_t^{labor} = \frac{L_t - L_{t-1}}{L_{t-1}} \tag{12}$$

, where L_t is the per capita labor income at time t, which proxies for the return on human capital.[10]

Following Jagannathan and Wang's methodology, it is assumed that the market return is a linear function of R_t^{vw} and R_t^{labor}. So, there are some constants $\emptyset_0$, $\emptyset_{vw}$, and $\emptyset_{labor}$ s.t.

$$R_{mt} = \emptyset_0 + \emptyset_{vw}R_t^{vw} + \emptyset_{labor}R_t^{labor} \tag{13}$$

The labor beta is defined as

$$\beta_i^{labor} = \frac{Cov(R_{it}, R_t^{labor})}{Var(R_t^{labor})} \tag{14}$$

[10] In the "Overview of conditional CAPM and the cross-section of expected returns", Katamadze introduced a different way to calculate the rate of return on per capita income, $(L_{t-1}+L_{t-2})/(L_{t-2}+L_{t-3})$ to smooth measurement errors. But here I introduce a universal way to calculate the growth rate.

By plugging (13) to (2), then with (5) and (14), β_i will be re-defined as below.

$$\beta_i = b_{vw}\beta_i^{vw} + b_{labor}\beta_i^{labor} \tag{15}$$

According to the assumptions used for (7), (8), (13), and (14) equations, the unconditional expected return on an asset i is assumed to be a linear function of its value-weighted beta, premium beta, and labor income beta. And by placing (15) into equation (11), the Premium-Labor model (P-L model) will be defined as below.

There are some constants c_o, c_{vw}, c_{prem}, and c_{labor} s.t.

$$\mathrm{E}(R_{it}) = c_o + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor} \text{ for every asset } i. \tag{16}$$

Presented and derived above, the P-L model will be tested to see how much the P-L model can fit with data. The main purpose of the tests is to detect if any other variable(s) is (are) necessary to explain the expected return which cannot be explained by the P-L model. Even though three considerably important macro-economic variables are introduced and betas for all three variables are applied to the three-beta model, any new macro-economic variable would be able to be introduced if this P-L model is not sufficient for the expected return predictability. Hence, first of all, the whole period data from June 1994 to December 2010 will be used to run all four models; unconditional model, conditional with risk premium, unconditional with income, and three-beta model. Then, secondly, a model which shows the best fit among provided models will test if the CAPM model can explain the expected return projection well before and after the surge of HFT. In other words, can CAPM interpret the impact of the HFT toward the financial market? Hence, the whole period will be divided into three sub-periods; From June 1994 to December 2000, January 2001 to May 2005, and from June 2005 to December 2010. It is because since there was the deregulation on the HFT in June 2005 which made the HFT be popular and dominate the financial market as the main tool for stock trading.

3) Statistical Method

To a) empirically compare the relative performance of four different specifications of CAPMs using information criteria and b) to estimate the model coefficients and accountability of a chosen best fitting model separately for three distinct time periods, Jagannathan and Wang's method is introduced. In Jagannathan and Wang's paper, they used the cross-sectional method to find the best fit and to estimate the capability of CAPM. However, in this paper, the mixed model, instead, will be applied. Byoun (2008) argues that the result from the cross-sectional method will generate approximately the same result as that from the mixed method. Furthermore, he comments that the dependence among the repeated observations for the same firms in the panel can lead to the covariance/correlation structures of error. The CAPMs are often estimated with the cross-sectional regression modeling approach (Black, Jensen, & Scholes, 1972 and Fama & MacBeth, 1973), in which the coefficients are estimated by the averages from year-over-year regressions under the assumption of independent residuals across different years. However, it is essential to recognize the interdependency of observations when the same firms are repeatedly measured in multiple time points (i.e., panel data). When the panel data are analyzed without regard to dependency among clustered observations, Type I error is inflated leading to unwarranted rejection of the null hypothesis (Dorman, 2009 and Hedges, 2007). Thus, the current study applied the mixed modeling approach that supports various cross-sectional heteroskedastic or time-wise (e.g., autoregressive) covariance structures of error.

The error covariance components (random effects) and the CAPM coefficients (fixed effects) were simultaneously estimated using restricted maximum likelihood (REML) estimation. ML estimation finds a set of parameter estimates that maximizes the likelihood of the data given a distributional assumption. In contrast to the earlier ML estimation, REML estimation can accommodate the observations missing at random (Little & Rubin, 1987) and, as a result, produces unbiased estimates of the error covariance parameters in unbalanced data[11].

The performance of the unconditional CAPM and conditional CAPMs were compared using two information criteria – Akaike Information Criterion (AIC; Akaike, 1977) and

[11] For more details on REML, refer to Patterson and Thompson (1971). For alternative and more general derivations of REML, refer to Harville (1974), Cooper and Thompson (1977), and Verbyla (1990).

Bayesian Information Criterion (BIC; Schwarz, 1978). When model parameters are estimated using ML estimation, it is possible to increase the likelihood by adding parameters, which may result in over-fitting. Information criteria resolve this problem by introducing a penalty term for the number of parameters in the model. Accordingly, AIC and BIC evaluated the CAPMs with regard to the goodness-of-fit (accuracy) as well as the principle of parsimony; the smallest number of parameters to adequately capture the structure of the data.

The AIC is given by

$$AIC = 2k - 2logL$$

, where k is the number of parameters in the model and $logL$ is the log-likelihood of the model.

Similarly, the BIC is given by

$$BIC = klog(n) - 2logL$$

, where n is the number of observations.

Given a set of candidate models for the data, the preferred model is the one with the minimum information criterion value(s). For example, as a preliminary step, the intercept-only model was fitted with different specifications of error covariance structure. The compound symmetry error structure provided smaller AIC and BIC than did the unstructured, first-order autoregressive, and variance component error structures and thus was chosen for the competing CAPMs.

The generalized *R*-square based on likelihood ratio test statistic (Magee, 1990)[12] was also used to evaluate goodness-of-fit of the CAPMs. The likelihood ratio test of *R*-square is given by

$$R_{LR}^2 = 1 - exp\left(-\frac{2}{n}(logL_m - logL_0)\right)$$

, where $logL_m$ is the log-likelihood of the model of interest (which would include fixed effects and random effects), $logL_0$ is the log-likelihood of the intercept-only model, and n is the number of observations.

[12] Xu (2003) also discussed several *R*-square measures for a specific type of mixed models used in panel studies.

One of the advantages of using the likelihood ratio R-square is that this measure has a direct relationship with the Kullback-Liebler distance and information gain,

$$IG = -log(1 - R_{LR}^2) \text{ (Kent, 1983).}$$

The mixed modeling was performed using SAS 9.2 (SAS Institute, 2002-2008).

4. Main Result

By the comparison of four unconditional and conditional CAPMs using restricted ML test, M3 model, the unconditional CAPM with the labor income, is chosen. Table 4 shows the result of restricted ML test. Compared with M2, M1 has the lower figures in log likelihood ratio (-2LL), AIC and BIC. So, M1 is preferred to M2. So, even one more variable, the risk premium, is introduced, it is hard to tell than M2 can explain better than M1. When M1 is compared with M3, AIC and BIC are both decreased in M3. By introducing the variable, labor income, M3 can explain better than M1 can do. So, M3 dominates M1. AIC and BIC figures in M4 are exactly the same as those in M3 and, by parsimony, M3 will be chosen as the best-fit model.

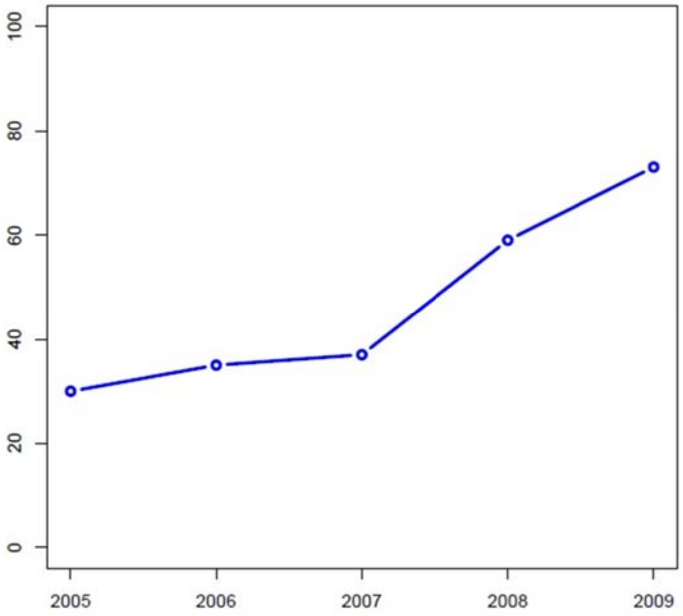
Through the test to estimate the model coefficients and accountability of a chosen best fitting model separately for three distinct time periods, the main result is generated as Table 5. Yet, the R-squares for each separated periods cannot indicate that the chosen M3 model can explain the projection of the expected returns. But, there is the tendency that the R-square is decreasing gradually and even goes to 0% for the third period. After the HFT method accounted for main portion of the total trading volume in the financial market, the CAPM seems to lose its power to generate the predictability of the returns. Truly, the labor income is one of the main resources to buy and sell equities in the financial market thus is employed in the CAPM model establishment, the CAPM model with the income still cannot explain the expected returns.

5. Conclusion

To find out the accurate predictability on the expected return has been a major issue in the modern financial field. Using a linear model like CAPM (Sharpe, 1964 and Litner, 1965) was a landmark and has been a proper way to predict the expected return on investment. To generate any more accurate predictions on the expected return, conditional CAPM and APT models are introduced and tested. Yet, it is not clear which model is a better predictor on the expected return. Jagannathan and Wang (1996) claim that the conditional model with macro-economic variables can predict the cross-section of the expected returns well. On the contrary, Lewellen and Nagel (2006) argue that the conditional CAPM does not show a more accurate result than the unconditional CAPM in explaining the macro-economic irregularities. Buss, Schlag, and Vilkov (2009) test the validity of a time-varying conditional CAPM and find out the conditional model is sufficient to explain the expected returns, especially when the economy is in the good shape. However, when there is a fluctuation in the financial market or when the whole economy is in a poor shape, the conditional model may need additional market factors. Febian and Herwany (2010) conclude that the accuracy of the asset price models depends on the macroeconomic characteristics and CAPM is not capable to explain the expected return during the economic crisis. Following these above arguments, the introduced P-L model cannot explain the expected return accurately, especially after the time period when the HFT is surged.

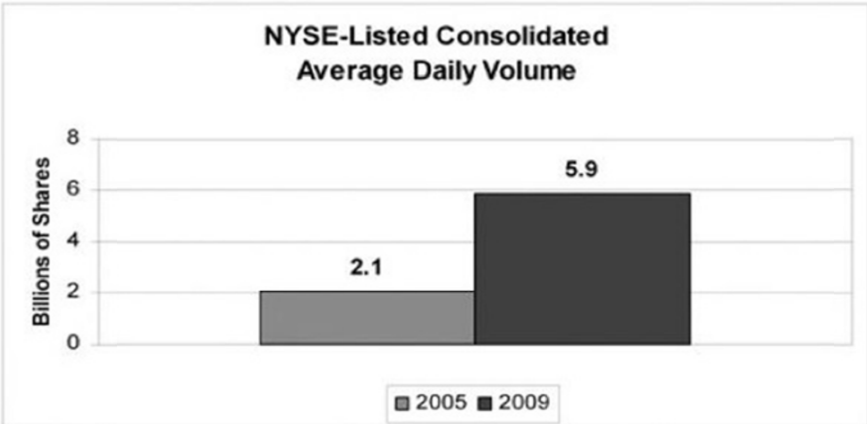
As the internet-based trading method has become popular, a return prediction model, the CAPM, cannot generate the accurate returns. It is even worse after the deregulation on the HFT in 2005. There are many claims that the HFT is not the major cause of the Flash Crash, the huge fluctuation in the U.S. stock market in June 2010, but as the trading volume by the HFT increases, the variance/covariance between an asset and the market will lose the power to predict the expected return because the HFT is a factor that cannot explain by the CAPM. As suggested by conditional CAPM and APT model claimers, there should be one good representation for the HFT in the Capital Asset Pricing Model to regain its power to predict the expected return.

Table 1. The proportion of all US equity trading through HFT



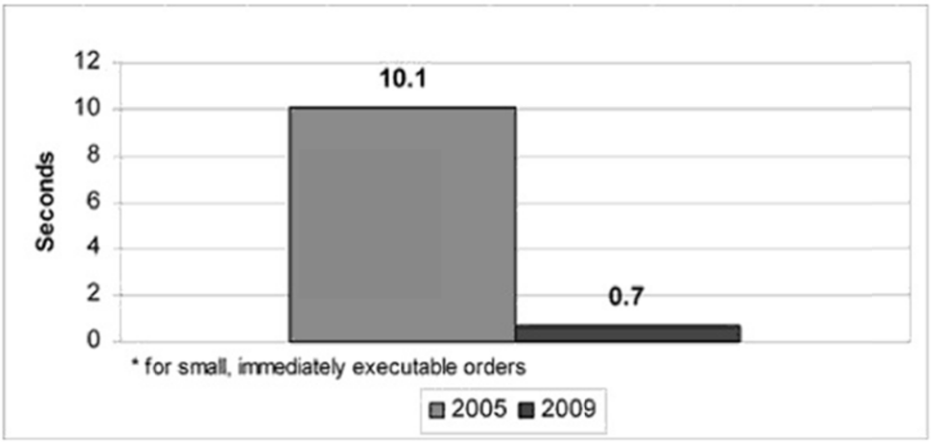
Source: Reginald Smith. Is high-frequency trading inducing changes in market microstructure and dynamics? (Sep 2010)

Table 2. Trading volume, amount, and size by HFT in NYSE



Source: Michael Durbin. All about the high frequency trading (2010), McGraw-Hill

Table 3. Execution time



Source: Michael Durbin. All about the high frequency trading (2010), McGraw-Hill

Table 4. Model Selection

Coefficient		Constant	VW	Prem	Income
M1	Estimate	0.08	0.46		
	Standard Error	0.32	0.25		
	t	0.24	1.84		
	p	0.81	0.08		
M2	Estimate	0.15	0.44	0.05	
	Standard Error	0.34	0.25	0.07	
	t	0.43	1.75	0.74	
	p	0.67	0.09	0.46	
M3	Estimate	0.20	-0.04		0.14
	Standard Error	0.29	0.29		0.05
	t	0.68	-0.13		2.62
	p	0.50	0.90		0.02
M4	Estimate	0.40	-0.21	0.12	0.18
	Standard Error	0.29	0.29	0.06	0.05
	t	1.36	-0.74	2.01	3.33
	p	0.19	0.47	0.06	0.00
		-2LL	AIC*	BIC*	R-Square
M1		41230	41234	41236	0.04%
M2		41233	41237	41240	-0.01%
M3		41228	41232	41234	0.09%
M4		41228	41232	41234	0.09%

* Smaller is better

Note

1. M1 is the unconditional CAPM without Human Capital (Income)
2. M2 is the conditional CAPM without Human Capital (Income)
3. M3 is the conditional CAPM with Human Capital (Income)
4. M4 is the conditional CAPM with Human Capital (Income) and Risk Premium.

Table 5. The result from the conditional with labor income model (M3)

Coefficient		Constant	VW	Prem	Income
Period 1	Estimate	0.93	-0.46		0.20
	Standard Error	0.57	0.57		0.11
	t	1.64	-0.81		1.91
	p	0.12	0.43		0.07
Period 2	Estimate	-0.73	0.13		0.27
	Standard Error	0.75	0.75		0.14
	t	-0.97	0.18		1.94
	p	0.34	0.86		0.06
Period 3	Estimate	0.07	0.32		-0.03
	Standard Error	0.67	0.67		0.12
	t	0.10	0.48		-0.23
	p	0.92	0.63		0.82
		-2LL	AIC	BIC	R-Square
Period 1		16785.6	16789.6	16792.1	0.53%
Period 2		11009.5	11013.5	11016	0.39%
Period 3		13281.2	13285.2	13287.8	-0.09%

Note

1. Period 1 spans from June 1994 to December 2000.
2. Period 2 spans from January 2001 to May 2005.
3. Period 3 spans from June 2005 to December 2010.

References

Akaike, H., 1977, On entropy maximization principle. In Krishnaiah, P.R. (Ed.). *Applications of statistics* (pp. 27-41). North-Holland, Amsterdam.

Bertholon, Henri, A. Monfort, and F. Pegoraro, 2008, Econometric Asset Pricing Modelling, *Journal of Financial Econometrics, Vol. 6, No. 4*, 407-458

Bhupathi, Tara, 2010, Recent Development: Technology's Latest Market Manipulator? High Frequency Trading: The Strategies, Tools, Risks, and Responses, *North Carolina Journal of Law & Technology*

Buss, Adrian, Christian Schlag and Grigory Vilkov, 2009, CAPM with Option-Implied Betas: Another Rescue Attempt

Byoun, Soku, 2008, How and When Do firms Adjust Their Capital Structures toward Targets?, *The Journal of Finance, Vol. LXIII, No. 6*

Caldarelli, Guido, Marina Piccioni, and Emanuela Sciubba, 2000, A Numerical Study on the Evolution of Portfolio Rules: Is CAPM Fit for NASDAQ?

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading Volume and Serial Correlation in Stock Returns, *The Quarterly Journal of Economics*

Chakravarty, Sugato, James Upson, and Robert Wood, 2010, The Flash Crash: Trading Aggressiveness, liquidity supply, and the impact of Intermarket Sweep Orders, *Journal of Economic Literature Working Paper Series*

Chen, Carl R., Peter P. Lung, and F. Albert Wang, 2009, Mispricing and the cross-section of stock returns, *Review of Quantitative Finance and Accounting* 32, 317-349

Dorman, J. P., 2009, Partitioning the variance in scores on classroom environment instruments. *Australian Journal of Educational & Developmental Psychology, 9*, 18-31.

Durbin, Michael, 2010, All About High Frequency Trading, McGraw-Hill

Febian, Erie and Aldrin Herwany, 2009, The performance of asset pricing models before, during, and after financial crisis in emerging market: Evidence from Indonesia, *Laboratory of Management and Business Faculty of Economics*

Filho, Elmo T., Fabio Gallo Garcia, and Joshua Onome Imoniana, 2009, Empirical test of conditional CAPM using expected returns of Brazillian, Argentinean, German and United States of American portfolio, *Corporate Ownership & Control, Vol. 7, Issue 2,* 269 – 279

Hedges, L. V., 2007, Correcting a significance test for clustering. *Journal of Educational and Behavioral Statistics, 32*, 151–179.

Jagannathan, Ravi and Zhenyu Wang, 1996, The Conditional CAPM and the Cross-Section of Expected Returns, *Federal Reserve Bank of Minneapolis, Research Department Staff Report 208*

Jayasinghe, Prabhath and Albert K. Tsui, 2009, Time-Varying Currency Betas: Evidence from Developed and Emerging Markets, *Singapore Centre for Applied and Policy Economics Working Paper Series*

Jules H. van Binsbergen, Michael W. Brandt, and Ralph S.J. Koijen, 2010, On the timing and pricing of dividends, *National Bureau of Economic Research Working Paper Series*
Katamadze, Eka, 2008, Overview of Conditional CAPM and the Cross-Section of Expected Returns

Kent, J. T., 1983, Information gain and a general measure of correlation. *Biometrika, 70*, 163-173.

Kirilenko, Andrei, Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2010, The Flash Crash: The Impact of High Frequency Trading on an Electronic Market, *Journal of Economic Literature Working Paper Series*

Kohn, Donald L., 2009, Monetary Policy and Asset Prices Revisited, *Cato Journal, Vol. 29, No. 1*

Lewellen, Jonathan and Stefan Nagel, 2006, The Conditional CAPM Does Not Explain Asset-Pricing Anomalies, *Journal of Financial Economics*

Little, R. J. A., & Rubin, D. B., 1987, *Statistical analysis with missing data.* New York: John Wiley.

McGowan, Michael, 2010, The Rise of Computerized High Frequency Trading: Use and Controversy, *Duke Law & Technology Review*

Magee, L., 1990, R^2 measures based on Wald and likelihood ratio join significance tests. *Journal of American Statistical Association, 44*, 250-253.

Merton, Robert C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica, Vol. 41, No. 5*, 867-887

Schwarz, G. E., 1978, Estimating the dimension of a model. *Annals of Statistics, 6*, 461-464.

Smith, Reginald, 2010, Is high-frequency trading inducing changes in market microstructure and dynamics?, *Journal of Economic Literature Working Paper Series*

Weinbaum, David, 2009, Investor Heterogeneity, asset pricing and volatility dynamics, *Journal of Economic Dynamics & Control* 33 1379-1397

Xu, R., 2003, Measuring explained variation in linear mixed effects models. *Statistics in*

Medicine, *22*, 3527-3541.

Yalcin, Atakan and Nuri Ersahin, 2010, Does the conditional CAPM work? Evidence from the Istanbul Stock Exchange, *Tusiad-Koc University Economic Research Forum Working Paper Series*