PhD in Business // No. 82 - 2020

Expected Returns: An Empirical Asset Pricing Study

Vu Le Tran

NORD UNIVERSITY BUSINESS SCHOOL

NORD University

www.nord.no

Expected Returns:
An Empirical Asset Pricing Study

Vu Le Tran

PhD dissertation submitted to
Nord University Business School
for the degree of PhD

Vu Le Tran
Expected Returns: An Empirical Asset Pricing Study

1. edition 2020
PhD nr 82 – 2020

ISBN: 978-82-92893-72-2
ISSN 2464-4331

Print: Trykkeriet Nord University

Nord University
8049 Bodø
www.nord.no

To my grandfather and father

Acknowledgement

The completion of this thesis is not a solitary work. Foremost, I am indebted to my supervisor, Thomas Leirvik. Without him, I cannot complete this Ph.D. journey. Thomas is not only my supervisor but also a colleague and a friend. I appreciate how he encourages me to discover new things and never give up on research ideas. I thank him for all his supports, suggestions during my Ph.D. years. I also thank Espen Sirnes, my co-supervisor for all his supports and research discussions.

I express my gratitude to Tarun Chordia, who invited me to visit Emory University, and to Costas Xiouros for the visiting at BI Norwegian Business School. These research-stays help me a lot in developing my research ideas. I thank Allison Gilmore, Lisa Johnson, from Emory University, and Eli Johanne Christensen, from BI Norwegian Business School, for all the warm administrative and logistic supports during these research stays.

I would like to also thank all the wonderful people that help me to complete this thesis by providing me lectures, feedback, and research discussions: Tomas Björk, Turan Bali, Tarun Chordia, Jay Shanken, Ilan Cooper, Costas Xiouros, Bruno Gerard, Stephen Geoffrey Dimmock, Francisco Santos, Alexandre Rubesam, Matti Suominen, Peter Schotman, Anatoli Bourmistrov, Oleg Nenadic, Salvatore Miglietta, Espen Henriksen, Jon H. Fiva, and Hilde C. Bjørnland. I am thankful to former and current Ph.D. students, who can spend hours to discuss research ideas with me and make my Ph.D.'s life more enjoyable. They are Kristian Støre, Andreas Mikkelsen, Muhammad Kashif, Yevheniia Antoniuk, Dinh-Vinh Vo, Duc-Hong Hoang, Sudip Tiwari, Obaidur Rehman, Jonas Barth-Stenersen, Wang Anne, Renata Silva de Almeida, Mahdieh Vosough Gohari, Rong Geng, Oleh Korop, Irena Kustec, Espen Seidel, and Andrew Sang Jun Lee. I express my thank to Grete Ingemann Knudsen and Ellen Abelgård, from Nord University, for all your valuable supports during my Ph.D. journey.

Finally, I am indebted to my family: my grandparents, parents, parents-in-law, sister for their unconditional love. Thank you, Giang, my wife, who always with me during good and bad times of this journey. Without my family's support, this thesis would not have been completed.

Oslo, June 2020
Vu Le Tran

Contents

Chapter I

Introduction

The goal of this thesis is to study the expected return of an asset and the predictability of this return in both the short-term and the long-term. The thesis consists of four papers.

In the first two papers, I investigate the predictability of stock returns and cryptocurrency returns with daily frequency. I develop a new measure of market efficiency ($AMIM$) which is easy to implement. $AMIM$ helps us to compare the level of efficiency of different assets at different points in time.

In the last two papers, I investigate the predictability of the expected stock returns in a long-term framework. In the first of these papers, I investigate how stable liquidity is priced in the stock market. In the last paper, I study the impact on stock returns when there are collective biases from the investors regarding future stock pay-offs.

In this chapter, I will first go through the theories that my research is based on: the efficient-market hypothesis (EMH), adaptive market hypothesis (AMH), stochastic discount factors, and factors model in asset pricing. These theories will elaborate the process of how the market prices an asset and hence help us to define the expected return of an asset. Then, I will give the reader a summary of each paper's contribution to my thesis. I also discuss the methodologies that I use for the four papers in this thesis.

1.1. Rationality, Arbitrage, and Stochastic Discount Factor

I will use this part to go through the theory of deriving a rational price and expected return. First, I will discuss how the process of smoothing the consumption of rational investors helps us to derive a fair price for an asset. Then, I move on to discuss the properties of the pricing kernel of an asset under the law of one price and the no-arbitrage condition. Finally, I discuss the equilibrium expected return of an asset.

1.1.1. Rationality, Maximization of Utility, and Consumption CAPM

Markets are the place where different parties interact with each other to determine trading prices and quantities. The central research topic in many studies is how markets incorporate information into prices, or how do we know if an asset is fairly priced or not? Neoclassical economics is one of the prominent paradigms to study these questions. The paradigm has a set of important assumptions; according to Weintraub (2002), these are:

- individuals have rational preferences.

- individuals maximize utilities and profits.
- individuals act independently on the basis of full and relevant information.

The neoclassical economics paradigm usually studies the equilibrium market stages when all market participants rationally and independently maximize their utility, expressed economically through an utility function. The traditional approach also assumes that every economic agent has the same utility function. This also implies that an equilibrium state can be reached if every economic agent individually maximizes their profits (or utility). If so, we can sum all participants' utility functions to become an overall market utility function, then optimize this single function to reach the efficient steady state of the market.

In a more realistic case, each individual has a different approach toward the markets, so there are many different utility functions. Therefore, there is no single common optimal-behavior for each economic agent in the market. However, in a competitive market, one can prove that by individually maximizing their utility function, all the agents can lead the market to an equilibrium stage. Therefore, we can treat the market (or all economic agents) as a representative agent.[1]

The use of one representative agent is convenient to study the quality of the market in equilibrium. This paradigm is useful to the extent that we can use mathematical calculus to solve the optimization problem, thus finding the optimal behavior of the markets and building theories on these findings. Then, we test these theories using empirical data.

Using the above approach, we solve an optimization problem on how to allocate consumption and investment into different assets overtime. The goal is to maximize the stream of utility from consumption. Solving this consumption maximization problem will help us to derive the equilibrium price of an asset.

Specifically, let p_t and p_{t+1} be the asset's price at time t and at time $t+1$ respectively. We also call d_{t+1} the dividend at time $t+1$, and $u(c_t)$ and $u(c_{t+1})$ the consumption utilities at time t and $t+1$. Finally, β is the subjective discount of utility. The equilibrium price of the asset will therefore be:

$$p_t = E_t\left[\beta \frac{u'(c_{t+1})}{u'(c_t)}(p_{t+1} + d_{t+1})\right] \tag{1.1}$$

In this above equation, $\beta \frac{u'(c_{t+1})}{u'(c_t)}$ is the marginal rate of substitution (MRS) in con-

[1]The representative agent is a common discussion in economics, especially in the market setting. For a detailed discussion, please consult Chapter 7 in Back (2010).

sumption utility. $p_{t+1} + d_{t+1}$ is the pay-off of the asset in the future. The price of an asset today depends on the size of the trade-off between the marginal utility in the future versus today, and the pay-off from the asset in the future. This is because a person needs to forgo some of the utility of consumption today to buy this asset. The asset will generate a pay-off in the future, hence increasing the future consumption.

If we divide the above equation by p_t, and define the gross rate of return as $R_{t+1} = \frac{(p_{t+1} + d_{t+1})}{p_t}$, we will come up with the below result:

$$1 = E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)} \cdot \frac{(p_{t+1} + d_{t+1})}{p_t}\right] = E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)} \cdot R_{t+1}\right] \tag{1.2}$$

Equation (1.2) is true for any asset return. These are the main results of the consumption capital asset pricing model (CCAPM).

1.1.2. Law of One Price, Absence of Arbitrage, and Stochastic Discount Factor

The law of one price entails that assets with the same pay-off should have the same price. If the law of one price holds, then there exists at least a stochastic discount factor (denoted as SDF or m_{t+1}) that can be used to price the asset at time t.[2] The reasoning for this claim is quite simple. Because all the assets or portfolios with the same pay-off should have a same price, hence there should be a common pricing kernel that links the pay-off to a common price. However, the pricing kernel need not be unique or positive. It is called SDF because it depends on the state in the future. Therefore, m_{t+1} is a abbreviation of $m_{t+1}(\omega)$, where ω is the state of the economy at time $t+1$. Therefore, m_{t+1} is not deterministic at time t. For example, with the consumption pricing model in equation (1.1), the SDF is $m_{t+1} = \beta\frac{u'(c_{t+1})}{u'(c_t)}$.

Next, we will present another concept called "Arbitrage". Arbitrage is just a free lunch. Following Björk (2009), an arbitrage opportunity arises when we have a portfolio or an asset that has zero value today (hence a zero price today) but will have positive value in the future. It is a free lunch in the sense that we can acquire this asset at zero cost but can still have a positive pay-off in the future.

If such an opportunity happens, it would be lucrative for an investor to buy as much as possible of this asset. When everyone is doing that, the demand for this asset rises.

[2]See Cochrane (2009) for detailed proof.

Hence, people are likely to pay something to acquire the asset. Therefore, its price will rise from zero. The price will rise to an equilibrium point when the price represents today's fair value of this asset's pay-off in the future. Hence, the arbitrage opportunity disappears when investors try to exploit it.

Therefore, the existence and persistence of an arbitrage opportunity is rare on the market. In financial economics, we use to assume a no-arbitrage condition. As the pay-off of an asset in the future is positive, the price of this asset should be positive under the no-arbitrage condition. This would mean that we have to pay something to get something in the future.

Now, we come back to our pricing kernel m_{t+1}. First, we know that the price of an asset can be stated as: $p_t = E_t\Big[m_{t+1}(p_{t+1} + d_{t+1})\Big]$. Second, we know that $(p_{t+1} + d_{t+1})$ is positive and p_t is positive under the no-arbitrage condition. Hence, the pricing kernel m_{t+1} is likely positive. It turns out this is the case. The reader can find a detailed proof in Cochrane (2009) or Björk (2009).

In summary, under the no-arbitrage condition, there should exist a positive SDF to price assets on the market. In the next part, I analyze the expected return, when we use a positive pricing kernel.

1.1.3. The Expected Return

With $1 = E_t\Big[m_{t+1}(R_{t+1})\Big]$ true for any return, including a risk-free rate return, the gross risk-free rate at time $t+1$, $R_{f,t+1}$, is deterministic at time t. Hence, we can make two majors observations:

- $E_t\Big[m_{t+1}\Big] = 1/R_{f,t+1}$. This has to be positive under the no-arbitrage condition.
- and $E_t\Big[m_{t+1}R^e_{t+1}\Big] = 0$, where R^e_{t+1} is the excess return from the risk-free rate.

From these observations, we can derive the expected return of an asset:

$$E_t[R^e_{t+1}] = -R_{f,t+1} \cdot cov(m_{t+1}, R^e_{t+1}) \tag{1.3}$$

Hence, the expected return of a stock is determined by the covariance of its return to the SDF. We used to call it the covariance risk. Given the same expected pay-off or cash flow of two assets, any investor would like to pay less to buy the riskier asset. Hence, the expected return of the riskier asset needs to increase to compensate for the risk.

We can represent the SDF in different ways, or decompose the SDF in different combining parts. Hence, as a result, the expected return of an asset will depend on the covariance between this return and all the combining parts of the SDF. We call all the SDF's components risk factors. Therefore, the expected return of an asset can be predicted based on its exposure to the risk factors. This is an important conclusion. However, to get an exact prediction of the expected return, it is essential that we use the right model (or the right specification of the SDF).

I will discuss here the two most common factors model in the literature, the consumption risk factor and the market factor. Other factors will be discussed in details in the below section.

First, if we apply the consumption CAPM model, then $m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)}$. As $u'(c_t)$ is deterministic, then $E_t[R^e_{t+1}] \propto -cov[u'(c_{t+1}), R^e_{t+1}]$. The marginal utility function decreases with consumption. Therefore, we can say that $E_t[R^e_{t+1}] \propto cov[c_{t+1}, R^e_{t+1}]$.

So an asset whose return is positively correlated with the consumption will be considered riskier. Why is that? Because this asset does not offer any hedging benefit to our smoothing consumption process. Indeed, the asset offers bad returns in bad times (when consumption is low). The asset has to be cheap enough for us to buy it. In other words, the expected return has to be high enough to compensate for the risk of taking it.

On the opposite side, an asset whose return is negatively correlated with our consumption is considered less risky because it offers a good return in bad times. In that sense, this asset can be considered as an insurance against bad things. Hence, we have to pay a premium to buy it. This would mean that the expected return of this asset is lower in the future.

If we specify the utility function as either a quadratic utility function or an exponential utility function plus a jointly-normal distribution of return, we will come up with the CAPM model. (Treynor, 1961, 1962; Sharpe, 1964; Lintner, 1965; Mossin, 1966).[3]

The CAPM model states that:

$$E_t[R_{t+1}] = R_f + \beta_t \cdot E_t[R_{t+1,M} - R_f] \tag{1.4}$$

where M is the market portfolio and $\beta_t = \frac{Cov_t\Big(R_{t+1}, R_{t+1,M}\Big)}{Var_t(R_{t+1,M})}$. The only risk factor in the CAPM model is the market return. Only the covariance part of the asset's return to the market return is compensated for in the expected return. Other idio-syncratic risk is not compensated for. This is because following CAPM, the idio-syncratic risk can be hedged away by holding a well-diversified portfolio. However, the systematic covariance

[3]See Cochrane (2009) for the derivation.

risk of the market cannot be hedged away.

Following the CAPM's point of view, the market portfolio offers the best risk-reward ratio, and hence the best Sharpe ratio. The market portfolio is also on the mean-variance efficient frontier. Given the same level of risk, portfolios on the mean-variance efficient frontier will have the highest expected return.

1.2. Efficient Market Hypothesis and Return Predictability

In the previous section, we discussed the SDF as an important building block of the expected return of an asset. In this section, I will present another important building block: the efficient-market hypothesis (EMH). I also discuss how the efficient-market hypothesis relates to the pricing kernel that we investigated in the previous section and the predictability of asset returns.

1.2.1. Efficient Market Hypothesis

According to Fama (1970), the efficient market hypothesis (EMH) states that an asset's price should rationally reflect all relevant information. Fama (1970) divides market efficiency into three levels:

- Weak form: Where all information from price, volume, and trading is included in the price.
- Semi-strong form: Where all public information, which includes all the information set in the weak-form plus fundamental finance information of the firm, is included in the price.
- Strong-form: Where all information, both private and public, is already included in the price.

If markets are efficient, we cannot systematically earn any abnormal return apart from the expected return, so it is not possible to beat the market. However, as discussed before, to know whether the market is efficient or not, one has to know the exact pricing model, because knowing the exact pricing model helps us to define the right expected return and hence to identify the abnormal return. In addition, to test or decide which asset pricing model is true to be used, we have to make sure that the market is efficient, because we need all information reflected in the asset pricing model so we can decide which model does the best pricing job. Fama (1991) mentions this as joint hypothesis testing in the sense that you test both the asset pricing model and the market efficiency hypothesis.

1.2.2. EMH, SDF, Random Walk, and Stock Return Predictability

EMH does not entail that a stock's return is not predictable. Indeed, risk plays a large part in the predictability of a stock's return. Following the discussion in the previous part, with equation (1.3), we can predict stocks' expected returns based on the covariance risk between return and risk factors in the SDF. In this framework, if we can construct the right SDF or identify the exact risk factors inside the SDF, we can construct the right expectation of return. Hence, we can test whether there exist any systematic abnormal returns apart from the model. If so, we can say that the market is inefficient.

However, with the joint hypothesis testing problem, we have to make sure that the asset pricing model that we use is right. Therefore, the testing of market efficiency becomes very hard to do.

Risk plays a big part in identifying the expected return. However, risk is usually important in a long-horizon. In a very short time frame, for example at daily or minute frequency, the covariance risks are substantially small and therefore can be seen as not important.

In a very short time frame, we can consider the SDF as a deterministic value. Hence, the covariance risk between return and SDF can be seen as zero. Following equation (1.3), this would lead the expected return of an asset to be equal to the risk-free rate: $E_t[R_{t+1}] = R_{f,t+1}$, which is a constant. So, in the short run, we can assume that the asset has a constant expected return. Of course, this assumption is not always correct. However, as long as the covariance risk between return and the SDF is small enough in the short run, the pricing error is very small and can be ignored.

In some cases, the gross risk-free rate can also be considered to be 1 also in the short run.[4] Hence, the gross expected return of an asset can be set to be equal to 1.[5] This is also equivalent to the assumption that SDF is constant and equal to 1.

With the pricing equation $p_t = E_t[m_{t+1}(p_{t+1} + d_{t+1})]$, $m_{t+1} = 1$, and assuming no dividend in the short run, then the expected asset price of the next period is today's price or $p_t = E_t[p_{t+1}]$. Hence, we can write the asset price as a random walk (RW)

[4] Hence, the net risk free rate is zero.

[5] Hence, the net return is zero.

process of

$$p_{t+1} = p_t + \varepsilon_{t+1} \tag{1.5}$$

where ε_{t+1} is the shock at time $t+1$. This shock is independent of (hence orthogonal with) the current price. Then, $E[p_t \cdot \varepsilon_{t+1}] = 0$. The shock is also independent of the previous shocks; hence, $E[\varepsilon_{t+1-j} \cdot \varepsilon_{t+1}] = 0, \forall j \leq t$. The RW process is the common process that goes alongside EMH. The logic behind this is quite simple. If the market is efficient enough and all information is included in the price, the best guess for tomorrow's price is today's price. The only thing that drives tomorrow's price is the information shocks happening tomorrow.

RW is the common process that is used to describe the market efficiency in the short-run. With RW process, we have some very interesting properties that can be used to test market efficiency.

The two most common tests with the RW process are the auto-correlation test, and the variance test. The former is based on the fact that $p_t = E_t[p_{t+1}]$, the net return $E_t[r_{t+1} = \frac{p_{t+1}}{p_t} - 1] = 0$. Hence, if we run an auto regression (AR) such as:

$$R_t = \beta_0 + \beta_1 \cdot R_{t-1} + \beta_2 \cdot R_{t-2} + \cdots + \beta_q \cdot R_{t-q} + \varepsilon_t \tag{1.6}$$

and if markets are efficient, with constant expected return, then the coefficient β_0 will capture the constant expected return. The other coefficients (β_1, β_2,..., β_q) should also be zero, or at least insignificantly different from zero. If EMH with constant expected return does not hold, the β coefficients are (significantly) non-zero. Fama (1965); Fama and Blume (1966) shows that these coefficients tend to be very small and the deviation from RW cannot be exploited with profit after the trading cost. So, in the 1970s, researchers were in favor of the EMH.

The other test is the variance ratio test of Lo and MacKinlay (1988, 1989). They use the fact that the variance of asset return under the RW process is proportional to the time horizon to test the EMH. For example, the ratio between the variance of the 3-days asset return and the variance of the 1-day asset return should be equal to 3, if the stock return follows an RW process. They found that this is not always the case. This finding is against EMH. In the first paper of this thesis (discussed below in detail), I derive a novel measure to test market efficiency, and as such contributes to the on-going debate in the literature on market efficiency.

1.3. Adaptive Market Hypothesis (AMH)

In the above section, we discussed the EMH, whose underlying premise is a rational pricing model. The discussion about market efficiency is long and controversial. The basic assumption of rationality is usually denied by other researchers in the behavior finance. The two schools of thought rarely agree with each other. There is a new theory named "adaptive market hypothesis" of Lo (2004, 2017), which is a synthesis of these two approaches, that can be used to describe the market. In this part, I will first go through some critiques of the rationality assumption. Then, I will present the main idea of AMH with illustrative examples. I will also explain why AMH is a good base for me to construct the first two papers of my thesis.

If we first just focus on the human aspects of the market, then the market is the collection of all economic agents who interact with each other. As we discuss above, the central property of the human in neoclassical-economics is rationality. This property ensures that we can portray human-economic-behavior in a structured and institutionalized way to better predict economic outcomes. However, some may also argue that human behavior is not always predictable. Therefore, the neoclassical approach is merely an abstraction, and an over-simplification of human behavior.

According to Hausman (2003) "The Philosophy of Economics", rationality is reflected through the preference among different choices. The common assumption in economics is that human preferences are either risk-neutral, risk-averse, or risk-loving. However, through different experimental studies, it can be seen that the human preference for risk is not *consitent*.

The two psychologists Kahneman and Tversky (1979), show that humans are usually risk-averse when it comes to gaining new opportunities, and risk-loving when it comes to loss. For example, people tend to prefer a certain gain of 40 USD over a lottery with a 50% chance of gaining 100 USD and a 50% chance of gaining nothing. In this case, even though the expectation value of the lottery is 50 USD, because people are risk-averse, they tend to choose a lower certain value gain (40 USD).

However, when it comes to loss, people also tend to prefer a lottery with a 50 % losing 100 USD or a 50 % chance of losing nothing over a certain loss of 40 USD. In this case, even though the lottery's expectation of loss is 50 USD, which is more severe than a certain loss of 40 USD, because people are risk-loving, they tend to gamble. Lichtenstein and Slovic (1971) show that people are also irrational when it comes to classifying the preference for a choice. These examples are the most typical in behavioral economics, which illustrates the uncertainty of human-behavior in economics. Therefore, it will be harder to study the market if we only have a predetermined set of assumptions about human-economic-behavior as in the neoclassical paradigm. This uncertainty about preferences and the rationality of economic-behavior leads to a new branch of economics/finance study called behavior economics and behavior finance.

Apart from the rational and behavioral approaches, Lo (2004, 2017) states that markets are not always rational but can be emotional and also have to adapt to a changing environment via simple heuristics. This is due to the fact that people are not always rational. The most important aim of the investor is to first survive in the market "environment" before maximizing profits. Therefore, a single equilibrium stage is hard to persist. If so, market efficiency is also dynamic and can change from time to time. Lo (2004, 2017) propose a concept called adaptive market hypothesis (AMH) and suggests that we can use evolutionary approach to study the dynamics of market efficiency. Indeed, if we treat participants as different species in a market environment, we can classify for example some "species" as the "prey", and some as the "predator".

One example of this environmental approach can be presented as follows. High frequency trading (HFT) is a new "species" in the market that emerged in the 2000s. These activities are now essential in the market. According to Miller and Shorter (2016), ***55 %*** of the trading volume in equity in the US is done by HFT. The percentage is roughly ***80%*** in foreign exchange. Hence, there exists a new strong species (HFT) that changes the market environment context.

Trades now happen in a couple micro-seconds with super computers. If the average traders do not adapt to this new environment, they are too slow and thus can become "prey" to HF traders, who are faster at acquiring information. Biais, Foucault, and Moinas (2015) prove that fast traders with more information can impose the cost of adverse selection to other slow traders.[6] This can cause a market failure. The authors also show that fast trading can compete with each other to better exploit information, thus generating an arms race of technology over-investment with the high temptation of to be fast rather than slow. A well-designed tax scheme (Pigovian), or two separate markets (fast and slow), according to the authors, can solve the problem. Different "species" in the market will lead to competition, adaptation, and natural selection among species.

To illustrate the above competition, consider a "front-running" case when one "slow"

[6] If market efficiency is weak, and some information is not fully reflected in prices, it is easy for the informed traders to beat the market. In contrast, the ignorant trader can make a bad decision in trading. This is an *adverse selection* situation.

investor would like to buy 100 000 shares of APPLE. Such a large volume order is partly filled in one exchange (for example : 10 000 shares in exchange A). The remaining volume will be routed to an other exchange (for example: 90 000 in exchange B). If the HF trader can somehow predict the order movement, he can quickly buy the APPLE shares in exchange B, then resell them at a higher price to the "slow" investor. Algorithms can perform the "front running" process in just a couple of milliseconds. Figure (1.1) illustrates the process.

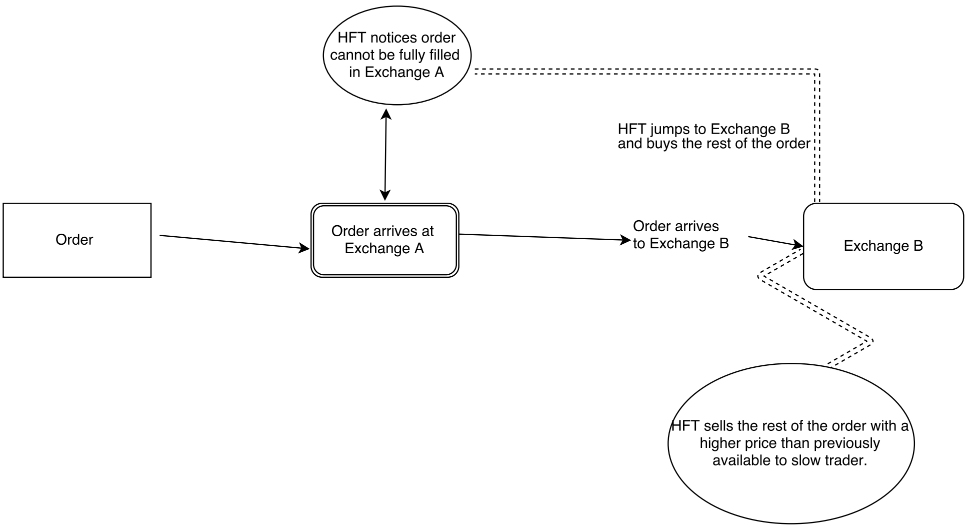


Figure 1.1. Front Running.

Following these above discussions, it is more realistic for AMH to say that, first, markets are not always efficient. Second, we cannot consider every economic agent to have optimal behavior, because they have first to survive before optimizing profit. Sometimes, the decision has to be made quickly enough to survive.

- So, the decision can be based on heuristics (which is not always optimal). Optimized behavior is not always in place due to emotional biases. Lo (2004, 2017) suggest that instead of "optimized behavior", "heuristics" behavior is more common in the market-place. Heuristics behavior is similar to the "bounded rationality" of Simon (1955). The concept explains that optimization is costly, so humans make choices that they find satisfying but not necessarily optimal. This is more reasonable, for example, in the HFT setting. The decisions are mainly made by computer algorithms. In this setting, different computers will compete with each other on speed. To have a completely optimized decision, one can reduce the speed of decision making, thus risks becoming the "prey" like in the above example.

- Second, the decision is also based on a limited information set (not always complete). This is a corollary of heuristics.
- Third, the decision of one agent is not always independent of an other agent's decision. One example of the correlation of decisions can be found in the Flash Crash on 6 May 2010. During the crash, stock indexes such as the S&P 500, the Dow Jones Industrial Average, and the Nasdaq Composite collapsed and recovered rapidly in half an hour. There was no specific economic reasons that could explain the crash. Kirilenko, Kyle, Samadi, and Tuzun (2015) argue that HFT algorithms increased the market volatility during the crash by repeatedly buying and selling amongst themselves.

The AMH is based on the idea that the market environment changes over-time and economic agents have to adapt to survive. Thus, the AMH will have some interesting implications when we compare it with the usual rationality approach that we discussed above.

- The risk-reward relationship is not stable over-time, because it is based on the market environment, which consists of different actors (or species) that evolve over-time. Tax, law, and regulation also change over-time. These changes create a dynamic market environment. Risk and return that are not stable overtime will lead to a changing risk premium.
  This time-varying risk premium can also be established with the rational expectation model that I present above when we have a change in risk preference. However, the new contribution of AMH is that the change in the risk preference can be due to a natural selection process. This means that when the environment changes, some types of investors survive and redefine the risk preferences and hence the risk-reward relationship.
- Arbitrage opportunities can appear from time to time. This is contrary to EMH, which assumes that all information is included in prices. If mispricing happens, then arbitrage force arises quickly to push the price toward equilibrium. So, there is no free lunch, or systematic abnormal return, under EMH. Through the arbitrage force, the market will have a strong trend toward strong efficiency. However, AMH views that arbitrage can happen in cycles, where it appears-disappears-reappears.

When an arbitrage opportunity appears, it can die out when everybody exploits it. However, new arbitrage opportunities will arise because of the changing of types of investors, changing regulation, and changing business conditions.

- The third implication is that the profitability of investment strategies will also go up and down because of the natural selection process. With EMH, by the arbitrage force, new abnormal returns will disappear when everybody exploits them. With AMH, Lo (2004) argues that certain good strategies can decline in certain environmental conditions. However, they can also rise again when the market conditions become favorable.

  The second and third implications would also imply that the market's efficiency will change and adapt over-times. There will be periods when we have significant abnormal returns along with efficient periods.
- The fourth implication is that the ultimate goal is survival, not optimization of utility or profit. Although utility or profit maximization is relevant, in a natural selection context, one has to survive first.
- The final implication is that innovation is the key to survival. With EMH, to achieve a certain level of expected return, we just need to be exposed to a certain level of risk. With AMH, this is not enough, because the environment, and the risk-reward relation, changes over-time. When the key goal is to survive, economic agents have to be adaptive. The key to being adaptive to a new environment is to innovate.

These observations of AMH are the background on which I construct the first two papers of my thesis. Based on the implication that market efficiency can change over-time and is not static, I construct a measure of market efficiency. Then, I apply this measure to investigate the market efficiency level across assets and time. I will summarize the findings of these two papers in the summary section below. In the next section, I will discuss the asset pricing factors model, which is the foundation of my last two papers.

1.4. Factor Model and Expected Return

The above sections set the ground for us to define the expected return. As shown above, the most common approach is to use the SDF to derive the rational expectation of an asset's future return. In the first section, the SDF can be described as the marginal rate of substitution of consumption. Another common approach is using the factor model for the SDF. In this approach, one usually constructs the SDF as a linear combination of different factors. Each factor is usually a portfolio return. The combination of these factors will give us a portfolio that has the maximum Sharpe ratio squared. Hence, the portfolio lies in the mean-variance efficient frontier. First, I will go through the Hansen and Jagannathan (1991) bound and the logic of why we can use portfolios on the mean-variance efficient frontier to present the SDF. Then, I will move on to how to use characteristics-managed portfolios to construct the SDF, and finally, the factor representation of an expected return.

1.4.1. Hansen Jagannathan Bound, SDF Presented with Portfolio Return.

Recall equation (1.3), where $E_t[R^e_{t+1}] = -R_{f,t+1} \cdot cov(m_{t+1}, R^e_{t+1})$. With $1/R_{f,t+1} = E_t[m_{t+1}]$, we can formulate the equation as:

$$\begin{aligned} E_t[R^e_{t+1}] &= -R_{f,t+1} \cdot cov(m_{t+1}, R^e_{t+1}) \\ \frac{|E_t[R^e_{t+1}]|}{\sigma(R^e_{t+1})} &= \frac{|\rho \cdot \sigma(m_{t+1})|}{E_t[m_{t+1}]} \end{aligned}$$

where ρ is the correlation between the SDF and excess return. Because the correlation is from -1 to 1, we come up with the Hansen and Jagannathan (1991) bound:

$$\frac{|E_t[R^e_{t+1}]|}{\sigma(R^e_{t+1})} \leq \frac{\sigma(m_{t+1})}{E_t[m_{t+1}]} \tag{1.7}$$

The left-hand side is the absolute value of the Sharpe ratio, and the right-hand side is the ratio of the volatility of the SDF to the SDF's expected value. The bound can be illustrated as follows in figure (1.2).

In the figure, the two slopes are bound by the value: $\pm\frac{\sigma(m_{t+1})}{E_t[m_{t+1}]}$. The two bounds create a mean-variance frontier. Returns in the upper part of the frontier are perfectly negatively correlated with the SDF and hence positively correlated with consumption. These asset returns are considered risky and thus should have higher expected returns.

The logic is the opposite with the asset returns in the lower part of the frontier. These returns are perfectly positively correlated with the SDF and hence negatively correlated with consumption. These asset returns have hedging benefit, thus, requiring lower expected returns.

As we can see, the maximum absolute Sharpe ratio of a portfolio has to be equal to $\frac{\sigma(m_{t+1})}{E_t[m_{t+1}]}$. This is possible if and only if this portfolio return is perfectly correlated with the SDF. This portfolio can hence be found on the mean-variance efficient frontier (the

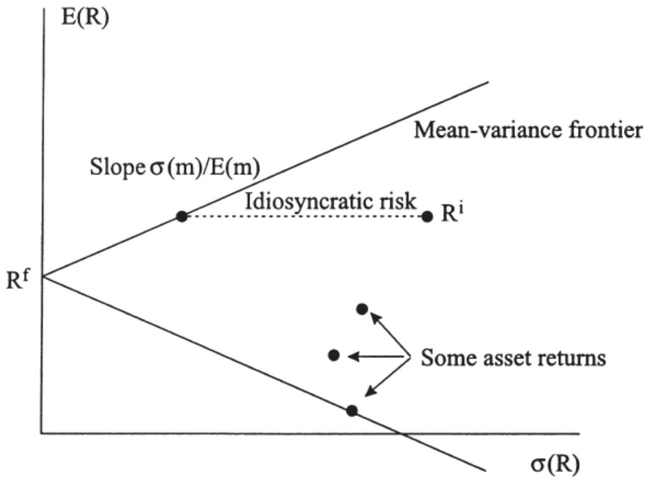


Figure 1.2. Hansen and Jagannathan (1991) bound. The figure is from Cochrane (2009).

upper part of the mean-variance frontier). If we call that portfolio excess return $R^{e,mv}_{t+1}$, then we can always represent the linear relationship between $R^{e,mv}_{t+1}$ and m_{t+1} as

$$\begin{aligned} m_{t+1} &= g + h \cdot R^{e,mv}_{t+1} \\ R^{e,mv}_{t+1} &= n + f \cdot m_{t+1} \end{aligned}$$

where n, f, g, and h are scale numbers. Therefore, we can represent the SDF as just a linear combination of the maximum Sharpe-ratio return on the mean-variance efficient frontier and a scale number. It turns out that not only the maximum Sharpe-ratio return, but also any portfolios' returns in the mean-variance efficient frontier can carry the information of the SDF. So, we can always find the right parameters g, h to make $m_{t+1} = g + h \cdot R^{e,mv}_{t+1}$ with any portfolios in the mean-variance efficient frontier. A proof can be found in Cochrane (2009).

Note that we can always use the asset's returns on the market to construct a portfolio on the mean-variance efficient frontier. We can also use different portfolio returns to construct a portfolio on the mean-variance efficient frontier. This is the essential point of different factor models in the literature. Indeed, many factors in the literature are portfolio returns. The underlying assumption of these factor models is that these factors help us to construct the maximum Sharpe ratio portfolio.

A new factor is usually discovered because its return cannot be explained by other existing factors' return. Hence, if we include that new factor among the old factors, we can get a higher maximum Sharpe ratio and hence a new SDF with less pricing error.

In the next part, I will first discuss the general form of a linear factor model in asset pricing. Then , I will move forward to the difference between conditional and unconditional models. And finally, the rationale for using the characteristics based factor models that are very popular in the literature.

1.4.2. Linear Factor Model

The covariance between return and SDF is considered a risk to consumption. In that sense, if SDF is a combination of different risk factors that can affect consumption, then the covariance between return and these risk factors should offer a reward also.

Generally speaking, we can assume that the combination of an SDF can be represented as: $m_{t+1} = a_t + \mathbf{\tilde{f}^T_{t+1}} \mathbf{d_t}$, where $\mathbf{\tilde{f}^T_{t+1}}$ is the transposition of risk factors vector, and $\mathbf{\tilde{f}^T_{t+1}}$ is usually a demeaned vector ($\mathbf{\tilde{f}^T_{t+1}} = \mathbf{f^T_{t+1}} - E_t[\mathbf{f^T_{t+1}}]$), hence having an expectated value of zero. The term $\mathbf{d_t}$ is the risk coefficient vector, and a_t is a scale number. We can also scale a_t and $\mathbf{d_t}$ up with the SDF such as $m_{t+1} = 1 - \mathbf{\tilde{f}^T_{t+1}} \mathbf{b_t}$. This SDF has an expectation of 1. From this SDF, we can derive a β pricing model for the expected excess return as follows:

$$
\begin{aligned}
E_t[R^e_{t+1}] &= -\frac{cov_t(m_{t+1}, R^e_{t+1})}{E_t[m_{t+1}]} \\
&= cov_t(R^e_{t+1}, \mathbf{\tilde{f}^T_{t+1}}) \cdot \mathbf{b_t} \\
&= cov_t(R^e_{t+1}, \mathbf{\tilde{f}^T_{t+1}}) var_t(\mathbf{\tilde{f}_{t+1}})^{-1} var_t(\mathbf{\tilde{f}_{t+1}}) \cdot \mathbf{b_t} \\
&= cov_t(R^e_{t+1}, \mathbf{f^T_{t+1}}) var_t(\mathbf{f_{t+1}})^{-1} var_t(\mathbf{f_{t+1}}) \cdot \mathbf{b_t} \\
E_t[R^e_{t+1}] &= \boldsymbol{\beta}_t^T \boldsymbol{\lambda}_{t+1} \qquad (1.8)
\end{aligned}
$$

So, the expected excess return of an asset will depend on the $\boldsymbol{\beta}_t$ exposure of this asset return to each and every risk factors, and on the risk premium vector $\boldsymbol{\lambda}_{t+1}$ of the factors.

Now, knowing from the previous sub-section that one can always represent the SDF as a linear combination between a scale number and a return from the mean-variance efficient frontier ($R^{e,mv}_{t+1}$), we can always write that portfolio as a linear combination of different portfolios. Therefore, we can consider these element-portfolio returns as factors such that $\tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}\mathbf{b_t} = h \cdot R^{e,mv}_{t+1}$. Hence, the above analysis lets us have a factor model when expected return depends on the exposure to different factors in the form of portfolio returns. In this case, when all factors are tradable assets (or portfolios), then in an arbitrage-free market, $\boldsymbol{\lambda}_{t+1} = E_t[\mathbf{f_{t+1}} - R_{f,t+1} \cdot \mathbf{1}]$. This factor structure of expected returns is similar to the factor structure of Ross (1976) arbitrage pricing theory.

1.4.3. Conditional and Unconditional Factor Model

The model in equation (1.8) is a conditional asset pricing model adaptive at time t. Thus, we have the expectation at time t. This conditional pricing model does not always imply the unconditional pricing model. This means that $E_t[m_{t+1}R^e_{t+1} = (1 - \mathbf{b_t^T}\tilde{\mathbf{f}}_{\mathbf{t+1}})R^e_{t+1}] = 0$, with $\mathbf{b_t}$ deterministic at time t, does not always lead us to find a constant risk coefficient vector $\mathbf{b}$ such that $E[m_{t+1}R^e_{t+1} = (1 - \mathbf{b^T}\tilde{\mathbf{f}}_{\mathbf{t+1}})R^e_{t+1}] = 0$. Or, in other words $E_t[R^e_{t+1}] = cov_t(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}) \cdot \mathbf{b_t}$ does not imply $E[R^e_{t+1}] = cov(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}) \cdot \mathbf{b}$. In fact, we can start by taking the unconditional expectation of the conditional model:

$$\begin{aligned} E[R^e_{t+1}] = E[E_t[R^e_{t+1}]] &= E[cov_t(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}) \cdot \mathbf{b_t}] \\ &= E[cov_t(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}})]E[\mathbf{b_t}] + cov(cov_t(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}), \mathbf{1^T} \cdot \mathbf{b_t}) \\ &= cov(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}})\mathbf{b} + cov(cov_t(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}), \mathbf{1^T} \cdot \mathbf{b_t}) \end{aligned}$$

The unconditional model holds if the covariance term $cov(cov_t(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}), \mathbf{1^T} \cdot \mathbf{b_t}) = 0$. This will happen when $\mathbf{b_t} = \mathbf{b}$, or when the risk coefficients on factors are not time-varying.

In brief, if $m_{t+1} = 1 - \mathbf{b^T}\tilde{\mathbf{f}}_{\mathbf{t+1}}$, then both conditional and unconditional factor pricing models hold.

1.4.4. Characteristics-Based Factor Models

So far, we have shown that the expected return can be presented as a factor structure, where every factor can be a return. In general, we can also present the SDF as follows:

$$m_{t+1} = 1 - \mathbf{b_t^T}\left[\mathbf{r^e_{t+1}} - E_t[\mathbf{r^e_{t+1}}]\right] \tag{1.9}$$

where $\mathbf{r^e_t}$ is the vector of the net excess return of all assets. As we showed above, $\mathbf{b_t}$ can be time-varying and will lead to the conditional pricing model. To test this conditional model, we have to gather asset return data and estimate $\mathbf{b_t}$ at every point in time. In practice, this is not always an easy procedure. In fact, we often can only test the unconditional model when we can use all the historical data. Hence, we need a better way to both use the unconditional model and allow $\mathbf{b_t}$ to be time-varying. The most prominent way is to use instrumental variables.

Instrumental variables in this context will be characteristics of assets. These characteristics can be accounting variables or other information about the assets. For example, some well-known characteristics are: book to market (BM), size, momentum, profitability, investment, liquidity, skewness of asset return, volatility of asset return, etc. In the literature, we gather these characteristics into a matrix called $\mathbf{Z_t}$. $\mathbf{Z_t}$ has a dimension of $N \times K$, where N is the number of assets and K is the number of characteristics. Hence, every column in $\mathbf{Z_t}$ contains the information of one characteristic of N assets. Each column in $\mathbf{Z_t}$ will be demeaned to have every column's sum equal to zero. This matrix is adaptive at time t, and we can easily gather the information of $\mathbf{Z_t}$. We use $\mathbf{Z_t}$ instrumentally in such a way as

$$\mathbf{b_t} = f(\mathbf{Z_t}) \cdot \mathbf{b} \tag{1.10}$$

where $f(\mathbf{Z_t})$ is a function that maps $\mathbf{Z_t}$ from $N \times K$ to $N \times L$, and $\mathbf{b}$ is an $L \times 1$ vector. As usual, this transformation will also ensure that every column's sum of $f(\mathbf{Z_t})$ equals zero. Hence, we can write $f(\mathbf{Z_t})^T \cdot \mathbf{r^e_{t+1}} = \mathbf{f_{t+1}}$, where each element in the vector $\mathbf{f_{t+1}}$ is a return from a net long-short portfolio. Therefore, each element in $\mathbf{f_{t+1}}$ is a factor in the form of a net long-short portfolio return. Then, we can rearrange our SDF as

$$\begin{aligned}
m_{t+1} &= 1 - \mathbf{b^T} \cdot f(\mathbf{Z_t})^T \left[\mathbf{r^e_{t+1}} - E_t[\mathbf{r^e_{t+1}}] \right] \\
&= 1 - \mathbf{b^T} \cdot \left[\mathbf{f_{t+1}} - E_t[\mathbf{f_{t+1}}] \right] \\
m_{t+1} &= 1 - \mathbf{b^T} \tilde{\mathbf{f}}_{\mathbf{t+1}} \qquad (1.11)
\end{aligned}$$

Following the discussion in the previous subsection, we can derive an unconditional factor model : $E[R^e_{t+1}] = cov(R^e_{t+1}, \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}) \cdot \mathbf{b} = \boldsymbol{\beta}^T \boldsymbol{\lambda}$. This unconditional model is more suitable for testing.

I will show some of the most-used functions of $f(\mathbf{Z_t})$. First, it is easy to let $f(\mathbf{Z_t}) = \mathbf{Z_t}$, hence $\mathbf{Z_t}^T \cdot \mathbf{r^e_{t+1}} = \mathbf{f_{t+1}}$. This is the well-known approach that can be found, for example, in Kozak, Nagel, and Santosh (2018, 2019); Kelly, Pruitt, and Su (2019); etc.

Second, there is another family of functions called sorting functions based on characteristics. Indeed, every characteristic column in $\mathbf{Z_t}$ will be sorted from low to high values, then assigned a weight. For example, we can sort the column of the characteristic $Size$ into deciles, and assign positive weight to the top decile of assets and negative weight to the bottom decile. Other assets in other deciles will get zero weight. The weight assignment also ensures that the total weight of this column is equal to zero. This process will give us a factor of net long-short portfolio return. This process is named "uni-variate portfolio sorting."

The sorting process can be more complicated by applying double sorting independently, or dependently to create different factors.[7] These approaches are quite common in the literature, for example, Fama and French (1993) three-factors model, Carhart (1997) momentum factor, Fama and French (2015) five factor model, Hou, Xue, and Zhang (2015) q-factor model, etc.

With these cited models, not only does the use of instrumental variables help us to derive the unconditional model while still letting $\mathbf{b_t}$ be time-varying, but they also help us to reduce the dimension of the problem. Usually, with the family of sorting functions, $K = L$, and K is usually small, below 10. Hence, we will come up with a reduced form factor-model with only a few important factors. The low number of factors will make any testing process simpler when one can just run a regression of the excess return on a reduced set of factors.

Indeed, a low number of K, the number of characteristics, also helps us to better estimate $\mathbf{b}$, and hence better estimate the risk premium ($var[\mathbf{f_t}] \cdot \mathbf{b}$). Note that $E[\mathbf{f_{t+1}} m_{t+1}] = E[\mathbf{f_{t+1}}(1 - \tilde{\mathbf{f}}^{\mathbf{T}}_{\mathbf{t+1}}\mathbf{b})] = 0$. Subsequently, $E[\mathbf{f_{t+1}}] = \Sigma \mathbf{b}$. Hence, $\mathbf{b} = \mathbf{\Sigma^{-1}} E[\mathbf{f_{t+1}}]$. The matrix $\mathbf{\Sigma^{-1}}$ is the inverted matrix of the covariance matrix of factors. The estimation of the covariance matrix will become very inaccurate when the number of factors increase substantially. When K is big, the estimation of the inverted this matrix is also computational costly and inaccurate.

Therefore, traditional approach tends to reduced the number of factors to alleviate these statistical, and computational obstacles. One can always raise the concern of omitted factors if we use just a handful of factors. Certainly, these reduced-factors models cannot explain all the variation in the cross-sectional asset returns, but empirical works show that the Fama French three and five factors, and the q-factor model are quite

[7] Bali, Engle, and Murray (2016) book gives a very good description of these sorting procedures.

powerful. These models can explain hundred of anomalies that CAPM cannot explain.

Since there is no strict restriction on which characteristics should be included in the characteristics matrix, the literature has witnessed a boom inn discovering factors. Harvey, Liu, and Zhu (2016), and Harvey (2017) document that this “factor zoo” consists of more than 300 factors. These factors easily prompt false positive discovery. Hence, there is an emerging literature on how to tame this factor zoo. The problem is sometimes enlarged to not just 300 but thousands factors and predictors of asset returns.

Recent approaches are to include all possible factors, then we will use different techniques from machine learning, such as LASSO, ridge regression, deep learning, tree regression, etc. to not just reduce the dimension of the problem but also account for non-linear relationships and interaction between factors. In our context, the function $f(\mathbf{Z_t})$ can be changed to $f(\mathbf{Z_t}, \mathbf{x_t})$ to map a dimension of $\mathbf{Z_t}$ of $N \times K$ to $N \times L$ when L can be up to thousands or hundred of thousands. This is possible when we let the characteristics interact. It is also possible when we introduce $\mathbf{x_t}$ as exogenous variables, for example, macro-economic variables. The macro variables can interact with other characteristics to form thousands of new characteristics. Some recent works in this field can be found in Kozak et al. (2018, 2019); Harvey et al. (2016); Harvey (2017); Kelly et al. (2019), Kozak (2019); Feng, Polson, and Xu (2018); Chen, Pelger, and Zhu (2019); Gu, Kelly, and Xiu (2020), etc.

My last two papers in this thesis are based on the characteristics-based factor-model framework. In the third paper, I investigate the stock’s stable liquidity as a characteristic that is priced in the market. I show that the factors built on liquidity and liquidity-volatility have risk premia. These risk premia cannot be explained by other common factors in the literature. In the fourth paper, I use more than 100 characteristics to build factors. These characteristics help investors to form their believes about the stock future returns. In the next section, I will discuss in details the contribution of these papers.

1.5. Summary of Four Papers

In the above sections, I went through the major theories and methodologies that my thesis is based on. In the last section of this chapter, I will give a summary of each paper. I will also discuss how these papers fit in with these above theories and methodologies. Finally, I highlight the main contribution of these papers.

1.5.1. Paper 1: "A Simple but Powerful Measure of Market Efficiency"

As discussed above, we can use the RW process to describe the EMH in the short term. Based on this finding, several works try to develop a test for market efficiency. In the 1970s, the common view was that the market is efficient. With Lo and MacKinlay (1988, 1989), we recognized that the market can be inefficient. Then, we came to the adaptive market hypothesis of Lo (2004, 2017). The main idea is that the market is not a static object. The market can sometimes be efficient and sometimes not. Based on this observation, we develop a new measure (AMIM) to quantify the level of market efficiency.

The measure is first based on the auto-correlation test on return, but then is improved to be robust against insignificant auto-correlation. This paper is a joint work between me and Thomas Leirvik and is published in *Finance Research Letter*. Our measure is not the first measure of the level of market efficiency. However, we show that our new measure offers new advantages compared with existing measures in the literature. Indeed, our measure is easy to compute. AMIM is also easy to use when giving a common ground for comparing the efficiency level across assets and time. We also show that AMIM's type 1 and type 2 error are quite small. Finally, AMIM is very sensitive to major events in the market. This confirms the AMH in the sense that market efficiency changes over time due to major changes in the market. In this paper, we also illustrate how to apply AMIM to different data.

1.5.2. Paper 2: "Efficiency in the Markets of Crypto-currencies"

This is a joint-work with Thomas Leirvik and is accepted for publishing in *Finance Research Letters*. In this paper, we apply the AMIM measure that we derive in the first paper to investigate the efficiency level in the crypto-currency markets. The crypto-currency markets have attracted a lot of attention from the public in recent years. We use AMIM to check the level of efficiency of five major crypto-currencies (Bitcoin, Litecoin, Ethereum, EOS, and Ripple) from 2013 to 2019.

We find that the efficiency level in these markets varies over-time. Before 2017, crypto-currency markets were mainly inefficient. This corroborates other results on the matter. However, in the period 2017-2019, crypto-currencies markets became more efficient. On average, in our sample, Litecoin is the most efficient currency, while Ripple is the least efficient currency.

1.5.3. Paper 3: "Stable Liquidity"

In the first two papers, I investigate the expected return and market efficiency in the short term. In papers 3 and 4, I investigate the expected return in the longer-run. As discussed above, bearing systematic risk is likely the main reward for a higher expected return in the future. In the third paper, I investigate the effect of liquidity and liquidity-volatility on the expected stock returns from 1993 to 2013. This is a joint-working-paper with Thomas Leirvik.

We define the effective spread as a measure of liquidity because it represents the true trading cost. We also argue that investors care not only about the level of liquidity but also about its variation. In fact, investors will likely prefer stocks that have a more stable liquidity level. By stable liquidity, we refer stocks that have both a high average level of liquidity and a low liquidity-volatility.

Therefore, we investigate if the level of liquidity and the liquidity-volatility are priced in the cross-sectional stock returns. We find that both the level and the volatility of liquidity are priced in the cross-sectional stock returns, even when we control for other well-known factors such as book to market, momentum, β market exposure, and reversal. Stocks with a high level of illiquidity (high spread) will have a higher expected return than the lower stable spread. We also find that a stock has a high liquidity-volatility level has a higher expected return than a stock with a lower liquidity-volatility level.

The positive relationship between the liquidity-volatility and the expected return is robust in our paper. This result contradicts Chordia, Subrahmanyam, and Anshuman (2001)'s results. Indeed, Chordia et al. (2001) document a negative relationship between the expected return and the liquidity-volatility, which they call the liquidity-volatility puzzle. We show the opposite results. There are two reasons for that: first, we use a more direct and precise measure of liquidity-volatility (the volatility of the effective spread), and second, Chordia et al. (2001)'s measure of liquidity-volatility is a noisy measure that has weak predicting power on the expected return.

1.5.4. Paper 4: "Mispricing Characteristics"

In this paper, I investigate the economy when investors have heterogeneous views on future pay-off and return. I derive a consumption equilibrium model and investigate the feature of the model. In my model, investors have biased views (or sentiment) about future pay-off and return. The aggregate biased belief of all investors in the market creates a sudden rise or fall in the stock's price. Hence, a correction will likely to occur in the future. This correction creates the dispersion in the cross-sectional expected stock return.

I also use characteristics as instrumental variables in my model. In that sense, investors will use stock's characteristics as information to form their biased view about future expected returns. Using characteristics as instrumental variables can create a factor structure for the expected return as we discussed in the Characteristics-Based Factor Model section. In that sense, the expected correction return due to mispricing today can be described as a characteristics based factor model. Therefore, the model in this paper builds a bridge that connects the characteristics-based model and the sentiment model.

In the paper, I use empirical analysis to confirm the model's implications. I also investigate which characteristics-based factors are important to the expected return. I use 101 characteristics with the data sample of US stocks from 1980 to 2018. It turns out that only a handful of characteristics matter both statistically and economically (4 to 11 out of 101).

Using ex-ante information, I also form a mispricing factor. This mispricing factor shorts the over-valued and longs the under-valued. The long-short portfolio generates an α from 0.7% to 1.38% a month after controlling for other common factors and mispricing measures. I also find that some well-known anomalies, and risk premium in the literature (β anomaly, size, illiquidity, maximum return, momentum, liquidity-volatility, return on equity) is only represented in either under-priced or over-priced stocks but not in the entire cross-section of stocks.

Finally, I show that the mispricing sentiment model in this paper can also be represented as a risk model. Indeed, I prove that the correction return due to mispricing can be represented as the covariance risk between return and a latent factor containing information of the aggregate bias belief in the market. Hence, the risk presentation and the sentiment presentation become just two sides of the same coin in this model framework.

REFERENCES

Back, Kerry, 2010, *Asset pricing and portfolio choice theory* (Oxford University Press).

Bali, Turan G, Robert F Engle, and Scott Murray, 2016, *Empirical asset pricing: the cross section of stock returns* (John Wiley & Sons).

Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2015, Equilibrium fast trading, *Journal of Financial Economics* 116, 292–313.

Björk, Tomas, 2009, *Arbitrage theory in continuous time* (Oxford University Press).

Carhart, Mark M, 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Chen, Luyang, Markus Pelger, and Jason Zhu, 2019, Deep learning in asset pricing, *Available at SSRN 3350138* .

Chordia, Tarun, Avanidhar Subrahmanyam, and V Ravi Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3–32.

Cochrane, John H, 2009, *Asset Pricing:(Revised Edition)* (Princeton university press).

Fama, E, and K French, 1993, Common Risk Factors in the Returns of Stocks and Bonds, *Journal of Financial Economics* 33, 3–56.

Fama, EF, 1970, Efficient capital markets: A review of theory and empirical work, *The Journal of Finance* 25, 383–423.

Fama, Eugene, and Marshall Blume, 1966, Filter rules and stock market trading, *Journal of Business, Security Prices: A Supplement* 39.

Fama, Eugene F, 1965, The behavior of stock-market prices, *The Journal of Business* 38, 34–105.

Fama, Eugene F, 1991, Efficient capital markets: Ii, *The Journal of Finance* 46, 1575–1617.

Fama, Eugene F, and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Feng, Guanhao, Nicholas G Polson, and Jianeng Xu, 2018, Deep learning in asset pricing, *ArXiv Preprint ArXiv:1805.01104* .

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *The Review of Financial Studies* 33, 2223–2273.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Harvey, Campbell R, 2017, Presidential address: The scientific outlook in financial economics, *The Journal of Finance* 72, 1399–1440.

Harvey, Campbell R, Yan Liu, and Heqing Zhu, 2016, . . . and the cross-section of expected returns, *The Review of Financial Studies* 29, 5–68.

Hausman, Daniel M, 2003, Philosophy of economics .

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *The Review of Financial Studies* 28, 650–705.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica: Journal of the Econometric Society* 263–291.

Kelly, Bryan T, Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* .

Kirilenko, Andrei A, Albert S Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2015, The flash crash: The impact of high frequency trading on an electronic market, *Available at SSRN 1686004* .

Kozak, Serhiy, 2019, Kernel trick for the cross section, *Available at SSRN 3307895* .

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *The Journal of Finance* 73, 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2019, Shrinking the cross-section, *Journal of Financial Economics* .

Lichtenstein, Sarah, and Paul Slovic, 1971, Reversals of preference between bids and choices in gambling decisions., *Journal of Experimental Psychology* 89, 46.

Lintner, J, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

Lo, Andrew W, 2004, The adaptive markets hypothesis, *The Journal of Portfolio Management* 30, 15–29.

Lo, Andrew W, 2017, *Adaptive markets: Financial evolution at the speed of thought* (Princeton University Press).

Lo, Andrew W, and A Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *The Review of Financial Studies* 1, 41–66.

Lo, Andrew W, and A Craig MacKinlay, 1989, The size and power of the variance ratio test in finite samples: A monte carlo investigation, *Journal of Econometrics* 40, 203–238.

Miller, Rena S, and Gary Shorter, 2016, High frequency trading: Overview of recent developments, *Washington: Congressional Research Service* .

Mossin, J, 1966, Equilibrium in a capital asset market, *Econometrica* 34, 768–783.

Ross, Stephen A, 1976, The arbitrage pricing of capital asset pricing1, *Journal of Economic Theory* 13, 341–360.

Sharpe, William F, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *The Journal of Finance* 19, 425–442.

Simon, Herbert A, 1955, A behavioral model of rational choice, *The Quarterly Journal of Economics* 99–118.

Treynor, Jack L, 1961, Market value, time, and risk, *Time, and Risk (August 8, 1961)* .

Treynor, Jack L, 1962, Jack treynor's' toward a theory of market value of risky assets', *Available at SSRN 628187* .

Weintraub, E Roy, 2002, Neoclassical economics, *The Concise Encyclopedia of Economics* 1, 1.

Chapter II

A simple but powerful measure of Market Efficiency

Published paper with Thomas Leirvik in Finance Research Letters 29 (2019): 141-151.

Finance Research Letters 29 (2019) 141–151

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

A simple but powerful measure of market efficiency

Vu Le Tran*, Thomas Leirvik

Nord University Business School, Universitetsaléen 11, Bodø 8049 Norway

ARTICLE INFO

Keywords:
Market efficiency
Adaptive market hypothesis

ABSTRACT

We construct a simple measure to quantify the level of market efficiency. We apply this measure to investigate the level of market efficiency and analyze its variation over time. The main contribution of the new measure is that it makes it easy to compare market efficiency across assets, time, regions, and data frequencies. We find that markets are often efficient, but can be significantly inefficient over longer periods. Our empirical results indicates that in many periods of major economic events, financial markets becomes less efficient. This corroborates earlier results on market efficiency, and simplifies interpretation and comparisons.

1. Introduction

In this paper, we derive a new measure to quantify the level of market efficiency. We use the term *Adjusted Market Inefficiency Magnitude* (*AMIM*). The *AMIM* increases as market efficiency decreases, and decreases as market efficiency increases. The maximum level of *AMIM* is 1, which implies a highly inefficient market. There is no lower boundary, but if the measure produces a negative number, the market is assumed to be efficient. This makes interpretation very simple: a positive *AMIM* indicates an inefficient market, and a negative *AMIM* indicates an efficient market. The *AMIM* is very easy to compute, and is computationally inexpensive. This implies that comparisons over time, assets, asset classes, and geographical regions are carried out with ease. We show that it has several advantages over existing measures of market efficiency, and are able to detect periods of the economy that is known for much uncertainty about prices and values.

The Efficient Market Hypothesis (EMH) is based on the idea that an asset's price should reflect all relevant information and that economic agents, and thus the financial markets, are rational. The EMH was introduced in the seminal paper by Fama (1970). Market efficiency is usually described in three levels: weak, semi-strong, and strong form. There is a vast amount of literature in the field to test if markets are efficient in both weak form and strong form, see for example the papers by Fama (1970), Fama (1991), and Yen and Lee (2008) for more details. The consequence of market efficiency is that future prices, and returns, are random and should not be possible to predict. This randomness can be modeled by a random walk, which is a mathematical description of a stochastic process where each increment is random and independent of earlier increments. That stock returns are not totally random has been shown in several empirical papers, for example (Reinganum, 1983; De Bondt and Thaler, 1985; Jegadeesh and Titman, 1993). In this paper we derive an estimator of the level of market efficiency. The measure, *AMIM*, makes it possible to quantify how efficient the market is, and determine whether it should be classified as inefficient or efficient.

According to Lo (2004), markets are not always rational, nor optimal, but sometimes heuristic, and emotional. Lo (2004, 2017) proposes a concept called *Adaptive Market Hypothesis* (AMH), and suggests that we can use evolutionary models for studying the markets. The assumption is that financial markets are not static objects, but adapt to a changing environment via simple heuristics. If so, market efficiency is also dynamic and can change over time. Tests of the AMH for different markets, assets, and frequencies of

* Corresponding author.
E-mail addresses: vu.l.tran@nord.no (V.L. Tran), thomas.leirvik@nord.no (T. Leirvik).

https://doi.org/10.1016/j.frl.2019.03.004
Received 30 April 2018; Received in revised form 30 January 2019; Accepted 1 March 2019
Available online 02 March 2019
1544-6123/

observations, has been carried out by numerous authors, see for example (Urquhart and Hudson, 2013; Urquhart and McGroarty, 2014; 2016), and the references therein. These papers largely conclude that markets are adaptive, thus market efficiency varies over time. This variation of market efficiency in time has been studied thoroughly, see, for example the paper by Lim and Brooks (2011) for a survey on the matter. To study the variation of market efficiency, one often applies moving windows that consists of daily, weekly or monthly data. These windows are then applied to investigate whether prices behave according to a random walk process, see for example (Choi, 1999; Kim and Shamsuddin, 2008). The usual focus is on answering the question: *Are markets always efficient?* We enlarge the discussion of market efficiency by addressing the questions: *How large is the inefficiency level, and how does it vary over time?* We find that market efficiency is indeed varying over time. The *AMIM*-measure captures periods with much uncertainty and hence difficulties in determining what information should be incorporated in prices.

The papers by Ito et al. (2014, 2016) and Noda (2016) derive a measure to quantify market efficiency. The authors investigate the variation of the efficiency level by estimating the auto-correlation in stock monthly return through a time variant auto-regressive (TV-AR) model, which is designed to capture a set of auto-correlation coefficients in each observation in time. In particular, Noda (2016)'s measure aim to capture the time-varying degree of market efficiency (*TIME*). From the return auto-correlation coefficients, *TIME* captures the time-varying degree of market efficiency, and hence aims to measure the inefficiency level of the market. In this paper, we extend the results in two main areas: first, our model does not depend on the frequency of the data in the sample, whereas the existing models are more suitable for low frequency data. Second, we do not choose the number of autocorrelation lags in advance. Indeed, Ito et al. (2016, 2014) and Noda (2016) model is challenging to apply to high frequency data when the number of estimations can be up to millions each day.

The *AMIM* is derived using both the autocorrelation coefficients of a time series of stock returns and the confidence intervals of these coefficients. The measure is thus robust against insignificant autocorrelation. Specifically, we start with a measure that we denote market inefficiency magnitude (*MIM*), and use its confidence interval to adjust the *MIM* to produce *AMIM*. Therefore, our measure is also a type of test of market efficiency. *MIM* builds upon (Noda, 2016)'s measure called time-varying degree of market efficiency (*TIME*). The *TIME* measure has many novel contributions, but is relatively difficult to compute. Moreover, a more serious drawback is that the denominator of *TIME* can be close to zero, equal to zero, or even change sign. Thus, there is a discontinuity that is likely to occur, and which will make inference troublesome. *MIM* addresses both drawbacks of *TIME*, and offers a simple solution to make analysis of market efficiency very simple. Our approach also provides a quicker way to find the confidence interval of the inefficiency magnitude. The main reason for this is because our confidence intervals can easily be computed from the sample under investigation. This is a major contribution, as comparable measures, for example the one applied in Noda (2016), relies on simulations and bootstrapping. Finally, *AMIM* helps us to easily compare the inefficiency magnitude between different assets, across different point in time.

We construct *AMIM* through 4 steps. The first step is to estimate the auto-correlation coefficients in the return series through standard regression methods, and then standardize them. The second steps is to derive a raw measure of the market inefficiency magnitude (*MIM*). The third step derives the confidence interval under the null hypothesis of efficient markets for *MIM*. In the final step, we adjust *MIM* with its confidence interval to derive our measure *AMIM*. The measure is a convenient test score of market inefficiency level; $AMIM > 0$ means that the market is significantly inefficient while $AMIM < 0$ means that we cannot reject the null hypothesis of efficient markets. By design, the inefficiency magnitude will positively correlated with *AMIM*.

Second, our measure can be tested easily on samples that consist of many different assets over time by computing a unique set of confidence intervals. This also decrease the computational burden, especially when analyzing big data. In contrast, even though *TIME* is a very good measure of the inefficiency magnitude, its design implies that it can only be tested sample by sample.

Third, our measure is uniformly continuous, meaning that there are no discontinuities, in all levels of auto-correlation. This is very important when conducting inference and interpreting the results. An inherent challenge with the measure applied in Ito et al. (2016, 2014) is that it is a fraction with sums of the autocorrelation coefficients included. Not only can the denominator be zero, but also the summation can make positive auto-correlation canceling out negative-correlation. In this paper, we address these issues and compute the absolute values of the auto-correlation coefficients before making any summation.

To test the performance of our measure, we estimate the *AMIM* for some US stock market indexes. Concerning robustness test and compatibility with other market efficiency estimators, we apply *AMIM* to the same dataset studied in Noda (2016). We show that *AMIM* can capture similar result in Noda (2016). We also do a simulation to check the power and the size of our test *AMIM* and make some computations to show that *AMIM* is very reasonable in terms of producing estimates that corresponds to financial theory. The results also show that market efficiency varies considerably over time, and reflects major economic events. This is also very important, as, according to the AMH, one can expect market efficiency to change over time. From an economic point of view, the changes should not be completely random, but be linked to economic conditions around the world. Indeed, for the sample under consideration in this paper, we can disentangle major economic events from the movement in the *AMIM*. *AMIM* also provides the main results of Noda (2016) for the Japanese markets.

2. Model and estimation methods

According to Fama (1970) stock prices should, under the Efficient Market Hypothesis (EMH), reflect all relevant information in the market. Therefore, if we are in period t, the return in the next period $t + 1$ should not be predictable. Hence, following the EMH, an auto-regressive process $AR(q)$ of returns (r_t) on its own lags cannot explain the dynamics of returns over time. For example, if EMH holds, then the $AR(q)$ model

 Finance Research Letters 29 (2019) 141–151

$$r_t = \alpha + \beta_1 r_{t-1} + \beta_2 r_{t-2} + \cdots + \beta_q r_{t-q} + \varepsilon_t \tag{1}$$

should have coefficients ($\beta_1, \beta_2, \dots, \beta_q$) that are all close to zero, or at least insignificantly different to zero. If the EMH does not hold, the β coefficients are (significantly) non-zero. Lo (2004) used the first auto-regressive coefficient to characterize the inefficiency level. If there are more lags with significant coefficients, then there is even more evidence against a strongly efficient market. Our aim is to construct a measure that takes the auto-correlation coefficients into account. The Adjusted Market Inefficiency Magnitude, $AMIM_t$, is constructed following four steps:

2.1. Normalizing the auto-correlation coefficients (Step 1)

Let $\hat{\beta}$ be a column vector which contains the estimated coefficients $(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_q)'$ from Eq. (1). $\hat{\beta}$ will be asymptotically distributed as follows:

$$\hat{\beta} \sim N(\beta, \Sigma). \tag{2}$$

Here β is the unknown true beta vector, i.e. the vector of auto-correlation coefficients. Σ is the asymptotic co-variance matrix of the estimated $\hat{\beta}$ vector, which can be separated into two triangular matrices by Cholesky decomposition as: $\Sigma = LL'$. The estimated coefficients have different standard errors and can be correlated. Therefore, we standardize the $\hat{\beta}$ vector multiplying it by the inverse of the triangular matrix L. Thus, the standardized beta is given as:

$$\hat{\beta}^{standard} = L^{-1}\hat{\beta}. \tag{3}$$

Under the null hypothesis that market is efficient ($\beta = 0$), then $\hat{\beta}^{standard}$ should be normally distributed as follows:

$$\hat{\beta}^{standard} \sim N(0, I) \tag{4}$$

Where I is an identity matrix. Therefore, the normalizing process in Eq. (3) helps us in two ways. First, by multiplying $L^{-1}\hat{\beta}$, it makes each component in $\hat{\beta}^{standard}$ independent. Second, the standardized coefficients are very convenient for testing any measures constructed from $\hat{\beta}^{standard}$.

2.2. The magnitude of market inefficiency (Step 2)

In this section, we construct the unadjusted, or raw, measure of market inefficiency. To calculate the inefficiency level, we first construct the Magnitude Market Inefficiency, MIM_t, as follows:

$$MIM_t = \frac{\sum_{j=1}^{q} |\hat{\beta}_{j,t}^{standard}|}{1 + \sum_{j=1}^{q} |\hat{\beta}_{j,t}^{standard}|} \tag{5}$$

As we are interested in violations of the assumption that the auto-regressive coefficients are zero, we use the absolute value to eliminate the sign effect. MIM_t is the Market Inefficiency Magnitude at time t whereas $\hat{\beta}_{j,t}^{standard}$ is the j^{th} auto-correlation coefficient in Eq. (1) after standardization. Following the above construction, the auto-correlation $\sum_{j=1}^{q} |\hat{\beta}_{j,t}^{standard}|$ is positively related to the Market Inefficiency Magnitude. The variation of MIM_t is smooth from 0 (very efficient market) to almost 1 (inefficient market). So when comparing two stocks, the one having a higher MIM_t will be more affected by the past than the one having lower MIM_t.

Noda (2016) has the similar approach of using the auto-regressive coefficients to compute the Market Inefficiency Magnitude, though different formula for market efficiency, $TIME_t$, is applied. $TIME_t$ is given as:

$$TIME_t = \left| \frac{\sum_{j=1}^{q} \hat{\beta}_{j,t}}{1 - \sum_{j=1}^{q} \hat{\beta}_{j,t}} \right| \tag{6}$$

Eq. (6) uses the non-standardized coefficients from Eq. (1). Hence, it will be inconsistent when $\sum_{j=1}^{q} \hat{\beta}_{j,t} \in [0, 1]$. Indeed (Ito et al., 2014)'s ratio will converge to ∞ when $\sum_{j=1}^{q} \hat{\beta}_{j,t}$ is around 1. An interesting implication of this is that sometimes markets are oddly more efficient when the auto-correlation level is high (i.e. $\sum_{j=1}^{q} \hat{\beta}_{j,t} = -2$) than when the auto-correlation level is low (i.e. $\sum_{j=1}^{q} \hat{\beta}_{j,t} = 0.6$). To see this, we get $TIME_t = \left|\frac{-2}{1-(-2)}\right| \approx 0.667$, indicating a level of market efficiency of 0.667. In a more efficient case, for example when the sum of auto-correlation coefficients equals 0.6, the $TIME$ measure is $TIME_t = |\frac{0.6}{1-0.6}| = 1.5$ which indicates a less efficient market even though the sum of autocorrelation coefficients having a very different meaning.

Although a simple example, the consequence is that $TIME_t$ cannot be used for inference in this case. Furthermore, a large-scale analysis using $TIME_t$ will need to be accompanied by an individual check of each case to make sure that the result makes economic sense. Furthermore, Eq. (6) sums all the raw coefficients. This can also make the measure inconsistent. For example, if we have two

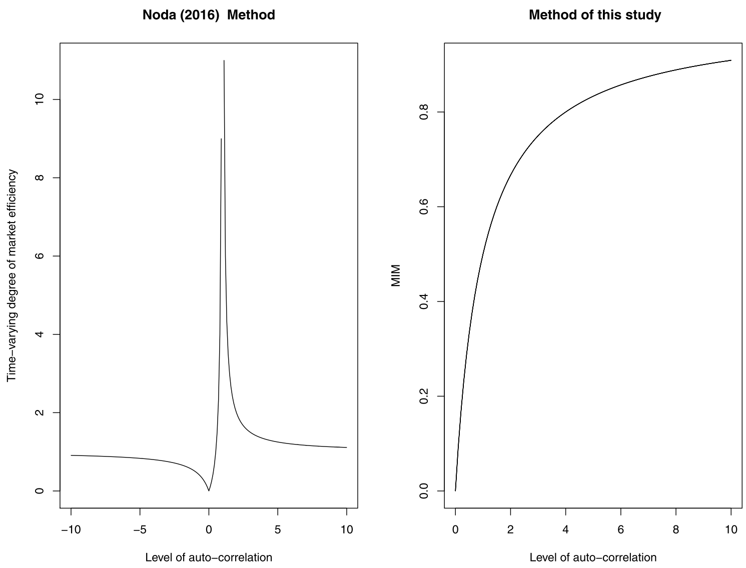


Fig. 1. Market Inefficiency Magnitude MIM_t with auto-correlation level $\sum_{j=1}^{q}\hat{\beta}_{j,t}$ (Noda (2016) methods), and $\sum_{j=1}^{q}|\hat{\beta}_{j,t}^{standard}|$ (our methods).

auto-correlation coefficients, $\beta_1 = -0.5$, $\beta_2 = 0.5$, the Noda (2016)'s measure will equal to zero which indicates an efficient market. Therefore, we take the absolute value of $\hat{\beta}_{j,t}^{standard}$ before sum up all the coefficients. This process will help to avoid the eliminating effect between positive and negative coefficients.

Moreover, we use the standardized $\hat{\beta}$ coefficients before compute MIM_t. This step will be crucial to compute the confidence interval in the following step. By standardizing the auto-correlation coefficients, we can derive a unique set of confidence intervals for *MIM* under the null hypothesis of efficient markets, thus reducing the computational burden.

Fig. 1 illustrates the difference of the two methods of computing the market efficiency level where the level of auto correlation is $\sum_{j=1}^{q}\hat{\beta}_{j,t}$ in Noda (2016) and $\sum_{j=1}^{q}|\hat{\beta}_{j,t}^{standard}|$ in our method.

Second, we use a non-overlapping window method to compute the auto-correlation coefficients of each time interval[1]. Ito et al. (2014, 2016) and Noda (2016) used a *time-varying auto-regressive model (TV-AR)* to compute the auto-correlation coefficients. The latter model will give a set of coefficients for each observation in time. For example, if we have 1 observation/second then Ito et al. (2014, 2016)'s model will have $3600 \cdot q$ coefficients for each hour, where q is a constant number of lags. In brief, the total number of coefficients is equal to the number of observation times q. Thus, this can be computationally intensive when the number of observations increases, in particular using high frequency data.

2.3. Building confidence intervals (step 3)

The Market Inefficiency Magnitude is by construction between 0 and 1. However, the raw value of the MIM_t can give us a false impression of the market efficiency. Due to an absolute process to eliminate the sign effect in step 2, the MIM_t will be, by construction, positively correlated with the number of lags in the Eq. (1). Even for markets that are very efficient, it is likely that MIM_t can be very high. This is undesirable.

To correct for this, we compute the confidence interval of MIM_t. To get the confidence interval under the null hypothesis of efficient markets ($\hat{\beta}_{j,t}^{standard} = 0$), we can either use convergence of random variables, or simulations. The former approach is quite tricky with a function as *MIM* while the latter is more reasonable.

Because all $\hat{\beta}_{j,t}^{standard}$ are standard normal, knowing the number of lags in Eq. (1), we can identify the confidence interval of MIM_t under the null hypothesis through simulation. We first simulate 100 000 observations for each $\hat{\beta}_{j,t}^{standard}$ following a standard normal

[1] In the empirical part, we also apply the rolling window methods and having the same results qualitatively.

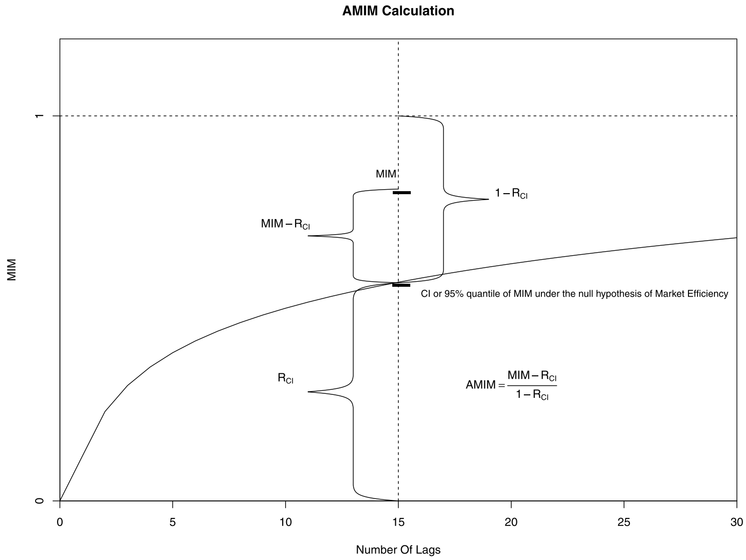


Fig. 2. Illustration of Adjusted Market Inefficiency Magnitude Calculation ($AMIM = \frac{MIM - R_{CI}}{1 - R_{CI}}$). The curvature line is the upper bound of 95% confidence interval of MIM under the null hypothesis of efficient markets.

distribution. Based on $\hat{\beta}_{j,t}^{standard}$, we compute *MIM*. For each number of lags we have 100 000 observations of *MIM* under the null hypothesis of market efficiency. After that, we find the 95th percentile of *MIM*. Because *MIM* is only varying in [0, 1), the interval between this 95th percentile and 0 is the 95% confidence interval of *MIM* under the null hypothesis. This confidence interval is thus unique for each number of lags. This again gives us a table of confidence intervals (CI) which can be used in a different context. See Table 3 in the appendix for details of the computation of the confidence intervals.

2.4. *The adjusted market inefficiency magnitude (step 4)*

In this section, we derive the adjusted market inefficiency magnitude, *AMIM*. From the previous step, we know the 95% confidence interval of MIM_t under the null hypothesis of efficient markets. First, we compute the range of the confidence interval, basically the distance between zero and the 95% quantile of *MIM* under the null hypothesis of market efficiency. We then adjust the *MIM* by first subtracting the range of the CI from the *MIM*; $MIM - R_{CI}$, then we divide this distance between MIM_t and R_{CI} with the distance between the theoretical maximum value of *MIM*, which is one, and R_{CI}. Mathematically, this is given as:

$$AMIM_t = \frac{MIM_t - R_{CI}}{1 - R_{CI}}. \tag{7}$$

Because $MIM_t < 1$, the estimates of $AMIM_t$ and R_{CI} are always less than one as well. MIM_t is also always greater or equal to zero, which implies that *AMIM* can be negative. In fact, whenever $AMIM > 0$ the market is inefficient, whereas when $AMIM < 0$ the market is efficient. Fig. 2 gives an illustration of *AMIM* formula. Loosely speaking, *AMIM* only stresses on the inefficient part of *MIM*, which passes the null hypothesis CI. The $AMIM_t$ is thus more reliable than MIM_t because it penalizes the mechanical variation of MIM_t due to high number of lags in $\hat{\beta}_{j,t}^{standard}$. We divide ($MIM - R_{CI}$) by the difference between one and R_{CI} to give a common ground for comparison between stocks. Indeed, different stocks will have different *MIM* values with again different R_{CI}'s at different point in time. Adjusting for R_{CI} gives us the same comparison criteria for all assets. By this construction when $AMIM_t < 0$, we cannot reject the null hypothesis that markets are efficient. If $AMIM_t > 0$ we can say, markets are significantly inefficient. Markets are more inefficient when $AMIM_t$ increases.

3. The size and power of *AMIM*

To investigate the size and power of *AMIM*, we carry out a Monte Carlo simulation. We simulate an AR(1) model for returns where ρ is the auto-correlation coefficient. We set the return innovation as normally distributed with mean 0.03 % a day and a daily standard deviation of 1 %. This is the typical long-run mean and standard deviation for the S&P 500 index. We set ρ to be (0, ± 0.3,

Table 1
Simulation of AMIM. We simulate an AR(1) model for returns where ρ is the auto-correlation coefficient. We set the return innovation as normal distributed with mean 0.03 % a day and a daily standard deviation of 1 %. For each ρ value, we simulate 100 000 batches. Each batch consists 200 observations. Each batch gives 1 value of *AMIM*. N is number of observations of *AMIM*. Q is the quantile of the *AMIM* distribution.

ρ	N	Q0.01	Q0.05	Q0.1	Q0.25	Q0.5	Q0.75	Q0.9	Q0.95	Q0.99
0	100 000	−0.22	−0.143	−0.061	0	0	0	0.052	0.127	0.235
0.3	100 000	0.104	0.206	0.251	0.316	0.392	0.467	0.525	0.555	0.603
0.5	100 000	0.296	0.359	0.391	0.445	0.506	0.561	0.604	0.627	0.667
0.7	100 000	0.345	0.402	0.432	0.481	0.534	0.582	0.622	0.644	0.681
0.9	100 000	0.322	0.38	0.41	0.458	0.51	0.558	0.599	0.621	0.66
−0.3	100 000	0.125	0.217	0.26	0.324	0.399	0.474	0.534	0.563	0.609
−0.5	100 000	0.297	0.359	0.391	0.446	0.506	0.562	0.605	0.629	0.671
−0.7	100 000	0.341	0.399	0.429	0.48	0.533	0.582	0.622	0.644	0.683
−0.9	100 000	0.319	0.377	0.407	0.457	0.509	0.558	0.599	0.622	0.661

± 0.5, ± 0.7, ± 0.9) respectively. With ρ = 0, we have the efficient market case. For this case, we expect *AMIM* to be smaller or equal 0. For other cases ρ ≠ *0*, we expect *AMIM* to be greater than 0. For each ρ value, we simulate 100 000 batches. Each batch consists 200 observations. Each batch gives one value of *AMIM*. We end up with 100 000 *AMIM* values for each ρ value. Table 1 gives the estimates of the simulation. Here, *N* is the number of observations of *AMIM*, and *Q* is the quantile of the *AMIM* distribution.

In the Efficient Market case $\rho = 0$, hence *AMIM* is supposed to be smaller than, or equal to zero. 90% of our simulated *AMIM* is smaller than 0.052. Subsequently, 95% *AMIM* is smaller than 0.127. These results show that in the efficient market case, we do not make a huge mistake with *AMIM*. Even in case we make a mistake, the error is not big because *AMIM* is wrongly positive, but it is small and very close to zero. Therefore, if we wrongly conclude that markets are inefficient instead of efficient, the wrong inefficient level is also small which makes less harm.

In the inefficient market case, for example $\rho = 0.3$, *AMIM* is supposed to be greater than zero. 99% of our simulated *AMIM* is greater than 0. The logic goes so on with different ρ values. These results show that when market is not efficient, the *AMIM* measure makes very little to no mistake of discovering it. In summary our *AMIM* measure performs quite well in simulation to detect the size and power of the test. Of course, our drawback is not considering all the available alternative hypotheses. Such a full analysis is out of scope of this research and is worth investigating in future research.

4. Data

To have a better comparison with Noda (2016)'s measure, we use the same dataset as they applied. The dataset is price levels for the Tokyo Stock Price Index (TOPIX) and the Tokyo Stock Exchange Second Section Stock Price Index (TSE2). Stocks in TOPIX and TSE2 indexes are different. The TOPIX index has a much higher market capitalization and trading volume than TSE2. The data source is Bloomberg. We compute the log return r_t from the daily prices, p_t, thus $r_t = \log(p_t/p_{t-1})$.

We also investigate the efficiency level of US stock markets for both small stocks and large stocks. We use S&P 500 index as a proxy for large stocks. For US small stocks, we first sort stocks belonging to AMEX, NYSE, NASDAQ exchanges from CRSP database into decile portfolios based on market capitalization, then taking the value weighted return of the first portfolio as an index portfolio for small stocks. We call this portfolio as CAP1.[2]

We use these datasets to compute *AMIM*. The frequency of the data is daily, and cover the period from 1962 through 2017. We compute the auto-correlation coefficients ($\hat{\beta}$) for each index in each year using all daily return. To identify the number of lags of each time interval, we use Akaike information criterion (AIC)[3] We required that each year having at least 200 observations to run the regression. For each year, the model estimates one Market Inefficiency Magnitude, MIM_t, and one Adjusted Market Inefficiency Magnitude, $AMIM_t$.

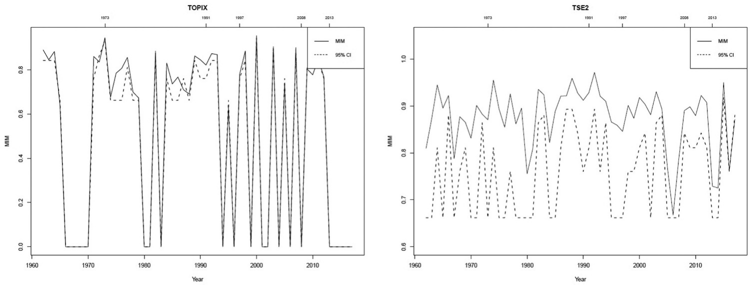
We also estimate *AMIM* using rolling-window data. In detail, we estimate *AMIM* daily using one year rolling-window data. Then every day we will have one value of *AMIM*.

5. Empirical results

Fig. 3 shows the value of *MIM* over year using non-overlapping window of TOPIX and TSE2. We can spot a clear fluctuation of *MIM* over time. However, it is hard to say that TOPIX is more efficient than TSE2 or the other way around. It is also hard to say which time the markets are more efficient than the other is. As discussed above, at each point in time with each asset we have a different confidence interval value of *MIM*. Therefore, we do not have a same base for comparison. Indeed, we can have a high *MIM* value but

[2] The sorting procedure is done through Wharton Research Data Service called "CRSP Stock File Indexes - Daily Index Built on Market Capitalization" at: https://wrds-web.wharton.upenn.edu/wrds/ds/crsp/indexes_a/mktindex/cap_d.cfm?navId=124

[3] *AMIM* is very flexible in choosing the number of lags in contrast with a fixed number of lags setting in TV-AR methods. *AMIM* does not depend only on AIC or any information criteria. With our construction, it is possible to apply other criteria to select the number of lags in the first step. In our paper, we only use AIC as a decision criterion, because the focus is more on introducing our measure *AMIM*, how to use it, and its feature to reflect major economic events.



(a) TOPIX MIM (b) TSE2 MIM

Fig. 3. MIM of TOPIX and TSE2 indexes. MIM is estimated with non-overlapping window. The solid line is MIM value while the dotted line is the 95% confidence interval under the null hypothesis of market efficiency.

being not significant and vice versa a low *MIM* value but showing a significant inefficiency level.

AMIM solve this issue by adjusting *MIM* with the confidence interval. Now it is enough to compare with the same base line of zero; remember that *AMIM* > 0 implies a significant inefficiency level, and $AMIM_t < 0$ implies an efficient market. Fig. 4 shows the evolution of *AMIM* for TOPIX, TSE2, S&P 500, and CAP1. For the Japan market, we can also confirm the major empirical findings of Noda (2016) with our measure. These are: i) first, market efficiency changes over time with both TOPIX and TSE2; ii) TOPIX has a lower and less volatile inefficiency level (mean (μ) $=0.08$, standard deviation (σ) $= 0.15$) than TSE2 ($\mu = 0.47$, $\sigma = 0.20$); iii) both TOPIX and TSE2 inefficiency level significantly decreases after 2010. Hence, both markets becomes more efficient after 2010. Table 2 gives the summary statistics of *AMIM* of the indexes.

For the US market, we find a similar difference in *AMIM* between S&P 500 and CAP1. The small stock index has a higher mean ($\mu = 0.38$), a higher standard deviation ($\sigma = 0.22$) than the ones of S&P 500 ($\mu = 0.09$, $\sigma = 0.17$). So for both US and Japan markets, large stocks indexes (TOPIX, S&P 500) are more likely efficient. Both TOPIX and S&P 500 have a low median near zero. This means that 50% of time, these indexes are efficient.

In addition, our measure also offers an additional feature. *AMIM* reflects very well important economic events, for example, it increases in periods of economic turbulence or crisis, then decreases after such periods. For Japan market, we can see this pattern with both TOPIX and TSE2 through the Oil-Crisis (1973–74), the bursting of Japanese asset bubble (1991–92), the Asian financial crisis (1997–99), and the financial crisis (2008). *AMIM* also decreases in 2013, which reflects the period of quantitative easing.

For US market, we catch the similar pattern. *AMIM* of both S&P 500 and CAP1 raises in the Oil-Crisis (1973–74) then decreases. *AMIM* raises again in the 1987 crisis and in the 2001 dot-com bubble burst. In these two crises (1987, 2001), *AMIM* of small stocks (CAP1) raises more than *AMIM* of large stock (S&P 500). These two crises could hit small stocks harder than large stocks. However in the financial crisis 2008, *AMIM* of large stocks (S&P 500) experienced a sharp increase and being almost as high as *AMIM* of CAP1. This can be due to the fact that a lot of large stocks (especially financial industry stocks) was smashed very hard during that crisis.

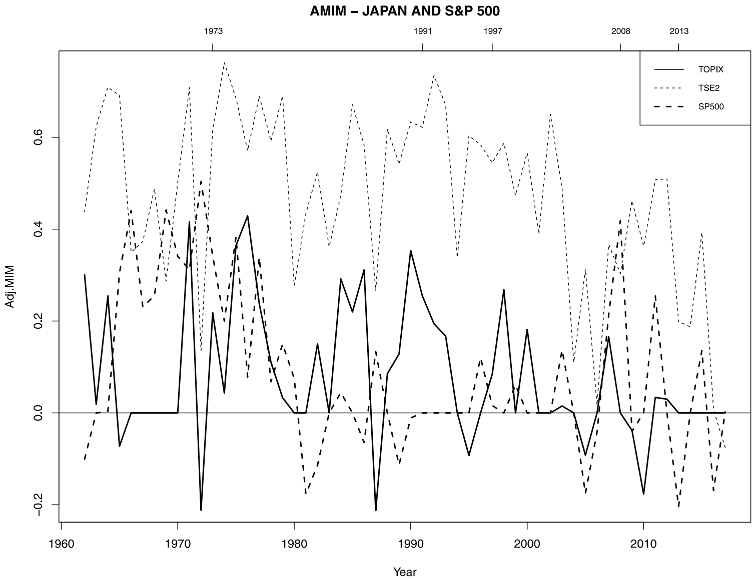
For a robustness check, we also estimate *AMIM* using rolling window. In detail, we estimate *AMIM* daily using a one year rolling window of data. So we will have one *AMIM* value per day. The results of the overlapping window estimates corroborates our earlier results. The summary statistics of *AMIM* using rolling window is in Table 4 in the appendix. For a clearer illustration of the trend of *AMIM*, we calculate a 100-day Moving Average (MA) of *AMIM* and plot it below in Fig. 5. As the figures illustrates, the S&P 500-index indicates an efficient market most of the time from 1980 through 2018, but with some periods of inefficiency. For example, during the oil-crisis in the early 70-ies, and during the more recent financial crisis of 2008-09, the market is inefficient. Smaller stocks are, not surprisingly, less efficient than large stocks, only being significantly efficient over small periods of time.

The fact that AMIM increases in times of crisis, and then decreases afterwards, confirms the results of Lo (2004, 2017) on the hypothesis of adaptive markets; financial markets are not always efficient, nor always inefficient but changing overtime. These results also indicate that our measure is a valid measure of the level of market efficiency.

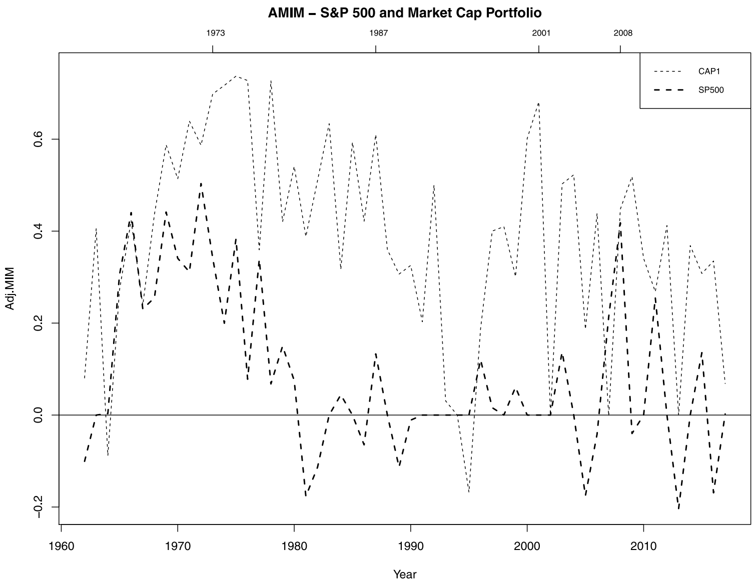
One caveat of our study is that we do not establish the causal effect between different factors (i.e. inflation, interest rates, unemployment rates, etc.) to market efficiency. Another caveat is that we do not have a full horse race between all efficiency measures in all markets. We recognize them as interesting subjects for further research. We herein focus more on developing *AMIM*, explaining how to use it, and showing its important features. Hence, we consider *AMIM* as a good alternative measure of efficiency that is easy to use, light in computation, alongside with other measures such as the variance ratio, *TIME*, etc.

6. Conclusion

This study derives a measure for the level of market efficiency, named Adjusted Market Inefficiency Magnitude *AMIM*. The measure is easy to be applied and computed via four steps. *AMIM* improves two challenges of the measures derived in Noda (2016). First, *AMIM* demands less computational effort and can easily be interpreted. Second, *AMIM* also provides a better foundation for

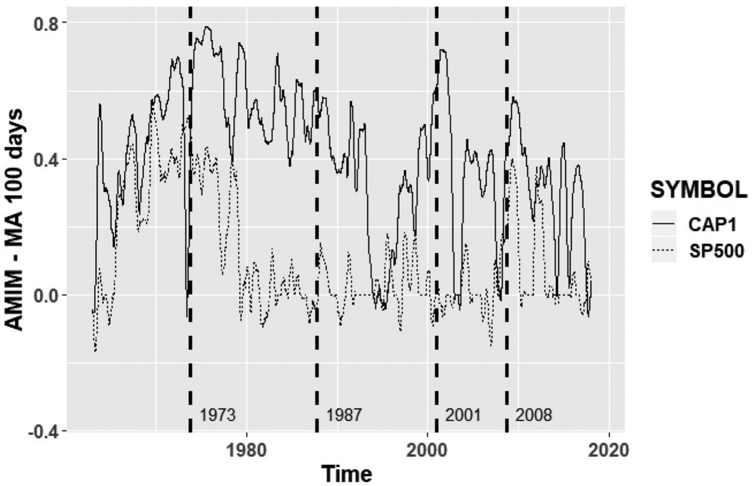


(a) AMIM of TOPIX, TSE2, S&P 500

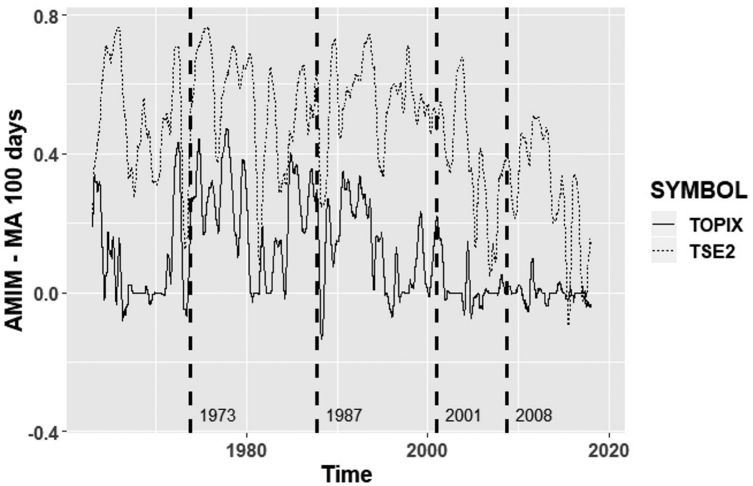


(b) AMIM of S&P 500 and CAP 1

Fig. 4. Adjusted Market Inefficiency Magnitude *AMIM*, using non-overlapping window, of TOPIX, TSE2, S&P 500, and CAP1. CAP1 is the portfolio containing 10% of small stocks on NYSE, AMEX, and NASDAQ exchanges. The data range is from 1962 to 2017.



(a) Moving Average 100-day of AMIM of CAP1, and S&P 500



(b) Moving Average 100-day AMIM of TOPIX and TSE2

Fig. 5. Moving average (MA) 100-day of *AMIM* of TOPIX and TSE2, S&P 500, and CAP1. *AMIM* is estimated daily using a 1 year rolling window data. CAP1 is the portfolio containing 10% of small stocks on NYSE, AMEX, and NASDAQ exchanges. The data range is from 1962 to 2017.

comparison the inefficiency level between different assets in different time. Applying our measure to the same dataset as in Noda (2016), we can confirm the major findings of Noda (2016)'s work. In addition, our measure also reflects very well major economic events in the US and Japanese economies. These empirical results shows that market efficiency is not constant over time, assets, or regions, which corroborates the Adaptive Market Hypothesis of Lo (2004, 2017).

Table 2
Summary statistic of AMIM measure using non-overlapping window for TOPIX, TSE2, S&P500, and CAP1 where CAP1 is the portfolio containing 10% of small stocks on NYSE, AMEX, and NASDAQ exchanges. The data range is from 1962 to 2017.

Index	n	mean	sd	median	min	max	skew	kurtosis	Q0.25	Q0.5	Q0.75
TOPIX	56	0.08	0.15	0.01	−0.21	0.43	0.52	−0.24	0.00	0.01	0.18
TSE2	56	0.47	0.20	0.50	−0.08	0.76	−0.81	0.07	0.36	0.50	0.62
CAP1	56	0.38	0.22	0.41	−0.17	0.74	−0.48	−0.38	0.27	0.41	0.53
SP500	56	0.09	0.17	0.00	−0.20	0.50	0.69	−0.33	0.00	0.00	0.20

Appendix A. Confidence Interval of *MIM* under the null hypothesis of Market Efficiency.

Table 3
The upper bound 95% Confidence Intervals (CI) under the null hypothesis of efficiency market for *MIM*, where: $MIM_t = \frac{\sum_{j=1}^{q} |\hat{\beta}_{j,t}^{standard}|}{1 + \sum_{j=1}^{q} |\hat{\beta}_{j,t}^{standard}|}$ with different number of lags in the Eq. (1). The lower bound of this interval is 0. To compute CI, we first simulate 100 000 observations for each $\hat{\beta}_{j,t}^{standard}$ following a standard normal distribution. Then based on these $\hat{\beta}_{j,t}^{standard}$, we compute *MIM*. For each number of lags we have 100 000 observations *MIM* under the null hypothesis of market efficiency. After that, we find the 95^{th}% quantile of *MIM*. So the interval between this 95^{th}% quantile and 0 is the 95% confidence interval of MIM under the null hypothesis.

Number of Lags	95% CI	Number of Lags	95% CI	Number of Lags	95% CI
1	0.6618747	16	0.9441565	31	0.9682057
2	0.7604725	17	0.9468745	32	0.9690909
3	0.81105	18	0.9493466	33	0.9699065
4	0.8423915	19	0.9516287	34	0.9706732
5	0.864342	20	0.9536607	35	0.9714273
6	0.8806096	21	0.9555671	36	0.9721273
7	0.8932211	22	0.9572666	37	0.9727706
8	0.903343	23	0.9588263	38	0.9734095
9	0.9115645	24	0.9603012	39	0.9740274
10	0.9184596	25	0.9616615	40	0.9745969
11	0.9243942	26	0.9629152	41	0.9751548
12	0.9293885	27	0.9641263	42	0.9756867
13	0.9338437	28	0.9652404	43	0.9761773
14	0.9376448	29	0.9662761	44	0.976653
15	0.9411291	30	0.9672589	45	0.9771318

Appendix B. Summary statistics of AMIM using rolling window

Table 4
Summary statistic of *AMIM* measure for TOPIX, TSE2, S&P500, and CAP1 where CAP1 is the portfolio containing 10% of small stocks on NYSE, AMEX, and NASDAQ exchanges. The data range is from 1962 to 2017. *AMIM* is estimated daily using a 1-year rolling window.

Index	N	mean	sd	median	min	max	skew	kurtosis	Q0.25	Q0.5	Q0.75
CAP1	13,898	0.411	0.218	0.432	−0.227	0.796	−0.579	−0.118	0.296	0.432	0.570
SP500	13,898	0.106	0.180	0.000	−0.366	0.610	0.738	−0.306	0.000	0.000	0.223
TOPIX	13,632	0.104	0.154	0.005	−0.395	0.550	0.542	−0.471	0.000	0.005	0.222
TSE2	13,619	0.468	0.194	0.495	−0.223	0.798	−0.683	0.158	0.339	0.495	0.620

References

Choi, I., 1999. Testing the random walk hypothesis for real exchange rates. J. Appl. Econom. 14, 293–308.
De Bondt, W.F.M., Thaler, R., 1985. Does the stock market overreact? J. Finance 40, 793–805. https://doi.org/10.1111/j.1540-6261.1985.tb05004.x.
Fama, E., 1970. Efficient capital markets: a review of theory and empirical work. J. Finance 25, 383–423.
Fama, E.F., 1991. Efficient capital markets: ii. J. Finance 46, 1575–1617.
Ito, M., Noda, A., Wada, T., 2014. International stock market efficiency: a non-bayesian time-varying model approach. Appl. Econ. 46, 2744–2754.
Ito, M., Noda, A., Wada, T., 2016. The evolution of stock market efficiency in the us: a non-bayesian time-varying model approach. Appl. Econ. 48, 621–635.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. J. Finance 48, 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x.
Kim, J.H., Shamsuddin, A., 2008. Are asian stock markets efficient? evidence from new multiple variance ratio tests. J. Empir. Finance 15, 518–532.
Lim, K.-P., Brooks, R., 2011. The evolution of stock market efficiency over time: a survey of the empirical literature. J. Econ. Surv. 25, 69–108.
Lo, A.W., 2004. The adaptive markets hypothesis. J. Portfolio Manag. 30, 15–29.
Lo, A.W., 2017. Adaptive Markets: Financial Evolution at the Speed of Thought. Princeton University Press, Princeton, NJ.
Noda, A., 2016. A test of the adaptive market hypothesis using a time-varying ar model in japan. Finance Research Letters 17, 66–71.
Reinganum, M.R., 1983. The anomalous stock market behavior of small firms in january. empirical tests for tax-loss selling effects. J. Financ Econ. 12, 89–104. https://doi.org/10.1016/0304-405X(83)90029-6.
Urquhart, A., Hudson, R., 2013. Efficient or adaptive markets? evidence from major stock markets using very long run historic data. Int. Rev. Financ. Anal. 28, 130–142.
Urquhart, A., McGroarty, F., 2014. Calendar effects, market conditions and the adaptive market hypothesis: evidence from long-run us data. Int. Rev. Financ. Anal. 35, 154–166.
Urquhart, A., McGroarty, F., 2016. Are stock markets really efficient? Evidence of the adaptive market hypothesis. Int. Rev. Financ. Anal. 47, 39–49.
Yen, G., Lee, C.-f., 2008. Efficient market hypothesis (emh): past, present and future. Rev. Pacif. Basin Financ. Market. Polic. 11, 305–329.

Chapter III

Efficiency in the Markets of Crypto-Currencies

Accepted to be published in Finance Research Journal.

ARTICLE IN PRESS

Finance Research Letters xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Efficiency in the markets of crypto-currencies

Vu Le Tran*, Thomas Leirvik

Nord University Business School, Universitetsaléen 11, Bodø 8049, Norway

ARTICLE INFO

Keywords:
Market efficiency
Adaptive market hypothesis
Crypto-currencies

ABSTRACT

We show that the level of market-efficiency in the five largest cryptocurrencies is highly time-varying. Specifically, before 2017, cryptocurrency-markets are mostly inefficient. This corroborates recent results on the matter. However, the cryptocurrency-markets become more efficient over time in the period 2017–2019. This contradicts other, more recent, results on the matter. One reason is that we apply a longer sample than previous studies. Another important reason is that we apply a robust measure of efficiency, being directly able to determine if the efficiency is significant or not. On average, Litecoin is the most efficient cryptocurrency, and Ripple being the least efficient cryptocurrency.

1. Introduction

In this paper we analyze the market efficiency of the five largest cryptocurrencies.[1] We find that the markets for these five currencies are currently mostly efficient, but has been significantly inefficient in the past. The market for cryptocurrencies has received much attention the last three years, both from regulators, the public, and traders. The trading volume in the largest such currencies has grown exponentially, and with this increase in liquidity, the prices increased as well. We analyze how this increased interest and trading volume affects the efficiency of the cryptocurrency markets. We find that the efficiency increases significantly during 2017 and remains efficient midway through 2018. Our results also shows that these markets are sensitive to various events. For example, in June 2016 the DAO hack leads to a separation of Ethereum into Ethereum and Ethereum Classic, which caused increased uncertainty in the market. The level of efficiency dropped significantly following this event. The markets also stayed highly inefficient for several months, before stabilizing at weakly inefficient in late 2016 and early 2017.

Market efficiency has received much attention since Fama (1970) and the follow-up paper by Fama (1991). In the papers, the Efficient Market Hypothesis (EMH) is introduced and and the author sorts the efficiency of the market into three segments: strong efficiency, semi-strong efficiency, and weak efficiency. Furthermore, the author argues that financial markets are, to a large extent, strongly efficient. This implies that all available information is reflected in the price of the security. The challenge was for a long time to quantify market efficiency. Lo and MacKinlay (1989) proposed a method to test if markets are efficient or not. Furthermore, Lo (2004) proposed an alternative to the static view of market efficiency, proposing that the efficiency evolves over time. This is denoted the *Adaptive Market Hypothesis, (AMH)*. The papers by Urquhart and Hudson (2013), Ito et al. (2014), Noda (2016), Ito et al. (2016), and Urquhart and McGroarty (2016) investigates the market efficiency with methods derived with the AMH in mind. Furthermore, Chu et al. (2019) investigates the AMH for the two largest cryptocurrencies, and find evidence that supports the hypothesis of a time-varying market efficiency. However, some specific measures of market efficiency has potential challenges. For example, the efficiency

* Corresponding author.
E-mail addresses: vu.l.tran@nord.no (V.L. Tran), thomas.leirvik@nord.no (T. Leirvik).

[1] The size is measured by market capitalization as of Feb 28th, 2019. The currencies includes Bitcoin (BTC), Ethereum(ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). The market capitalization varies substantially over time, so other currencies might be larger in other periods.

https://doi.org/10.1016/j.frl.2019.101382
Received 9 September 2019; Received in revised form 15 November 2019; Accepted 25 November 2019

estimator can be given as a fraction, where the denominator can be zero, or fluctuate between positive, zero, and negative. This can lead to discontinuities in the estimates, yielding unreliable estimates, and implies crucial challenges for applications in testing market efficiency. We apply a novel measure of the level of market efficiency, derived in Tran and Leirvik (2019). This measure allows for any values for both nominator and denominator. The estimator is continuous for all input parameter values.

The cryptocurrency market, and in particular the market for Bitcoin, is found to be largely inefficient, see, for example, Urquhart (2016); Vidal-Tomás and Ibañez (2018); Jiang et al. (2018); Wei (2018); Hu et al. (2019); Caporale et al. (2018); Zargar and Kumar (2019); Al-Yahyaee et al. (2018), and Nan and Kaizoji (2019). However, the cryptocurrency market can be efficient in certain periods as well, see, for example, Kristoufek (2018); Kristoufek and Vosvrda (2019), or a power transformation of Bitcoin return can be weakly efficient Nadarajah and Chu (2017). Other studies such as (Omane-Adjepong and Alagidede, 2019; Omane-Adjepong et al., 2019; Sifat et al., 2019; Antonakakis et al., 2019; Katsiampa et al., 2019) also show that cryptocurrencies are strongly interlinked reflecting by volatility spill-over, volatility co-movement, lead-lag effect, market co-movement. The systematic risks involved in such markets are also thoroughly investigated in Corbet et al. (2019). One reason for risks and inefficiencies can be that the markets has been difficult to trade in, and hence liquidity has been relatively low compared to other markets. The ease of trading one cryptocurrency can be significantly different from the ease of trading another such currency, thus the liquidity in various such currencies varies substantially, see Phillip et al. (2018). Liquidity and market efficiency is closely related, and different markets show different levels of liquidity, see for example Amihud (2002); Chordia et al. (2008); Leirvik et al. (2017); Wei (2018); Brauneis and Mestel (2018), and de la Horra et al. (2019). In this paper we show that the level of market efficiency is varying over both time and individual currencies. In particular, we show that the Bitcoin market is largely inefficient until early 2017. In contrast with other cited studies, we find that Bitcoin becomes significantly efficient after 2017. The other currencies under investigation shows similar time-varying patterns.

2. Data

The markets for cryptocurrencies is relatively new, and our sample covers the period April 29th, 2013 through February 28th, 2019. Bitcoin, Litecoin and Ripple enter the sample from the very beginning (2013). Some of the currencies has been developed after April 29th, 2013 (Ethereum from 2015; EOS from 2017), but has shown to rely on a solid technology compared to other currencies, and has quickly become some of the largest currencies by market capitalization. We apply freely available data from Coinmarketcap.com, and import all available data via a statistical package named “crypto” in the software R. The data is at daily frequency which contains open, high, low, close prices, volume, and market capitalization.[2] Table 1 shows the descriptive statistics for the simple returns of the five currencies we analyze. We use simple returns because log-returns might give unreliable estimates for assets with extremely high volatility. In fact, for our sample the minimum daily log-return is −130.2%.[3] This is clearly not economically sound. To eliminate the chance of using uneconomic reasonable estimates for returns, we exclusively use simple returns as inputs to our calculations.

Not surprisingly, there is a significant variation between the various cryptocurrencies. The differences seems to be very heterogeneous. Fig. 1 shows time series plots of the normalized prices (Panel a) and trading volume in Billions USD of the five currencies (Panel b) during the last 3 years. As is evident from those illustrations, the price and volume increases tremendously at the end of 2017, with significant variation over time. This exceptional rise in prices and the corresponding high volatility has attracted much attention in media. We will analyze whether the prices can be considered efficient, and whether the level of efficiency varies over time.

3. Methodology

To estimate the level of efficiency, we apply a recently derived method to quantify the level of market efficiency, see Tran and Leirvik (2019). In this paper, the authors derive a measure for the level of Adjusted Market Inefficiency Magnitude (AMIM). In short, to compute the AMIM, we start with representing the returns of a currency as

$$r_{i,t} = \beta_0 + \beta_1 \cdot r_{i,t-1} + \beta_2 \cdot r_{i,t-2} + \cdots + \beta_q \cdot r_{i,t-q} + \varepsilon_{i,t}. \quad (1)$$

If markets are efficient then the coefficients ($\beta_1, \beta_2, ..., \beta_q$) should be zero, or at least insignificantly different to zero. If not, the β coefficients are (significantly) non-zero. Lo (2004) used the first auto-regressive coefficient to characterize the inefficiency level. One can argue that the Efficient Market Hypothesis (EMH) is based on a random walk or martingale dynamics of the price or the log-price. This has a direct implication in which the future price differences and log-price differences (log-returns) cannot be predicted. In this study, we mainly use the simple return, not log-returns. That can potentially be problematic. However, it turns out that if price follows a RW process then the future simple return cannot be predicted either.[4] Hence a regression of simple return on its lag should

[2] By the end of Feb.2019, the market capitalization of the top 5 crypto-currencies in billions USD are: BTC (67.70), ETH (14.37), XRP (13.03), EOS (3.21), LTC (2.81). Just after the top 5 are Bitcoin Cash (2.33 Billions USD) and Tether (2.04 Billions USD).

[3] Reader can find the summary statistics using log return in the appendix.

[4] Consider a RW process: $y_{t+1} = y_t + \varepsilon_{t+1}$. Where ε_{t+1} is a shock at time $t+1$ in the future which cannot be predicted, hence $E[\varepsilon_{t+1}] = E_t[\varepsilon_{t+1}] = 0$, and ε_{t+1} is independent with y_t. The simple return at time t+1 is : $r_{t+1} = \frac{\varepsilon_{t+1}}{y_t}$. We will show that $E[r_{t+1}] = E_t[r_{t+1}] = 0$, which means both the conditional and unconditional expectation of simple returns cannot be predicted. First, it is clear that $E_t[r_{t+1}] = \frac{E_t[\varepsilon_{t+1}]}{y_t} = 0$. Second $E[r_{t+1}] = E[E_t(r_{t+1})] = E[E_t(\varepsilon_{t+1})]. E[1/y_t] + cov(E_t[\varepsilon_{t+1}], 1/y_t)$. With $E[\varepsilon_{t+1}] = E_t[\varepsilon_{t+1}] = 0$, in addition ε_{t+1} is independent with y_t hence we can believe that $cov(E_t[\varepsilon_{t+1}], 1/y_t) = 0$. Therefore $E[r_{t+1}] = 0$.

Table 1
Summary Statistics of daily *simple returns* for the 5 crypto-currencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). The sample is from 29 April 2013 to 28th February 2019. *n* is the number of observations, *mean* is the sample average of the returns, *sd* is the sample standard deviation of the returns, *min* and *max* are the minimum and maximum daily returns, respectively. *skew* is the skewness of returns, *kurtosis* measures the thickness of the tails of the return distributions, and *se* is the standard error of the means.

Crypto	*n*	mean	sd	median	min	max	skew	kurtosis	se
BTC	2132	0.25%	4.36%	0.18%	−23.37%	42.97%	0.50	9.94	0.00
EOS	607	0.70%	11.28%	−0.20%	−31.96%	168.32%	5.99	80.83	0.00
ETH	1301	0.58%	7.29%	−0.09%	−72.80%	51.03%	0.27	13.13	0.00
LTC	2132	0.35%	7.34%	0.00%	−40.19%	129.10%	4.77	65.90	0.00
XRP	2034	0.51%	8.75%	−0.29%	−46.00%	179.37%	6.12	99.47	0.00

yield a non-significant coefficient if the market is efficient.

We estimate Eq. (1) and use the Akaike information criterion (AIC) to choose the number of lags. We denote $\hat{\beta}$ is the vector of these autocorrelation coefficients and $\Sigma = LL'$ is the asymptotic co-variance matrix of the estimated vector. The sum of the absolute value of the autocorrelation coefficients, after being standardized ($\hat{\beta}^{standard} = L^{-1}\hat{\beta}$), are divided by the sum of absolute values of parameter estimates plus one. This measures the Market Inefficiency Magnitude (*MIM*), and is related to the measure applied in Noda (2016) and Ito et al. (2016). Specifically, the *MIM* is given by:

$$MIM_t = \frac{\sum_{j=1}^{q} |\hat{\beta}_{j,t}^{standard}|}{1 + \sum_{j=1}^{q} |\hat{\beta}_{j,t}^{standard}|} \tag{2}$$

As Eq. (2) sums up the standardized auto-regression coefficients of Eq. (1), it should be statistically equal to zero in a strongly efficient market. To reduce the impact of insignificant parameter estimates, we subtract the range of the confidence interval under the null hypothesis of efficient market from the *MIM* and divide by one minus the range of the confidence interval under the null hypothesis of efficient market. We call this the *AMIM*. The measure is thus robust against insignificant autocorrelation. In short, we estimate the *AMIM* for any financial asset price by the estimator

$$AMIM_t = \frac{MIM_t - R_{CI}}{1 - R_{CI}} \tag{3}$$

The R_{CI} is the range of the confidence interval for the *MIM* under the null hypothesis of efficient market. For further explanations and derivations, the reader is encouraged to read the article by Tran and Leirvik (2019).

Because the *MIM* is constrained between zero and one, the $R_{CI} < 1$. Thus, Eqs. (2) and (3) makes sure that both *MIM* and *AMIM* are continuous functions. Accounting for Eq. (3), the $AMIM_t$ cannot be larger than one. It can, however, be zero, or negative. A positive value, i.e. $AMIM_t > 0$, indicates an inefficient market. If $AMIM_t$ is less than zero, i.e. $AMIM_t \leq 0$, then the market is efficient. Hence, the measure is simple to compute, and very easy to interpret. In addition, it is also very easy to use *AMIM* to compare the level of efficiency for different assets in different points in time. Another quality of *AMIM* that we want to exploit is that it reflects very well economic events influencing the assets.

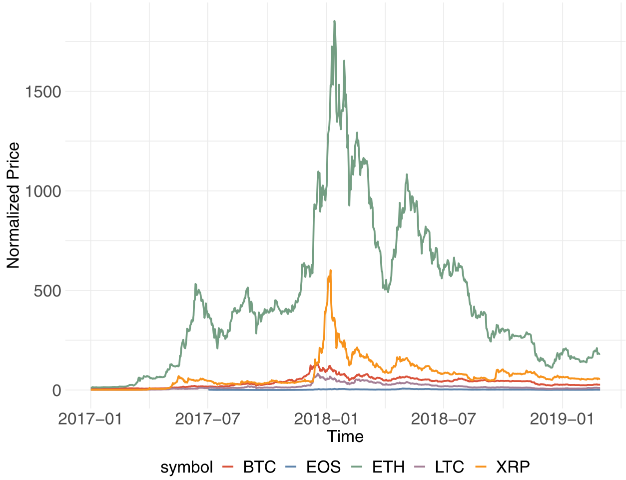
4. Empirical results

Table 2 shows some descriptive statistics of the estimation of Eq. (3). In our study we have applied daily observations of cryptocurrency prices. We largely follow Tran and Leirvik (2019) by estimating the *AMIM* daily using over-lapping window data of 1 year. Tran and Leirvik (2019) shows that the non-overlapping and overlapping window approaches give the same AMIM results.

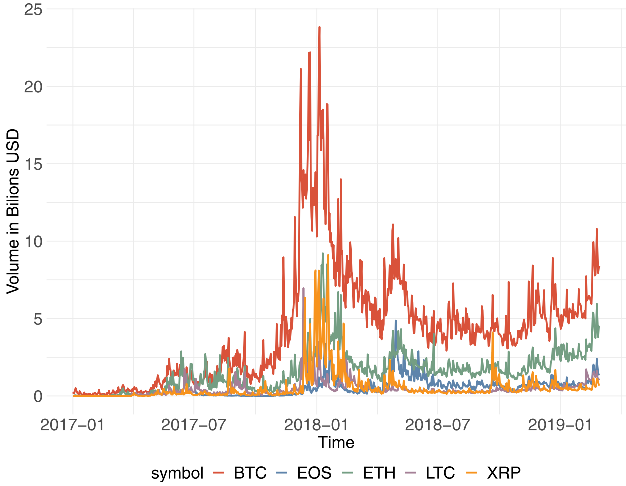
Table 2 shows that on average the prices for BitCoin, EOS, Ethereum, Litecoin, and Ripple are all inefficient ($\overline{AMIM} > 0$). This is statistically significant, as seen by the small size of the standard error. The median, however, is zero for BTC, ETH, and LTC. This indicates that the efficiency of these currencies experiences substantial periods of efficiency.

Fig. 2 shows a time series plot of the *AMIM* for all cryptocurrencies. For a clearer illustration of the trend of AMIM, we calculate a 30-day Moving Average (MA) of AMIM. As one can see from this graph, the level of market efficiency varies substantially over time. In particular, in the early stage of the sample period, the prices of all currencies we analyze are relatively inefficient. Ripple has a bump in efficiency in early 2015, which corresponds to the timing of some large banks announcing that they would apply this currency in new real-time international transactions. Moreover, the creator of *XRP* was fined by US authorities in late 2015 for violations of the Bank Secrecy Act. The price inefficiency saw an immediate spike in the last quarter of that year.

The Bitcoin (BTC) efficiency level also varies substantially, and follows various market events. For example in the end of 2013 and beginning of 2014, the *AMIM* for BTC increases significantly. This is due to a lot of uncertainty in markets when accusations of illegal activities (drug trading, money laundering, etc.) was related to BTC. For example, on early October 2013 the Federal Bureau of Investigation (FBI) arrest Ross Ulbricht and shuts down Silk Road, a black market trading illicit goods using BTC.



(a) Normalized Price



(b) Volume

Fig. 1. Trading Volume in Billions USD and Normalized Price of the top 5 crypto-currencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). We normalize the price for each crypto-currency by dividing the actual price to the first observed price in our data sample.

At the end of Q1 2014, Mt. Gox, one of the largest BTC exchanges, files for bankruptcy protection in Japan. Mt. Gox reports that 744,000 Bitcoin (about 350 millions USD) is stolen.[5] In addition, around the same time, there were also a lot of uncertainty about whether China would ban Bitcoin or not.[6] These events led *AMIM* of BTC spiked again in Q2 2014. In June 2016, the hack on the DAO project of Ethereum casted doubt on crypto-currencies. *AMIM* of Bitcoin did not raise significantly in this period. However, as expected, the *AMIM* of Ethereum shows a significant increase following this event. The *AMIM* is very high but does not last long because Ethereum community responded very quickly to eliminate doubt and uncertainty. Indeed, to save Ethereum from the hack,

[5] Details: See BBC (2014) MtGox Bitcoin exchange files for bankruptcy http://www.bbc.com/news/technology-25233230
[6] Details: See Coin-Desk (2014) Price of Bitcoin Falls Under 500 USD Amid Uncertainty in China https://www.coindesk.com/price-bitcoin-remains-500-amid-china-uncertainty/

Table 2
Summary Statistics of AMIM for 5 crypto-currencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). The AMIM is computed daily with one year overlapping data using *simple return*. The sample is from 29 April 2013 to 28th February 2019. *n* is the number of observations, $\overline{AMIM}$ is the sample average of AMIM, *sd* is the sample standard deviation of AMIM, *min* and *max* are the minimum and maximum AMIM, respectively. *skew* is the skewness of AMIM, *kurtosis* measures the thickness of the tails of the AMIM distributions, and *se* is the standard error of the mean.

Crypto	*n*	$\overline{AMIM}$	sd	median	min	max	skew	kurtosis	se
BTC	1933	0.083	0.132	0.000	−0.349	0.370	0.532	−0.609	0.003
EOS	408	0.086	0.089	0.110	−0.200	0.195	−0.839	−0.201	0.004
ETH	1102	0.061	0.095	0.000	−0.305	0.281	0.700	−0.202	0.003
LTC	1933	0.011	0.117	0.000	−0.324	0.434	0.090	0.458	0.003
XRP	1835	0.251	0.151	0.268	−0.250	0.535	−0.705	0.298	0.004

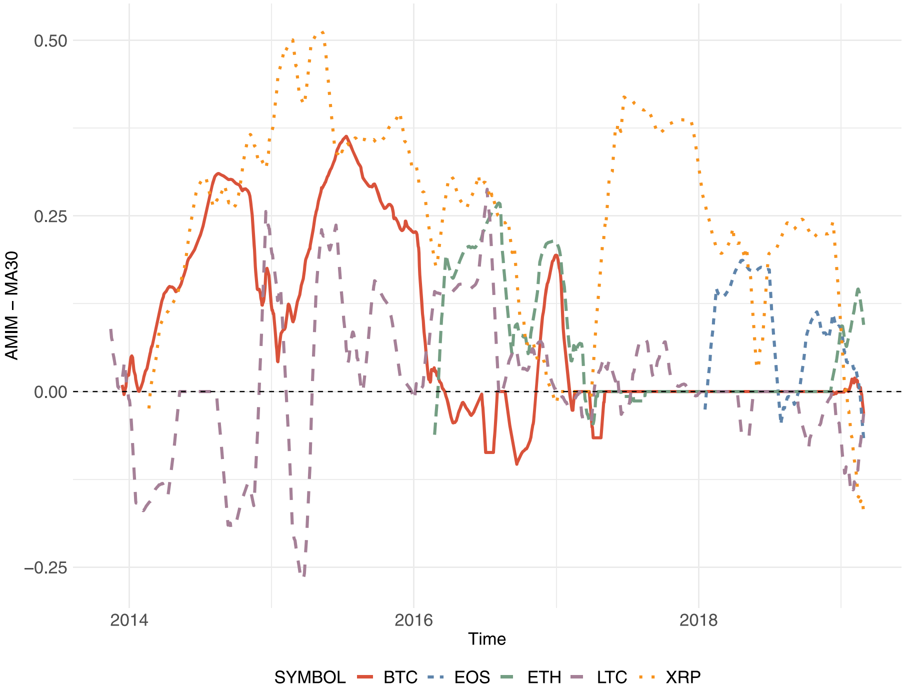


Fig. 2. The time series plot of Moving-Average 30 days (MA30) of AMIM for 5 crypto-currencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). The AMIM is computed daily with 1 year overlapping data using *simple return*.

the Ethereum community decided to "hard-fork", which are upgrades to the programming code that add new rules to the Ethereum software that are incompatible with earlier versions. Basically, the Ethereum community rewrote the ledger, which is the common transaction book, to eliminate all the hacked transactions.[7]

All in all, many of the spikes and drops in market inefficiency, as shown in Fig. 2, can be related to idiosyncratic events. This is particularly true for the time early in the sample period of each cryptocurrency. In the later stages of our sample period, all currencies show a significant improvement in efficiency, as shown by a negative estimate of the *AMIM*. This means that the currencies are significantly efficient. However, it seems that the prices are turning less efficient in the last half of Q1-2018, and the first half of Q2-

[7] This process violates the core motivation of using blockchain technology which prevents any modification of past information. This process requires a lot of computing power and is a collective work. Therefore debates were going at that time. This leads to a separation of Ethereum into Ethereum and Ethereum Classic. The Ethereum Classic does not accept the hard fork, thus accept all the hacked transactions, and being a separate community from Ethereum. See Leising (2017) for more details.

Table 3
Summary Statistics of daily *log returns* for the 5 crypto-currencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). The sample is from 29 April 2013 to 28th February 2019. *n* is the number of observations, *mean* is the sample average of the returns, *sd* is the sample standard deviation of the returns, *min* and *max* are the minimum and maximum daily returns, respectively. *skew* is the skewness of returns, *kurtosis* measures the thickness of the tails of the return distributions, and *se* is the standard error of the mean.

Crypto	*n*	mean	sd	median	min	max	skew	kurtosis	se
BTC	2132	0.16%	4.35%	0.18%	−26.62%	35.75%	−0.19	7.88	0.00
EOS	607	0.21%	9.54%	−0.20%	−38.50%	98.70%	2.18	20.11	0.00
ETH	1301	0.30%	7.70%	−0.09%	−130.21%	41.23%	−3.38	65.28	0.00
LTC	2132	0.11%	6.68%	0.00%	−51.39%	82.90%	1.74	25.18	0.00
XRP	2034	0.20%	7.66%	−0.29%	−61.63%	102.74%	2.01	27.55	0.00

Table 4
Summary Statistics of AMIM for 5 crypto-currencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). The AMIM is computed daily with one year overlapping data using *log return*. The sample is from 29 April 2013 to 28th February 2019. *n* is the number of observations, $\overline{\text{AMIM}}$ is the sample average of AMIM, *sd* is the sample standard deviation of AMIM, *min* and *max* are the minimum and maximum AMIM, respectively. *skew* is the skewness of AMIM, *kurtosis* measures the thickness of the tails of the AMIM distributions, and *se* is the standard error of the mean.

Crypto	*n*	$\overline{\text{AMIM}}$	sd	median	min	max	skew	kurtosis	se
BTC	1933	0.081	0.128	0.000	−0.364	0.376	0.515	−0.412	0.003
EOS	408	0.039	0.071	0.048	−0.167	0.174	−0.479	−0.382	0.004
ETH	1102	0.043	0.071	0.000	−0.311	0.253	0.712	0.695	0.002
LTC	1933	0.039	0.108	0.000	−0.364	0.415	0.123	0.851	0.002
XRP	1835	0.234	0.152	0.253	−0.367	0.545	−0.517	0.737	0.004

2018 but return to the efficient level at the end of 2018. These results corroborates the main idea of the Adaptive Market Hypothesis of Lo (2004), where market efficiency is changing over time, and reacting to events in the market.

We also redo the above exercises with AMIM using log-return. The results are given in the appendix (see Tables 3, 4 and Fig. 3). The estimated AMIM using log-returns is qualitatively the same as the AMIM results using simple returns. However, we should be aware that the log-return series will be mechanically smoother than the simple return series in the positive return region. For example, a simple return of 5% will have a log-returns of about *log*(1.05) = 4.88%, and a simple return of 10% will be about *log*(1.1) = 9.53% in log return. For negative simple returns, the corresponding log-return will be larger in absolute terms. For large negative deviations, log-returns might in fact be less than -100%. For example, a price moving from 50 to 10, yielding a simple return

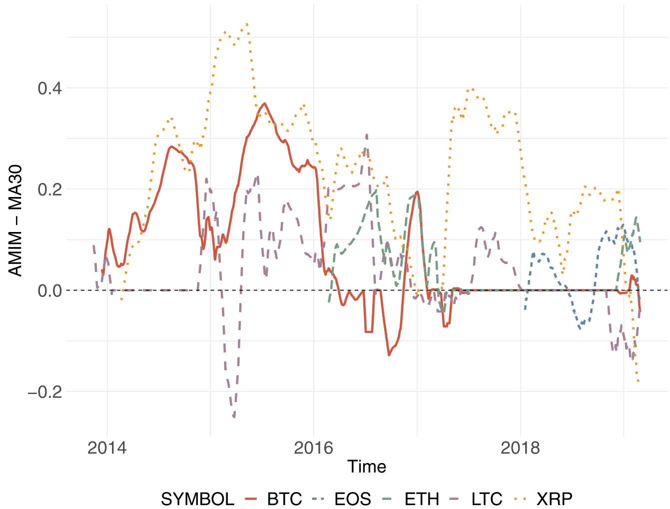


Fig. 3. The time series plot of Moving-Average 30 days (MA30) of AMIM for 5 crypto-currencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), EOS (EOS). The AMIM is computed daily with 1 year overlapping data using *log return*.

of -80%, will give a log-return of about −160%, which is not economically sound. Therefore, we need to be more cautious on the interpretation of any result from log-returns with highly volatile financial assets.

5. Conclusion

In this paper we investigate the inefficiency of the prices of five different cryptocurrencies. Using a rolling window, we find that the prices has been significantly inefficient during our sample. However, there are signs that the efficiency of all cryptocurrencies are improving, with all having significant drop in *AMIM* in the last 6 quarters. These results are consistent with recent research on the topic. The markets for cryptocurrencies are improving at an exceptional pace, with volume improving and becoming less volatile. This invites more research in the near future, both on the topic of efficiency of these markets, but also other aspects, such as for example price-return volatility, liquidity, and the relationship to other assets.

CRediT authorship contribution statement

Vu Le Tran: Conceptualization, Methodology, Software, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Thomas Leirvik:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Validation.

Appendix

References

Al-Yahyaee, K.H., Mensi, W., Yoon, S.-M., 2018. Efficiency, multifractality, and the long-memory property of the Bitcoin market: a comparative analysis with stock, currency, and gold markets. Finance Res. Lett. https://doi.org/10.1016/j.frl.2018.03.017.
Amihud, Y., 2002. Illiquidity and stock returns cross-section and time-series effects. J. Financ. Mark. 5 (1), 31–56.
Antonakakis, N., Chatziantoniou, I., Gabauer, D., 2019. Cryptocurrency market contagion: market uncertainty, market complexity, and dynamic portfolios. J. Int. Financ. Mark. Inst. Money 61, 37–51. https://doi.org/10.1016/j.intfin.2019.02.003.
BBC, 2014. Mtgox Bitcoin exchange files for bankruptcy. http://www.bbc.com/news/technology-25233230[Published: 28 February 2018] [Accessed: 31 May 2018].
Brauneis, A., Mestel, R., 2018. Price discovery of cryptocurrencies: Bitcoin and beyond. Econ Lett 165, 58–61. https://doi.org/10.1016/j.econlet.2018.02.001.
Caporale, G.M., Gil-Alana, L., Plastun, A., 2018. Persistence in the cryptocurrency market. Res. Int. Bus. Finance 46, 141–148. https://doi.org/10.1016/j.ribaf.2018.01.002.
Chordia, T., Roll, R., Subrahmanyam, A., 2008. Liquidity and market efficiency. J. Financ. Econ. 87 (2), 249–268.
Chu, J., Zhang, Y., Chan, S., 2019. The adaptive market hypothesis in the high frequency cryptocurrency market. Int. Rev. Financ. Anal. 64, 221–231. https://doi.org/10.1016/j.irfa.2019.05.008.
Coin-Desk, 2014. Price of Bitcoin falls under 500 USD amid uncertainty in China. https://www.coindesk.com/price-bitcoin-remains-500-amid-china-uncertainty/ [Published: 28 March 2018][Accessed: 31 May 2018].
Corbet, S., Lucey, B., Urquhart, A., Yarovaya, L., 2019. Cryptocurrencies as a financial asset: a systematic analysis. Int. Rev. Financ. Anal. 62, 182–199. https://doi.org/10.1016/j.irfa.2018.09.003.
Fama, E., 1970. Efficient capital markets: a review of theory and empirical work. J. Finance 25, 383–423.
Fama, E.F., 1991. Efficient capital markets: II. J. Finance 46 (5), 1575–1617.
de la Horra, L.P., de la Fuente, G., Perote, J., 2019. The drivers of Bitcoin demand: a short and long-run analysis. Int. Rev. Financ. Anal. 62, 21–34. https://doi.org/10.1016/j.irfa.2019.01.006.
Hu, Y., Valera, H.G.A., Oxley, L., 2019. Market efficiency of the top market-cap cryptocurrencies: further evidence from a panel framework. Finance Res. Lett. 31, 138–145. https://doi.org/10.1016/j.frl.2019.04.012.
Ito, M., Noda, A., Wada, T., 2014. International stock market efficiency: a non-Bayesian time-varying model approach. Appl. Econ. 46 (23), 2744–2754.
Ito, M., Noda, A., Wada, T., 2016. The evolution of stock market efficiency in the US: a non-Bayesian time-varying model approach. Appl. Econ. 48 (7), 621–635.
Jiang, Y., Nie, H., Ruan, W., 2018. Time-varying long-term memory in Bitcoin market. Finance Res. Lett. 25, 280–284. https://doi.org/10.1016/j.frl.2017.12.009.
Katsiampa, P., Corbet, S., Lucey, B., 2019. High frequency volatility co-movements in cryptocurrency markets. J. Int. Financ. Mark. Inst. Money. https://doi.org/10.1016/j.intfin.2019.05.003.
Kristoufek, L., 2018. On Bitcoin markets (in)efficiency and its evolution. Phys. A 503, 257–262. https://doi.org/10.1016/j.physa.2018.02.161.
Kristoufek, L., Vosvrda, M., 2019. Cryptocurrencies market efficiency ranking: not so straightforward. Phys. A 531, 120853. https://doi.org/10.1016/j.physa.2019.04.089.
Leirvik, T., Fiskerstrand, S.R., Fjellvikås, A.B., 2017. Market liquidity and stock returns in the norwegian stock market. Finance Res. Lett. 21, 272–276.
Leising, M., 2017. The ether thief. https://www.bloomberg.com/features/2017-the-ether-thief/Bloomberg Markets. [Published: 13 June 2017,][Accessed: 4 June 2018].
Lo, A.W., 2004. The adaptive markets hypothesis. J. Portf. Manag 30 (5), 15–29.
Lo, A.W., MacKinlay, A.C., 1989. The size and power of the variance ratio test in finite samples: a Monte Carlo investigation. J. Econom. 40 (2), 203–238.
Nadarajah, S., Chu, J., 2017. On the inefficiency of Bitcoin. Econ. Lett. 150, 6–9. https://doi.org/10.1016/j.econlet.2016.10.033.
Nan, Z., Kaizoji, T., 2019. Market efficiency of the Bitcoin exchange rate: weak and semi-strong form tests with the spot, futures and forward foreign exchange rates. Int. Rev. Financ. Anal. 64, 273–281. https://doi.org/10.1016/j.irfa.2019.06.003.
Noda, A., 2016. A test of the adaptive market hypothesis using a time-varying ar model in japan. Finance Res. Lett. 17, 66–71.
Omane-Adjepong, M., Ababio, K.A., Alagidede, I.P., 2019. Time-frequency analysis of behaviourally classified financial asset markets. Res. Int. Bus. Finance 50, 54–69. https://doi.org/10.1016/j.ribaf.2019.04.012.
Omane-Adjepong, M., Alagidede, I.P., 2019. Multiresolution analysis and spillovers of major cryptocurrency markets. Res. Int. Bus. Finance 49, 191–206. https://doi.org/10.1016/j.ribaf.2019.03.003.
Phillip, A., Chan, J., Peiris, S., 2018. A new look at cryptocurrencies. Econ. Lett. 163, 6–9.
Sifat, I.M., Mohamad, A., Shariff, M.S.B.M., 2019. Lead-lag relationship between Bitcoin and Ethereum: evidence from hourly and daily data. Res. Int. Bus. Finance 50, 306–321. https://doi.org/10.1016/j.ribaf.2019.06.012.
Tran, V.L., Leirvik, T., 2019. A simple but powerful measure of market efficiency. Finance Res. Lett. 29, 141–151. https://doi.org/10.1016/j.frl.2019.03.004.

Urquhart, A., 2016. The inefficiency of Bitcoin. Econ. Lett. 148, 80–82.
Urquhart, A., Hudson, R., 2013. Efficient or adaptive markets? evidence from major stock markets using very long run historic data. Int. Rev. Financ. Anal. 28, 130–142.
Urquhart, A., McGroarty, F., 2016. Are stock markets really efficient? Evidence of the adaptive market hypothesis. Int. Rev. Financ. Anal. 47, 39–49.
Vidal-Tomás, D., Ibañez, A., 2018. Semi-strong efficiency of Bitcoin. Finance Res. Lett. https://doi.org/10.1016/j.frl.2018.03.013.
Wei, W.C., 2018. Liquidity and market efficiency in cryptocurrencies. Econ. Lett. 168, 21–24.
Zargar, F.N., Kumar, D., 2019. Informational inefficiency of Bitcoin: a study based on high-frequency data. Res. Int. Bus. Finance 47, 344–353. https://doi.org/10.1016/j.ribaf.2018.08.008.

Utgitt i ph.d. serie ved Handelshøgskolen:

Nr. 1 – 2003 Lars Øystein Widding
Bygging av kunnskapsreservoarer i teknologibaserte nyetableringer

Nr. 2 – 2005 Pawan Adhikari
Government Accounting in Nepal: Tracing the Past and the Present

Nr. 3 – 2005 Tor Korneliussen
The Relationship between Initation, Barriers, Product Quality and Internationalization

Nr. 4 – 2005 Bjørn Willy Åmo
Employee innovation behavior

Nr. 5 – 2005 Odd Birger Hansen
Regnskap og entreprenørskap. En fortolkende studie av hvordan to entreprenører bruker regnskap

Nr. 6 – 2006 Espen John Isaksen
Early Business Performance
- Initial factors effecting new business outcomes

Nr. 7 – 2006 Konstantin Timoshenko
Russian Government Accounting:
Changes at the Central level and at a University

Nr. 8 – 2006 Einar Rasmussen
Facilitating university spin-off ventures
-an entrepreneurship process perspective

Nr. 9 – 2007 Gry Agnete Alsos
Portfolio Entrepreneurship - general and farm contexts

Nr. 10 – 2007 Elsa Solstad
Tre sykehus - to verdener - en fusjon.
En studie av reorganisering i et helseforetak

Nr. 11 – 2007 Levi Gårseth-Nesbakk
Experimentation with accrual accounting at the central government level in Norway - how a global phenomenon becomes a local practice

Nr. 12 – 2007 Tatiana Iakovleva
Factors Associated with new venture performance:
The context of St. Petersburg

Nr. 13 – 2007 Einar Lier Madsen
Utvikling av dynamiske kapabiliteter i små og mellomstore bedrifter

Nr. 14 – 2008 Anne Haugen Gausdal
'Network Reflection' – a road to regional learning, trust and innovation

Nr. 15 – 2008 Lars Rønning
Social capital in farm-based entrepreneurship and rural development

Nr. 16 – 2008 Terje Andreas Mathisen
Public Passenger Transport in Norway – Regulation, Operators' Cost Structure and Passengers' Travel Costs

Nr. 17 – 2008 Evgueni Vinogradov
Immigrant Entrepreneurship in Norway

Nr. 18 – 2008 Elin Oftedal
Legitimacy of Creative Destruction

Nr. 19 – 2009 Frode Kjærland
Valuation of Generation Assets – a Real Option Approach

Nr. 20 – 2009 Tatiana Maximova-Mentzoni
Marketization of the Russian University: Origins, Features and Outcomes

Nr. 21– 2009 Hugo Skålsvik
Studies of Market led Processes influencing Service Performance:
-Case Studies on the Norwegian Coastal Voyage

Nr. 22– 2009 Svein Oskar Lauvsnes
Determinants of a shifting effective demand equilibrium.
An explorative investigation of the interaction between psychological, financial and real factors

Nr. 23– 2010 Frode Fjelldal-Soelberg
Entreprenøriell markedsføring. En studie av entreprenørskap og markedsføring som overlappende fenomen

Nr. 24– 2010 Heidi Rapp Nilsen
From Weak to Strong Sustainable Development
An analysis of Norwegian economic policy tools in mitigating climate change

Nr. 25– 2010 Gowindage Chamara Jayanath Kuruppu
Development of Central Government Accounting in Sri Lanka:
Three perspectives on the accounting changes

Nr. 26– 2010 Marina Z. Solesvik
Interfirm collaboration: The context of shipbuilding.

Nr. 27– 2010 Jan Terje Henriksen
Planning, Action and Outcome
- Evaluation of the Norwegian Petroleum System:
A Structuration Approach to Ripple Effect Studies

Nr. 28– 2010 May Kristin Vespestad
Empowered by Natures – Nature-based High North Tourism Experiences in an International Context

Nr. 29– 2011 Andrei Mineev
How has the petroleum supply industry developed in The Russian Barents Sea Region? Institutional and managerial aspects

Nr. 30– 2011 Jorunn Grande
Entrepreneurship in small rural firms - the case of agriculture

Nr. 31– 2011 Thomas Johansen
Paradigms in Environmental Management Research:
Outline of an Ecosophical-Hermeneutic Alternative

Nr. 32– 2011 Elena Dybtsyna
Accountant in Russia: changing times, changing roles.

Nr. 33– 2012 Harald Fardal
Information Systems Strategy in Practice
A Social Process Perspective

Nr. 34– 2012 Kristin Haugland Smith
Hva er bedrifters samfunnsansvar?
- En empirisk tilnærming av bedrifters ansvar overfor samfunnet

Nr. 35– 2012 Are Branstad
The management of entrepreneurship support
– Organisation and learning in corporate incubation, technology transfer and venture capital

Nr. 36– 2012 Victoria Konovalenko
A “coordination kaleidoscope”:
The role of a “Corporate University” as a coordinator of knowledge flows in a Russian transnational corporation

Nr. 37– 2012 Thor-Erik Sandberg Hanssen
Essays in Transport Economics with application to Transport Policy

Nr. 38– 2013 Are Severin Ingulfsvann
Verdiforskyvning i friluftslivet i lys av økologisk økonomi

Nr. 39– 2013 Natalia Andreassen
Sustainability Reporting in a Large Russian Oil Corporation.
Production Safety Issues

Nr. 40– 2013 Elena Panteleeva
Contemporary Management Accounting Practices in Russia:
The Case of a Subsidiary in a Russian Oil Company

Nr. 41– 2013 Thusitha S.L.W.Gunawardana
Impact of Power Sources on Channel Members' Performance

Nr. 42– 2013 Nadezda Nazarova
Mastering Nature and Managing Frictions: Institutional Work and Supply Chain Management in the High North

Nr. 43– 2013 Inge Hermanrud
Managed Networks of Competence in Distributed Organizations
- The role of ICT and Identity Construction in Knowledge Sharing

Nr. 44– 2013 Kari Djupdal
Sustainable entrepreneurship:
outcomes associated with an environmental certification resource

Nr. 45– 2013 Imtiaz Badshah
Federal government accounting in The Islamic Republic of Pakistan

Nr. 46– 2014 Muhammad Arif
Inter-organizational Exchange Relationships
– Exchange Relationships between Local Service Suppliers and Tour Operators in the Tourism Distribution Channel

Nr. 47– 2014 Wondwesen Tafesse
The Marketing Functions of the Trade Show System

Nr. 48– 2014 Fritz J. Nilssen
Erfaringsutveksling som grunnlag for mestring og livskvalitet
Diagnoseoverskridende samtalegrupper for familier med barn som har nedsatt funksjonsevne og eller kronisk sykdom.

Nr. 49– 2014 Ingebjørg Vestrum
The Resource Mobilisation Process of Community Ventures
-The Case of Cultural Events in Rural Communities

Nr. 50– 2014 Ragnhild Johnson
The Practice of Project Management
- A qualitative analysis of complex project-based organizations

Nr. 51– 2014 Ann Heidi Hansen
Memorable moments
Consumer immersion in nature-based tourist experiences

Nr. 52– 2014 June Borge Doornich
Entry modes and organizational learning during internationalization
An analysis of Norwegian supply companies' entering and expanding in the Russian oil and gas sector

Nr. 53– 2014 Kjersti Karijord Smørvik
Opplevelsesskaping i dynamiske opplevelsesrom:
En studie av turisters opplevelser på Hurtigruten

Nr. 54– 2015 Marianne Terese Steinmo
How Firms use University-Industry Collaboration to Innovate:
The role and Development of Social Capital and Proximity Dimensions

Nr. 55– 2015 Eva J.B. Jørgensen
Border Firms: Norway and Russia

Nr. 56– 2015 Krister Salamonsen
Exogenous Shocks as Drivers of Growth in Peripheral Regions.
- A Multilevel Approach to Regional Development

Nr. 57– 2015 Hindertje Hoarau Heemstra
Practicing open innovation in experience-based tourism:
the roles of knowledge, values and reflexivity

Nr. 58– 2015 Elena Zhurova
Environmental Performance Reporting of Russian Oil and Gas Companies

Nr. 59– 2016 Siri Jakobsen
Environmental innovation cooperation:
The development of cooperative relationships between Norwegian firms

Nr. 60– 2016 Antonina Tsvetkova
Supply Chain Management in the Russian Arctic:
An institutional perspective

Nr. 61– 2017 Kjersti Granås Bardal
Impact of Adverse Weather on Road Transport:
Implications for Cost-Benefit Analysis

Nr. 62– 2017 Kristian Støre
Methodological contributions and applications in real options analysis

Nr. 63– 2017 Thomas André Lauvås
The dynamics of university-industry collaboration:
A longitudinal case study of research centers

Nr. 64– 2017 Sølvi Solvoll
Development of effectual and casual behaviors:
Exploring new venture creation in the tourism industry

Nr. 65– 2017 Evgenii Aleksandrov
The changing role of accounting from reformees' perspective:
A study of public sector reforms in Russia

Nr. 66– 2017 Igor Khodachek
Budget, Strategy and Accounting.
Managing institutional change in Russia's governments

Nr. 67– 2018 Vivi Marie Lademoe Storsletten
Quality as flourishing
A study of quality based upon leadership in kindergartens with
implications for Ecological Economics

Nr. 68– 2018 Olga Iermolenko
The human side of accounting:
The bonds between human agency and management accounting
practices' changes in the transitional economy

Nr. 69– 2018 Karin Wigger
Mobilization of Collective Resources for Entrepreneurship:
Case Studies in Nordic Peripheries

Nr. 70 – 2018 Andreas Mikkelsen
Trading fast and slow: algorithmic trading in the Nordic region

Nr. 71 – 2018 Asbjørn Veidal
Strategic entrepreneurship in farm businesses

Nr. 72 – 2018 Are Jensen
Early imprints in and on new technology-based firms

Nr. 73 – 2018 Marianne Arntzen-Nordqvist
The financing process of new technology-based firms
- The entrepreneur’s perspective

Nr. 74 – 2019 Irina Nikolskaja Roddvik
Deprivation of control: A driving force to gain influence during
the internationalization process of MNC

Nr. 75 – 2019 Petter Gullmark
Unraveling the Building Blocks of Local Government Organizations’
Innovativeness: Insights from a Dynamic Capabilities Perspective

Nr. 76 – 2019 Hanne Stokvik
Knowledge for Innovation

Nr. 77 – 2019 Anastasiya Henk
Between the Devil and the Deep Blue Sea: Managing Business Processes
in Turbulent Environments

Nr. 78 – 2019 Tadeu Fernando Nogueira
Entrepreneurial Learning: An Exploration of the Learning of
New Venture Founders

Nr. 79 – 2020 Veronika Vakulenko
Public Sector Reforms in Ukraine: Roles Played by Global and Local
Agents in Implementing Converging and Diverging Changes

Nr. 80 – 2020 Lars Hovdan Molden
Adapting to Change - On the Mechanisms of Dynamic Capabilities

Nr. 81 – 2020 Sudip Kranti Tiwari
Navigating International Entrepreneurship in a Developing Economy
Context: Lessons from Nepal

Nr. 82 – 2020 Vu Le Tran
Expected Returns: An Empirical Asset Pricing Study

The goal of this thesis is to study the expected return of an asset and the predictability of this return in both the short-term and the long-term. The thesis consists of four papers.

In the first two papers, the thesis investigates the predictability of stock returns and crypto-currency returns with daily frequency. The author develops a new measure of market efficiency (AMIM) which is easy to implement. AMIM helps us to compare the level of efficiency of different assets at different points in time. Using this new measure, the author shows that market efficiency is dynamic, changing over time, and being usually worse in the crisis periods.

In the last two papers, the author investigates the predictability of the expected stock returns in a long-term framework. In the first of these two papers, the thesis studies how stable liquidity is priced in the stock market. The results show that on average, investors discount the price of illiquid, and liquidity volatile stocks over stable liquidity stocks.

In the last paper, the author studies the impact on stock returns when there are collective biases from the investors regarding future stock pay-offs. The investors use stock characteristics to form their own biased view regarding future stock pay-offs. Therefore they create stock's mispricing. The data shows that this mispricing will be corrected in the future.

NORD
University

ISBN: 978-82-92893-72-2
ISSN 2464-4331
Trykk: Trykkeriet, Nord universitet
www.nord.no