THREE ESSAYS IN SEMI-PARAMETRIC MODELLING OF TIME-VARYING DISTRIBUTION

Minjoo Kim

Submitted in accordance with the requirements for the degree of Ph.D.

The University of Leeds
CASIF
Leeds University Business School

June 2011

ISBN:

To my family and fiancée.

Acknowledgements

I wish to express my deep and sincere gratitude to my main supervisor: Yongcheol Shin. His lessons and enthusiasm for my study have always been a source of inspiration to me over the years. I am also deeply indebted to another supervisor: Charlie X. Cai. He has given me important suggestions and new ideas about how to improve my research. I am deeply indebted to two professors: Inmoo Kim and Joon Y. Park who taught me attitude appropriate to a scholar. It is always the polar star for me not to lose a direction.

I am very grateful to Kausik Chaudhuri, Viet Ahn Dang, Mattew Greenwood-Nimmo and Qi Zhang who have provided excellent co-works. I wish also to thank Martin Anastasov, Jongseok Lee, Viet Hoang Nguyen and Kevin Reilly for useful comments and suggestions that have helped the development of my thesis. I would like to express my gratitude to Kevin Keasey, Michelle Dickson and directors of the Centre for Advanced Studies in Finance (CASIF) for their generous support during my PhD programme. Jongseok Lee and Jungsoo Shin whose encouragement and friendship have been a great source of strength for me.

I dedicate much of my work to my father and mother who have been waiting for my degree and thesis with endless support and trust. I wish also to express my deep and sincere love to my sister and brother. My lovely fiancée was the cheerleader as I ran the marathon of the Ph.D. programme. Her unfaltering love and encouragement have been truly heart-warming. I would not have come to where I am without her constant love and trust.

Oh, how can I forget my buddies - those who always encourage me to study and provide timeless friendship: Jeonghun Seo, Yangki Jung and Heejoon Hwang. I am so lucky to get to know all of them.

Abstract

During the last century we have been frustrated by the number of economic crises which trigger extreme uncertainty in the global economic system. Economic agents are sensitive to the uncertainty of inflations, as well as to asset values, for survival in such circumstances. Hence, modern finance and monetary economics emphasise that risk modelling of asset values and inflations are key inputs to financial theory and monetary policy. The risk is completely described by the distribution which is verified to be time-varying and non-normal. Although various parametric and non-parametric approaches have been developed to model the time-varying nature and the non-normality, they still suffer from intrinsic limitations. This study proposes the dynamic modelling of the non-parametric distribution (*Functional Autoregressive Model* (FAR) and *Spatial Distribution Analysis*) in order to overcome the limitations.

Firstly, we apply FAR to the Value-at-Risk analysis. It forecasts an intraday return density function by the functional autoregressive process and calculates a daily Value-at-Risk by the Normal Inverse Gaussian distribution. It reduces economic cost and improves coverage ability in the Value-at-Risk analysis.

Secondly, we apply FAR to forecasting the cross-sectional distribution of sectoral inflation rates, which holds the information of the heterogeneous variation across sectors. As a result, it improves the aggregate inflation rate forecasting. Further, the heterogeneous variation is utilised for constructing the uncertainty band of the aggregate inflation forecast, like the fan-chart of the Bank of England.

Thirdly, we apply the spatial distribution analysis to rank investment strategies by comparing their time aggregated utilities over the investment horizon. To this end, we use a spatial dominance test. Since a classical stochastic dominance approach considers only the return distribution at the terminal time point of the investment horizon, it cannot properly evaluate the risk, broken out exogenously or endogenously, in the middle of the investment horizon. However, the proposed spatial dominance approach considers completely the interim risk in evaluating alternative investment strategies.

Contents

List of Tables

List of Figures

Chapter 1

Overviews

1.1 Distribution of Asset Return

Over the last century, we have experienced a number of financial crises:

- the Wall Street Crash in 1929, followed by the Great Depression.
- the Oil price shock in 1973, causing the 1973-74 stock market crash.
- the Latin American debt crisis in the 1980's.
- the Japanese asset price bubble collapse in the 1990's.
- the Economic crisis in Mexico in 1994-95; speculative attacks and default on the Mexican debt.
- the Asian financial crisis in 1997; devaluations and banking crises across Asia.
- the Russian financial crisis in 1998.
- the Bursting of the dot-com bubble in 2001; speculations concerning crashed Internet companies.
- the global financial crisis in 2007-10, followed by the late 2000's recession and the 2010 European sovereign debt crisis.

The risk of the asset price is fully described by the asset return distribution. Bachelier (1900), considered as a pioneer in the study of financial mathematics and stochastic processes, is credited with being the first academic to discuss the use of Brownian motion to evaluate stock options. Almost 60 years later, Osborne (1959) formally specifies the asset return generating process as Brownian motion. They both assume that price changes from transaction to transaction in the individual security are random drawings from the same distribution, implying that successive price changes are independent and identically distributed (IID). The central limit theorem leads us to expect that the distribution of the sum of IID random drawings generally approaches a normal distribution as the number of items in the sum increases. Thus the distributions of daily, weekly, and monthly price changes are approximately normal in the Bachelier-Osborne model. Strictly speaking, continuously compounded single-period returns are IID normal, which implies that single-period gross simple returns are distributed as IID log-normal variates. The IID (log) normal distribution assumption is then widely accepted in modern financial theories: portfolio theory Markowitz (1952), capital asset pricing model (Sharpe, 1964; Lintner, 1965; Mossin, 1966; Black, Jensen, and Scholes, 1972), option pricing (Black and Scholes, 1973; Merton, 1973), arbitrage pricing theory (Ross, 1976) and risk management (RiskMetrics, 1996). But as attractive as the log-normal model is, it is not consistent with all of the properties of historical stock returns. Historical returns show the weak evidence of the skewness and the strong evidence of the excess kurtosis on a short horizon. Sample estimates of the skewness for daily US stock returns tend to be negative for stock indexes, but close to zero or positive for individual stocks, whilst those of the excess kurtosis for daily US stock returns are large and positive for both indexes and individual stocks, indicating that returns have more mass in tail areas than would be predicted by a normal distribution. Hence, the initial approach does not satisfy any requirements of the empirical asset return distribution.

A fat-tailed distribution is first captured by Mandelbrot (1963) who criticises the extensive reliance on the normal assumption for asset pricing and the investment theory. Fama (1963, 1965), Clark (1973) and Blattberg and Gonedes (1974) argue the non-normality of stock returns and develop their modelling as IID draws from fat-tailed distributions (e.g. non-normal stable distributions). The non-normal stable distribution has more probability mass in tail areas than the normal distribution. Although stable distributions were popular in 1960's and early 1970's, they have fallen out of favour, partly because they make a theoretical modelling so difficult. Although standard financial theories always require the finite second moment of asset returns and often finite higher moments as well, stable distributions have some counter-factual implications:

Firstly, they imply that sample estimates of second and higher order moments will tend to increase as a sample size increases, whereas these estimates seem to converge in practice.

Secondly, they imply that long-horizon returns will be just as non-normal as short-horizon returns. In practice, the evidence for the non-normality is much weaker for long-horizon returns than for short-horizon returns.

Various candidates for the asset return distribution have been suggested in the literature later.[1] Recommended symmetric distributions are Student-t and Generalised Error Distribution (GED). The Student-t has fatter tails compared with the standard normal distribution and the kurtosis is determined by the degrees of freedom parameter. The GED of Subbotin (1923) is first used by Nelson (1991) for modelling a stock return distribution. It possesses a moment generating function for the straightforward derivation of moments that are used to price derivative assets. Moreover, Theodossiou (2000) demonstrates that the parameter estimation of GED is no less sensitive to outliers in the data than the Generalised-t distribution, and it is much more flexible in specifying tails via a shape parameter.

Recent research studies the asymmetry and higher order moments of the asset return distribution. The latest finance theory extends the classical modern finance theory based on the mean-variance normality, into the theory associated with higher order moments such as skewness and kurtosis. The asset pricing theory concerns the pricing of systematic higher order co-moment risk, like co-skewness or co-kurtosis risk (Harvey and Siddique, 2000; Chung, Johnson, and Schill, 2006; Vanden, 2006; Post, Van Vliet, and Levy, 2008; Adrian and Rosenberg, 2008). The importance of including the higher order co-moments is also found in the portfolio theory as being essential factors for optimal asset allocations (Mitton and Vorkink, 2007; Guidolin and Timmermann, 2008). The option pricing study of Bakshi, Kapadia, and Madan (2003) provides several insights into economic significances of the skewness and it also provides an explanation for the presence and the evolution of the risk-neutral skewness over time and across cross-section. Although these theories fundamentally concern the existence of higher-order (co-)moments of asset returns and the including them as systematic factors for the efficient pricing of the risk, they do not much concern the choice of the distribution and the estimation of higher order moments.

Risk management is most sensitive to the underlying distribution of the asset return, since the risk is usually defined as a maximum loss given the probability. For instance, Value-at-Risk (VaR) depends entirely on the quantile of the distribution. Hence, the risk management devotes all efforts to the accurate modelling of the return distribution. Early VaR models employ a symmetric distribution determined by the first two moments, e.g. normal distribution. The RiskMetrics of the JP Morgan company is a typical example. However, such models fail to improve accuracy and efficiency because of the biased choice

[1] The exponentially generalised beta distribution of type II (McDonald and Xu, 1995), the generalised Student's t (McDonald and Newey, 1988) and its skewed generalisation (Theodossiou, 1998), the local distribution (Gourieroux and Jasiak, 2001), the generalised hyperbolic distribution (Eberlin and Keller, 1995), the stable Paretian distribution (Panorska, Mittnik, and Rachev, 1995), and Levy and Duchin (2004) tested more distributions; beta, extreme value, gamma, logistic, skewed normal and Weibull.

of the distribution. Recent studies deal with the choices of distribution that allow for the asymmetry and higher order moments. Mittnik and Paolella (2000) shows that the asymmetric modelling of the second oreder moment improves the VaR forecast. Furthermore, Venkataraman (1997), Fernández and Steel (1998) and Giot and Laurent (2003) improve the VaR forecast by employing the family of skewed distributions. Most recently, Bali, Mo, and Tang (2008) demonstrate that the GARCH specification of the first four moments with a skewed generalised Student's t distribution improves the accuracy of the VaR forecast in line with Hansen (1994), Harvey and Siddique (1999), Jondeau and Rockinger (2003) and Brooks, Burke, Hervi, and Persand (2005).

In addition to the non-normality of the asset return distribution, the time-varying nature is another key factor in modelling and forecasting the asset return distribution. Hence, the modelling of the return distribution usually comprises two procedures: (i) the specification of the return distribution and (ii) the time-dependence specification of the return distribution. Procedures are methodologically embodied in a parametric, a non-parametric, or a hybrid manner.

The parametric approach is the one most dominant in the literature. A distribution form is pre-specified and a time-varying structure is then applied to moments of the asset return distribution. As discussed above, there are a number of distribution family for describing the asset return distribution. However, the goodness-of-fit of the selected distribution family varies over time horizons and across asset returns. The time-varying structure is mainly specified by the autoregressive process. The autoregressive model of the first order moment (mean) has a long history in the time series model. The time-varying second order moment is generally modelled by the GARCH process (Engle, 1982; Bollerslev, 1986) where a volatility depends on its own past values as well as past squared errors. There are a number of extended GARCH version. ARCH-M (Engle, Lilien, and Robins, 1987) specifies a relationship between the excess return and the risk premium in the term structure, but the empirical evidences for the relationship are not robust. The exponentially weighted moving average (EWMA) of RiskMetrics (1996) is the special case of the nonstationary version of the GARCH which is introduced as IGARCH (Engle and Bollerslev, 1986). RiskMetrics is the first step in the industry to combine risk management with the GARCH family. IGARCH assumes the infinite horizon of decaying horizon of the volatility, but empirical findings describe that shock on the volatility seems to have a long, but not infinite memory. Hence, RiskMetrics tends to overestimate the persistence of the volatility. Over persistence of the volatility stimulates the fractionally integrated GARCH model (FIGARCH, Bollerslev and Mikkelsen, 1996). A fractional parameter controls the rate of the hyperbolic decay in the autocorrelation of the conditional variance. The generalisation of FIGARCH is hyperbolic GARCH (HYGARCH, Davidson, 2004). Component GARCH (CGARCH, Ding and Granger, 1996; Engle and Lee, 1999) also captures slow decay in the conditional variance.

In the second-order moment (volatility) specification, the asymmetric effect on the conditional variance is an important issue in the empirical finance. Exponential GARCH (EGARCH, Nelson, 1991) captures the widely accepted notion that negative shocks lead to larger conditional variances than positive shocks. GJR-GARCH (Glosten, Jagannathan, and Runkle, 1993) is another such extension which sought to capture such asymmetry. These two models therefore reflect the leverage effect in stock return (Black, 1976). Engle and Ng (1993) propose the feedback mechanisms of positive and negative shocks (good and bad news) using the news impact curve to evaluate asymmetric GARCH models. Their empirical results of daily Japanese stock return data suggest that GJR-GARCH is the best parametric model. EGARCH also can capture most of the asymmetry, but there is evidence that the variability of the conditional variance implied by EGARCH is too high. Threshold GARCH (TGARCH, Zakoian, 1994), smooth transition GARCH (STGARCH, González-Rivera, 1998) and asymmetric power ARCH (APARCH, Ding, Granger, and Engle, 1993) are also included in the asymmetric GARCH models.

Early GARCH family assume fundamentally that time dependent structure of asset return can be described completely by the first two conditional moments. However, the non-normality of asset return cannot be explained completely by the first two moments. It needs higher order moments than second order moment. At least skewness and kurtosis must be included because skewness describes the asymmetry and kurtosis the degree of fat-tail. Moreover, the time-dependence of these two moments should be specified for time-varying nature. Therefore, volatility specification alone by the classical GARCH-family is not enough to describe the time-varying non-normal asset return distribution. Harvey and Siddique (1999) present the methodology for estimating time-varying conditional skewness with the non-central Student's t distribution (GARCHS). They find that the conditional skewness is important in daily, weekly and monthly stock returns. They show that the evidence of asymmetric variance is consistent with conditional skewness, and inclusion of conditional skewness also impacts the persistence in conditional variance. Brooks, Burke, Hervi, and Persand (2005) propose the model for the autoregressive conditional heteroscedasticity and kurtosis with the standard Student's t distribution (GARCHK). They find significant evidence of the presence of autoregressive conditional kurtosis in a set of four daily financial asset return series comprising US and UK stocks and bonds. Hansen (1994) generalises ARCH model to specify the higher order moments parameters on the past information set. He uses the specialised version of generalised Student's t distribution. He finds the substantial empirical evidence of significantly time-varying higher order moments: the skewness and degrees-of-freedom parameters, that are used to model the standard residuals on the one-month excess holding yield on U.S. Treasury securities and the monthly Dollar/France exchange rate, are statistically significant. Following Hansen (1994), Jondeau and Rockinger (2003) provide the different econometric specification of the time-varying higher order moments (GARCHSK). They specify the time-dependence of higher order moments, as well as contemporary relationship between moments (skewness parameter and kurtosis parameter) which is not found in Hansen (1994). They provide evidence that time-varying skewness, and degrees-of-freedom parameters of the generalised skewed Student's t distribution for standardised residuals on a set of daily stock indexes and foreign exchange rate are statistically significant. In most cases, they find a time-dependence of the asymmetry parameter, whilst the degree-of-freedom parameter is generally found to be constant over time. They also provide evidence that skewness is strongly persistent, whilst kurtosis is much less persistent. The model of Jondeau and Rockinger (2003) is most popular among parametric time-varying higher order moment models because of its general specification. However, as presented in their paper, there are many possible specifications (6 cases) compared with other models, since there are more parameters that must be considered.[2]

In spite of intensive efforts for developing parametric models, they still suffer from several uncertainties. Most well established uncertainties associated with parametric econometric models can be summarised as follows:

Firstly, there is an uncertainty in selecting the best approximation to the true (unknown) asset return distribution, which also tends to vary over time horizons and across asset returns.[3] The choice of an appropriate distributional form is also likely to be affected by the nature of the markets, the assets and the regions. Moreover, we must test the goodness-of-fit in order to select the best one. Simply put, the use of an incorrect distribution will clearly affect the entire empirical analysis.

Secondly, there are uncertainties in modelling the time-dependent structure of the underlying parameters. Especially, when modelling dynamic interactions among the first four

[2]When we include higher order moments, we should consider the combination of 3 possible cases. First is the time-dependence of higher order moments. Second is the contemporary relationship between moments (e.g. skewness and kurtosis, variance and kurtosis). Third is the time dependent relationship between moments (e.g. skewness (t) and kurtosis $(t-1)$, variance (t) and kurtosis $(t-1)$). Hence, more specifications should be considered.

[3]Among various parametric distributions, we must choose the best for empirical modelling of asset return distribution. There is however no dominant distribution over various asset returns, which is found in the results of Levy and Duchin (2004) who test the goodness of fit of 11 theoretical distributions including the normal distribution, fat-tailed distributions, and skewed distributions for investment horizons ranging from one day to four years. Dominant distribution varies through horizons (short and long) and assets (stock, corporate bond, government bond, treasury bill). Their test results propose that we have to choose different distributions for different assets and different horizons.

moments considered by Jondeau and Rockinger (2003), the misspecification error will be likely to be more substantial, potentially resulting in misguided empirical findings.

Finally, computational complexity and burden will be non-negligible. The distribution is determined by parameters which are estimated by MLE using an numerical optimisation algorithm. Most parametric models employ MLE technique, mainly using the numerical optimisation algorithm to deal with the potential non-linearity and asymmetry of the likelihood function. Indeed modelling the time-varying interactions among the higher order moments obtained from the underlying asymmetric distribution function (e.g. skewed Student's t-distribution) makes it much more complicated to optimise the likelihood function.

In this regard, the non-parametric modelling of the asset return distribution may be an alternative approach. It is parameter free and should be a robust approach when we have enough observations to estimate an empirical distribution. The use of the non-parametric approach is found in the historical simulation technique for the VaR analysis and the stochastic dominance evaluation. Even though the non-parametric approach is easier, less uncertain and more robust than the parametric approach, there is a lack of dynamics on the asset return distribution. Ignoring the time-varying nature is more likely to provide misleading results. In order to mitigate the drawbacks of the non-parametric approach, a hybrid approach has been developed. It takes advantages from the parametric and the non-parametric approaches to improve the modelling of the asset return distribution. A filtered historical simulation (Barone-Adesi, Giannopoulos, and Vosper, 1999, 2002) strongly supports the advantage of the hybrid approach to the VaR analysis. Therefore, it is prudent to develop the hybrid approach which is essentially required in modelling the time-varying non-normal asset return distribution.

1.2 Distribution of Inflation Rates

Empirical findings on inflation rate show persistences in mean and variance that are nearly nonstationary. Friedman (1977) conjectures, in his Nobel lecture, that a higher inflation rate is generally associated with its higher volatility. Conversely, Cukierman and Meltzer (1986) postulate that the higher level of inflation uncertainty leads to the higher rate of inflation. Therefore, the monetary economics is more interested in a relationship between the inflation rate and its higher order moments. This implies that the non-normality of inflation rate distribution is generally accepted in the modern inflation study. On the other hand, several theoretical frameworks have been developed to explain the relationship between the inflation rate and the higher order moments of relative price changes: the exogenous form of downward nominal rigidity in product market (Tobin, 1972), the menu cost explanation

(Ball and Mankiw, 1995) and technology shock (Balke and Wynne, 2000). This issue is important for Economist to forecast the future inflation that is a key input in the monetary policy associated with inflation. If the forecasting model cannot correctly specify the underlying structure between the inflation rate and the higher moments of the relative price changes, it should have a low forecasting power. Therefore, recent studies on the inflation focus on how construct the uncertainty of inflation rate forecasting by utilising the time-varying nature of inflation rate and heterogeneous information across sectoral inflation rates. Moreover, the accurate forecasting of inflation rate with uncertainty is practically very urgent, since monetary authorities around the World began to adopt inflation targeting as the monetary policy framework in the 1990s.

Early studies have employed an autoregressive process for modelling the inflation rate, and GARCH for its uncertainty. However, it is well established that the models that are employed neglect the importance of accounting for higher order moments.[4] In this regards, researchers have attempted to develop a parameter free approach for avoiding the limitations of two moments based parametric approach. They use the information heterogeneity across forecasters by using the forecasts from the Survey of Professional Forecasters (SPF) (Zarnowitz and Llambros, 1987; Lahiri and Liu, 2006; Lahiri and Sheng, 2008). However, the approach based on the SPF is likely to suffer from a small-sample bias regarding the cross-sectional dimension. Its probability density function is reported as a histogram, and thus it is not straightforward to evaluate moments from a histogram (Clements, 2010; Engelberg, Manski, and Williams, 2009). Additionally, the heterogeneity of forecasters and the panel composition change over time, which makes the interpretation of temporal variation in an aggregated forecast more difficult (Manski, 2010). As an alternative approach, this study focuses on developing the model utilising the cross-sectional variation over sectors in a non-parametric manner.

1.3 Econometric Analysis

The common findings from distributions of asset returns and inflation rates are that they are time-varying and non-normal. Furthermore, it is proven that the accurate estimation and forecasting of the distributions are key inputs for the finance and monetary policy. Hence, the dynamic modelling of the time-varying and non-normal distribution is the most important issue in this study. The parametric approach is successful in modelling the time-varying nature, but inherently limited to estimating and specifying higher order moments. On the

[4] Bryan, Cecchetti, and Wiggins (1997) document fat-tailed properties and Roger (2000) provides evidence of right skewness in inflation data.

other hand, the non-parametric approach is free from the specification of parameters and the numerical estimation such as MLE, and can better capture the non-normality than the parametric approach. However, it suffers from neglecting the dynamics of the distribution. Hence there is a trade-off between the parametric and the non-parametric approaches in modelling the time-varying non-normal distribution.

In this regard we need to develop the dynamic modelling of non-normal distribution in a semi-parametric manner. It is a variation of the hybrid model that takes the dominant genes from both the parametric and the non-parametric approaches. It is a reasonable expectation that the hybrid model improves the estimation and the forecasting performance of the time-varying non-normal distribution. In particular, this study suggests two novel econometric approaches for dealing with non-normal and time-varying distributions: Functional Autoregressive Model and Spatial Distribution Analysis. The models that have been introduced by Park (2007) and Park and Qian (2007, 2011) for the econometrics, have been well developed in a stochastic process (See Bosq, 2000; Revuz and Yor, 1994). The functional autoregressive model is the autoregressive modelling of the functional data that is obtained by density function or cross-sectional distribution approximated non-parametrically. The spatial distribution analysis is the non-parametric approach, suitable for evaluating both stationary and non-stationary time series.

1.3.1 Functional Autoregressive Model

The functional autoregressive model (FAR) assumes that the sequence of density functions (or cross-sectional distributions) follows an autoregressive process. In practice, it first non-parametrically estimates an empirical density function and then applies the autoregressive process to the estimated density function. There are two possibilities. The density function can be estimated using the high-frequency data (e.g. asset returns are recorded in 5-minute or 30-minute intervals) within a fixed time interval (e.g. 1-day or 1-week). On the other hand, the cross-sectional distribution can be estimated at each point of time using the cross-sectional data. The time-dependence of the empirical density at a fixed time interval is then specified by the FAR. Notice that the FAR nests all the autoregressive specifications of conditional higher order moments considered in the literature (Bollerslev, 1986; Brooks, Burke, Hervi, and Persand, 2005; Engle, 1982; Engle and Manganelli, 2004; Hansen, 1994; Harvey and Siddique, 1999; Jondeau and Rockinger, 2003). Moreover, the recent advances in the database of finance (e.g. WRDS) and national economic statistics (e.g. Office for National Statistics, UK; BLS, US) and computing technology will make the proposed approach more plausible.

This study applies FAR to VaR and inflation forecasting. For the VaR forecasting, the intraday return density function is predicted by the FAR using the intraday returns (e.g. 5-minute returns). Then, we construct the daily return density from which we evaluate a daily VaR. To this end, we first derive the first four moments (mean, variance, skewness, and kurtosis) from the intraday density function and extend them to those of the daily return distribution. These daily return moments are used to construct the daily return distribution using the Normal Inverse Gaussian (NIG) distribution, which is quite a promising approach for modelling financial returns (Bandorff-Nielsen, 1997). This hybrid approach uses both the non-parametric estimation of the density function and the explicit dynamic estimation by the FAR, and thus it is expected to improve VaR analysis. Inflation forecasting is a good illustration of the FAR approach to an analysis of the time-varying cross-sectional distribution. This application aims to model the cross-sectional distribution. It uses the cross-sectional variation over sectors, and its dynamics, for constructing the uncertainty band of the aggregate inflation forecasting.

1.3.2 Spatial Distribution Analysis

Financial decisions depend critically on current and past information. Most information can be obtained from the distribution of asset returns in theory and practice. Hence, the time-varying nature and the non-normality of the distribution should be carefully considered when making efficient financial decisions. However, many financial models are empirically conducted under the strong assumption that the asset return distribution is time-invariant. For example, a stochastic dominance is known as the most efficient financial decision framework. It ranks the expected utilities obtained from alternative investments by comparing their distributions. To this end, it (parametrically or non-parametrically) estimates the underlying distribution using the historical asset returns, assuming that the asset return distribution is time-invariant. It is however clearly contradictory to the empirical finding that the asset returns distribution is time-varying. Hence, the (static) stochastic dominance is not the most efficient framework under the time-varying nature. Even if the time-varying nature of the asset return is allowed, it takes the specific form of the parametric specification, such as GARCH family. It is also limited by the inherent disadvantages of the parametric approach. In this regard, it would be more desirable to develop an efficient non-parametric framework, fully incorporating the time-varying nature of the underlying distribution.

The spatial distribution analysis will be a promising candidate for this requirement. Park (2007) defines a spatial distribution as the temporal aggregation of the time-varying distribution over a certain time horizon (e.g. the investment holding period). Moreover, it can be fairly available for both stationary and non-stationary time-series. Following recent econometric and statistical advances, the non-parametric estimation of any spatial distribution is quite straightforward. Further, the inferences on their distributional characteristics can be conducted easily without invoking any limitations of parametric models (Park, 2007). The spatial distribution is flexible enough to apply it easily to financial applications. For example, we can define a VaR based on the spatial distribution. Since it is the time aggregation of time-varying distributions, the VaR is interpreted as a time aggregated VaR that measures the maximum loss of asset values occurring over the (fixed) horizon. The advantages of the time aggregated VaR are that it is available for both the stationary and the non-stationary time series, and it can capture interim risk such as political risk, economic risk and exogenous jumps that are likely to occur in between. In addition, the time aggregated expected utility generated by the asset value is determined solely by its spatial distribution, suggesting that the spatial distribution will play the central role in many financial problems involving a dynamic decision-making based on the utility maximisation (Park, 2007). When comparing the time aggregated expected utilities of the alternative investments over a certain holding period, we can fruitfully utilise the spatial dominance, which is the dynamic version of the (static) stochastic dominance. The key advantages are that it considers all possible risks during the investment period, and it is robust and available for both the time-invariant and time-varying distribution. If the underlying time series follows a stationary distribution, the spatial dominance reduces to the stochastic dominance. Therefore, it is the general version of stochastic dominance allowing general time-varying nature.

1.4 Summary of Applications

This study employs novel econometric approaches to the dynamic modelling of non-parametric distributions and demonstrates their advantages by applying them to finance and monetary economic applications.

Firstly, FAR is applied to the VaR analysis with the high-frequency data. The basic idea is that an intraday return density function is non-parametrically estimated, and the dynamics of the density function are specified by FAR. Given the autoregressive structure, the forecast of the intraday density function is easily obtained. To obtain a daily return density function, the Normal Inverse Gaussian Distribution (Bandorff-Nielsen, 1997) is adopted with a suitable regularity assumption. This framework aims to improve the efficiency and the robustness of the VaR model by employing a hybrid approach that takes the dominant advantages of purely parametric and non-parametric approaches.[5]

[5]The efficiency implies that the model can minimise the economic cost, and the robustness implies that the

Secondly, FAR is applied to modelling and forecasting the cross-sectional distribution of the inflation rate. This application requires that we model the cross-sectional distribution, holding the information of the heterogeneous variation across sectors. Then the heterogeneous variation is used for constructing the uncertainty band of the aggregate inflation forecasting.

Finally, the spatial dominance approach is applied to rank alternative investment strategies by comparing the time aggregate utility from investments. The classical stochastic dominance approach evaluates utilities from alternative investments at the terminal point of the investment horizon. Hence, it depends entirely on the return distribution of the final time point of investment. We may consider the case in which two investment markets are indifferent in terms of the stochastic dominance but there are exogenous shocks or jumps in one market. Then rational investors prefer the market with no shocks or no jumps. This common and rational inference cannot be derived from the stochastic dominance approach. But this expectation can be demonstrated by the spatial dominance approach in a straightforward manner.

This study comprises three applications with proposed econometric methods. Chapter 2 develops the improved VaR analysis by FAR using the Dow Jones intraday return data. Chapter 3 applies FAR to forecasting a UK inflation rate using the disaggregated UK CPI data. Chapter 4 evaluates time aggregated utilities from emerging and developed stock markets by the spatial dominance approach using the MSCI indexes.

In what follows, we provide the summary of three applications.

Application I: Hybrid Functional Autoregressive Modelling of Non-parametric Density for Improved Value-at-Risk Analysis

FAR modelling of the non-parametric density function is proposed to improve Value-at-Risk (VaR) analysis by taking into account the relative advantages of parametric and non-parametric models in a hybrid manner. In particular, this approach enables us to use the intraday information for forecasting the time-varying daily return density function. It is well-established that VaR forecasts obtained from parametric and non-parametric models involve a trade-off between minimising the associated economic costs and providing the valid coverage of VaR. The Monte Carlo simulation study and empirical evaluations of VaR, based on thirty components of the Dow Jones Industrial Average and their equal weighted portfolio, clearly demonstrate that the overall performance of the proposed hybrid model is superior to those of both the parametric and the non-parametric models in terms of sev-

model can keep the violation ratio very close to the nominal rate.

eral (sometimes conflicting) criteria. In particular, the proposed hybrid approach is shown to simultaneously increase coverage ability, reduce economic costs and enhance the statistical validity. Hence, it is recommended to use the hybrid functional autoregressive nonparametric density approach (along with another hybrid model called the filtered historical simulation approach) for improving an internal VaR model for both regulators and banks in a fair and satisfactory manner.

Application II: Forecasting Distributions of Inflation Rates: Functional Autoregressive Approach

This application uses the functional autoregressive approach to model the time-varying distribution of UK monthly inflation rates using sectoral inflation rates. This approach is free of any assumptions as to the structure of the distributions, or the number of dimensions in which the distributions may vary. This framework provides an alternative, for use by independent researchers, to the forecasts made by the central bank, and allows us to incorporate complex and time-varying responses by the policy makers to the disaggregate shocks. The in-sample forecasting evaluation results provide a superior performance from the proposed approach, compared with the benchmark autoregressive models, in forecasting both the cross-sectional distribution of sectoral inflation rates and the density function of mean inflation rates. Furthermore, out-of-sample forecasting results suggest that the mean is projected to be stable at around 2.3%-2.6%, whilst the uncertainty bands stay between 1.5% and 4.5%, over the period March 2008 - February 2010, correctly accommodating inflation rates observed during this period. The forecasts also suggest that the probability of achieving the inflation target of less than 2% is fairly low (around 20-25%), while the probabilities of maintaining inflation between 1% and 3% keep decreasing as the forecast horizon increases (around 65-70%).

Application III: Time Aggregate Utility from Emerging and Developed Stock Market Investments: Spatial Dominance Approach

This application compares investment opportunities in the emerging and the developed stock markets, which is an important issue in international portfolio management. However, it is well-established that the stock market returns are non-normal and have time-varying distributions. This creates a challenge in ranking alternative investment strategies, especially in a dynamic setting. To this end, this application proposes a distribution-free test, based on the spatial dominance approach, which is more general than the stochastic dominance approach and allows us to compare the time aggregate utilities obtained from alternative investments accumulated over the planned investment horizon, instead of at the terminal point of the investment. Applying the proposed test, it is found that the investments in the emerging markets are indifferent to their developed market counterparts over all the investment horizons ranging from 3 months to 5 years, but only if the currency risk is explicitly taken into account. This suggests an integration between the two markets according to the definition of Bekaert and Harvey (2003). It is also found that the returns of emerging markets, denominated in the local currency, dominate those in US dollar related markets over 1- and 5-year investment horizons, implying that there is still an insufficient interaction between equity prices and foreign exchange rates in the emerging markets over the longer-term period. As expected, the currency risk is found to be mostly irrelevant for developed market investments.

Chapter 2

Hybrid Functional Autoregressive Modelling of Non-parametric Density for Improved Value-at-Risk Analysis

2.1 Introduction

Value-at-Risk (VaR) has been used as the central measure of risk in the banking regulatory framework. VaR was popularised by the so-called RiskMetric of J. P. Morgan. It was first established in 1989, when Dennis Weatherstone, the chairman of J. P. Morgan, asked for a daily report measuring and explaining the risk of his firm and nearly four years later in 1992, J. P. Morgan launched RiskMetrics. Moreover, VaR was more popular as a risk measure chosen by investment banks to measure their portfolio risk for the benefit of banking regulators (RiskMetrics, 1996). VaR was also strongly recommended by regulatory organisations. In 1997, the U.S. Securities and Exchange commission ruled that public corporations must disclose quantitative information about their derivatives activity. Major banks decided to comply with this rule by including VaR information in their financial statements. Furthermore, in 1999, the worldwide adoption of the Basel II Accord stimulated financial institutions to use VaR. Nowadays, most major banks report the VaR forecast to the regulator who supervises the bank's risk management.

The urgent need of an accurate internal VaR model was demonstrated during the 2007–2009 global financial crisis. The role of financial risk management has become a notable concern for survival at both the micro- (firm) and macro-level (nationwide). Furthermore, it is critical for banks to accurately measure and disclose the level of risk in their business that is of interest to their investors, creditors and regulators. While the use of VaR as a summary financial risk measure is appealing in theory and practice, difficulties arise in searching for the best VaR model which minimises economic cost and maximises coverage ability.

There have been major developments of the VaR model in the last two decades (see Jorion, 2006 for an excellent survey). However, challenges remain due to the difficulty of modelling a time-varying non-normal return distribution as well as forecasting an extreme value. The previous literature on the VaR modelling falls into three groups: parametric, non-parametric and hybrid approaches. First, there is the parametric approach such as GARCH models and extreme value distribution models (hereafter, EVT models). However, it is well known that the parametric approach is intrinsically bound to suffer from the following limitations. Particularly, there are uncertainties in (i) the distribution assumption, (ii) the time dependent specification of the higher order moments, and (iii) the computational complexities of the numerical optimisation algorithm. Extensive empirical evidence shows that stock return distributions are fat-tailed.[1] Hence, a normal distribution always underestimates the risk of the fat-tailed distributions observed in financial time series. For example, it is well-established that Gaussian GARCH underestimates VaR (Netftci, 2000). RiskMetrics (hereafter RM) is a typical example using the normal distribution (Johansson, Seiler, and Michael, 1999). The Student's t-distribution is thus a preferable choice. It is however a symmetric distribution so that it cannot capture the asymmetry of the asset return distribution (Giot and Laurent, 2003; Mittnik and Paolella, 2000). Recent studies model the fat-tail and the asymmetry of the asset return distribution by allowing for higher-order moments with a specific parametric distribution such as the skewed generalised error distribution (Theodossiou, 1998). Furthermore, some progress has been made in modelling the dynamics of the higher order moments within the GARCH framework; GARCH with skewness (GARCHS, Harvey and Siddique, 1999), GARCH with kurtosis (GARCHK, Brooks, Burke, Hervi, and Persand, 2005) and GARCH with both skewness and kurtosis (GARCHSK, Jondeau and Rockinger, 2003). Bali, Mo, and Tang (2008) demonstrate that the accuracy of a VaR forecast improves significantly when conditional higher order moments are modelled by adopting an asymmetric distribution function. However, the contemporaneous and the time-dependent structure among the moments are too complicated to be specified completely (Jondeau and Rockinger, 2003). EVT models consider the extreme value distribution given a block size.[2] The Generalised Extreme Value (GEV) distribution

[1]Mandelbrot (1963) first captured the fat-tail of stock returns distribution and criticised the extensive reliance on the normality assumption for asset pricing and investment theory. Fama (1963, 1965), Clark (1973) and Blattberg and Gonedes (1974) proposed allowing the stock returns to be non-normal and developing the statistical modelling of stock returns as draws from a fat-tailed distribution.

[2]EVT draws the worst event out of the considered period (such as 1 week or 1 month) then uses them as an

and the Generalised Pareto Distribution (GPD) are popular and frequently employed in VaR modelling. Since EVT models provide a conservative VaR, they are likely to overestimate VaR, leading to an expensive cost.

Second, a non-parametric model is popular among practitioners. A simple historical simulation (hereafter, HS) is found to be the industry standard (Perignon, Deng, and Wang, 2008; Perignon and Smith, 2009). In spite of its popularity, severe problems are reported in the literature (Barone-Adesi, Giannopoulos, and Vosper, 1999, 2002; Jorion, 2006) due to its ignoring the time-varying nature of the asset return distribution. Since HS is intrinsically sensitive to the existing state, it cannot provide a valid VaR when the underlying regime changes.

The limitations associated with both parametric and non-parametric models lead to an alternative hybrid model. The hybrid model is to mix the dominant genes taken from both the parametric and non-parametric models to minimise the economic cost and maximise the coverage ability. For instance, a filtered historical simulation (hereafter, FHS) combines the non-parametric density function with the GARCH filtration of the return data. It is reported that FHS generally outperforms the parametric and the non-parametric models (Barone-Adesi, Giannopoulos, and Vosper, 1999, 2002; Kuester, Mittnik, and Paolella, 2006; Pritsker, 2001). Hence, a more flexible hybrid approach is called for, in order to obtain a more accurate forecast of a time-varying non-normal return distribution.

Recently, a conditional autoregressive quantile model (hereafter CAViaR, Engle and Manganelli, 2004) utilises an empirical quantile without any assumption on the distribution for VaR modelling. However, the performance has been mixed, so far (Engle and Manganelli, 2004; Kuester, Mittnik, and Paolella, 2006).

In line with recent developments in risk modelling, this paper contributes to this growing literature by introducing a novel hybrid technique which is capable of modelling and forecasting a non-parametric density function of the asset return in a flexible manner. In the context of developing hybrid models for the density function, our study utilises the functional autoregressive modelling of the non-parametric density function (Bosq, 2000; Cardot, Mas, and Sarda, 1999; Cardot, Ferraty, Mas, and Sarda, 2003; Cardot, Mas, and Sarda, 2007; Park and Qian, 2007, 2011, hereafter FAR) to forecast a daily VaR using intraday data.[3] We develop a two-stage approach for VaR forecasting. First, we construct the sequence of empirical intraday density functions by a kernel density estimation using the high-frequency data and then we model the time-dependence of the density function by

extreme sample.

[3]This functional approach is also in line with modelling the dynamics of economic functions recently documented by Bowsher and Meeks (2008), employing a functional signal plus noise, (FSN)-ECM, and Kargin and Onatski (2008), employing a functional autoregressive model for forecasting a yield-curve.

an autoregressive process in the functional space. Based on the estimated model, we can easily obtain a density forecast conditional on the past information set. This first-stage is well established in Park and Qian (2007, 2011) but it needs an additional stage to extend the intraday return density function to the daily return density function to obtain a daily VaR. In the second-stage, therefore, we employ the Normal Inverse Gaussian function developed by Bandorff-Nielsen (1997). While it is a parametric extension, it has the advantage of utilising the first four moments obtained from the intraday density function.

FAR approach has three important advantages in density forecasting with the VaR application. First, it models the dynamics of the non-parametric density function in the functional space. By estimating the density function non-parametrically, we can avoid uncertainties from both misspecification and the estimation errors in the parametric approach. On the other hand, the dynamics of the density function is too complicated to be specified in a parametric way. In general, all moments and the locations of the density function are contemporaneously related and their structure is time-dependent. A time-varying conditional moments and quantile approach is unable to identify their true time-dependent structure. But FAR can specify the contemporaneous and time-dependent structure in a reduced form via an autoregressive operator. It is therefore a much more attractive approach. Second, our proposed approach provides a framework for a robust VaR application, successfully minimising the economic cost and maximising the coverage ability. Existing parametric and non-parametric models are unsatisfactory in the following sense: a parametric model can minimise the economic cost but tends to underestimate the VaR (e.g., the GARCH family). A non-parametric model can provide robust and conservative coverage but fails to minimise the economic cost because the dynamics of the return distribution are neglected. The FAR approach overcomes these limitations and provides the desirable features of VaR modelling. Third, it allows us to utilise intraday information for financial risk management. The key issue is estimating and forecasting the daily risk in the investor's or bank's asset portfolio. A return is calculated daily and the volatility of the return is the key measure of risk. However, it has been documented in the market microstructure literature that the use of the high-frequency data provides a better account of the daily volatility:(Andersen, Bollerslev, Diebold, and Ebens, 2001; Bollen and Inder, 2002). Andersen and Bollerslev (1998) show how the high-frequency intraday returns contain valuable information for the measurement of volatility at the daily level. Intraday returns encompass important information which is relevant to market participants in forming their future expectations. This is especially true for large scale institutional trading with active market movements. The intraday return distribution is a key input to access the accumulated results of the daily return. The true interaction among traders happens at the intraday level in all the liquid markets. Therefore, these arguments and the increased availability of the high-frequency financial data motivate us to examine the daily return distribution using the high-frequency data. Given that high-frequency financial data is often characterised as extremely dispersed and non-normally distributed (Hasbrouck, 2007), FAR offers an ideal setting to abstract the daily volatility and the higher order moments from high-frequency data.

We demonstrate the methodological advantage of FAR through extensive evaluation schemes using both actual and simulated data. We first find that FAR is the best predictor for intraday return density forecasting, outperforming alternative functional models such as AVE (IID density functions) and LAST (martingale density function). Employing divergence criteria such as the Hilbert norm, uniform norm and entropy measure for an analysis of the thirty stock components of the DJIA and their equal weighted portfolio over 2000–2008, we find that the density forecast estimated by FAR is closest to the realised density in terms of 'closeness' in a functional space. We also conduct a Monte Carlo simulation study in which we generate the sample based on the empirical findings, and evaluate the forecasting precision by comparing a VaR forecast with a true VaR. Employing Bias, RMSE (Root Mean Square Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), and PMAD (Percentage Mean Absolute Deviation), we find that FAR provides the best precision of VaR forecasting for all precision measures.

Moreover, we demonstrate its advantage for VaR applications by evaluating its performance against existing popular models using a wide range of assessment criteria with 30 stocks and the equal-weighted DJIA portfolio over 2000–2008. The validation of VaR models is investigated with the set of backtesting tools. The backtesting is essential to the Basel committee's decision to allow internal VaR models for capital requirements. We adopt various backtesting tools developed in the academic literature and mentioned in the regulatory documents. Backtesting is a formal framework to verify whether the actual loss is in line with the projected loss. This involves systematically comparing the history of VaR forecasts with their associated (portfolio) returns. When a model is well specified, the number of observations falling outside VaR should be in line with the given nominal level. The number of excesses is known as the 'number of exceptions' (Jorion, 2006, p. 140). The number of exceptions exceeding the cut-off indicates that the model underestimates VaR. If so, too little capital is allocated to the risk-taking units. The regulators impose penalties on banks that adopt such an underestimated model. On the other hand, a number of exceptions lower than the cut-off indicates that the model overestimates VaR, implying that the capital allocation is inefficient. Consider the following assessment. First, coverage ability and economic cost are examined. The empirical coverage probability and the predictive quantile loss are the conventional quantitative measures. We also examine the performance by checking against the Basel Committee's recommendations such as the Basel penalty zone and the market risk capital requirement (hereafter, MRCR). The Basel Committee provides those measures from sufficient simulation studies as well as empirical analyses of financial institutions. The empirical coverage probability and the Basel penalty zone evaluate the coverage ability while the predicted quantile loss and the MRCR evaluate economic cost. Second, regulators also require that formal statistical tests should be conducted so as to evaluate the accuracy of the VaR forecast. To this end, we consider the unconditional test (Kupiec, 1995), conditional coverage tests (Christoffersen, 1998; Engle and Manganelli, 2004) and dynamic quantile tests (Engle and Manganelli, 2004).

From the empirical evaluations of alternative VaR models, we observe that hybrid models such as FAR and FHS outperform parametric models such as GARCH, EVT and CAViaR. The hybrid models obtain high scores in all backtestings. The GARCH models tend to underestimate the VaR, as reported in Johansson, Seiler, and Michael (1999) and Netftci (2000). They perform badly in coverage ability evaluations. They are more often rejected by all the conditional coverage tests but do well for the economic cost analysis. On the contrary, the HS performs well in the coverage ability evaluations but badly in the economic cost analysis. It is however often rejected by all the conditional coverage tests. This clearly demonstrates that there is a trade-off between the two approaches. The EVT models mostly overestimate the VaR, as expected. They perform badly for both the coverage ability evaluations and the economic cost analysis. They are often rejected by the conditional coverage tests while the filtered EVT models are relatively less rejected by the dynamic quantile tests. CAViaR is the worst performing model for all the backtestings. Therefore, the evaluation results support that the dynamic modelling of non-parametric density functions outperforms both the static non-parametric models and the dynamic parametric models. Moreover, empirical findings that the hybrid models outperform other models are consistently repeated for the alternative positions (the short and the long position) and the alternative window sizes (of 250 and 500 business days).

The rest of this chapter is organised as follows. In Section 2, we describe the functional autoregressive modelling for VaR forecasting, along with alternative VaR models. Section 3 provides an overview of the existing backtestings in terms of quantitative measures and statistical tests. Section 4 presents the intraday data used in the analysis. Section 5 provides the Monte Carlo simulation study and the empirical evaluation of functional models and alternative VaR models using 30 components for DJIA and their equal weighted portfolio. Section 6 gives some concluding remarks.

2.2 VaR Models

Quantitative financial risk management is generally performed on a daily basis: an asset return is calculated daily and its volatility is measured as the core of the daily risk. Recently a better account of daily volatility by utilising high-frequency data has been well documented in the market microstructure literature (Andersen and Bollerslev, 1998; Andersen, Bollerslev, Diebold, and Ebens, 2001; Bollen and Inder, 2002). They explain that the intraday data contains valuable information which is relevant to market participants for forming their future expectations. Especially, those arguments are true for large institutional trading with active market movements. Hence, it is key to access the daily return risk to understand the behaviour of the intraday return distribution. Also, the increasing availability of high-frequency financial data motivates us to model the dynamics of the intraday return distribution first and utilise it for forecasting the daily return risk. Furthermore, given that high-frequency financial data often possesses extreme dispersion and is non-normally distributed (Hasbrouck, 2007), FAR offers an ideal way to abstract the daily volatility and higher moments from the high-frequency data. We first present our proposed model below.

2.2.1 Hybrid Functional Autoregressive Modelling of Non-parametric Density

We assume that trading prices are observed in a fixed interval (e.g., 5-minute)[4] and define an intraday return by the first difference of the log price,

$$r_{ti} = \ln P_{ti} - \ln P_{t,i-1},\ t = 1,\dots,T;\ i = 1,\dots,m \tag{2.1}$$

where t and i denote the tth day and its ith intraday observation, respectively. Especially, $P_{t,0}(= P_{t-1,m})$ denotes the close price yesterday, so we include overnight information in the first observation (r_{t1}) at the open time today. Further, the daily return is defined as the sum of the intraday returns:

$$r_t = \sum_{i=1}^{m} r_{ti},\ t = 1,\dots,T. \tag{2.2}$$

In sum, we have the sequence of time series such that

$$X = (X_t)_{t=1}^{T} \text{ and } X_t = (r_{ti})_{i=1}^{m}, \tag{2.3}$$

where X is the set of intraday return paths and X_t is the intraday return path at the tth day.

[4]In practice, the trading action is not regular in time so that the frequency is generally irregular. We thus filter observations and draw prices in a fixed time base, e.g., 9:35 AM, 9:40 AM, ..., 4:00 PM.

First, we assume that the intraday return series, $r_{ti} \in X_t$, is strictly stationary within a given day and the intraday density function can be estimated for each time point t (day) given the sample observation. Once we denote the intraday density function by f_t, the sequence $(f_t)_{t=1}^{T}$ can be defined in the functional space. Second, we assume that this local stationarity does not carry over to a longer horizon. This implies that the time series $(X_t \in X)$ is nonstationary, that is, the intraday density function is time-varying. These assumptions are known as the "piecewise stationarity" of a stochastic process.[5] In the general framework, "piece" is taken as any appropriate time unit such as a day, week or month. Thus a day is a "piece" in our study. FAR has been developed based on the "piecewise stationarity" condition for the observed stock return series. The intraday return series is stationary in a day but it will be nonstationary in the long-run, that is, each day would have a different intraday return distribution.

We assume that the time-varying intraday density function (f_t) follows an autoregressive process with lag order one[6]

$$w_t = Aw_{t-1} + \varepsilon_t,\ t = 2, \ldots, T, \tag{2.4}$$

where $w_t = f_t - \mathbb{E}f$ is the fluctuation of the density function from the well-defined common expectation of the density function $(\mathbb{E}f)$, A is an autoregressive operator on the Hilbert space $(\mathscr{H})$ satisfying $\|A\| < 1$, and $(\varepsilon_t)_{t=1}^{T}$ is the sequence of the functional white noise process. Eq. (2.4) can be rewritten for the density function:

$$f_t = \mathbb{E}f + Aw_{t-1} + \varepsilon_t,\ t = 2, \ldots, T. \tag{2.5}$$

Thus the one-step ahead forecast of the density function is obtained by the conditional expectation on the past information set $(\mathscr{F}_{t-1})$:

$$\mathbb{E}\left[f_t | \mathscr{F}_{t-1}\right] = \mathbb{E}f + Aw_{t-1}. \tag{2.6}$$

FAR is a very general modelling of the conditional autoregressive moments. For example, the centred mean in (2.4), $\tilde{\mu}_t = \langle x, w_t \rangle$, can represented by

$$\tilde{\mu}_t = \langle x, Aw_{t-1} \rangle + \langle x, \varepsilon_t \rangle = \langle A^* x, w_{t-1} \rangle + \eta_t, \tag{2.7}$$

[5]Piecewise stationarity is well demonstrated by "Short-Time Fourier Transform" (Gabor, 1946) in the signal processing literature. The powerful spectral analysis, invented for stationary process, can be applied to nonstationary signals under this assumption.

[6]It is trivial to extend the model to longer lag orders such as an $AR(p)$ model in a vector space. Furthermore, the higher lag orders should be available for higher order FAR modelling.

where A^* is the adjoint of A, $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathscr{H}$, $\langle \nu, u \rangle = \int_{\mathbb{R}} \nu u(u)\, du$, and $(\eta_t)_{t=1}^{T}$ becomes the white noise process under the regular assumptions. The right hand side of (2.7) can be represented by the infinite sum of polynomials, $\sum_{k=1}^{\infty} c_k \left\langle x^k, w_{t-1} \right\rangle = \sum_{k=1}^{\infty} c_k \tilde{\mu}_{k,t-1}$, where $\tilde{\mu}_{k,t-1}$ is the first-lagged centered k-th order moment. This representation implies that the centred mean is specified by the linear combination of all the first-lagged higher order moments. Hence, it includes ARCH-M as a very specific case. Moreover, the centred second order moment, $\tilde{\mu}_{2t} = \left\langle x^2, w_t \right\rangle$, is analogously expressed by all the first-lagged higher moments. It is thus the generalised conditional autoregressive process of the second order moment. These special cases appear in Park and Qian (2007) who derive ARCH and ARCH-M from FAR as special cases. Furthermore, a conditional autoregressive quantile process could be represented as a special case of FAR too. Intuitively, (2.5) can be applied to the quantile process. Let q_t satisfy $\alpha = \int_{-\infty}^{q_t} f_t(s)\, ds$. Then the quantile process (q_t) tends to revert to the α-quantile of $\mathbb{E}f\ (\bar{q})$, as long as the mean reverting property holds in (2.5). This can be modelled by two prats; $\bar{q}$ and the adjustment of the deviation between q_{t-1} and $\bar{q}$. Hence, it follows that $q_t = g(\bar{q}, q_{t-1} - \bar{q}) + \xi_t$, where g is the general form of the nonlinear function. Even though the nonlinear functional form is implicit for the conditional quantile process, the autoregressive adjustment is true as long as FAR holds. Consequently, the ARCH process and the conditional quantile process can be regarded as special cases under some restrictions on the structure of the time-dependence. Usually those parametric models suffer from a misspecification of the dynamic structure when their underlying assumptions are not valid. On the contrary, FAR represents general conditional autoregressive moments for all higher orders and conditional quantile process by the autoregressive operator which is the reduced form for all such conditional autoregressive process. Therefore, FAR can be more flexible and improve the modelling of a time-varying asset return distribution.

Moreover, FAR reasonably explains the finance theories and findings of asset returns. Consider that the mean of the return obtained from f_t is constant and equivalent to that from $\mathbb{E}f$ in (2.4). This implies that there is no expected excess returns since $\langle x, w_t \rangle = \langle x, w_{t-1} \rangle = 0$ in (2.4). Hence, FAR can present the weak-form efficiency in the ‘Efficient Market Hypothesis’ with time-varying higher order moments such as the ARCH process. If the mean of the return has a long-memory, the mean would described by the AR process. Hence, FAR supports ‘behavioural finance’ arguments such as overconfidence, overreaction, and momentum (gradual adjustment to a new equilibrium) with time-varying higher order moments.

If A is the zero operator, the density function of the intraday return is independently and identically distributed in the functional space.

$$f_t = \mathbb{E}f + \varepsilon_t. \tag{2.8}$$

The best predictor is therefore the common expectation of the density function. Since the expected return is equivalent to the average return obtained from the $\mathbb{E}f$ and higher order moments are time constant, this model implies the 'Efficient Market Hypothesis' with time constant higher order moments. We denote this model by AVE.

If A the identity operator, the density function follows the functional martingale process:

$$f_t = f_{t-1} + \varepsilon_t. \tag{2.9}$$

The last observation is therefore the best predictor. All conditional moments including the expected return are nonstationary. Hence, any theories and arguments from the finance literature cannot be applicable to this case. It may be locally observed under the circumstance of "Black Swan" such as the Wall Street Crash (1929), Black Monday (1987) and the Financial Crisis (2007), since the shock used to be continued or diverge in the short term.

To apply FAR in practice, we first need to estimate the empirical intraday density functions at each point of time t. To this end, we suggest using the kernel density estimation. A standard kernel estimator is typically defined by

$$\hat{f}_t(x) = \frac{1}{mh_t} \sum_{i=1}^{m} K\left(\frac{x - r_{it}}{h_t}\right),\ t = 1, \dots, T \tag{2.10}$$

where K is a kernel, m is the number of observations and h_t is a bandwidth (smoothing parameter or window width). One practically important issue is the selection of an appropriate kernel and the bandwidth (usually selected by cross-validation criteria). We follow Silverman (1986) and use a Gaussian kernel with an optimal bandwidth given by $1.06\hat{\sigma}_t m^{-1/5}$ where $\hat{\sigma}_t$ is the sample standard deviation of r_{it}.[7] Given the sequence of the estimated density functions, $\left(\hat{f}_t\right)_{t=1}^{T}$, we estimate $\mathbb{E}f$ by the sample average of $\hat{f}_t$ such that $\bar{f} = T^{-1}\sum_{t=1}^{T}\hat{f}_t$. Then the sequence of the fluctuation $(\hat{w}_t)_{t=1}^{T}$ is obtained by $\hat{w}_t = \hat{f}_t - \bar{f}$.

We then estimate the autoregressive operator (A) which can be obtained by utilising the autocovariance operators of order 0 (C_0) and 1 (C_1) in infinite dimensions,

$$C_s = \mathbb{E}\left(w_t \otimes w_{t-s}\right),\ s = 0, 1,$$

[7] Various other kernels are also available in the literature including Epanechnikov, Bi-weight, Triangular and Rectangular.

where $\otimes$ denotes a tensor product in the infinite dimensional space.[8] Using the relationship $C_1 = AC_0$, we obtain an autoregressive operator of order one:

$$A = C_0^{-1} C_1. \tag{2.11}$$

Since autocovariance operators are consistently estimated by $\hat{C}_s = (T-1)^{-1} \sum_{t=2}^{T} (\hat{w}_t \otimes \hat{w}_{t-s})$ for $s = 0, 1$, we can consistently estimate the autoregressive operator of order one by $\hat{A} = \hat{C}_0^{-1} \hat{C}_1$. Using the spectral representation for a compact and self-adjoint operator C_0,

$$C_0 = \sum_{\ell=1}^{\infty} \lambda_\ell (v_\ell \otimes v_\ell), \tag{2.12}$$

where (λ_ℓ, v_ℓ) are the pair of eigenvalue and eigenfunction of C_0, the inverse of C_0 can be easily obtained by

$$C_0^{-1} = \sum_{\ell=1}^{\infty} \lambda_\ell^{-1} (v_\ell \otimes v_\ell). \tag{2.13}$$

Since C_0 is defined on an infinite dimensional space in principle, there is an ill-posed inverse problem. In other words, it needs an infinite number of eigenvalues and their corresponding eigenfunctions. To avoid this ill-posed inverse problem, we project A on a finite subspace of $\mathscr{H}$, define V_L as the subspace of $\mathscr{H}$ spanned by the L-eigenfunction, $v_1, \ldots, v_L$, and let $C_{0,L} = \Pi_L C_0 \Pi_L$, where Π_L is the projector on V_L.[9] Then the inverse of C_0 is approximated by

$$C_{0,L}^{+} = \sum_{\ell=1}^{L} \lambda_\ell^{-1} (v_\ell \otimes v_\ell), \tag{2.15}$$

which is defined on V_L. Hence, the estimator of the autoregressive operator on the subspaceV_L of $\mathscr{H}$ is estimated by

$$\hat{A}_L = \hat{C}_{0,L}^{+} \hat{C}_1. \tag{2.16}$$

Under some regularity conditions, Park and Qian (2007, 2011, Theorem 5) show that $\hat{A}_L$ is

[8]It is defined as $(u \otimes v) = \langle v, \cdot \rangle u$ which is equivalent to the outer product uv' in a finite dimensional vector space.

[9]In practice, the choice of L is guided by applying a functional principle component analysis (FPCA) and a cross validation (CV) method. FPCA explains the variation of the fluctuation and CV chooses an optimal dimension $L (\leq L_{max})$ by minimising the following criterion:

$$\sum_{i=1}^{N_{cv}} \left\| \hat{w}_{T-i+1}^{L} - \hat{w}_{T-i+1} \right\|^2 = \sum_{i=1}^{N_{cv}} \int \left[\hat{w}_{T-i+1}^{L}(x) - \hat{w}_{T-i+1}(x) \right]^2 dx, \tag{2.14}$$

where N_{cv} is the number of the last observations used in CV and $\hat{w}_{T-i+1}^{L}$ are the in-sample forecasts of w_{T-i+1} on L-dimensional subspace. We set $L_{max} = 20$ through our study and find that the cross-validation procedure given by (2.14) selects the optimal value of L ranging between 3 and 10.

a consistent estimator (See also Bosq, 2000; Cardot, Mas, and Sarda, 1999; Cardot, Ferraty, Mas, and Sarda, 2003; Cardot, Mas, and Sarda, 2007) and that the forecast errors follow the normal distribution asymptotically (see Theorem 8 in Park and Qian, 2007):

$$\left\|\hat{A}_L - A\right\| \overset{a.s.}{\to} 0 \text{ and } \sqrt{T/L}\left(\hat{A}_L \hat{w}_T - A w_T\right) \overset{d}{\to} \mathbb{N}\left(0, \Sigma\right),$$

where L satisfies the condition $LT^{-1/4}\log T \to 0$ requiring that L should not increase too fast with T and Σ is the covariance operator of ε_t, $\mathbb{E}\left(\varepsilon_t \otimes \varepsilon_t\right)$.

The one-step ahead forecast of the density function is therefore evaluated by

$$\hat{f}_{T+1} = \bar{f} + \hat{A}_L \hat{w}_T. \tag{2.17}$$

Figure 2.1 illustrates the forecasting mechanism of FAR using the equal weighted DJIA portfolio sample. Panel (a) and (b) present the density forecast and its forecasting error. Panel (c) shows that the fluctuation gradually disappears over time. Equivalently, Panel (d) shows that the forecasting mechanism of the FAR is the mean reversion of the density function.

To reduce the required computing time, $\hat{w}_t$ is often approximated by the Fast Fourier Transformation or the Wavelet transformation. Those approximations reduce the computing time substantially through shrinking the dimension of the function.[10] Second, they produce accurate approximations of the tail behaviour so that the VaR estimates can be improved. After obtaining the forecast of the transformed $\hat{w}_{T+1}$, the function is finally inverted to the original one (Antoniadis and Sapatinas, 2003; Besse, Cardot, and Stephenson, 2000; Lee and Ready, 1991).[11]

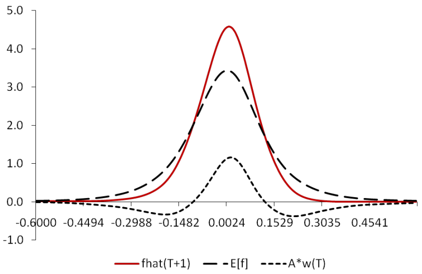
For the purpose of the daily VaR forecast using the intraday density forecast, we first transform the intraday density function to the daily density function by deriving the first four moments of the intraday density function and then construct the daily density function by the Normal Inverse Gaussian (NIG) distribution that is one of the most popular distribution families proposed in the literature (Bandorff-Nielsen, 1997) for describing return distributions. Let μ_t, v_t, s_t, and k_t be, respectively, the mean, variance, skewness, and kurtosis of the intraday density function. Then we calculate four parameters, $(\alpha_t, \beta_t, \gamma_t, \delta_t)$, that determine

[10] It is usually thirty times faster than the non-transformed case in our study.

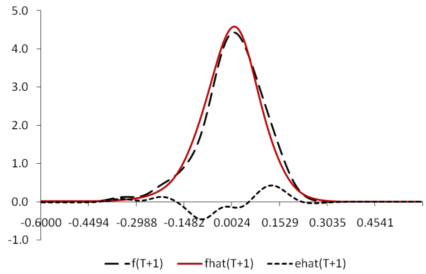
[11] We employ the Fast Fourier Transformation in the empirical evaluation below. The Wavelet transformation provides quantitatively similar results, which are available upon request.

Figure 2.1: Functional Autoregressive Modelling of Non-parametric Intraday Density Function

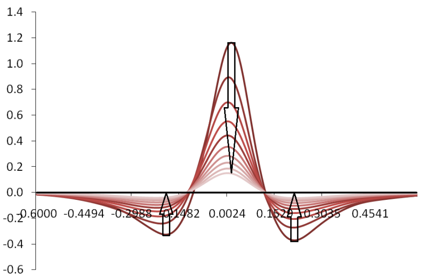
The figures present the intraday return density forecast of the equal weighted DJIA portfolio on the 24th May 2001. 250 past densities are used for the estimation of FAR. Panel (a) presents that the density forecast $\left(\hat{f}_{T+1}\right)$ consists of the mean of the density $\left(\bar{f}\right)$ and the correction$\left(\hat{A}_L w_T\right)$ in (2.17). Panel (b) shows the forecasting error $\left(\hat{\varepsilon}_{T+1}\right)$. Panel (c) shows that the fluctuation $\left(w_T\right)$ gradually disappears as time period increases and is completely gone at the end. Panel (d) shows the mean reversion of the density function such that the density forecast converges to the expectation of the density function as the time period increases.



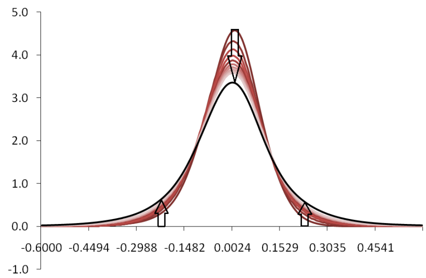
(a) $\hat{f}_{T+1} = \bar{f} + \hat{A}_L w_T$



(b) $f_{T+1} = \hat{f}_{T+1} + \hat{\varepsilon}_{T+1}$



(c) Correction of Fluctuation



(d) Mean Reversion of Density Function

the shape of the NIG distribution using the following formulas

$$\alpha_t = v_t^{-\frac{1}{2}} \left(3k_t - 4s_t^2 - 9\right)^{\frac{1}{2}} \left(k_t - \frac{5}{3}s_t^2 - 3\right)^{-1}, \ \beta_t = s_t v_t^{-\frac{1}{2}} \left(k_t - \frac{5}{3}s_t^2 - 3\right)^{-1},$$
$$\gamma_t = \mu_t - 3 s_t v_t^{\frac{1}{2}} \left(3k_t - 4s_t^2 - 9\right)^{-1}, \ \delta_t = 3^{\frac{3}{2}} \left\{ v_t \left(k_t - \frac{5}{3}s_t^2 - 3\right)\right\}^{\frac{1}{2}} \left(3k_t - 4s_t^2 - 9\right)^{-1}, \tag{2.18}$$

where the kurtosis should satisfy $k_t > 3 + (5/3)\, s_t^2$. Furthermore, α_t determines the tail heaviness, β_t the asymmetry, γ_t the location and δ_t the scale of the distribution. We assume that the intraday returns are strictly stationary on the tth day and the intraday density function can be consistently approximated by the following NIG density function:

$$g_{,t}(x) = \frac{\alpha_t \delta_t J_1 \left(\alpha_t \sqrt{\delta_t^2 + (x-\gamma_t)^2}\right)}{\pi \sqrt{\delta_t^2 + (x-\gamma_t)^2}} \exp\left[\delta_t \lambda_t + \beta_t (x - \gamma_t)\right], \tag{2.19}$$

where $\lambda_t = \sqrt{\alpha_t^2 - \beta_t^2}$ and J_1 denotes the modified Bessel function of the second kind. Under the strict stationarity assumption for the intraday return, the density function of the daily return, $r_t \,(= \sum_{i=1}^m r_{it})$, can be consistently approximated by the NIG density function such that

$$r_t \sim \mathbb{NIG}\left(\alpha_t, \beta_t, m\gamma_t, m\delta_t\right), \tag{2.20}$$

where m is the number of intraday return observations.[12] We then finally obtain the daily VaR from the inverse cumulative density function of NIG distribution:

$$VaR_t(\alpha) = G_t^{-1}(\alpha), \tag{2.21}$$

where $G(\cdot)$ denotes the cumulative density function of the NIG distribution.

In practice, we first calculate the first fourth empirical moments $\hat{\mu}_{T+1}$, $\hat{v}_{T+1}$, $\hat{s}_{T+1}$ and $\hat{k}_{T+1}$ by numerical integration of the intraday density forecast, $\hat{f}_{T+1}$. Then, the four moments are utilised to obtain a daily VaR forecast following the procedures (2.18)–(2.21).

In sum, FAR utilises a non-parametric density function, which overcomes many of the limitations of the parametric approach. Furthermore, imposing an autoregressive structure

[12] The NIG distribution is close under convolution for independent random variables X and Y

$$X \sim \mathbb{NIG}\left(\alpha, \beta, \mu_X, \delta_X\right), Y \sim \mathbb{NIG}\left(\alpha, \beta, \mu_Y, \delta_Y\right) \Rightarrow X + Y \sim \mathbb{NIG}\left(\alpha, \beta, \mu_X + \mu_Y, \delta_X + \delta_Y\right).$$

on the non-parametric density function improves the defect of the static non-parametric approach. Hence, it is expected that FAR can reduce the economic cost and improve accuracy in VaR analysis. Furthermore, it enjoys the rich information of the high-frequency intraday returns which are helpful in forecasting the daily risk. Therefore, FAR would be one of the most efficient and accurate models for VaR analysis.

2.2.2 Alternative Existing Models

We now overview popular VaR models developed by practitioners and academics including HS, FHS, GARCH models, EVT models, and CAViaR.

HS requires no explicit assumption on the asset return distribution. It non-parametrically estimates the empirical distribution from observations given the rolling window that generally ranges from 6 months to two years. Returns in the given window are sorted in ascending order and the α-quantile of interest is given by the return that leaves α percent of the observation on its left side. The rolling approach however implicitly assumes that the distribution does not change within each window. Hence, the empirical quantile estimator is consistent only if the window size goes to infinity under the time-variant condition for the distribution. In practice, the length of the window must satisfy two contradictory properties: it must be large enough in order to make statistical inference significant while it must not be too large to avoid the risk of taking observations outside of the current volatility cluster. VaR estimates based on the HS will underestimate (overestimate) the true VaR when the underlying state is moving from a low (high) to a high (low) volatility cluster. In addition, VaR estimates based on HS may present predictable jumps due to the discreteness of the extreme values.[13]

In order to mitigate the drawbacks of the HS, FHS was introduced (Barone-Adesi, Giannopoulos, and Vosper, 1999, 2002). FHS first standardises the returns using the estimated conditional mean and standard deviation obtained from GARCH models.[14] Then it applies HS to the filtered returns which are independently and identically distributed.[15] Hence, FHS is a kind of hybrid non-parametric model combining HS with GARCH. By absorbing

[13]For example, let the window size be 180 days and suppose there is an extreme observation today. It is easy to predict that VaR jumps upward. The reverse effect will appear after 180 days, when the large observation will drop out of the window.

[14]Even if the underlying distribution is far from Gaussian, a quasi maximum likelihood estimator may be consistent and asymptotically normal (Bollerslev and Wooldridge, 1992; Lee and Hansen, 1994; Lumsdaine, 1996).

[15]Barone-Adesi, Giannopoulos, and Vosper (1999, 2002) and Pritsker (2001) compute VaR from simulated sample paths using draws from the filtered returns. Alternatively, our paper computes VaR by applying the HS to the filtered returns without generating sample paths. The final VaR is computed by $VaR_t(\alpha) = \mu_t + \sigma_t VaR_t^*(\alpha)$, where $VaR_t^*(\alpha)$ is the VaR estimate of the filtered returns.

the advantages of both HS and GARCH, it improves the performance of VaR forecasting (Barone-Adesi, Giannopoulos, and Vosper, 1999, 2002; Pritsker, 2001; Kuester, Mittnik, and Paolella, 2006).

GARCH models simultaneously model a time-varying conditional mean and variance. We consider two GARCH models with differences in the distribution of the error term and the estimation of the first two conditional moments: the RM and GARCH with Student's t-distribution. The RM describes the asset return by the normal distribution. The major criticism of the RM is its symmetric distribution assumption for all asset returns and the lack of a description for the fat-tails of financial time series. Johansson, Seiler, and Michael (1999) find that the RM significantly underestimates the true risk exposures for small and undiversified portfolios. Assuming the normal distribution when its true distribution has a fat-tail will also underestimate the risk of extreme losses (Netftci, 2000). Thus GARCH utilises Student's t-distribution, possessing a fatter tail than that of normal distribution.[16]

EVT models consider probabilities associated with extreme events such as minimum or maximum values given a block size. It is useful for modelling crashes or stress on financial assets (see Embrechts, Klüppelberg, and Mikosch (1997) for a comprehensive review of EVT models). However, it tends to overestimate the VaR in a normal period and faces many challenges including the limited number of extreme data,[17] the choice of parameter estimation and the dynamic specification of the time-dependence on those parameters. Since the distribution of extreme values is not demonstrated by the central limit theorem, we employ non-standard distribution models: the generalised extreme value (GEV) distribution and generalised Pareto distribution (GPD). Their parameters are estimated by the maximum likelihood estimator (MLE) using extreme observations. The estimation is frequently limited by the small number of extreme observations. Filtered EVT (FEVT) models are suggested to control for time-varying volatility (Diebold, Gunter, and Tay, 1998; McNeil and Rudiger, 2000). Analogous to the FHS, it first filters the returns by the GARCH estimates, then applies the standard EVT procedure to the filtered returns. The filtered versions of GEV (FGEV) and GPD (FGPD) are also considered in our paper.

Engle and Manganelli (2004) introduce CAViaR that directly models the evolution of the quantile over time instead of modelling the evolution of the asset return distributions. The advantage of CAViaR is that it relaxes the assumption on the underlying return distribution. Thus it reduces the risk from a misspecification of the parametric distribution.

[16]Recently, the challenge has been how to allow for conditional higher order moments in the GARCH framework, in order to improve the accuracy of VaR forecasting (Bali, Mo, and Tang, 2008).

[17]Since minimum and maximum values are draws in a fixed block size, i.e., 100 days or 180 days, it requires sufficiently large a sample of observations to obtain the relevant size of minimum or maximum values. Hence, there is always a trade-off between the level of extremeness and the number of observations in practice.

Table 2.1: Summary of VaR Models

There are ten VaR models which compete each other in our study. They are categorised by modelling style: (i) parametric approach, (ii) non-parametric approach, (iii) empirical quantile approach and (iv) hybrid non-parametric approach. Our suggested VaR model, FAR, is included in the hybrid non-parametric approach.

Model	Name	Category
(1) FAR	Functional Autoregressive Modelling of Densities	Hybrid approach
(2) FHS	Filtered Historical Simulation	Hybrid approach
(3) HS	Historical Simulation	Non-parametric approach
(4) RM	RiskMetrics	Parametric approach
(5) GARCH	GARCH with t-distribution	Parametric approach
(6) GEV	Generalised Extreme Value Distribution	Parametric approach
(7) GPD	Generalised Pareto Distribution	Parametric approach
(8) FGEV	Filtered Generalised Extreme Value Distribution	Parametric approach
(9) FGPD	Filtered Pareto Distribution	Parametric approach
(10) CAViaR	Conditional Autoregressive Quantile	Empirical quantile approach

According to a recent study of alternative VaR strategies (Kuester, Mittnik, and Paolella, 2006), the FEVT models and the FHS outperform GARCH and CAViaR. The inadequate performance of GARCH and CAViaR could be attributed to a misspecification in the time-dependence of the distribution's parameters and quantiles as well as the choice of the distribution family. On the other hand, the outperformance of the filtered approaches associated to the FHS and the FEVT models suggests the advantages of a hybrid approach. In this context, we expect FAR to enjoy a similar advantage since it models the time-varying density function in the functional space without any parametric specifications on the distribution and the autoregressive operator contains the complicated contemporaneous and time-dependent structure on the moments and locations of the distribution in a reduced form. The models considered are summarised in Table 2.1.

2.3 Backtestings

In this section, we summarise the backtesting procedures, which are organised into two general headings: quantitative and statistical. The former intuitively evaluate the accuracy of the VaR models whilst the latter provide the formal statistical tests for both unconditional and conditional accuracy.

Before unfolding the backtesting procedures, we preliminarily define a violation of a VaR forecast. Given the nominal coverage probability, α, a violation is defined by the indicator function

$$H_s = \mathbf{1}\left\{r_s < \widehat{VaR}_s(\alpha)\right\},\ s = 1,\ldots,N, \tag{2.22}$$

where $\widehat{VaR}_s(\alpha)$ is the VaR forecast given the information set available at $s-1$ ($\mathscr{F}_{s-1}$) with the nominal coverage probability, α. An accurate VaR forecast thus satisfies

$$\mathbb{E}\left[H_s|\mathscr{F}_{s-1}\right] = \alpha, \tag{2.23}$$

which implies that H_s is independent of any function of the variables in $\mathscr{F}_{s-1}$. Hence, accuracy implies that the conditional binomial process (H_s) on the information set available at $s-1$ should follow the IID Bernoulli distribution (Lemma 1; Christoffersen, 1998):

$$H_s|\mathscr{F}_{s-1} \sim iid\ Bernoulli\left(\alpha\right), \tag{2.24}$$

where $Var\left(H_s|\mathscr{F}_{s-1}\right) = \alpha\left(1-\alpha\right)$.

2.3.1 Quantitative Evaluations

The empirical conditional coverage probability is calculated by the sample average of the N violations of H_s, that is $\hat{\alpha} = N^{-1}\sum_{s=1}^{N} H_s$ which is the consistent estimator of the conditional coverage probability under the true forecasting model assumption. Hence, the regulator prefers a VaR model with a empirical conditional coverage probability that is closest to its nominal value. Since the empirical conditional coverage probability is random, its significance needs to be tested through formal statistical tests which will be introduced in the next subsection.

The Basel penalty zone is found in BCBS (1996). It describes the strength of an internal model through the test of failure rate, which records the number of daily violations of the 99 percent VaR in the previous 250 business days. One may expect, on average, 2.5 violations out of the previous 250 VaR forecasts under the correct forecasting model. The Basel Committee rules that up to four violations are acceptable for banks and defines the range as a “Green” zone. If the number of violations is five or more, the banks fall into a “Yellow” (5–9) or “Red” (10+) zone, where the penalty is cumulatively imposed on the bank by the multiplicative factor (κ) from 3 to 4.[18] It is used for calculating the market risk capital requirement. If a bank falls into “Red” zone, the penalty is automatically generated. Whereas, if a bank is in “Yellow”, the supervisor will decide the penalty depending on the reason for the violation. Jorion (2006, p. 149) summarises the categories for the reasons suggested by the Basel Committee.[19]

The MRCR (market risk capital requirement) originated from the Basel II Accord. Ac-

[18] The multiplicative factor corresponds to the number of violation: 0–4 (3), 5 (3.4), 6 (3.5), 7 (3.65), 8 (3.75), 9 (3.85) and 10 (4), respectively.

[19] There are four categories: (i) Basic integrity of the model, (ii) Model accuracy could be improved, (iii) Intraday trading, (iv) Bad luck.

cording to the applications of the Basel II Accord, it requires the minimum capital requirement for market risk as well as credit risk and operating risk. The Basel II Accord provides two approaches to measure the MRCR: (i) Standardised approach and (ii) Internal models approach. We choose the second approach to evaluate the VaR forecast. Providing that a bank satisfies the qualitative requirements (the bank has a sound risk management system and an independent risk-control unit as well as external audits), the MRCR is summarised by the following four factors: (i) quantitative parameters, (ii) treatment of correlations, (iii) market risk charge, and (iv) plus factor.[20] The MRCR thus can be formulated by

$$MRCR_s = \max\left(\kappa \frac{1}{60}\sum_{i=1}^{60} VaR_{s-i}(\alpha), VaR_{s-1}(\alpha)\right) + SRC_s,\ s = 251, \ldots, N, \tag{2.25}$$

where SRC is the additional capital charge for the specific risk (BCBS, 1996, 2004) but it is ignored in our evaluation, since it is assumed to be a common value applied to all VaR models. κ is the Basel penalty factor. A model carrying the minimum MRCR is preferred by the regulator.

Predictive quantile loss evaluates a VaR forecast in terms of economic cost. We employ the predictive quantile loss using the "check" function of Koenker and Bassett (1978) which can be regarded as a "predictive" quasi-likelihood (Bertail, Haefke, Politis, and White, 2004; Komunjer, 2004). The expected loss of VaR for a given α is

$$Q_s(\alpha) = \mathbb{E}\left[(\alpha - H_s)\left(r_s - \widehat{VaR}_s(\alpha)\right)\right], \tag{2.26}$$

which can provide the measure of the lack-of-fit for a quantile model and be interpreted as the economic cost to carry a VaR model. The expected check function, $Q(\alpha)$, is estimated from the sequence of N violation indicates (H_s):

$$\hat{Q}(\alpha) = \frac{1}{N}\sum_{s=1}^{N}(\alpha - H_s)\left(r_s - \widehat{VaR}_s(\alpha)\right). \tag{2.27}$$

A model providing the minimum value of $\hat{Q}(\alpha)$ is preferable from the regulator's point of view.

2.3.2 Statistical Evaluations

The unconditional coverage test was first developed by Kupiec (1995). If we assume that VaR forecasts are independent over time, the violations of a VaR forecast can be regarded

[20] See BCBS (1996, 2004) for details and Jorion (2006) for a compact summary.

as the realisation of an independent binomial random variable whose probability of a realised return's exceeding the corresponding VaR forecast is equal to the nominal coverage probability, α. Therefore, accurate VaR forecasts should satisfy the requirement that their empirical coverage probability equals the nominal coverage probability. The test thus describes the null and alternative hypotheses by

$$\mathbb{H}_0 : \mathbb{E}[H_s] = \alpha \text{ against } \mathbb{H}_1 : \mathbb{E}[H_s] \neq \alpha. \tag{2.28}$$

Under the null hypothesis, (2.24) can be employed for constructing the likelihood ratio test statistic

$$LR_{uc} = 2[\ell(\hat{\alpha}; H_1, \ldots, H_N) - \ell(\alpha; H_1, \ldots, H_N)] \overset{d}{\sim} \chi^2_1, \tag{2.29}$$

where $\ell(\cdot)$ denotes the log-likelihood for the Bernoulli distribution such that

$$\begin{aligned} \ell(\hat{\alpha}; H_1, \ldots, H_N) &= N_1 \log \hat{\alpha} + N_0 \log(1 - \hat{\alpha}), & (2.30) \\ \ell(\alpha; H_1, \ldots, H_N) &= N_1 \log \alpha + N_0 \log(1 - \alpha). & (2.31) \end{aligned}$$

The MLE of $\hat{\alpha}$ is the ratio of the number of violation, N_1, to the total number of observation, $N_0 + N_1 = N$, that is, $\hat{\alpha} = N_1/(N_0 + N_1)$. Kupiec (1995) however reports that this test has a low power to distinguish among alternative hypotheses, even in sufficiently large samples. Hence, the regulator would have the low confidence in the bank's VaR report when its internal model can not be rejected by this test.

In the presence of time-dependence, volatility clustering and persistence are often found in financial time series. Therefore, the conditional accuracy of VaR forecasts is highly important. If a VaR model ignores such underlying state, its VaR estimates may have an incorrect conditional coverage probability whilst its unconditional coverage probability is still correct.[21] The regulator thus requires that the VaR model exhibit an accurate coverage probability regardless of the underlying state. In order to test the conditional accuracy of a VaR forecast, Christoffersen (1998) has developed the conditional coverage test by combining the unconditional coverage test and the independence test: resulting log-likelihood ratio test statistic is the summation of two log likelihood ratio test statistics:

$$LR_{cc} = LR_{uc} + LR_{ind} \overset{d}{\sim} \chi^2_2, \tag{2.32}$$

where LR_{uc} and LR_{inc} are the log likelihood ratio test statistics of the unconditional coverage

[21] An HS ignores the time-dependent structure of the underlying probability law of financial time series. It is frequently rejected for the independence test whilst it provides an accurate unconditional coverage by the unconditional coverage test.

test and the independence test, respectively. The independence test utilises the sequence of violation indicates (H_s) as a binary first-order Markov chain with transition probability matrix,

$$\Pi = \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix}, \tag{2.33}$$

where $\pi_{ij} = \mathbb{P}(H_s = j | H_{s-1} = i)$ for $i, j = 0, 1$. Under the null hypothesis of independence, the likelihood ratio test statistic is given by

$$LR_{ind} = 2\left[\ell\left(\hat{\Pi}; H_2, \ldots, H_N | H_1\right) - \ell\left(\hat{\pi}_1; H_2, \ldots, H_N | H_1\right)\right] \overset{d}{\sim} \chi_1^2. \tag{2.34}$$

The first joint log likelihood conditional on the first observation is

$$\ell(\Pi; H_2, \ldots, H_N | H_1) = N_{00} \log \pi_{00} + N_{01} \log \pi_{01} + N_{10} \log \pi_{11} + N_{11} \log \pi_{11}, \tag{2.35}$$

where N_{ij} denotes the number of transitions from state i to j, $N_{ij} = \sum_{s=2}^{N} \mathbf{1}(H_s = j | H_{s-1} = i)$, and the MLEs of π_{01} and π_{11} under the alternative hypothesis are $\hat{\pi}_{01} = N_{01}(N_{00} + N_{01})^{-1}$ and $\hat{\pi}_{11} = N_{11}(N_{10} + N_{11})^{-1}$, respectively. Under the null hypothesis of independence, we have $\pi_{01} = \pi_{11} \equiv \pi_1$, from which we obtain the second conditional log likelihood as follows:

$$\ell(\pi_1; H_2, \ldots, H_N | H_1) = (N_{00} + N_{10}) \log(1 - \pi_1) + (N_{01} + N_{11}) \log \pi_1, \tag{2.36}$$

where the MLE of π_1 is $\hat{\pi}_1 = (N_{01} + N_{11})(N_{00} + N_{10} + N_{01} + N_{11})^{-1}$. Hence, the LR test statistic in (2.32) is alternatively expressed by combining (2.29) and (2.34):

$$LR_{cc} = 2\left[\ell\left(\hat{\Pi}; H_2, \ldots, H_N | H_1\right) - \ell(\alpha; H_2, \ldots, H_N | H_1)\right] \overset{d}{\sim} \chi_2^2, \tag{2.37}$$

where the second log likelihood conditional on the first observation is

$$\ell(\alpha; H_2, \ldots, H_N | H_1) = (N_{00} + N_{10}) \log(1 - \alpha) + (N_{01} + N_{11}) \log \alpha. \tag{2.38}$$

Therefore, the regulator requires that the accuracy of VaR forecasts reported by a bank's internal model should be evaluated by the conditional coverage test rather than the single unconditional coverage test or the independence test. If the VaR model is rejected by the conditional coverage test, the regulator would recommend that the bank improve its VaR model to eliminate the time-dependent coverage ability in VaR forecasts. The proposed test however has a limitation, since it allows only the time-dependence of the first lag-order in the sequence of violation indicators (Gaglianone, Lima, Linton, and Smith, 2009). The testing procedure may be extended by allowing more lag orders.

The dynamic quantile test (Engle and Manganelli, 2004) is a general extension of the conditional coverage test allowing for more time-dependent information of $(H_s)_{s=1}^N$. It regresses H_s on some carefully selected explanatory variables in $\mathscr{F}_{s-1}$ such that

$$H_s = \alpha_0 + \sum_{i=1}^{p} \beta_i H_{s-i} + \beta_{p+1} \widehat{VaR}_s(\alpha) + u_s,\ s = p+1, \ldots, N. \tag{2.39}$$

Under the null hypothesis, the explanatory variables should have no power on H_s, that is

$$\mathbb{H}_0 : \alpha_0 = \alpha \text{ and } \beta_i = 0 \text{ for } i = 1, \ldots, p+1. \tag{2.40}$$

By subtracting α from both sides of (2.39), we have

$$(H_s - \alpha) = (\alpha_0 - \alpha) + \sum_{i=1}^{p} \beta_i H_{s-i} + \beta_{p+1} \widehat{VaR}_s(\alpha) + u_s,\ s = p+1, \ldots, N. \tag{2.41}$$

The simple vector notation of (2.41) is given by

$$y_s = Z_s \beta + u_s,\ s = p+1, \ldots, N, \tag{2.42}$$

where $y_s = H_s - \alpha$, $Z_s = \left(1, H_{s-1}, \ldots, H_{s-p}, \widehat{VaR}_s(\alpha)\right)'$ and $\beta = (\alpha_0 - \alpha, \beta_1, \ldots, \beta_{p+1})$. We can therefore express the null hypothesis in (2.40) as

$$\mathbb{H}_0 : \beta = 0. \tag{2.43}$$

Under the null hypothesis, the least squares estimator converges to the normal distribution such that

$$\hat{\beta} = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1} \mathbf{Z}'\mathbf{y} = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1} \mathbf{Z}'(\mathbf{H} - \alpha\iota) \overset{d}{\sim} N\left(\mathbf{0}, \left(\mathbf{Z}'\mathbf{Z}\right)^{-1} \alpha(1-\alpha)\right). \tag{2.44}$$

Then we can deduce the test statistics from Engle and Manganelli (2004) by

$$DQ = \frac{\hat{\beta}'\mathbf{Z}'\mathbf{Z}\hat{\beta}}{\alpha(1-\alpha)} \overset{d}{\sim} \chi^2_{p+2}. \tag{2.45}$$

The Monte Carlo simulation study by Berkowitz, Christofferse, and Pelletier (2011) shows that the dynamic quantile test appears to be the best backtest for a 99% VaR whilst other tests generally have much lower power against incorrect VaR models.[22]

[22] Berkowitz, Christofferse, and Pelletier (2011) make the criticisism that estimating the autoregression (2.42)

In summary, regulators require banks to evaluate the accuracy of VaR forecasts by the bank's internal models using statistical tests. There are two popular tests: the conditional coverage test (Christoffersen, 1998) and the dynamic quantile tests (Engle and Manganelli, 2004). The VaR models in our study are therefore mainly evaluated by these two conditional tests. Since VaR forecasts are generally evaluated in very small coverage probability (1%), testing power is usually less than those for the central statistics such as mean or variance. Hence it would be undesirable to evaluate the VaR model at a lower significance level (e.g., 1%). The regulator should place more confidence in VaR models not rejected by the conditional coverage test at a more restrictive significance level.

2.4 Data

We choose the 30 companies of the DJIA as the sample of this study. This Wall Street benchmark index contains the 30 largest stocks listed in the US market. These corporations are deemed to be too big to fail even during the financial crisis. Among these companies, three of them were rescued by the U.S. government in 2008.[23] Understanding the risk profiles of these 30 companies is essential for monitoring and safeguarding the financial system. Furthermore, unlike small companies, these companies are actively traded in a trading section which generates enormous amount of information. It is an interesting empirical question to ask whether the financial market revealed any insight regarding these company's risks during the intraday trading.

The list of components changes annually, and we choose the list in year 2005, because this year is in the middle of our whole sample period, which is from 2000 to 2008. Intraday transaction data are collected from Trade and Quote (TAQ) dataset. Following traditional studies which deal with intraday data, such as Lee and Ready (1991) and Hvidkjaer (2006), a filtering procedure has been applied to exclude any data likely to be erroneous. Specifically, all the trades (quotes) with condition codes, A, C, D, G, L, N, O, R, X, Z, 8, 9 (4, 5, 7–9, 11, 13–17, 19, 20) are removed from the sample. The trades with a correction code which is greater than 2 are also removed from the sample.[24] Quotes are excluded if ask is equal

by ordinary least squares can not completely treat heteroscedasticity for valid inference since the sequence of violations is binary. They instead recommend estimating a logit model by assuming that the error term u_s has a logistic distribution and then testing the null hypothesis by a likelihood ratio test. Another possibility is a duration based approach developed by Christoffersen and Pelletier (2004) and further by Berkowitz, Christofferse, and Pelletier (2011).

[23]Bailouts for American International Group, Inc. and Citi group Inc. were announced by the Federal Reserve Bank's Board of Governors. Though the case of General Motors Corporation is not technically a bailout, a bridge loan was given to the auto manufacturers by the U.S. government in 2008, which is referred to by most as a bailout.

[24]For the definitions of the codes, refer to the TAQ manual.

or less than bid, or bid–ask spread is above 75% of mid-quote, or ask (bid) is more than double or less than half of the previous ask. Only trades reported from 9:30 AM to 4:00 PM are included. According to Lee and Ready (1991), trades happen five seconds earlier than reported time. Therefore, trade times are calculated as the reported time minus 5 seconds. Trades are deleted if the trade price is more than double or less than half of the previous trade. After filtering data, from the opening time 9:30 AM of the exchange, total trading volume, close trade and quote prices of every 5 minutes have been calculated based on the intraday transaction data. Close trade and quote prices in intervals are defined as the price of the last trade in intervals and corresponding quote price when trade happens. If there is no trade happening in the interval, the closest trade and quote prices of the previous interval will be used as the one for current interval. In addition, a portfolio return is constructed by the equal weight of the 30 company's returns. It is used for the proxy of the Dow Jones Industrial Average index. A detailed data description is provided by Table 2.2.

Table 2.3 presents the descriptive statistics for the intraday returns and the daily returns for the 30 companies and their equally weighted portfolio. In Panel A, the statistics for the intraday return are obtained by averaging the intraday statistics from 2000 to 2008. There are 78 five-minute-interval return observations in each day. The mean and median are very close to zero which is consistent with the previous market microstructure literature. The maximum return (0.92%) is observed in Alcoa Inc. and the minimum return (-0.94%) is observed in Intel Corporation. These statistics suggest the sudden increase and decrease of 1% in a five minute interval. The standard deviation is also biggest in Intel Corporation. The asymmetry is insignificant since skewness is small, close to zero. On the contrary, the kurtosis is significantly larger than that of the normal distribution: this describes the typical fat-tail of a financial time series.

In Panel B, the statistics for the daily return are obtained from 2,263 daily observations whilst those for Hewlett-Packard Company, AT&T, and Verizon Communication Inc. are from 1,678, 2,256 and 2,173 observations, respectively. The mean returns are distributed around zero. The means of American International Group, Inc. (-0.17%) and Citigroup (-0.08%) are relatively smaller than the others which is not a surprise given that they needed rescues by the government in 2008. The maximums (minimums) and standard deviations of those companies are also relatively much larger (smaller) than others. The asymmetry of the return distribution is generally not serious, whilst AIG (-6.49) and The Procter & Gamble Company (-5.17) are very negatively skewed. A fat-tail is also strongly observed for all companies. In particular, the kurtosis of AIG (158.5), Citigroup (47.82) and the Procter & Gamble Company (120.3) are significantly bigger than the others. The return of the portfolio shows the average property in both intraday and daily return except for

Table 2.2: Data Description for 30 Components of DJIA and Portfolio

30 companies of DJIA are collected from Trade and Quote (TAQ) and the portfolio is constructed by equally weighting their returns.

	Company	From	To	Open	Close
(1)	Alcoa Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(2)	American International Group, Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(3)	American Express Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(4)	The Boeing Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(5)	Citi group Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(6)	Caterpillar Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(7)	E.L. Du Pont de Nemours & Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(8)	Walter Disney Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(9)	General Electric Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(10)	General Motors	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(11)	Home Depot, Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(12)	Honeywell International Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(13)	Hewlett-Packard Company	06-May-02	31-Dec-08	9:30 AM	4:00 PM
(14)	International Business Machine Corp.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(15)	Intel Corporation	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(16)	Johnson & Johnson	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(17)	J.P. Morgan Chase & Co.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(18)	The Coca-Cola Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(19)	McDonald's Corporation	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(20)	3M Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(21)	Altria Group Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(22)	Merck & Co, Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(23)	Microsoft Corporation	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(24)	Pfizer, Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(25)	The Procter & Gamble Company	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(26)	AT&T	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(27)	United Technology Corporation	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(28)	Verizon Communication Inc.	03-Jul-00	31-Dec-08	9:30 AM	4:00 PM
(29)	Wal-Mart Stores, Inc.	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
(30)	Exxon Mobil Corporation	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM
	Portfolio	03-Jan-00	31-Dec-08	9:30 AM	4:00 PM

the standard deviation. The standard deviation of the portfolio is least among the returns which reflects the diversification effect of the equal weighted portfolio. Overall, the returns present the typical nature of financial time series and some extreme statistics are observed especially for the bailout companies in the 2008 global crisis.

2.5 Evaluations

We evaluate the forecasting performance with a simulated sample, 30 individual stocks and their equal weighted portfolio. First, we evaluate the density forecasting of functional models such as FAR, AVE and LAST based on divergence criteria. This evaluation investigates which functional model is the best predictor of the density. Second, we evaluate the forecasting precision of the VaR models through a Monte Carlo simulation study in which we generate a sample based on a specification from empirical findings and we evaluate the forecasting precision by comparing a VaR forecast with a true VaR. Third, we evaluate the out-of-sample performance of VaR forecasting for alternative VaR models with 30 individual stocks and the portfolio. In the first part, the evaluations are performed for long and short positions with a window size of 250 business days, that is, we investigate the left-tail and the right-tail extreme events of the return distribution. Hence, this provides VaR forecasting evaluations for buying and selling activities. In the second part, we provide the effect of a longer window size on the performance of VaR forecasting, using for a 500 (business days) window size. These evaluations are also performed for long and short positions. Hence, it will better support our empirical evaluation results at the end.

2.5.1 The Density Forecasting of Alternative Functional Models

We first evaluate the density forecasting of alternative functional models; FAR, AVE and LAST—see (2.8) and (2.9). This evaluation demonstrates which functional model is the best predictor of the intraday density function using 30 individual stocks and portfolio. To this end we employ divergence criteria that measure the distance between the forecasted and the true density function; namely, the Hilbert norm (D_H), the uniform norm (D_U) and the generalised entropy (D_E). D_H and D_U are given by

$$D_H\left(\hat{f},f\right) = \frac{\int\left(\hat{f}(x)-f(x)\right)^2 dx}{\int\hat{f}(x)^2 dx+\int f(x)^2 dx} \text{ and } D_U\left(\hat{f},f\right) = \frac{\sup_x\left|\hat{f}_t(x)-f_t(x)\right|}{\sup_x f_t(x)}, \tag{2.46}$$

Table 2.3: Descriptive Statistics of Intraday Return and Daily Return for 30 Components of DJIA and Portfolio

Panel A is the average statistics of the intraday statistics. We assume that the intraday return is identically and independently distributed in the given trading day whilst the intraday return distribution is varies over time. We thus first calculate the intraday statistics for each day then we average the intraday statistics. The statistics for the daily returns are calculated using 2,263 trading days, except Hewlett-Packard Company, AT&T and Verizon Communication Inc. are calculated using 1678, 2256 and 2137 trading days, respectively. The statistics present the number of observations (T, N), mean (Mean), median (Med.), maximum (Max.), minimum (Min.), standard deviation (Stdev.), skewness (SK) and kurtosis (K) separately for the intraday return and the daily return.

	Panel A: Intraday Return								Panel B: Daily Return							
Company	T	Mean	Med.	Max.	Min.	Stdev.	SK	K	N	Mean	Med.	Max.	Min.	Stdev.	SK	K
(1)	78	0.00	0.00	0.92	-0.88	0.26	0.14	9.13	2263	-0.06	-0.03	20.88	-18.16	2.71	-0.06	10.77
(2)	78	0.00	0.00	0.83	-0.86	0.24	0.11	8.66	2263	-0.17	-0.06	53.90	-86.22	3.99	-6.49	158.5
(3)	78	0.00	0.00	0.77	-0.76	0.23	0.12	8.59	2263	-0.04	-0.02	16.49	-17.31	2.41	-0.37	9.26
(4)	78	0.00	0.00	0.72	-0.71	0.21	0.12	8.95	2263	0.00	0.03	14.53	-19.58	2.10	-0.33	9.37
(5)	78	0.00	0.00	0.86	-0.80	0.24	0.18	8.89	2263	-0.08	-0.05	47.00	-32.08	2.97	0.58	47.82
(6)	78	0.00	0.00	0.74	-0.73	0.22	0.12	8.82	2263	0.03	0.04	11.66	-15.33	2.08	-0.21	7.28
(7)	78	0.00	0.00	0.66	-0.66	0.20	0.07	8.16	2263	-0.04	-0.06	9.82	-12.27	1.88	-0.22	8.96
(8)	78	0.00	0.00	0.74	-0.78	0.22	-0.03	8.67	2263	-0.01	0.00	14.31	-20.12	2.22	-0.05	11.87
(9)	78	0.00	0.00	0.66	-0.65	0.19	0.12	8.67	2263	-0.05	-0.06	12.67	-13.88	2.03	-0.22	9.47
(10)	78	0.00	0.00	0.58	-0.57	0.18	0.12	7.89	2263	0.03	0.09	16.79	-15.92	1.79	0.15	14.63
(11)	78	0.00	0.00	0.79	-0.78	0.23	0.06	8.94	2263	-0.05	-0.05	12.79	-33.52	2.40	-1.11	22.44
(12)	78	0.00	0.00	0.84	-0.81	0.24	0.12	9.09	2263	-0.03	-0.01	19.42	-20.55	2.38	-0.48	13.30
(13)	78	0.00	0.00	0.79	-0.81	0.23	-0.05	10.1	1678	0.04	0.02	13.74	-16.21	2.25	0.13	9.91
(14)	78	0.00	0.00	0.65	-0.61	0.19	0.14	9.10	2263	-0.01	0.00	12.26	-17.01	1.94	-0.05	10.35
(15)	78	0.00	0.00	0.96	-0.94	0.28	0.13	9.66	2263	-0.05	0.00	18.20	-24.81	2.95	-0.49	9.49
(16)	78	0.00	0.00	0.49	-0.49	0.14	0.06	8.44	2263	0.01	0.00	8.94	-17.71	1.43	-0.84	18.52
(17)	78	0.00	0.00	0.84	-0.83	0.25	0.07	8.97	2263	0.00	0.00	16.95	-20.14	2.64	0.03	11.91
(18)	78	0.00	0.00	0.51	-0.54	0.16	-0.05	7.92	2263	-0.01	-0.02	9.94	-10.51	1.54	0.08	9.12
(19)	78	0.00	0.00	0.66	-0.67	0.20	0.05	8.41	2263	0.02	0.03	8.88	-13.20	1.78	-0.19	7.61
(20)	78	0.00	0.00	0.57	-0.54	0.17	0.13	7.84	2263	0.01	-0.03	9.99	-9.42	1.59	0.19	7.99
(21)	78	0.00	0.00	0.64	-0.59	0.18	0.18	8.57	2263	0.04	0.06	15.12	-14.37	1.88	-0.02	12.94
(22)	78	0.00	0.00	0.66	-0.69	0.19	0.04	9.10	2263	-0.04	0.00	11.84	-31.14	2.05	-1.83	31.32
(23)	78	0.00	0.00	0.72	-0.68	0.21	0.12	8.51	2263	-0.05	-0.04	17.77	-16.94	2.25	-0.08	11.07
(24)	78	0.00	0.00	0.65	-0.65	0.19	0.09	8.64	2263	-0.03	-0.04	9.48	-11.46	1.81	-0.24	7.34
(25)	78	0.00	0.00	0.48	-0.51	0.15	-0.05	7.61	2263	0.01	0.02	9.83	-37.44	1.66	-5.19	120.3
(26)	78	0.00	0.00	0.78	-0.79	0.23	0.04	8.90	2256	-0.05	-0.05	22.18	-21.00	2.44	0.28	14.90
(27)	78	0.00	0.00	0.65	-0.62	0.19	0.12	8.25	2263	0.03	0.03	12.31	-15.75	1.91	-0.16	9.55
(28)	78	0.00	0.00	0.66	-0.66	0.20	0.03	8.54	2137	-0.02	-0.02	13.28	-12.22	1.88	0.17	8.28
(29)	78	0.00	0.00	0.63	-0.62	0.19	0.11	8.43	2263	-0.01	-0.03	10.20	-8.31	1.79	0.26	6.88
(30)	78	0.00	0.00	0.58	-0.57	0.18	0.12	7.89	2263	0.03	0.09	16.79	-15.92	1.79	0.15	14.63
Portfolio	78	0.00	0.00	0.39	-0.39	0.12	0.17	9.43	2263	-0.02	0.02	10.95	-16.45	1.39	-0.75	18.66

Table 2.4: Forecasting Performance of Alternative FAR models by Three Divergence Criteria

D_H, D_U and D_E denote the Hilbert norm, uniform norm and entropy, respectively. Figures are evaluated by the mean and the median value between one-step-ahead density forecast and the true density for each of the three models.

	Mean			Median		
Models	D_H	D_U	D_E	D_H	D_U	D_E
Panel A: Average of 30 Components of Dow Jones Industry						
FAR	0.0284	0.2664	0.0406	0.0160	0.2044	0.0304
AVE	0.0473	0.3812	0.0601	0.0235	0.2493	0.0393
LAST	0.0353	0.2872	0.0557	0.0218	0.2335	0.0431
Panel B: Portfolio						
FAR	0.0276	0.2634	0.0465	0.0150	0.1992	0.0344
AVE	0.0568	0.4438	0.0916	0.0246	0.2534	0.0552
LAST	0.0320	0.2706	0.0600	0.0203	0.2245	0.0470

where $\hat{f}$ denotes the density forecast and f the true density. Following Ullah (1996), we define D_E by

$$D_E\left(\hat{f}, f\right) = \int f(x)\, g\left(\frac{\hat{f}(x)}{f(x)}\right) dx, \tag{2.47}$$

where $g(y) = (\gamma - 1)^{-1}(y^{\gamma} - 1)$ with $\gamma > 0$ and $\gamma \neq 1$. We follow Park and Qian (2007) and set $\gamma = 1/2$. If g is the natural log, this becomes the Kullback–Liebler divergence measure. All three quantities are non-negative and give a zero value if $\hat{f}_t = f_t$ and so we may call them global errors. D_H is useful for evaluating the goodness-of-fit of the model, D_U is informative for comparing the closeness of the function shape, and D_E assesses the difference in information content between the forecast and the true density function.

For the 30 stocks and the portfolio, we apply the rolling one-step-ahead forecast of the intraday density function based on the 250 window size from 03/01/2001 to 31/12/2008 for three functional models. In Table 2.4, the mean and median divergence measures are calculated based on the 2,011 daily forecasts (except for Hewlett-Packard, AT&T and Verizon Communication Inc., which are calculated using 1,416, 2,004 and 1,885 forecasts, respectively). Panel A presents the average divergence values of 30 stocks and Panel B presents the divergence value of the portfolio. The FAR has the minimum values for the three divergence criteria in mean and median in Panels A and B. The LAST has smaller values than the AVE for the three divergence measures in mean and median. The results show that FAR outperforms other functional models in forecasting the intraday density function, demonstrating that FAR is the best predictor of the intraday density function in our sample.

2.5.2 The Precision of VaR Forecasting

In this section, we evaluate the precision of the VaR forecasting by comparing it with the true VaR via Monte Carlo simulation. To this end, we make the following assumptions on the data generating process:

Assumption 1. Intraday returns are independently and identically distributed.
Assumption 2. The distribution of intraday returns is determined by the first four moments such that it is a member of NIG family.
Assumption 3. The time-dependence of the first four moments of the intraday density function follows a stationary vector autoregressive process of lag order 1.
Assumption 4. The daily return is the sum of the intraday returns and follows a NIG distribution.

First, we construct the time-varying four moments of the intraday density function by a stationary VAR(1) process

$$\mathbf{m}_t - \mathbf{m} = \Phi(\mathbf{m}_{t-1} - \mathbf{m}) + \xi_t,\ t = 1, \ldots, T, \tag{2.48}$$

where $\mathbf{m}_t = (\mu_t, v_t, s_t, k_t)'$, $\mathbf{m} = (\mu, v, s, k)'$ and Φ is 4×4 autoregressive coefficient matrix satisfying $\left\|\Phi^k\right\| < 1$ for any $k \geq 1$. $\xi_t = (\xi_{1t}, \xi_{2t}, \xi_{3t}, \xi_{4t})'$ is independently and normally distributed with zero mean and covariance such that

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}.$$

In practice, we generate the sample by

$$\mathbf{m}_t = \mathbf{c} + \Phi \mathbf{m}_{t-1} + \xi_t, \tag{2.49}$$

where $\mathbf{c} = (\mathbf{I} - \Phi)\mathbf{m}$ and $\mathbf{I} - \Phi$ is invertible. Furthermore, the moments must satisfy the following conditions (see the definition of NIG distribution):

$$v_t > 0 \text{ and } k_t > \frac{5}{3} s_t^2. \tag{2.50}$$

Let $\mathbf{m}_0 = \mathbf{0}$ and generate the first observation by

$$\mathbf{m}_1 = \mathbf{c} + \xi_1. \tag{2.51}$$

If the sample does not satisfy (2.50), then we regenerate it until it does. Other samples are sequentially generated by (2.49) in the same manner. We discard the first 100 observations to reduce the effect of the initial values.

Next, we calculate the four parameters $(\alpha_t, \beta_t, \gamma_t, \delta_t)'$, that determine the NIG distribution, using the four moments $(\mu_t, v_t, s_t, k_t)'$. Then we draw an intraday return from the NIG distribution and let $X_t := (r_{ti})_{i=1}^{m}$ and $X := (X_t)_{t=1}^{T+N}$ (See (2.3) for X_t and X). Under Assumptions 1 and 4, an NIG distribution for a daily return $(r_t = \sum_{i=1}^{m} r_{it})$ is determined by four parameters $(\alpha_t, \beta_t, m\gamma_t, m\delta_t)'$. Hence, we can calculate a true daily VaR for $(1-\alpha)\%$ from an inverse cumulative density function corresponding to α probability, $G_t^{-1}(\alpha)$.

We iterate the experiment 5,000 times and evaluate the precision of the VaR forecast for FAR and alternative models by comparing the VaR forecast and the true VaR. We employ Bias, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Percentage Mean Absolute Deviation (PMAD) as precision measures.[25]

The simulation uses a 250 window size $(T = 250)$ and one-step ahead forecast $(N = 1)$. We choose parameter values for $\mathbf{m}$, Φ and Σ by estimating (2.49) for the 30 stocks and the portfolio. Out of 31 estimations, we focus on three representative cases. We first consider the aggregate information of 30 stocks and select the portfolio that presents an equal weight composite of the 30 stocks, and obtain the following specifications for $\mathbf{m}$, Φ and Σ:

$$\mathbf{m} = \begin{bmatrix} -0.0003 \\ 0.0212 \\ 0.1666 \\ 9.4345 \end{bmatrix}, \Phi = \begin{bmatrix} -0.0495 & 0.0075 & 0.0000 & 0.0000 \\ -0.3850 & 0.5546 & 0.0019 & -0.0006 \\ -2.7238 & -0.3703 & -0.0623 & 0.0011 \\ 34.0111 & 12.2691 & -0.3776 & -0.0027 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 0.0003 & 0 & 0 & 0 \\ 0 & 0.0023 & 0 & 0 \\ 0 & 0 & 3.2681 & 0 \\ 0 & 0 & 0 & 88.6670 \end{bmatrix}, \tag{2.52}$$

where $\mathbf{I} - \Phi$ is invertible and the maximum eigenvalue of Φ is 0.5337. Hence, this specification guarantees the stationarity of the VAR(1) process.

Second, we consider a medium case for 30 stocks. Mean or median values of parameters

[25] $Bias = \frac{\sum_{t=1}^{N}\left[\hat{G}_t^{-1}(\alpha) - G_t^{-1}(\alpha)\right]}{N}$, $RMSE = \sqrt{Bias^2 + Var\left(\hat{G}_t^{-1}(\alpha)\right)}$, $MAE = \frac{\sum_{t=1}^{N}\left|\hat{G}_t^{-1}(\alpha) - G_t^{-1}(\alpha)\right|}{N}$, $MAPE = \frac{\sum_{t=1}^{N}\left|\frac{\hat{G}_t^{-1}(\alpha) - G_t^{-1}(\alpha)}{G_t^{-1}(\alpha)}\right|}{N}$ and $PMAD = \frac{\sum_{t=1}^{N}\left|G_t^{-1}(\alpha) - \hat{G}_t^{-1}(\alpha)\right|}{\sum_{t=1}^{N}\left|G_t^{-1}(\alpha)\right|}$.

are considered but the mean is severely affected by the outliers of the bailout companies such as AIG, Citi and GM. Hence, we select the median values of $\mathbf{m}$, Φ and Σ out of 30 estimations such that

$$\mathbf{m} = \begin{bmatrix} -0.0002 \\ 0.0581 \\ 0.1117 \\ 8.6073 \end{bmatrix}, \Phi = \begin{bmatrix} -0.0282 & 0.0027 & 0.0003 & -0.0001 \\ -0.2797 & 0.3434 & 0.0027 & -0.0016 \\ -2.1938 & 0.1015 & 0.0185 & -0.0015 \\ 5.2755 & 1.2691 & -0.0042 & 0.0286 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 0.0007 & 0 & 0 & 0 \\ 0 & 0.0218 & 0 & 0 \\ 0 & 0 & 2.6799 & 0 \\ 0 & 0 & 0 & 75.7837 \end{bmatrix}, \qquad (2.53)$$

where the maximum eigenvalue of Φ is 0.3355, guaranteeing the stationarity of the VAR(1) process.

Third, we consider the most persistent estimation case among 30 stocks. We search for a case presenting the maximum eigenvalue of Φ and finally select the case of Exxon Mobil Corporation where the values of $\mathbf{m}$, Φ and Σ are given by

$$\mathbf{m} = \begin{bmatrix} 0.0004 \\ 0.0420 \\ 0.1203 \\ 7.8954 \end{bmatrix}, \Phi = \begin{bmatrix} -0.1627 & 0.0260 & 0.0009 & 0.0000 \\ -0.4836 & 0.6382 & 0.0022 & -0.0009 \\ -4.3404 & -0.0316 & 0.0354 & -0.0011 \\ 4.8516 & 2.8674 & -0.1539 & 0.0427 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 0.0005 & 0 & 0 & 0 \\ 0 & 0.0046 & 0 & 0 \\ 0 & 0 & 2.1028 & 0 \\ 0 & 0 & 0 & 51.0363 \end{bmatrix}, \qquad (2.54)$$

where the maximum eigenvalue of Φ is 0.6165 which also guarantees the stationarity of the VAR(1) process.

Table 2.5 presents the evaluation results for alternative measures. Ten models are considered for each experiment and alternative positions (short and long) are taken into account for each model. For all experiments and measures, the results strongly conclude that FAR forecasts the true VaR most precisely. It provides the smallest values in the three experiments. RM, GARCH and CAViaR are less accurate than FAR while they are slightly more accurate than HS and FHS. The EVT models perform the worst in terms of precision. Their values are much greater than those of the other models. All models overestimate a true VaR and the biases are more or less symmetric for both positions. MAPE is quite sensitive to the extreme cases. The values of the long position are much greater than those of the short position. This implies that there are extreme negative realisations leading to the large forecasting error. Therefore, Monte Carlo simulation study provides the evidence that FAR can precisely estimate and forecast a true VaR.

2.5.3 The Out-of-sample Performance of VaR Forecasting

Given the success with the forecasting precision of FAR compared with the true intraday density function and VaR, we evaluate the empirical VaR forecasting performance of FAR compared with alternative models. We apply various backtestings to the performance evaluation, since a true VaR is unobservable in practice. To this end, we apply the rolling-forecast of the daily VaR based on a 250 window size from 03/01/2001 to 31/12/2008 to the 30 stocks and the equal weighted portfolio. We also consider the evaluations of the alternative positions (long and short). The quantitative measures and the test statistics for the 30 stocks and the portfolio are calculated based on the 2,011 daily forecasts (except those for Hewlett-Packard, AT&T and Verizon Communication Inc., which are calculated using 1,416, 2,004 and 1,885 forecasts, respectively).

2.5.3.1 Long Position

We first evaluate the backtestings for the long position with 250 window size and the results are summarised in Table 2.6. Panel A reports the quantitative backtestings and Panel B the statistical backtesting for the 30 stocks and the portfolio. In Panel A, the rows of (A) empirical coverage probability, (C) MRCR and (D) predictive quantile loss present the average quantitative values of the 30 stocks. The (B) Basel penalty zone counts the number of stocks falling into the zone determined by the average violation number for the 99 percent VaR in the previous 250 business days. The (A) empirical coverage probability and the (B) Basel penalty zone generally evaluate the coverage ability of the model. Furthermore, the (C) MRCR and the (D) predictive quantile loss evaluate the economic cost for employing the model. The figures of (E)–(I) count the number of stocks that are rejected at the 5% significance level for each test.[26] Panel B presents the evaluation results for the portfolio. The row (A), (C) and (D) present the same figures as those in Panel A whilst the row (B) presents the penalty zone based on the average number of daily violations of the 99 percent

[26]We have obtained quantitatively similar results for all cases using the 10% significance level.

Table 2.5: Monte Carlo Simulation Study for Evaluating The Precision of VaR Forecast

Intraday returns are generated by the NIG distribution using four moments that follow VAR(1) process (see Assumptions 1–4 for details). We generate 78 observations for each day and 251 days for conducting one-step-ahead forecast for both long- and short-position conditional on the past 250 days. The simulation study replicates this forecasting 5,000 times with ten models and evaluates the forecasting precision based on Bias, RMSE, MAE, MAPE and PMAD. Three representative experiments are considered out of 31 estimations. Experiment I considers the aggregate information of 30 stocks and select the portfolio that presents an equal weight composite of 30 stocks. Experiment II considers a medium case for 30 stocks. Mean or median values of parameters are considered but the mean is severely affected by outliers of bailout companies such as AIG, Citi and GM. Experiment III consider the most persistent estimation case among 30 stocks, Exxon Mobil Corporation.

	Bias		RMSE		MAE		MAPE		PMAD	
Models	Long	Short	Long	Short	Long	Short	Long	Short	Long	Short
Experiment I										
FAR	-0.63	-0.63	1.04	0.98	1.79	1.81	1.80	1.74	0.39	0.40
FHS	-1.62	-1.17	1.85	1.41	2.23	2.00	2.17	2.01	0.49	0.44
HS	-1.65	-1.21	1.83	1.39	2.23	2.01	2.20	2.03	0.49	0.44
RM	-0.96	-0.95	1.26	1.26	1.92	1.94	1.87	1.93	0.42	0.42
GARCH	-1.03	-1.02	1.15	1.15	1.89	1.90	1.92	1.95	0.41	0.42
GEV	-2.96	-3.04	3.20	3.28	3.17	3.25	2.87	2.87	0.69	0.71
GPD	-2.87	-2.97	3.12	3.22	3.10	3.20	2.81	2.83	0.68	0.70
FGEV	-2.90	-2.97	3.18	3.25	3.14	3.19	2.80	2.83	0.69	0.70
FGPD	-2.82	-2.91	3.12	3.20	3.08	3.15	2.75	2.80	0.67	0.69
CAViaR	-1.10	-1.13	1.97	2.02	2.25	2.30	2.15	2.00	0.49	0.50
Experiment II										
FAR	-0.57	-0.52	0.90	0.86	2.83	2.85	2.03	1.87	0.38	0.38
FHS	-2.66	-1.89	3.01	2.26	3.64	3.32	2.54	2.27	0.48	0.45
HS	-2.68	-1.92	2.98	2.22	3.63	3.31	2.54	2.28	0.48	0.44
RM	-1.51	-1.50	2.01	2.01	3.17	3.20	2.17	2.13	0.42	0.43
GARCH	-1.60	-1.60	1.78	1.78	3.09	3.12	2.25	2.15	0.41	0.42
GEV	-4.90	-4.94	5.27	5.32	5.24	5.31	3.20	3.21	0.70	0.71
GPD	-4.74	-4.78	5.16	5.19	5.12	5.19	3.14	3.17	0.68	0.70
FGEV	-4.85	-4.90	5.27	5.33	5.22	5.29	3.19	3.18	0.69	0.71
FGPD	-4.70	-4.75	5.15	5.22	5.11	5.18	3.15	3.14	0.68	0.69
CAViaR	-1.73	-1.83	3.29	3.37	3.76	3.80	2.24	2.42	0.50	0.51
Experiment III										
FAR	-0.72	-0.65	1.10	1.01	2.11	2.20	7.71	1.70	0.37	0.37
FHS	-2.06	-1.48	2.34	1.79	2.79	2.59	9.71	2.02	0.49	0.44
HS	-2.11	-1.55	2.32	1.78	2.79	2.61	9.73	2.07	0.49	0.44
RM	-1.31	-1.24	1.67	1.62	2.45	2.49	8.40	1.95	0.43	0.42
GARCH	-1.36	-1.30	1.52	1.47	2.39	2.45	8.72	1.93	0.42	0.41
GEV	-3.74	-3.86	4.02	4.15	3.99	4.15	12.37	2.95	0.71	0.70
GPD	-3.62	-3.74	3.92	4.06	3.90	4.06	11.87	2.92	0.69	0.69
FGEV	-3.67	-3.74	4.02	4.07	3.95	4.05	12.31	2.91	0.70	0.68
FGPD	-3.56	-3.64	3.93	4.01	3.87	3.99	11.84	2.91	0.68	0.67
CAViaR	-1.41	-1.47	2.53	2.68	2.83	2.94	9.18	1.98	0.50	0.50

VaR in the previous 250 business days. Furthermore, (E)–(I) present the test statistics and * , **, and *** denote the rejection of the null hypothesis for each test at, respectively, the 10%, 5% and 1% significance level. Those five test statistics employed in our evaluation are described in Section 3.2.

The quantitative backtestings of Panel A provide the following findings. For the (A) empirical coverage probability, FAR (1.07%) is closest to the 1% nominal probability. The FHS (1.10%) and the HS (1.19%) slightly underestimate the VaR. The GARCH models (1.60%–1.82%) considerably underestimate VaR whilst the EVT models (0.49%–0.55%) greatly overestimate VaR as expected. CAViaR (2.83%) shows the worst coverage ability by seriously underestimating the VaR. Overall, the hybrid non-parametric models demonstrate powerful coverage ability. For the (B) Basel penalty zone, the non-parametric models and the EVT models achieve the "Green" zone for all stocks. The non-parametric models obtain the "Green" zone due to their good coverage probability. The EVT models are however classified in the same zone due to their considerable overestimation of VaR. Since the Basel penalty zone is only concerned with underestimations (violation) of VaR, the overestimation receives better scores by the Basel penalty rule given their prudential principles. The GARCH models fail to achieve the "Green" zone for some stocks because of their underestimations of VaR. GARCH achieves the "Green" zone for 20 stocks whilst the RM achieves it for seven stocks. The worst case, CAViaR, fails to obtain the "Green" zone for any of stocks. There are 3 and 27 stocks falling into the "Red" and the "Yellow" zone, respectively.

For the (C) MRCR, the GARCH models (39.59%–39.78%) demand the smallest capital requirement. FAR (40.80%) requires a similar level of capital whilst FHS (45.29%%) and HS (45.28%) demand much more capital. The capital requirement of adopting CAViaR (47.43%) is quite high although the highest level is found in the EVT models (58.09%–60.75%). For the (D) predictive quantile loss, FAR (6.76%) faces the smallest loss. The losses of GARCH (6.80%), FHS (6.87%) and RM (6.89%) are more or less similar to those of FAR. On the other hand, HS (7.74%), the EVT models (7.61%–8.21%) and CAViaR (8.25%) require banks to endure big losses.[27]

For the (E) unconditional coverage test, FHS is not rejected for any of stocks. HS and FAR are rejected for 3 and 7 stocks, respectively. Other models are rejected for over 20

[27] For example, suppose that a bank manages $1 bn. and operates an internal VaR model. Then the CFO will report the economic cost using the MRCR and the predictive quantile loss every day. If the bank operates the EVT models, they should allocate $580.9 mm.–$607.5 mm. against the maximum loss for 10-day or endure $76.1 mm.–$82.1 mm. of loss every day. Then the CFO will decide to replace the EVT models with FAR, since the EVT models are too expensive to operate. The FAR reduces the capital requirement to $408 mm. and the loss to $67.6 mm. Therefore, FAR cuts down the capital requirement by $172.9 mm.–$199.5 mm. and the loss by $35.3 mm.–$41.3 mm.

stocks. These results imply that the hybrid models utilising a non-parametric distribution have statistically significant coverage ability whilst others have very poor ability. For the (F) independence test, the coverage ability of HS and CAViaR is statistically more dependent on the underlying states than other models. HS and CAViaR are rejected for 9 and 10 stocks. Hence, the violations of VaR forecasts are considerably time dependent on the underlying state. Regulators would therefore prefer the models allowing for a time-varying distribution which would prevent the violations from being abnormally clustered. The FAR would be the best candidate among such models since it is rejected for only 2 stocks. RM is rejected for 3 stocks, GARCH, FHS, and unfiltered EVT models are rejected for 4 stocks and the FEVT models are rejected for 6 stocks.

From the above two tests, the hybrid non-parametric models achieve statistically powerful coverage ability that is independent of the underlying state.[28] Other models fail to pass the two tests jointly. These findings are further supported by the (G) conditional coverage test. The hybrid non-parametric models are rejected for 4 or fewer stocks whilst the others are rejected for 9 or more stocks (ranging from HS: 9 stocks to CAViaR: 30 stocks). There are alternative types of conditional tests such as the (H) dynamic quantile test 1 and the (I) dynamic quantile test 2. The (H) dynamic quantile test 1 includes the past violations in the past information set and the (I) dynamic quantile test 2 additionally includes the past VaR forecast in the past information set. In the (H) dynamic quantile test 1, the hybrid non-parametric models and the FEVT models are rejected for a relatively small number of stocks whilst the others are rejected for over the half of the stocks. Especially, GARCH (22 stocks), HS (24 stocks), RM (28 stocks) and CAViaR (30 stocks) are rejected for over 70% of the 30 stocks. These results resemble the (I) dynamic quantile test 2. Overall, the hybrid non-parametric models statistically perform better than other models.

Panel B presents the evaluation results for the portfolio. The (A) empirical coverage probabilities of non-parametric models (0.80%–1.24%) are relatively closest to the 1% nominal probability. This finding confirms that the models utilising the non-parametric distribution have stronger coverage ability than those of the models utilising the parametric distribution. The GARCH models (1.79%–1.84%) and CAViaR (2.09%) considerably underestimate VaR whilst the EVT models (0.45%–0.55%) greatly overestimate it. These lead to that the adoption of the GARCH models and the CAViaR will be classified as the "Yellow" penalty zone. For the (C) MRCR, the GARCH models (24.64%–25%) demand the smallest capital requirements and the non-parametric models (25.89%–28.84%) and the CAViaR (26.66%) require slightly more capital. Banks would be required to pay an expen-

[28] This means that the models provide the closest coverage probability to the 1% nominal probability regardless of the underlying state such as business cycle or volatility clustering.

sive cost to adopt the EVT models (35.47%–35.56%). The (D) predictive quantile loss of FAR and the GARCH models (4.38–4.44%) are smallest whilst the EVT models (4.98%–5.18%) face the biggest loss. Those of the FHS, HS and CAViaR (4.48%–4.61%) are in the intermediate level. These results confirm that the non-parametric models are superior to the others in terms of their coverage ability and economic cost. The GARCH models and CAViaR reduce the economic cost well whilst they fail to increase the coverage ability. The EVT models present very poor performance in both coverage ability and economic cost analysis. Moreover, the statistical evaluations from (E)–(I) show that the non-parametric models are not rejected for all the tests whilst CAViaR is rejected for all of them. The GARCH models also fail for all the tests except for the (F) independence test. The FEVT models are not rejected for all the conditional tests at the weak significance level whilst the unfiltered EVT models are rejected for all the tests except for the (F) independence test. These results further confirm that the non-parametric models have statistically significant and robust coverage ability that is independent of the underlying state.

The results of the quantitative backtestings are strongly in favour of FAR for both regulatory purposes and bank risk management. The FAR would be preferred by the regulator since it has excellent coverage ability. Moreover, it would be preferred by banks as it requires less capital than other models require. Similarly, FHS is recommended given its ideal coverage ability with an acceptable cost advantage. The GARCH models fail to increase the coverage ability whilst they are successful in reducing the economic cost. On the contrary, the HS is successful in improving the coverage ability whilst it requires much capital to operate. The EVT models and CAViaR fail for both evaluations, neither accurate nor cheap. The results of the statistical backtestings strongly support those of the quantitative backtestings. The non-parametric models have more powerful coverage ability that is independent of the underlying state. However, the time-varying nature of financial time series should be applied to the non-parametric distribution. It is confirmed that the conditional tests of HS are rejected for many stocks whilst those of the hybrid non-parametric models are rejected for a small number of stocks.

2.5.3.2 Short Position

The results for the short position, summarised in Table 2.7, are generally consistent with those for the long position. The results for the quantitative backtestings in Panel A provide the following. For the (A) empirical coverage probability, the hybrid non-parametric models (0.96%–1.03%) are closest to the 1% nominal probability. FAR (0.96%) slightly overestimates the VaR whilst FHS (1.03%) slightly underestimates it. HS (1.20%) underes-

Table 2.6: Evaluations of Value-at-Risk for Long Position (Window Size = 250)

In Panel A, (A) the empirical coverage probability, (C) market risk capital requirement and (D) predictive quantile loss are evaluated by averaging 30 companies. The (B) Basel penalty zone is evaluated by the frequency of how many companies fall in each zone. Five tests, (E)–(I), are evaluated by the rejection frequency at the given 10% significance level so that the model with low rejection frequency is preferred by the regulator. Panel B evaluates the portfolio constructed by the equal weighted combination of 30 companies. The (B) Basel penalty zone is evaluated by the colour of the zone. Green is the most acceptable one whilst red is the worst one. *, ** and *** denote that the test statistic is rejected at the 1%, 5% and 10% significance level, respectively. In both panels, the figures in $(\cdot)$ denote the asymptotic distribution for each test statistic. Furthermore, all quantitative measures and statistics for each company and portfolio are calculated using 2,011 VaR forecasts (except Hewlett-Packard Company , AT&T and Verizon Communication Inc. are calculated using 1,416, 2,004 and 1,885 VaR forecasts, respectively).

VaR Models	FAR	HS	FHS	RM	GARCH	GEV	GPD	FGEV	FGPD	CAViaR
Panel A: 30 Components of Dow Jones Industry										
(A) Empirical Coverage Probability	1.07%	1.19%	1.10%	1.82%	1.60%	0.54%	0.55%	0.49%	0.51%	2.83%
(B) Basel Penalty Zone										
Green	30	30	30	7	20	30	30	30	30	0
Yellow	0	0	0	23	10	0	0	0	0	27
Red	0	0	0	0	0	0	0	0	0	3
(C) Market Risk Capital Requirement	40.80%	45.28%	45.29%	39.59%	39.78%	58.09%	60.75%	58.22%	60.21%	47.43%
(D) Predictive Quantile Loss	6.76%	7.44%	6.87%	6.89%	6.80%	7.99%	8.21%	7.61%	7.82%	8.25%
(E) Unconditional Coverage Test (χ_1^2)	7	3	0	27	20	23	21	28	26	30
(F) Independence Test (χ_1^2)	2	9	4	3	4	4	4	6	6	10
(G) Conditional Coverage Test (χ_2^2)	4	12	1	25	19	18	19	19	20	30
(H) Dynamic Quantile Test 1 (χ_5^2)	10	24	10	28	22	14	14	10	9	30
(I) Dynamic Quantile Test 2 (χ_6^2)	11	24	12	29	22	13	13	8	9	30
Panel B: Portfolio										
(A) Empirical Coverage Probability	0.80%	1.24%	1.14%	1.79%	1.84%	0.55%	0.45%	0.55%	0.65%	2.09%
(B) Basel Penalty Zone	Green	Green	Green	Yellow	Yellow	Green	Green	Green	Green	Yellow
(C) Market Risk Capital Requirement	28.84%	27.35%	25.89%	24.64%	25.00%	34.99%	34.54%	35.47%	35.56%	26.66%
(D) Predictive Quantile Loss	4.44%	4.61%	4.48%	4.38%	4.42%	5.05%	4.98%	5.09%	5.18%	4.59%
(E) Unconditional Coverage Test (χ_1^2)	0.91	1.11	0.40	10.27***	11.48***	4.99**	7.81***	4.99**	2.90*	18.32***
(F) Independence Test (χ_1^2)	0.26	0.63	1.24	0.17	1.77	0.12	0.08	0.12	0.17	3.36*
(G) Conditional Coverage Test (χ_2^2)	1.17	1.75	1.65	10.46***	13.27**	5.10*	7.88**	5.10*	3.06	21.70***
(H) Dynamic Quantile Test 1 (χ_5^2)	1.26	7.51	5.85	34.50***	30.31***	12.53**	16.81***	4.28	2.74	78.43***
(I) Dynamic Quantile Test 2 (χ_6^2)	1.38	8.24	9.41	41.20***	38.10***	12.62**	17.26***	5.49	4.59	107.52***

timates even more the VaR. The GARCH models (1.46%–1.66%) underestimate VaR whilst the EVT models (0.48%–0.56%) considerably overestimate it. CAViaR (2.31%) seriously underestimates the VaR. For the (B) Basel penalty zone, FHS achieves the "Green" zone for all stocks. FAR and HS achieve the "Green" zone for 29 stocks and the "Yellow" zone for 1 stock. The EVT models obtain "Green" zone for 30 stocks, since the models much overestimate VaR. The GARCH models fail to achieve the "Green" zone for some stocks because of their severe underestimations of VaR. GARCH achieves the "Green" zone for 21 stocks whilst RM achieves it for 14 stocks. The worst case is CAViaR, which obtains the "Green" zone for only one stock and the "Yellow" zone for 28 stocks. For the (C) MRCR, the GARCH models require the smallest capital (38.64%–39.04%) and the hybrid non-parametric models require a bit more capital (41.20%–42.88%). The EVT models (52.53%–56.51%) are the most expensive models for banks to adopt as their internal VaR models. The capital requirements for HS (45.13%) and CAViaR (44.13%) are somewhat higher. For the (D) predictive quantile loss, FAR (5.93%) has the smallest loss followed by GARCH (5.97%), FHS (6.05%) and RM (6.09%). On the other hand, HS (7.06%), the EVT models (6.60%–7.44%), and CAViaR (7.10%) require banks to endure big losses. The results of the quantitative backtestings show that the hybrid non-parametric models obtain a powerful coverage ability at small economic cost. The GARCH models and HS fail to obtain a strong coverage ability and a small economic cost jointly.

For the (E) unconditional coverage test, FHS is not rejected for any of the stocks. HS and FAR are rejected for three and five stocks, respectively. Other models are rejected for over 20 stocks except for GARCH, rejected for 14 stocks. For the (F) independence test, overall, the coverage ability of VaR forecasts are independent of the underlying state. The hybrid non-parametric models and the GARCH models are rejected for one company at most. Other models are rejected for slightly more stocks but less than four stocks. These results imply that the non-parametric models have statistically powerful coverage ability that is more or less statistically robust to the underlying states but other models fail to pass the two tests jointly. These findings are further confirmed in the (G) conditional coverage test. FHS is rejected for the none of 30 stocks and FAR is rejected for four stocks. HS is rejected for seven stocks, which is relatively small compared with other models (ranging from GARCH: 12 stocks to CAViaR: 30 stocks). Moreover, for the (H) dynamic quantile test 1, the hybrid non-parametric models and the FEVT models are rejected for five or fewer stocks (ranging from FHS: one company, to FGEV: five stocks) whilst others are rejected for more stocks. Especially, HS (16 stocks), RM (20 stocks) and CAViaR (30 stocks) are rejected over the half of the 30 stocks. These results are consistently observed in the (I) dynamic quantile test 2. Therefore, the hybrid non-parametric models provide statistically the most powerful coverage ability, which is independent of the underlying state.

In Panel B, the (A) empirical coverage probability of FAR (0.99%) is closest to the 1% nominal probability. GARCH (1.09%) and FHS (1.14%) somewhat underestimate VaR. HS (1.79%) and CAViaR (1.59%) much underestimate the VaR whilst the FEVT models (0.25%–0.35%) greatly overestimate it. These lead to considering that adoption of the HS and CAViaR models will be classified in the "Yellow" zone. The "Green" zone of the EVT models is completely due to the considerable overestimation of VaR, which is misleading. For the (C) MRCR, the GARCH models (23.98%–24.42%) require the smallest capital. The hybrid non-parametric models (25.19%–27.92%), CAViaR (24.77%), and the FEVT models (28.38%–28.62%) require slightly more capital. The unfiltered EVT models (33.44%–33.50%) have the most expensive cost of adoption. The (D) predictive quantile loss of the GARCH models (3.31%–3.34%) are the smallest whilst the losses are largest for HS and the unfiltered EVT models (4.43%–4.49%). Those of the hybrid non-parametric models, the FEVT models and CAViaR (3.35%–3.65%), are slightly more than those of the GARCH models. These results confirm that the hybrid non-parametric models are superior to the others in terms of coverage ability and reasonable economic cost. The GARCH models, the FEVT models and CAViaR are successful at reducing the economic cost but fail to achieve adequate coverage ability. Furthermore, the unfiltered EVT models are undesirable from the results of both evaluations. Moreover, the statistical evaluations from (E) to (I) show that the hybrid non-parametric models are successful for all the tests. HS, the EVT models and CAViaR are strongly rejected for two tests at least. The GARCH models fail for only (I) dynamic quantile test 2. Hence, the statistical backtestings further confirm that the coverage ability of the hybrid non-parametric models are accurate and state-independent.

The results for the short position are almost consistent with those for the long position. The results of quantitative backtestings are strongly in favour of the hybrid non-parametric models for both regulatory purposes and bank risk management. Those models would be recommended by the regulator since the models have excellent coverage ability. Moreover, they would be welcomed by banks as they require less capital than other models. The GARCH models, the FEVT models and CVAiaR fail to gain in coverage ability whilst they are successful in reducing the economic cost. HS and the unfiltered EVT models fail in both evaluations. From the results of the statistical backtestings, the results from the quantitative backtestings are confirmed again. The hybrid non-parametric models have the more powerful coverage ability that is also more time-independent of the underlying state.

Table 2.7: Evaluations of Value-at-Risk for Short Position (Window Size = 250)

In Panel A, (A) the empirical coverage probability, (C) market risk capital requirement and (D) predictive quantile loss are evaluated by averaging 30 companies. The (B) Basel penalty zone is evaluated by the frequency of companies' falling into each zone. Five tests, (E)–(I), are evaluated by the rejection frequency at the given 10% significance level so that the model with a low rejection frequency is preferred by the regulator. Panel B evaluates the portfolio constructed by the equal weighted combination of 30 companies. The (B) Basel penalty zone is evaluated by the colour of the zone. Green is the most acceptable whilst red is the worst. *, ** and *** denote that the test statistic is rejected at the given 1%, 5% and 10% significance level, respectively. In both panels, the figures in $(\cdot)$ denote the asymptotic distribution for each test statistic. Furthermore, all quantitative measures and statistics for each company and portfolio are calculated using 2,011 VaR forecasts (except Hewlett-Packard Company , AT&T and Verizon Communication Inc., which are calculated using 1,416, 2,004 and 1,885 VaR forecasts, respectively).

VaR Models	FAR	HS	FHS	RM	GARCH	GEV	GPD	FGEV	FGPD	CAViaR
Panel A: 30 Companies										
(A) Empirical Coverage Probability	0.96%	1.20%	1.03%	1.66%	1.46%	0.56%	0.54%	0.48%	0.48%	2.31%
(B) Basel Penalty Zone										
Green	29	29	30	14	21	30	30	30	30	1
Yellow	1	1	0	16	9	0	0	0	0	28
Red	0	0	0	0	0	0	0	0	0	1
(C) Market Risk Capital Requirement	41.20%	45.13%	42.88%	38.64%	39.04%	55.40%	56.51%	52.53%	53.77%	44.13%
(D) Predictive Quantile Loss	5.93%	7.06%	6.05%	6.09%	5.97%	7.44%	7.51%	6.60%	6.73%	7.10%
(E) Unconditional Coverage Test (χ_1^2)	5	3	0	23	14	23	22	27	28	30
(F) Independence Test (χ_1^2)	1	3	1	1	0	3	4	2	2	4
(G) Conditional Coverage Test (χ_2^2)	4	7	0	19	12	13	18	20	24	30
(H) Dynamic Quantile Test 1 (χ_5^2)	4	16	1	20	11	11	9	5	4	30
(I) Dynamic Quantile Test 2 (χ_6^2)	6	18	3	19	9	10	9	5	9	30
Panel B: Portfolio of 30 Companies										
(A) Empirical Coverage Probability	0.99%	1.79%	1.14%	1.34%	1.09%	0.70%	0.75%	0.35%	0.25%	1.59%
(B) Basel Penalty Zone	Green	Yellow	Green	Green	Green	Green	Green	Green	Green	Yellow
(C) Market Risk Capital Requirement	27.92%	30.31%	25.19%	24.42%	23.98%	33.44%	33.50%	28.62%	28.38%	24.77%
(D) Predictive Quantile Loss	3.65%	4.56%	3.35%	3.34%	3.31%	4.49%	4.43%	3.50%	3.50%	3.59%
(E) Unconditional Coverage Test (χ_1^2)	0.00	10.27***	0.40	2.15	0.17	2.10	1.44	11.53***	16.42***	6.02**
(F) Independence Test (χ_1^2)	0.40	0.17	0.53	0.74	0.49	2.97*	2.71*	0.05	0.02	1.04
(G) Conditional Coverage Test (χ_2^2)	0.40	10.46***	0.94	2.90	0.66	5.06*	4.14	11.57***	16.43***	7.07**
(H) Dynamic Quantile Test 1 (χ_5^2)	0.84	19.96***	3.73	7.23	6.09	14.46**	12.74**	8.63	11.44**	10.44*
(I) Dynamic Quantile Test 2 (χ_6^2)	1.38	8.24	9.41	41.20***	38.10***	12.62**	17.26***	5.49	4.59	107.52***

Stylised Facts

We can draw the following conclusions from the quantitative and the statistical backtestings. First, the hybrid non-parametric models, FAR and FHS, perform better in modelling the tail behaviour than purely non-parametric and parametric models do. Second, the GARCH models with normal/Student's t-distributions considerably underestimate VaR whilst the EVT models much overestimate it. The normal distribution generally has a thinner tail than the distribution of financial time series, which limits its ability in capturing extreme risk. Furthermore, both the normal distribution and Student's t-distribution are symmetric so that they cannot detect the skewness which is frequently observed in financial time series (Theodossiou, 1998). The GARCH models building on the assumption of these two distributions therefore tend to underestimate the tail risk. Furthermore, the GARCH models perform worse for a long position than for a short position, since the models cannot detect the negative skewness frequently observed in individual stocks. On the other hand, the EVT models estimate VaR by capturing the extreme of the extreme values which usually tends to be greater than the true extreme values. Hence, the EVT models can cover the loss very well in the period of a financial crash whilst it often much overestimates the potential loss in a normal period. Furthermore, there is always the trade-off between the number of observations and the degree of extremeness in applying EVT models.[29] Third, the dynamic models considering the time-varying distributions perform better than the static models such as HS and the unfiltered EVT models. Especially, the coverage ability of the dynamic models is much more independent of the underlying state than those of the static models. Furthermore, HS and the unfiltered EVT models are improved by the GARCH filtering of time-series. Finally, the above conclusions are consistently observed in both long and short positions. Hence, our proposed hybrid non-parametric models are quite robust in forecasting VaR.

2.5.4 The Effect of Window Size

We repeat the previous analysis for the window size of 500, as a robustness check. The quantitative measures and test statistics for the 30 stocks and the portfolio are calculated based on the 1,763 daily forecasts (except for Hewlett-Packard, AT&T and Verizon Communication Inc., which are calculated using 1,178, 1,756 and 1,637 forecasts, respectively). The evaluation results are presented in Tables 2.8 and 2.9. Table 2.8 presents the results

[29] The EVT estimates are affected by a small number of observation. If we set a wide block size for extreme observations, the number of extreme observations is not enough to estimate the parameters accurately. If we obtain enough observations by setting a narrow block size for extreme observations, the degree of extremeness is weaker.

for the long position and Table 2.9 presents the results for the short position. Overall, the results are consistent with those for the shorter window size.

2.5.4.1 Long Position

Table 2.8 reports the results for the long position. The quantitative backtesting results, summarised in Panel A, are strongly consistent with those of the shorter window size. The hybrid non-parametric models outperform other models in terms of the coverage ability and the economic cost. The GARCH models reduce the economic cost but fail to improve the coverage ability. The EVT models fail both at improving the coverage ability and at reducing the economic cost. The notable effects of the longer window size are observed in HS and the EVT models. First, the EVT models greatly overestimate VaR compared with the shorter window size. This is driven by the degree of extremeness becoming stronger as the window size increases. On the other hand, HS considerably underestimates VaR with the longer window size. As we use the longer history to estimate a non-parametric distribution, the weight on recent regime changes such as volatility clustering is reduced. Hence, the underestimation would be frequently observed during the "Recession".[30] Second, these inherent limitations of the EVT models and HS also make for an increase in the economic cost. Those of the EVT models most significantly increase and the magnitudes are much larger than those of the shorter window size. The performance of CAViaR improves a bit and the hybrid non-parametric models performs similarly with the shorter window size. The statistical evaluations of Panel A are also almost consistent with those of the shorter window size. The hybrid non-parametric models provide statistically desirable performance for banks and regulators to adopt the models as their internal VaR models. Generally, the rejection frequency decreases for all the tests except for HS. However, there are no significant changes in that the hybrid non-parametric models still outperform other models.

Panel B presents the evaluations for the portfolio. Based on the quantitative evaluations, the hybrid non-parametric models and CAViaR present good coverage ability and small economic cost. The GARCH models achieve a small economic cost whilst they still suffer from poor coverage ability by reason of their considerable underestimation of VaR. The EVT models obtain neither powerful coverage ability nor small economic cost by reason of their huge overestimation of VaR. Hence, the results of the quantitative evaluations are more or less consistent with those for the shorter window size. Moreover, the statistical evaluations show that the hybrid non-parametric models and CAViaR are successful for all

[30] For example, the recent regime indicates the "Recession" but HS includes the past information of the "Boom" and gives the same weight to all the historical events. Hence, HS is likely to underestimate VaR for the "Recession".

tests. The FEVT models are strongly rejected for the (E) unconditional coverage test and the (G) conditional coverage test. Other models are rejected for all tests except for the (F) independence test. The results of the statistical evaluations are consistent with those of the quantitative evaluations. Furthermore, the results are more or less consistent with those for the shorter window size.

In sum, the quantitative and the statistical evaluations confirm that the hybrid non-parametric models provide powerful coverage ability which is independent of the underlying state. Generally, the performance of the models somewhat improves as the window size increases whilst the HS and the EVT models still suffer from their inherent limitations.

2.5.4.2 Short Position

The results for the short position are summarised in Table 2.9. The quantitative backtestings, summarised in Panel A, are overall consistent with those of the long position. Furthermore, those are strongly consistent with the shorter window size. The hybrid non-parametric models outperform other models in terms of coverage ability and economic cost. The GARCH models are successful at reducing the economic cost but fail to improve the coverage ability. The EVT models fail to improve the coverage ability and yield a big economic cost by the reason of their considerable overestimation of VaR (see Section 5.2.2 for the detailed explanation of the longer window size effect). The statistical evaluations also present that the hybrid non-parametric models provide powerful coverage ability that is independent of the underlying state, which is consistently observed in the three conditional tests. The hybrid non-parametric models provide the statistically desirable performance for banks and the regulators in adopting the models as their internal VaR models.

The evaluations for the portfolio, summarised in Panel B, presents that the hybrid non-parametric models, GARCH and CAViaR have good coverage ability and small economic cost. The FEVT models obtain more or less a small economic cost but fail to obtain a good coverage ability. The unfiltered EVT models get neither a powerful coverage ability nor a small economic cost. Moreover, the statistical evaluations show that the hybrid non-parametric models, the GARCH models and CAViaR are successful for all the tests. The EVT models are strongly rejected by two or more tests and HS is strongly rejected by four tests. The overall results of the quantitative and statistical evaluations are more or less consistent with those for the shorter window size.

Overall, the quantitative and the statistical evaluations confirm that the hybrid non-parametric models provide a powerful coverage ability that is independent of the underlying state. Other models such as the GARCH models and CAViaR provide an improved

Table 2.8: Evaluations of Value-at-Risk for Long Position (Window Size = 500)

In Panel A, (A) the empirical coverage probability, (C) market risk capital requirement and (D) predictive quantile loss are evaluated by averaging 30 companies. The (B) Basel penalty zone is evaluated by the frequency of companies' falling into each zone. Five tests, (E)–(I), are evaluated by the rejection frequency at the given 10% significance level so that the model with low rejection frequency is preferred by the regulator. Panel B evaluates the portfolio constructed by the equal weighted combination of 30 companies. The (B) Basel penalty zone is evaluated by the colour of the zone. Green is the most acceptable one whilst red is the worst one. *, ** and *** denote that the test statistic is rejected at the given 1%, 5% and 10% significance level, respectively. In both panels, the figures in $(\cdot)$ denote the asymptotic distribution for each test statistic. Furthermore, all quantitative measures and statistics for each company and portfolio are calculated using 2,011 VaR forecasts (except Hewlett-Packard Company , AT&T and Verizon Communication Inc., which are calculated using 1,416, 2,004 and 1,885 VaR forecast, respectively).

VaR Models	FAR	HS	FHS	RM	GARCH	GEV	GPD	FGEV	FGPD	CAViaR
Panel A: 30 Companies										
(A) Empirical Coverage Probability	1.05%	1.55%	1.13%	1.82%	1.61%	0.29%	0.34%	0.20%	0.21%	1.76%
(B) Basel Penalty Zone										
Green	29	26	30	10	20	30	30	30	30	11
Yellow	1	4	0	20	10	0	0	0	0	19
Red	0	0	0	0	0	0	0	0	0	0
(C) Market Risk Capital Requirement	38.77%	42.09%	40.88%	36.84%	37.25%	77.46%	74.37%	70.93%	72.42%	42.12%
(D) Predictive Quantile Loss	6.59%	7.85%	6.37%	6.57%	6.46%	9.51%	9.39%	8.75%	8.92%	6.88%
(E) Unconditional Coverage Test (χ^2_1)	5	13	0	27	18	30	28	29	30	26
(F) Independence Test (χ^2_1)	3	10	2	3	4	5	6	1	2	7
(G) Conditional Coverage Test (χ^2_2)	6	18	2	22	16	28	28	29	30	22
(H) Dynamic Quantile Test 1 (χ^2_5)	9	27	6	25	20	19	16	15	16	24
(I) Dynamic Quantile Test 2 (χ^2_6)	11	27	7	25	21	17	9	10	9	30
Panel B: Portfolio of 30 Companies										
(A) Empirical Coverage Probability	0.62%	1.70%	0.85%	1.70%	1.64%	0.34%	0.28%	0.23%	0.23%	1.19%
(B) Basel Penalty Zone	Green	Green	Green	Green	Yellow	Green	Green	Green	Green	Green
(C) Market Risk Capital Requirement	27.84%	27.23%	24.23%	22.18%	23.12%	42.30%	38.84%	39.69%	42.00%	24.13%
(D) Predictive Quantile Loss	3.67%	4.75%	3.49%	3.52%	3.50%	5.64%	5.37%	5.29%	5.77%	3.63%
(E) Unconditional Coverage Test (χ^2_1)	2.91*	7.24***	0.42	7.24***	6.20**	10.40***	12.75***	15.50***	15.50***	0.61
(F) Independence Test (χ^2_1)	0.14	1.04	2.48	0.38	2.85*	0.04	0.03	0.02	0.02	0.51
(G) Conditional Coverage Test (χ^2_2)	3.04	8.30**	2.89	7.64**	9.07**	10.43***	12.76***	15.50***	15.50***	1.12
(H) Dynamic Quantile Test 1 (χ^2_5)	2.67	46.35***	5.94	19.09***	26.09***	24.50***	29.72***	10.61*	10.61*	4.28
(I) Dynamic Quantile Test 2 (χ^2_6)	2.96	47.31***	7.61	23.22***	29.94***	24.57***	30.08***	10.81*	10.85*	7.53

performance especially for the portfolio but they still provide inferior performance for the 30 stocks on average. The performance of the models slightly improves as the window size increases for the short position. However, the finding that the hybrid non-parametric models outperform other models is consistently observed.

Stylised Facts

We can draw the following conclusions for the case with the larger window size. First, the hybrid non-parametric models, FAR and FHS, perform better in modelling the tail behaviour than purely non-parametric and parametric models. This finding is consistently observed for both the long and the short position. Second, the performance generally improves as the window size increases for the long and the short position except for HS and the EVT models. HS more considerably underestimates VaR with the longer window size. If we use the longer history to estimate a non-parametric distribution, the weight on the recent regime such as volatility clustering is reduced. Hence, such underestimation would be frequently observed during the "Recession" by construction. On the other hand, the EVT models more greatly overestimate VaR with the longer window size, which is driven by the degree of extremeness becoming stronger as the window size increases. Therefore, we confirm that the hybrid non-parametric models outperform other models for the case of the longer window size.

2.6 Conclusion

This study introduces a novel econometric technique in order to improve VaR analysis in terms of coverage ability and economic cost. Our proposed model complements pure non-parametric and parametric approaches by applying FAR to estimating and forecasting the intraday return density function.

We perform extensive evaluations on the performance of FAR modelling for the density function. First, we demonstrates that FAR is the best predictor of the intraday density function compared with alternative functional models such as AVE and LAST, with 30 stocks of DJIA and an equal-weighted portfolio. Second, a Monte Carlo simulation study provides the evidences that FAR can more precisely estimate and forecast a true VaR than other existing models: HS, FHS, GARCH, RM, EVT models, FEVT models, and CAViaR. Third, a wide range of quantitative and statistical backtestings are applied to the alternative VaR models and the results point strongly favour the hybrid non-parametric approaches which includes our FAR approach and FHS.

Table 2.9: Evaluations of Value-at-Risk for Short Position (Window Size = 500)

In Panel A, (A) the empirical coverage probability, (C) market risk capital requirement and (D) predictive quantile loss are evaluated by averaging 30 companies. The (B) Basel penalty zone is evaluated by the frequency of companies' falling into each zone. Five tests, (E)–(I), are evaluated by the rejection frequency at the given 10% significance level so that the model with low rejection frequency is preferred by the regulator. Panel B evaluates the portfolio constructed by an equal weighted combination of 30 companies. The (B) Basel penalty zone is evaluated by the colour of the zone. Green is the most acceptable one whilst red is the worst one. *, ** and *** denote that the test statistic is rejected at the given 1%, 5% and 10% significance level, respectively. In both panels, the figures in $(\cdot)$ denote the asymptotic distribution for each test statistic. Furthermore, all quantitative measures and statistics for each company and portfolio are calculated using 2,011 VaR forecasts (except Hewlett-Packard Company , AT&T and Verizon Communication Inc., which are calculated using 1,416, 2,004 and 1,885 VaR forecast, respectively).

VaR Models	FAR	HS	FHS	RM	GARCH	GEV	GPD	FGEV	FGPD	CAViaR
Panel A: 30 Companies										
(A) Empirical Coverage Probability	0.99%	1.49%	1.10%	1.72%	1.44%	0.33%	0.39%	0.22%	0.23%	1.67%
(B) Basel Penalty Zone										
Green	29	27	30	11	23	30	30	30	30	14
Yellow	1	3	0	19	7	0	0	0	0	16
Red	0	0	0	0	0	0	0	0	0	0
(C) Market Risk Capital Requirement	38.69%	43.70%	40.20%	36.38%	36.75%	71.05%	68.11%	61.35%	62.18%	39.90%
(D) Predictive Quantile Loss	5.79%	7.62%	5.78%	5.96%	5.76%	8.70%	8.55%	7.30%	7.36%	6.32%
(E) Unconditional Coverage Test (χ^2_1)	3	11	1	24	11	27	25	30	30	19
(F) Independence Test (χ^2_1)	0	6	1	1	1	1	4	1	0	1
(G) Conditional Coverage Test (χ^2_2)	3	11	1	21	8	27	24	29	29	16
(H) Dynamic Quantile Test 1 (χ^2_5)	4	22	1	19	7	9	11	20	14	19
(I) Dynamic Quantile Test 2 (χ^2_6)	7	24	1	19	7	6	9	9	9	21
Panel B: Portfolio of 30 Companies										
(A) Empirical Coverage Probability	0.91%	2.10%	1.13%	1.47%	1.08%	0.17%	0.68%	0.11%	0.28%	1.25%
(B) Basel Penalty Zone	Green	Yellow	Green	Green	Green	Green	Green	Green	Green	Green
(C) Market Risk Capital Requirement	27.17%	29.64%	22.96%	22.49%	22.82%	43.74%	38.90%	28.88%	28.65%	22.44%
(D) Predictive Quantile Loss	3.65%	5.20%	3.13%	3.22%	3.14%	5.23%	5.00%	3.48%	3.45%	3.21%
(E) Unconditional Coverage Test (χ^2_1)	0.16	16.33***	0.31	3.50*	0.10	18.76***	2.05	22.69***	12.75***	1.01
(F) Independence Test (χ^2_1)	0.29	1.42	0.46	0.78	0.41	0.01	3.32*	0.00	0.03	0.56
(G) Conditional Coverage Test (χ^2_2)	0.45	17.78***	0.77	4.29	0.52	18.75***	5.36*	22.68***	12.76***	1.58
(H) Dynamic Quantile Test 1 (χ^2_5)	0.70	33.36***	4.13	9.21	4.12	12.23**	17.06***	13.96**	9.12	4.57
(I) Dynamic Quantile Test 2 (χ^2_6)	1.23	36.25***	5.60	10.70*	8.68	12.54*	17.67***	13.97**	9.53	5.22

Overall, our extensive horse-racing of VaR models contributes to the understanding of VaR analysis in the following ways. First, FAR describes the dynamics of the intraday return density well. It is a generalisation of all autoregressive specifications. Hence, it considerably removes the uncertainty of existing parametric autoregressive models. Second, a non-parametric density function associated with a dynamic structure is superior to a parametric density function in estimating and forecasting risk. This is driven by the strong coverage ability of the non-parametric approach being independent of the underlying state in time-varying modelling. Third, a hybrid scheme reduces the economic cost and improves the coverage ability. The coverage ability is mainly determined by the robust estimation of the VaR and the economic cost by the time-independence from the underlying state. The hybrid scheme absorbs the robustness from the non-parametric approach and the time-independence from the dynamic modelling. Fourth, intraday information is helpful in forecasting the daily risk. The intraday data possesses important information which is relevant to market participants in forming their future expectations. Understanding the intraday return distribution is key to access the accumulated results of the daily return distribution. Hence, an accurate and well organised modelling of the intraday return distribution helps to estimate and forecast the daily risk.

There are several unanswered questions raised by this study which cannot be explained further due to the limited space and the focus of our study. First, we will show that individual VaR measures perform differently under normal and extreme market condition. It would be possible to develop a systematic approach of switching the use of measures conditional on the market condition. This switching will provide the jump necessary for the sudden change of a company's risk profile when they are approaching bankruptcy. This will require a Bayesian learning approach or Markov switching modelling of FAR. Second, VaR models capture risk spillovers indirectly to the extent that institutions are exposed to common risk factors, but they do not provide explicit information about the co-dependence of the risk. Given the increased interest in systemic risk, it would be desirable to extend our analysis into the co-dependence of risk between individual institutions and the financial system. We can forecast CoVaR (Adrian and Brunnermeier, 2010) by applying FAR to the conditional quantile function of the financial system conditional on the individual institution. Hence, this will contribute to forecasting the systemic risk. Furthermore, Copula will be required for analysing the co-dependence of the risk with FAR. Finally, VaR analysis will call for a robust statistical test, since the existing tests usually suffer from low power.

Chapter 3

Forecasting Distributions of Inflation Rates: Functional Autoregressive Approach

3.1 Introduction

Monetary authorities around the world, including those of Australia, Canada, Finland, New Zealand, Spain, Sweden and the UK, began to adopt inflation targeting as the monetary policy framework in the 1990s. Since 1992, inflation targets have been set explicitly in the UK, and the Chancellor's stated objective was to achieve an average annual inflation rate of 2%. In June 1997, the Bank of England (BOE) was granted operational independence and established the Monetary Policy Committee (MPC), charging it with sole responsibility for maintaining inflation within the target range of 1% to 4% per annum. In this policy framework, an inflation forecast is a key input into the decision making process of the MPC since it signals that a potential change in policy may be required to ensure that inflation does not move outside its target range. In this study we aim to develop an inflation forecasting framework using the sectoral inflation rates that utilises both the informational contents in the higher order moments as well as the time-variation of the cross-sectional distribution.

Importance of providing further information on the uncertainty surrounding forecasts of key macroeconomic variables has been increasingly recognised (Giordani and Söderlind, 2003). The knowledge of the precision of forecasts enables policy makers to motivate actions based on these forecasts, and helps in achieving a more balanced evaluation of both forecasts and policy in the public arena (Casillas-Olvera and Bessler, 2006). Point forecasts of inflation can work well only in restrictive cases (Granger and Pesaran, 2000). Monetary policy decisions should explicitly accommodate the uncertainty surrounding point forecasts. The BOE itself produces a quarterly Inflation Report in which this uncertainty is conveyed using fan charts over a two-year horizon, with bands of various shades of red illustrating the range of likely inflation outcomes. Similarly, the European Forecasting Network and the Fed provide such forecasts for a set of key macroeconomic variables. Although this current practice is clearly an important step toward acknowledging the significance of forecast uncertainties in the decision making process, it is difficult for independent researchers to replicate them due to the subjective manner in which the central bank accommodates forecast uncertainty.

Non-normality along with the time-varying distribution of inflation rates may have dramatic consequences for the optimal monetary policy-making. Accordingly, the reliable forecasts of inflation rates have to be computed dynamically to account for changes in the distribution. Previous studies have focused on modelling a national inflation as an autoregressive (AR) process or the uncertainty in inflation as a generalised autoregressive conditional heteroskedastic (GARCH) process (Engle, 1982; Bollerslev, 1986). Here, the conditional distributions are assumed to be time-varying only in their first two moments, thus neglecting the importance of accounting for higher-order moments. Recently, there have been a number of studies examining the role of time-varying higher-order moments in the analysis of financial and macroeconomic data (e.g. Hansen, 1994; Harvey and Siddique, 2000; Jondeau and Rockinger, 2003). Also, in inflation data, Bryan, Cecchetti, and Wiggins (1997) document fat-tailed properties and Roger (2000) provides evidence of right skewness. A natural extension is to consider the temporal dependence structure between conditional higher-order moments of inflation.

Analysis based on national aggregate data may be severely biased given the presence of informational heterogeneity across the sectors as documented by Pesaran and Smith (1995) and Hsiao, Shen, and Fujiki (2005). Inflation can be viewed as a weighted average of different commodities across different periods, and thus, modelling time-varying cross-sectional distributions is relevant utilizing information available at the sectoral level. Suppose that we examine the dynamics of cross-sectional distribution. Current inflation of any sector depends not only on its own past moments but also on the moments of the other sectors. Our approach of defining the national inflation rate as a weighted average of sectoral inflation rates and developing a flexible framework for utilizing both informational contents in the higher-order moments and the time-variation of cross-sectional distribution of sectoral inflation rates could thus plausibly act as a crucial mechanism for designing an optimal monetary policy. To this end, we follow Bosq (2000) and Park and Qian (2007, 2011), and introduce a semi-parametric functional autoregressive (FAR) model for forecasting a time-varying distribution of the sectoral inflation rates. In particular, the cross-sectional distributions are treated as time series of the functional data and their dynamic nature is estimated via an autoregressive model in a functional space. As the dependence of the variations over sectors is being modelled non-parametrically, we do not impose any assumptions on the class or structure of the distributions, or on the number of dimensions in which the distributions may vary over time. Our approach is in line with the recent developments on the statistical analysis of the functional data, e.g. Cardot, Mas, and Sarda (1999, 2007), Mas (2007), Bowsher and Meeks (2008) and Kargin and Onatski (2008).

Researchers in the past have attempted to incorporate the importance of informational heterogeneity across forecasters by using the data from the Survey of Professional Forecasters (SPF) (e.g. Zarnowitz and Llambros, 1987; Lahiri and Liu, 2006; Lahiri and Sheng, 2008). Instead of using the data like the SPF, we exploit the variation across a large number of sectors for constructing our forecasts. This is important since the SPF data are likely to suffer from a small-sample bias regarding the cross-sectional dimension. The probability density function of the SPF data is reported as a histogram, and thus it is not straightforward to evaluate moments from a histogram (Engelberg, Manski, and Williams, 2009; Clements, 2010). Different measures of inflation uncertainty can reveal different association between expected inflation and uncertainty (Rich and Tracy, 2010). In SPF data, two additional problems exist: heterogeneity of forecasters and change in the panel composition over time. This makes the interpretation of temporal variation in an aggregated forecast more difficult (Manski, 2010). Our framework can easily handle the unbalanced nature of the panel and does not suffer from any attrition bias.

Based on a detailed exploratory analysis of the time-varying cross-sectional moments using the UK sectoral CPI inflation rates over the period January 1997 – February 2008, we observe that the first four cross-sectional moments are highly persistent, all close to being nonstationary. Average of sectoral inflation is correlated with both cross-sectional standard deviation and skewness, though the direction of association is contingent on the location of average inflation rates relative to the expected ones. The mean inflation is positively correlated with the standard deviation and skewness when the actual inflation rate is higher than the expected and negatively correlated otherwise. Our findings thus provide a partial support to Friedman (1977) and Cukierman and Meltzer (1986) and favour the menu cost explanation of Ball and Mankiw (1995) regarding the sluggish adjustment of prices in response to aggregate shocks only under the higher inflation regime. On the contrary, negative associations observed in the low inflation regime, support the exogenous form of downward nominal rigidity as advanced by Tobin (1972). Our finding of the U-shaped relationship between mean inflation and cross-sectional inflationary uncertainty is generally consistent with the recent studies by Chen, Shen, and Xie (2008) and Choi (2010).

The persistence and temporal dependence of moments are the strongly observed characteristics of cross-sectional distribution in our study. The persistence is conventionally modelled by the AR and and GARCH-M processes mostly involving the first and the second moments, which implies that such parametric models are quite restrictive in the selection of the number of moments included. Our approach employs the autoregressive operator to specify the time-dependence of the distribution function and thus allows all the moments to depend on all the past moments. Hence, the FAR model can be viewed as the general autoregressive model with all the higher moments, rendering our approach free from the particular moment specifications as imposed in the conventional parametric models.

To conduct a number of forecasting evaluation exercises, we consider a total of thirteen different models. These include three basic models, denoted FAR, AVE and LAST, where AVE utilises an average of all distributions while LAST uses the last distribution. Further, in order to explicitly deal with the presence of (possible) structural breaks and\or with near-nonstationary nature of inflation rates and their higher moments, we adopt similar methodologies employed by Clements and Hendry (2002, 2006), and develop modified FAR and AVE models. The in-sample forecasting evaluation results show that the performance of (unmodified) FAR and AVE models is rather poor and sometimes dominated even by the benchmark AR model. Importantly, the models incorporating the time-varying mean corrections clearly outperform both AR and ARCH-M models. Overall, we find that the first-differenced FAR model (DFAR), which accounts for near-nonstationary behaviour of inflation rates, outperforms other models in forecasting both the cross-sectional distribution of sectoral inflation rates and the density function of mean inflation rate. Our results show that utilisation of disaggregate data provides more useful information in forecasting the aggregate inflation, than an aggregate parametric model which is likely to result in a misleading forecast. Furthermore, we evaluate uncertainty bands with mean forecasts (similar to the fan-chart type) via the bootstrap simulations, confirming that the realised inflation rates are well within the bands over the whole forecast horizon for modified FAR models. In particular, the uncertainty band is shown to be much tighter with the DFAR model.

Out-of-sample forecasting results obtained from the modified FAR models suggest that the mean is projected to be stable around 2.4%-2.6% over the forecast horizon, March 2008 - February 2010, though the associated uncertainty bands quickly grows within 3-4 months and stays between 1.5% and 4.5% over the 24-month forecast horizon. The inflation rate as announced by BOE is 3.5% in January 2010, which also lie well within our predicted band. In addition, to further relate our empirical findings to the inflation targeting policy pursued by BOE, we also conduct the out-of-sample probability event forecasting of inflation targets through the bootstrap simulations. The probability of achieving the inflation target of less than 2% over the forecasting period is fairly low (around 20-25%) while the probabilities of maintaining inflation between 1% and 3% keep decreasing as the forecast horizon increases (around 65-70%). Our finding suggests that the extent to which CPI inflation would deviate from the target level of 2% in the medium term is highly uncertain, and is generally consistent with the recent evidence as documented in the BOE reports: the previous report predicting that CPI inflation will be below the 2% target in 2009 whereas the February 2010 report documenting that the probability of inflation above the 2% in the near term is high due to the continuing impact of Sterling's depreciation.

The rest is organised as follows. Section 2 introduces the FAR methodology. The exploratory data analysis of the sectoral inflation rates are reported in Section 3. The main estimation and forecasting results based on the FAR models are provided along with the policy implications in Section 4. Section 5 concludes.

3.2 Functional Autoregressive Distribution

The statistical analysis of the functional data has been developed by Cardot, Mas, and Sarda (1999, 2007), Bosq (2000) and Mas (2007). Recently, the FAR modelling has been applied explicitly to the density function of the high frequency stock returns (Park and Qian, 2007, 2011) and the economic function such as the yield-curve (Bowsher and Meeks, 2008; Kargin and Onatski, 2008). In the line with these developments, we develop the hybrid FAR modelling to an analysis of the time-varying cross-sectional distribution of sectoral inflation rates.

Let $\{f_t\}_{t=1}^{T}$ be the sequence of the cross-sectional distribution of sectoral inflation rates (π_{it}). Notice that the national inflation rate (π_t) is constructed as the weighted average of N sectoral inflation rates at each time t:

$$\pi_t = \sum_{i=1}^{N} v_{it}\pi_{it},\ t = 1,...,T, \tag{3.1}$$

where v_{it} is the weight on π_{it} with $\sum_{i=1}^{N} v_{it} = 1$. Then we define the fluctuation of the cross-sectional distribution, w_t that is a deviation from the well-defined common expectation of the distribution, $\mathbb{E}[f_t]$:

$$w_t = f_t - \mathbb{E}[f_t],\ t = 1,...,T. \tag{3.2}$$

We assume that this fluctuation quickly or slowly disappears as a time period increases and this adjustment mechanism is specified by an autoregressive process. Hence, $\{w_t\}_{t=1}^T$ is generated by an autoregressive process of order one in a functional space:

$$w_t = Aw_{t-1} + \varepsilon_t,\ t = 1,\ldots,T, \tag{3.3}$$

where A is an autoregressive operator in the Hilbert space ($\mathscr{H}$) and $\{\varepsilon_t\}_{t=1}^T$ is the sequence of the functional white noise process. This model is referred to the FAR of order one in a functional space. Combining (3.2) and (3.3), we get:

$$f_t = \mathbb{E}[f_t] + Aw_{t-1} + \varepsilon_t. \tag{3.4}$$

This portrays that the cross-sectional distribution at time t consists of the common expectation ($\mathbb{E}[f_t]$) and the correction of the fluctuation at time $t-1$ (Aw_{t-1}).

Given our need of estimating the non-parametric cross-sectional distribution at each time to construct the FAR, we require a weighted kernel density estimator (Marzio and Taylor, 2004; Hazelton and Turlach, 2009):

$$\hat{f}_t(z) = \frac{1}{h_t}\sum_{i=1}^{N} v_{it} K\left(\frac{z-\pi_{it}}{h_t}\right),\ t = 1,...,T \tag{3.5}$$

where K is a kernel, N the number of sectors and h_t a bandwidth. One important issue lies on the selection of the appropriate kernel and the corresponding bandwidth. We follow Silverman (1986) and use a Gaussian kernel with an optimal bandwidth given by $h_t = 1.06\sigma_t N^{-1/5}$ with the standard deviation σ_t of π_{it} at each time t.[1] We set 1,024 grids for estimating an empirical cross-sectional distribution in the empirical analysis, $z = \{z_1, z_2, \ldots, z_{1,024}\}$. Given the sequence of the estimated cross-sectional distributions, $\left\{\hat{f}_t\right\}_{t=1}^T$, we estimate the sequence of fluctuation by $\hat{w}_t = \hat{f}_t - \bar{f}$, where $\bar{f} = \frac{1}{T}\sum_{t=1}^T \hat{f}_t$ is the consistent estimate of $\mathbb{E}[f_t]$.

To estimate (3.3), we need to estimate an autoregressive operator, A by utilizing the autocovariance operators,

$$C_s = \mathbb{E}(w_t \otimes w_{t-s}),\ s = 0,1,2,.., \tag{3.6}$$

where $\otimes$ denotes a tensor product.[2] Considering two autocovariance operators of order 0

[1] Various other kernels are also available in the literature including Epanechnikov, Bi-weight, Triangular and Rectangular. Alternatively, we may also allow the bandwidth parameter to be time-varying and estimate it using the cross-validation technique.

[2] In the infinite-dimensional space, it is defined as $(u \otimes v) = \langle v, \cdot \rangle u$ that is equivalent to outer product uv' in the finite-dimensional vector space.

and 1, denoted C_0 and C_1, and using the relationship, $C_1 = AC_0$, we obtain an autoregressive operator of order 1:

$$A = C_0^{-1} C_1. \tag{3.7}$$

Autocovariance operators are consistently estimated by $\hat{C}_s = T^{-1} \sum_{t=1}^{T} (\hat{w}_t \otimes \hat{w}_{t-s})$ for $s = 0, 1, 2, ...$ Hence, the autoregressive operator can be consistently estimated by $\hat{A} = \hat{C}_0^{-1} \hat{C}_1$. Using the spectral representation for a compact and self-adjoint C_0:

$$C_0 = \sum_{\ell=1}^{\infty} \lambda_\ell (v_\ell \otimes v_\ell), \tag{3.8}$$

where (λ_ℓ, v_ℓ) are the pair of eigenvalue and eigenfunction of C_0, the inverse of C_0 can be easily obtained by

$$C_0^{-1} = \sum_{\ell=1}^{\infty} \lambda_\ell^{-1} (v_\ell \otimes v_\ell). \tag{3.9}$$

However, notice that there is an ill-posed inverse problem since C_0 is defined on the infinite dimension, that is, it needs the infinite number of eigenvalues and corresponding eigenfunctions. To avoid this problem, we restrict A to be in the finite-dimensional subspace of $\mathscr{H}$, define V_L as the subspace of $\mathscr{H}$ spanned by the L-eigenfunction, $v_1, \dots, v_L$, and let $C_{0,L} = \Pi_L C_0 \Pi_L$, where Π_L is the projector on V_L. Then we approximate the inverse of C_0 by

$$C_{0,L}^{+} = \sum_{\ell=1}^{L} \lambda_\ell^{-1} (v_\ell \otimes v_\ell), \tag{3.10}$$

which is defined on V_L. The practical choice of L is guided by the footnote 9 of Section 2.2.1. [3]

Therefore, the autoregressive operator on the subspace V_L of $\mathscr{H}$ is now estimated by

$$\hat{A}_L = \hat{C}_{0,L}^{+} \hat{C}_1. \tag{3.11}$$

Then, an m-step ahead forecasts of the cross-sectional distribution is evaluated by

$$\hat{f}_{T+m} = \bar{f} + \hat{A}_L^m \left(\hat{f}_T - \bar{f} \right), \; m = 1, 2, ..., \tag{3.12}$$

where $\hat{f}_T$ is the estimate of the cross-sectional distribution at time T. Finally, the m-step ahead forecasts of the national inflation rates are obtained by integrating out the cross-

[3] We set $L_{max} = 20$ through our study and find that the cross-validation procedure given by (2.14), selects the optimal value of L ranging between 5 and 10.

sectional distribution for x, which is the continuous version of (3.1):

$$\hat{\pi}_{T+m} = \int_{-\infty}^{\infty} x \hat{f}_{T+m}(x)\, dx, \tag{3.13}$$

where the integral operator is numerically approximated by the middle Riemann sum given grid set.

3.3 Data and Analysis of time-varying Moments

We use a set of sectoral inflation rates (defined as the annual percentage change) based on Consumer Price Index (CPI, 2005 =100) and their respective weights from the Office for National Statistics. The data spans over January 1997 to February 2008, thus giving us a total of 134 monthly observations.[4] The overall average inflation rate (across the sectors and over the time periods) stands at 1.5 % with a median rate of 1.6 %. Considerable heterogeneity exists across the sectors: Liquid Fuels experiencing the highest inflation (10.5 %) and Information Processing Equipment the lowest (-22.3 %). The standard deviation ranges from 0.4 % (Social Protection) to 26.5 % (Liquid Fuels). Out of 85 sub-sectors, inflation is positively skewed for 49 sub-sectors while excess kurtosis is present for 36 sub-sectors. The Jarque-Bera statistic is significant at the 5% significance level, providing a strong evidence against the normality.

To understand the time-varying nature along with the degree of association between the first four cross-sectional moments, we provide the time series plots of four moments (mean, standard deviation, skewness, kurtosis) evaluated from the cross-sectional distribution of sectoral inflation rates that are obtained by apply the kernel estimation to the sectoral inflation rates (Figure 3.2). The maximum mean of sectoral inflation rates was in December 2006 and the minimum in June 2002. In 51 out of 134 months, inflation is positively skewed. Cross-sectional uncertainty of inflation substantially differs across time periods: deviating by more than 1 percentage point from the target only thrice, but around 11 times by more than 0.5 percentage point in recent years (September 2006 to February 2008).

Table 3.1 documents the persistent nature of the first four cross-sectional moments for the whole sample. The mean (the national inflation rate by construction) is highly persistent with an AR(1) coefficient being 0.93. Interestingly, both variance and skewness also exhibit high persistence with an AR(1) coefficient of 0.90 and 0.89, respectively. These findings

[4] In some sectors we have less observations as the disaggregation has started at a later date. Sub-sector consists of 79 sectors (January 1997 - November 2000), 84 sectors (December 2000 - November 2001) and 85 sectors (December 2001 - February 2008). Further, sub-sector weights change annually. The data can be downloaded from `www.statistics.gov.uk/statbase/TSDSeries1.asp`.

Figure 3.1: Time-varying Moments of Sectoral Inflation Rates.

All moments are evaluated from the estimated cross-sectional distribution of the sectoral inflation rates. See also Table 3.1 for details.

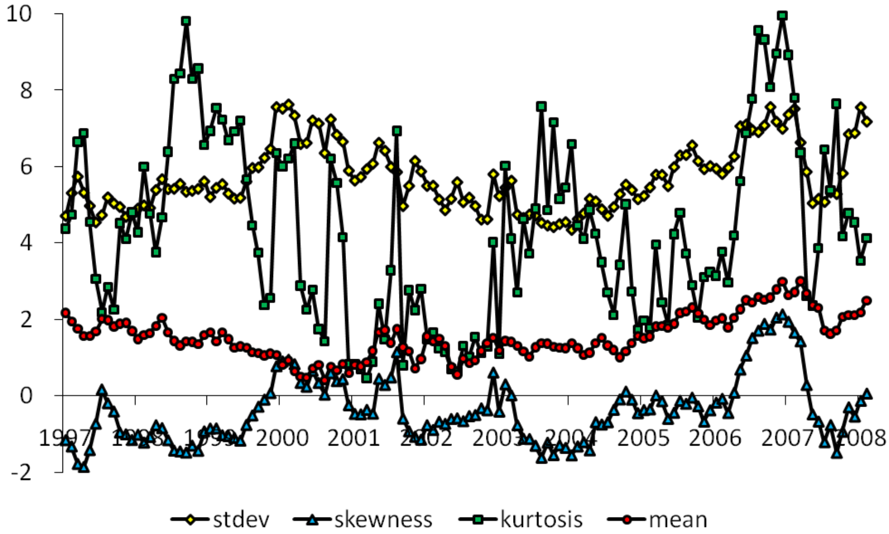


Table 3.1: Persistence of Moments of Sectoral Inflation Rates

All moments are computed from the estimated cross-sectional distribution of the sectoral inflation rates by the middle Riemann sum. The null hypothesis of unit root is tested using the Augmented Dickey-Fuller (ADF) specification (a) with intercept only and (b) with intercept and trend. $[\cdot]$ contains the 95% confidence interval of AR(1) coefficient. Figures in $(\cdot)$ denote the *p*-value of the ADF test obtained using the MacKinnon approximation.

Moments	AR(1) Coefficient	Unit Root Test	
Mean	0.933 [0.862,1.004]	-1.225 (0.663)	-2.059 (0.569)
Stdev	0.901 [0.826,0.976]	-2.224 (0.198)	-2.304 (0.432)
Skewness	0.892 [0.806,0.978]	-2.891 (0.046)	-3.196 (0.085)
Kurtosis	0.765 [0.656,0.873]	-2.384 (0.146)	-2.374 (0.394)

suggest that mean inflation rate as well as cross-sectional inflation uncertainty (measured by cross-sectional standard deviation and/or skewness) are highly persistent and close to being nonstationary, which is formally confirmed by the Augmented Dickey-Fuller test reported in the third column in Table 3.1.

We now turn to analyse the correlation patterns among the cross-sectional moments. This may provide an insight to understanding the relationship between mean inflation and cross-sectional inflation uncertainty in general. Friedman (1977) in his Nobel prize lecture postulates that an increase in inflation uncertainty exerts a dampening effect on economic efficiency and possibly on output growth. Hence, we expect that there is a positive relationship among inflation rate, volatility and inflation forecast uncertainty. In the context of

Table 3.2: Asymmetric Relationships Between Moments of Sectoral Inflation Rates

We decompose moments into those under the high and the low inflation regimes and apply a threshold regression. μ_t, σ_t, s_t and k_t denote the mean, the standard deviation, the skewness and the kurtosis of the cross-sectional distribution of sectoral inflation rates. The high (+) and the low (-) inflation regimes are defined for the case that a mean inflation rate is over and under the overall mean inflation rate (1.5%). Figures in $(\cdot)$ and $[\cdot]$ denote the standard errors of coefficients and the p-value of F-test for the null of symmetry, respectively.

Model	β^+	β^-	
$\mu_t = (\alpha^+ + \beta^+ \sigma_t^+) + (\alpha^- + \beta^- \sigma_t^-) + \varepsilon_t$	0.32 (0.00)	-0.21 (0.00)	[0.00]
$\mu_t = (\alpha^+ + \beta^+ s_t^+) + (\alpha^- + \beta^- s_t^-) + \varepsilon_t$	0.29 (0.00)	-0.20 (0.00)	[0.00]
$\mu_t = (\alpha^+ + \beta^+ k_t^+) + (\alpha^- + \beta^- k_t^-) + \varepsilon_t$	0.07 (0.00)	0.04 (0.00)	[0.00]
$\sigma_t = (\alpha^+ + \beta^+ s_t^+) + (\alpha^- + \beta^- s_t^-) + \varepsilon_t$	0.47 (0.00)	0.19 (0.00)	[0.00]

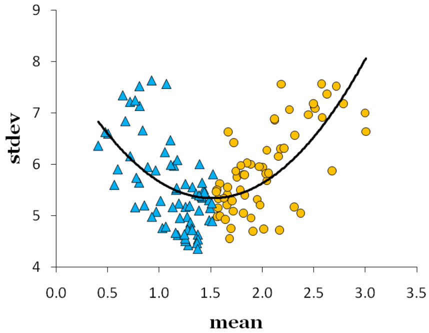
a reverse causation, Cukierman and Meltzer (1986) posit that higher inflation uncertainty increases the inflation rate while Holland (1995) predicts an opposite effect of uncertainty on inflation. Furthermore, the skewness could result from an exogenous form of downward nominal rigidity in product markets (Tobin, 1972) or endogeneity as suggested by the menu cost model (Ball and Mankiw, 1995). The former would imply a negative relation between mean inflation rate and skewness whereas the latter a positive one. This is important as the central banks need to differentiate between these two sources as they would exert different implications for an optimal monetary policy. For example, with downward nominal rigidity, lower inflation rates are as harmful as they complicate the (downward) adjustment of relative prices whereas it is desirable in other case as it decreases the costs associated with changing prices.

We investigate the relationships using a simple two-regime threshold model, where the high and the low inflation regimes are classified according to whether actual inflation rates are greater and less than the expected rates. Figure 3.2 displays the scatter plots between the pair of the cross-sectional moments for the whole period. From the panels (a) and (b), we observe that the relationships between mean and standard deviation and between mean and skewness are remarkably asymmetric (strong U-shaped relationship), though there is no asymmetry between standard deviation and skewness in the high- and the low- inflation-regimes. Table 3.2 reports the formal threshold regression results, showing that the impacts of standard deviation (skewness) on mean are measured at 0.32 and -0.21 (0.29 and -0.20), respectively, in the high- and low-inflation-regime. Moreover, the null hypotheses of symmetry are strongly rejected (except for the mean-kurtosis relationship). Interestingly, we find a positive association between standard deviation and skewness which is robust across different regimes, though the association is considerably stronger in the high-inflation regime.

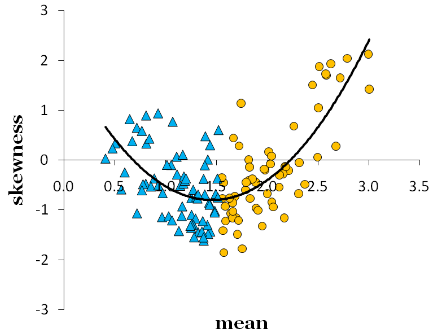
A positive association between the level of realised inflation and uncertainty measured

Figure 3.2: The Scatter Plots between Moments of Sectoral Inflation Rates.

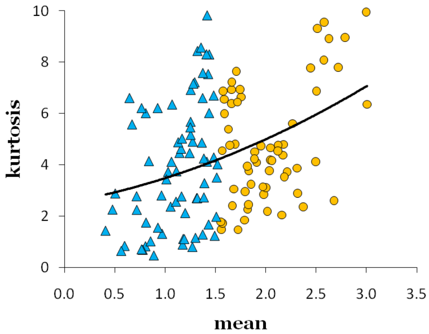
The circle denotes moments under the higher inflation regime while the triangle denotes moments under the lower inflation regime. The high (low) inflation regime is defined for the case that a mean inflation rate is over (under) the overall mean inflation rate (1.5%).



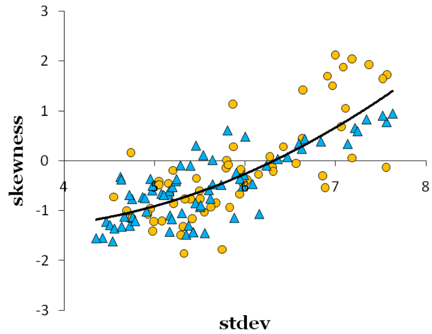
(a) mean-stdev



(b) mean-skewness



(c) mean-kurtosis



(d) stdev-skewness

by either standard deviation or skewness in the high-inflation-regime confirms the predictions by Friedman (1977) and Cukierman and Meltzer (1986). Such a positive association may support the menu cost explanation proposed by Ball and Mankiw (1995), suggesting sluggishness in the adjustment of individual prices in response to aggregate shocks. On the other hand, interestingly, a negative association in the low-inflation-regime may support the predictions by Tobin (1972) and Holland (1995). Recently, Chen, Shen, and Xie (2008) also document the U-shaped relationship between inflation and inflation uncertainty for Hong Kong, Singapore, South Korea and Taiwan. In particular, Choi (2010) finds that the relationship between inflation and relative price variability is approximately U-shaped for the US and Japan, and offers a calibration-based explanation for the observed asymmetric relationship using the modified Calvo (1983) model with the sectorally heterogeneneous price rigidities. However, Choi's approach is somewhat limited in general and cannot accommodate the asymmetric relationship between the inflation and skewness in particular.

3.4 Forecasting The UK Inflation Rates

Our goal in this section is to forecast the density function of the national inflation rate utilizing informational contents in the sectoral inflation rates by combining the kernel density estimation and the functional autoregressive modelling. This hybrid approach enables us to take into account the relative advantage of parametric and non-parametric approach in a robust manner.

We then suggest to use the non-parametric bootstrap technique in order to generate the empirical distribution of the national inflation that is constructed as the weighted average of N sectoral inflation rates at each time period as follows: We draw the prediction errors, $\hat{\varepsilon}^{(b)}_{T+h}$ for $h = 1,...,H$, and $b = 1,...B$, with replacement from the pool of the residuals, $\{\hat{\varepsilon}_1,...,\hat{\varepsilon}_T\}$, estimated from (3.4). Then, the h-step forecast of cross-sectional distribution $\left(\hat{f}_{T+h}\right)$ is resampled by

$$\hat{w}^{(b)}_{T+h} = \hat{w}_{T+h} + \sum_{i=1}^{h} \hat{A}_L^{h-i} \hat{\varepsilon}^{(b)}_{T+i},\ h = 1,...,H,\ b = 1,...B, \tag{3.14}$$

where $\hat{w}_{T+h} = \hat{f}_{T+h} - \bar{f}$ and $\hat{A}_L$ is the consistent estimator of the autoregressive operator defined in (3.7). It is easily seen that the forecast of the national inflation rate, denoted $\hat{\pi}_{T+h}$, is obtained as the mean of $\hat{f}_{T+h}$ using (3.13) while the density forecast of the national inflation rate can be obtained from the empirical distribution of $\left\{\hat{\pi}^{(b)}_{T+h}\right\}_{b=1}^{B}$. Then, the event probability forecasts can be evaluated by

$$\Pr\left(\hat{\pi}_{T+h} < a\right) = \frac{1}{B}\sum_{b=1}^{B} I\left(\hat{\pi}_{T+h}^{(b)} < a\right), \tag{3.15}$$

where a is the threshold of interest.

Following Park and Qian (2007), we consider three basic models, denoted FAR, AVE and LAST, where AVE utilises an average of all observed distributions while LAST uses the last observation for forecasting purpose. In practice, however, the standard FAR is unlikely to be robust to the presence of the structural breaks and/or the regime switching events. Hence, we consider two modifications. First, we follow Clements and Hendry (2006), who identify structural instability as a key factor behind poor forecasting performance, and consider the intercept shift modification by allowing the time-varying mean, $\mathbb{E}[f_t]$ in FAR, where $\mathbb{E}[f_t]$ is estimated by 3-, 6-, 9- and 12-month moving averages of f_t. These models are referred to as FAR3M, FAR6M, FAR9M, FAR12M, AVE3M, AVE6M, AVE9M and AVE12M, respectively. Secondly, we consider the differenced version of the FAR model, denoted DFAR, in order to accommodate near nonstationary nature of mean inflation rate and cross-sectional higher-order moments as discussed in Section 3. By construction, it removes $\mathbb{E}[f_t]$ in FAR, and thus eliminate any effects of structural breaks on the conditional mean distribution. This is similar to the methodology recommended by Clements and Hendry (2002), who suggest a double-differencing VAR in order to avoid the forecasting difficulties in the presence of structural breaks, see also Kapetanios, Labhard, and Price (2007). Finally, we consider the combination of all twelve models using equal weights (referred as EWMAVE). For example, Elliot and Timmermann (2008) find that this simple combination modelling performs relatively better than the more sophisticated model based on the optimal weights. See Table 3.3 for details of 13 different models.

3.4.1 In-sample Forecasting Evaluation

We conduct four evaluation exercises in a recursive manner to evaluate the forecasting performance of all the models considered above. This practice of holding out sample is called “pseudo real time” experiments, see Elliot and Timmermann (2008) for details.

Our first emphasis is on forecasting mean inflation rates. As benchmark models, we consider an autoregressive model (AR) with a maximum lag order of 12, and an ARCH in mean model (ARCH-M) adding a time-varying conditional variance term modelled by ARCH(1) in AR(1). Those two models utilise only national inflation rates while our proposed models utilise disaggregate sectoral inflation rates. We first estimate all the thirteen

Table 3.3: Forecasting Models

The time-varying mean functions; $\bar{f}_{3M}$, $\bar{f}_{6M}$, $\bar{f}_{9M}$, $\bar{f}_{12M}$, are obtained by $\bar{f}_{3M} = \frac{1}{3}\sum_{i=0}^{2} f_{t-i}$, $\bar{f}_{6M} = \frac{1}{6}\sum_{i=0}^{5} f_{t-i}$, $\bar{f}_{9M} = \frac{1}{9}\sum_{i=0}^{8} f_{t-i}$ and $\bar{f}_{12M} = \frac{1}{12}\sum_{i=0}^{11} f_{t-i}$ at each time point t.

Model		Equation
(1)	FAR	$w_t = Aw_{t-1} + \varepsilon_t$, where $w_t = f_t - \bar{f}$
(2)	FAR3M	$w_t = A_3 w_{t-1} + \varepsilon_t$, where $w_t = f_t - \bar{f}_{3M}$
(3)	FAR6M	$w_t = A_{6M} w_{t-1} + \varepsilon_t$, where $w_t = f_t - \bar{f}_{6M}$
(4)	FAR9M	$w_t = A_{9M} w_{t-1} + \varepsilon_t$, where $w_t = f_t - \bar{f}_{9M}$
(5)	FAR12M	$w_t = A_{12M} w_{t-1} + \varepsilon_t$, where $w_t = f_t - \bar{f}_{12M}$
(6)	AVE	$f_t = \bar{f} + \varepsilon_t$
(7)	AVE3M	$f_t = \bar{f}_{3M} + \varepsilon_t$
(8)	AVE6M	$f_t = \bar{f}_{6M} + \varepsilon_t$
(9)	AVE9M	$f_t = \bar{f}_{9M} + \varepsilon_t$
(10)	AVE12M	$f_t = \bar{f}_{12M} + \varepsilon_t$
(11)	LAST	$f_t = f_{t-1} + \varepsilon_t$
(12)	DFAR	$w_t = Bw_{t-1} + v_t$, where $w_t = f_t - f_{t-1}$
(13)	EWMAVE	$f_t = \frac{1}{12}\sum_{i=1}^{12} f_t^{(i)}$, $i = FAR, \ldots, DFAR$

models over the period from January 1997 to December 2002, and compute one month to twelve month-ahead distribution forecasts of sectoral inflation rates. We repeat this process moving forward one month at a time, ending with forecasts for March 2007 - February 2008 based on models estimated over the period January 1997 – February 2007. Hence, we obtain 51 observations for each of m-month-ahead distribution forecasts, $m = 1, ..., 12$, giving us a total of 612 experiments. We use four evaluation criteria: mean absolute error (MAE), mean absolute percentage error (MAPE), percent mean absolute deviation (PMAD) and mean square error (MSE). MAE and MSE are popular measures for 'cancelling out' effects by averaging the absolute value and the squares of forecasting errors, though MAE is less sensitive to outliers. Furthermore, MAPE and PMAD also take into account the different scale effects. The results reported in Table 3.4 clearly demonstrate that the performance of DFAR (and FAR3M) is the best. The performance of FAR and AVE without accommodating the time-varying mean correction is poor and dominated by the benchmark AR. The models incorporating the time-varying mean corrections such as FAR3M-FAR12M and AVE3M-AVE12M are shown to outperform AR and ARCH-M. This may reflect that it is important for the selected models to be able to accommodate both the time-varying and nonstationary behaviour of inflation. Furthermore, the results clarify that utilising the information of disaggregate data improves forecasting performance and that the use of aggregate data is likely to result in severely biased forecast as it ignores the true underlying dynamics of the constituent micro units (Pesaran and Smith, 1995; Hsiao, Shen, and Fujiki, 2005).

Second, we evaluate the cross-sectional distribution forecasting for alternative models by comparing the divergence criteria that measure the distance between the forecasted and

Table 3.4: Mean Forecasting Performance of FAR Models against AR and ARCH-M Models

Measures for forecasting errors are defined by $MAE = N^{-1}\sum_{t=1}^{N}|\hat{\varepsilon}_t|$, $MAPE = N^{-1}\sum_{t=1}^{N}|\hat{\varepsilon}_t/\pi_t|$, $PMAD = \sum_{t=1}^{N}|\hat{\varepsilon}_t|/\sum_{t=1}^{N}|\pi_t|$ and $MSE = N^{-1}\sum_{t=1}^{N}\hat{\varepsilon}_t^2$. By minimizing aggregate forecasting errors, the different AR models are selected for each measure as follows:

$$(MAE)\,\pi_t = c+\phi_3\pi_{t-3}+\phi_7\pi_{t-7}+\phi_8\pi_{t-8}+\varepsilon_t,$$
$$(MAPE)\,\pi_t = c+\phi_3\pi_{t-3}+\phi_5\pi_{t-5}+\phi_7\pi_{t-7}+\phi_8\pi_{t-8}+\phi_9\pi_{t-9}+\phi_{10}\pi_{t-10}+\phi_{12}\pi_{t-12}+\varepsilon_t,$$
$$(PMAD)\,\pi_t = c+\phi_3\pi_{t-3}+\phi_5\pi_{t-5}+\phi_7\pi_{t-7}+\phi_8\pi_{t-8}+\phi_9\pi_{t-9}+\phi_{10}\pi_{t-10}+\phi_{12}\pi_{t-12}+\varepsilon_t,$$
$$(MSE)\,\pi_t = c+\phi_1\pi_{t-3}+\phi_7\pi_{t-7}+\varepsilon_t.$$

The conditional mean and variance of ARCH-M is specified by $\pi_t = c+\phi\pi_{t-1}+\gamma\sigma_t^2+u_t$ and $\sigma_t^2 = \kappa+\theta\sigma_{t-1}^2$, respectively.

Model	MAE	MAPE	PMAD	MSE
FAR	0.468	0.221	0.226	0.367
FAR3M	0.336	0.178	0.175	0.190
FAR6M	0.337	0.176	0.174	0.197
FAR9M	0.340	0.175	0.173	0.201
FAR12M	0.347	0.177	0.177	0.202
AVE	0.586	0.269	0.277	0.562
AVE3M	0.333	0.173	0.171	0.188
AVE6M	0.342	0.174	0.174	0.191
AVE9M	0.354	0.178	0.179	0.199
AVE12M	0.368	0.184	0.187	0.211
LAST	0.331	0.175	0.173	0.187
DFAR	0.327	0.173	0.171	0.182
EWMAVE	0.345	0.173	0.174	0.196
AR	0.438	0.212	0.217	0.294
ARCH-M	0.793	0.416	0.430	0.910

the true distribution; namely, the Hilbert norm (D_H), the uniform norm (D_U) and the generalised entropy (D_E). D_H is useful for evaluating the goodness-of-fit of the model, D_U is informative for comparing the closeness of the function shape and D_E assesses the difference in information contents between the forecasted and the true density function. (see Appendix for definitions). Further we evaluate the density forecasting for the mean inflation rate using the probability integral transformations (PIT) proposed by Diebold, Gunter, and Tay (1998).[5] We estimate all thirteen models over the period from January 1997 to December 2002, compute one month-ahead forecast of cross-sectional distribution in each model, and repeat the process moving forward one month at a time, ending with forecasts for February 2008 based on models estimated over the period January 1997 - January 2008. We calculate divergence criteria between the forecasted and the actual distribution both in mean and median. Table 3.5 presents these evaluation results. Mean and median divergence criteria show that LAST and DFAR have minimum values while AVE has the maximum difference for all the divergence criteria. The variants of FAR models and EWMAVE have moderate values. To conduct the PIT test, we use a bootstrapping method introduced above. In this connection we report the Kolmogorov-Smirnov statistic for testing the null hypothesis that the forecasted and the actual density functions are equal for each of the models. The null hypothesis is rejected for FAR, FAR12M, AVE, AVE9M, AVE12M and EWMAVE at different levels of significance. The Kolmogorov-Smirnov statistic takes the minimum value for DFAR followed by the LAST. Combining these results with our previous results, we confirm that DFAR outperforms other models in forecasting both the cross-sectional distribution of sectoral inflation rates and density function of mean inflation rate.

Given the success with forecasting evaluation exercises, we extend mean forecasting exercise with associated forecasting uncertainty; namely the uncertainty bands similar to the fan-chart type. We estimate all thirteen models over the period from January 1997 to June 2006, and compute one-month to twenty-month-ahead forecasts of cross-sectional distributions. Uncertainty bands and event probabilities are obtained by the bootstrap method describe above.[6] Figure 3.3 presents in-sample mean forecasting (dash line) and actual national inflation rate (solid line) with the confidence bands (shades) that present 12 intervals from 25th permiles (bottom) to 975th permiles (top) out of empirical distribution. A clear conclusion emerges: the realised inflation rates are well within the bands over the whole

[5]Clements (2004) uses PIT to evaluate the BOE's density forecast of year-ahead inflation over 1997Q3-2002Q1. Further, Mitchell and Hall (2005) evaluate the performance of BOE's and NIESR's density forecasts using Kullback-Leibler information criterion.

[6]We use bootstrap method to construct the uncertain bands of the central forecasts using the residuals obtained from each of the estimated models. Drawing residuals with replacement we re-estimate the density functions with 1,000 iterations. Using 1,000 sample paths of the density functions of national inflation rates we draw uncertainty bands for each month by computing the respective quantiles.

Table 3.5: Density Forecasting Performance of FAR Models

Hilbert norm (D_H), uniform norm (D_U) and entropy (D_E) are employed for divergence criteria. The uniform norm (D_U) and Hilbert norm (D_H) are formulated by

$$D_U\left(\hat{f}_t, f_t\right) = \frac{\int \left(\hat{f}_t(x) - f_t(x)\right)^2 dx}{\int \hat{f}_t(x)^2 dx + \int \hat{f}_t(x)^2 dx}, D_H\left(\hat{f}_t, f_t\right) = \frac{\sup_x \left|\hat{f}_t(x) - f_t(x)\right|}{\sup_x f_t(x)}$$

where $\hat{f}_t$ is the distribution forecast and f_t is the realised distribution at t. Following Ullah (1996), we also use the generalised entropy measure (D_E) defined as

$$D_E\left(\hat{f}_t, f_t\right) = \int \hat{f}_t(x)\, g\left(\frac{\hat{f}_t(x)}{f_t(x)}\right) dx,$$

where $g(y) = (\gamma - 1)^{-1}(y^\gamma - 1)$ with $\gamma > 0$ and $\gamma \neq 1$. We follow Park and Qian (2007) and set $\gamma = 1/2$. If g is a natural log function, it becomes the Kullback-Liebler divergence measure. All three quantities are non-negative and give zero value if $\hat{f}_t = f_t$ and so we may call them global errors. Further, the Kolmogorov-Smirnov (K-S) statistics is adopted for testing the equality of empirical distribution of PITs with uniform distribution. The test statistics (K-S) is obtained by $\max\{D_+, D_-\}$, where $D_+ = \max_x\left\{\hat{F}(x) - U(x)\right\}$ and $D_- = \max_x\left\{U(x) - \hat{F}(x)\right\}$. $\hat{F}(x)$ is the empirical cumulative probability function of PIT and $U(x)$ is the uniform cumulative probability distribution on the closed interval $[0,1]$. The divergence criteria are evaluated for the mean and the median value of the 62 closeness between one-step-ahead cross-sectional distribution forecasts and the true distributions for each of the thirteen models. The PIT test is also evaluated using 62 PITs. *, ** and *** denote the rejection of test statistics at the 10%, 5% and 1% significance levels.

	Mean			Median			PIT Test
Model	D_H	D_U	D_E	D_H	D_U	D_E	K-S
FAR	0.0027	0.0741	0.0077	0.0018	0.0670	0.0062	0.402***
FAR3M	0.0025	0.0720	0.0061	0.0020	0.0668	0.0044	0.092
FAR6M	0.0025	0.0717	0.0062	0.0019	0.0668	0.0051	0.101
FAR9M	0.0025	0.0730	0.0067	0.0022	0.0732	0.0051	0.141
FAR12M	0.0026	0.0750	0.0067	0.0023	0.0743	0.0056	0.151*
AVE	0.0082	0.1244	0.0184	0.0057	0.1114	0.0176	0.363***
AVE3M	0.0028	0.0776	0.0071	0.0021	0.0702	0.0057	0.086
AVE6M	0.0037	0.0877	0.0090	0.0025	0.0736	0.0076	0.133
AVE9M	0.0045	0.0998	0.0108	0.0037	0.0941	0.0100	0.181**
AVE12M	0.0051	0.1086	0.0124	0.0043	0.1101	0.0117	0.210***
LAST	0.0022	0.0652	0.0054	0.0016	0.0574	0.0046	0.087
DFAR	0.0024	0.0697	0.0061	0.0017	0.0601	0.0043	0.066
EWMAVE	0.0024	0.0705	0.0064	0.0020	0.0700	0.0050	0.407***

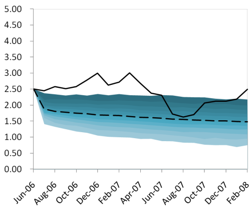
forecast horizon for FAR3M, FAR6M, AVE3M, LAST and DFAR. The realised inflation rates are well outside the bands for the AVE and the FAR due to the failure to accommodate time-varying and nonstationary behaviour of the moments. The confidence band associated with the combination model (EWMAVE) is much smaller compared to the other models and hence underestimates the true uncertainty.

Figure 3.4 plots the in-sample event probability forecasting results over the 20-month forecasting horizon from June 2006 to February 2008. We consider two events: the probability that a national inflation rate is less than 2% and the probability that a national inflation rate lies between 1% and 3%. These two events are closely related to the current remit of the BOE inflation target. Looking at the left panel ($\pi < 2\%$), we find that the probabilities of this event generated by FAR3M and DFAR models are starting quite low (almost close to 0), steadily moving up and reaching around 0.25 only as forecast horizon increases. Indeed, these probabilities seem to track the realised pattern of inflation rates fairly well especially before June 2007 when the mean inflation rates are well above 2%, except for a number of periods when the inflation rates are unexpectedly too low during June and October 2007 caused by the onset of the financial crisis. Even though the realised event probability cannot be observed in practice, this pattern is clear visible in Figure 3.3. The event probabilities produced by FAR and AVE, however, range between 0.88 and 0.98, and are quite misleading. This once again highlights the overall poor forecasting performance of unmodified FAR models. Turning to the right panel ($1\% < \pi < 3\%$), the event probabilities generated by FAR3M and DFAR seem to track fairly well the realised patterns of inflation rates (see also Figure 3.3); the event probabilities are initially close to unity but keep decreasing as the horizon increases, reaching around 75% at the end of horizon, which clearly reflects the increasing forecasting uncertainty associated with the longer horizon.

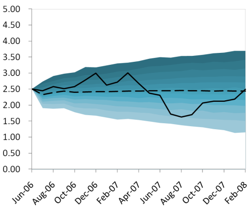
3.4.2 Out-of-sample Forecasting

We now provide two out-of-sample forecasting results; the mean forecasting with uncertainty bands and the probability event forecasting of inflation targets. We estimate the best four models, namely DFAR, FAR3M, AVE3M and LAST over the period from January 1997 to February 2008 and compute one to twenty four month ahead cross-sectional distribution forecasts of sectoral inflation rates for March 2008 – February 2010. Figure 3.5 presents out-of-sample mean forecasting (dash line) with the confidence bands (shades) for these models. The patterns of out-of-sample forecasting results are somewhat similar to those of in-sample-based forecasting. Further, notice that the fan-chart of DFAR is strikingly similar to the fan-chart provided by the BOE obtained using the information up

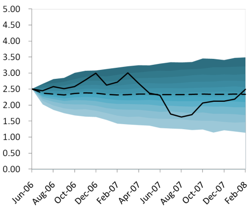
Figure 3.3: In-sample Pseudo Real Time Forecasting of Mean Inflation Rate.



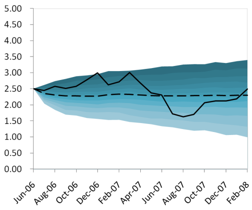
(a) FAR



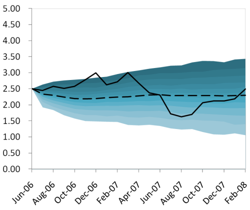
(b) FAR3M



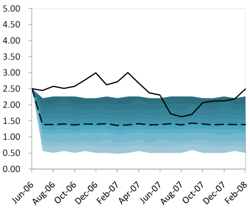
(c) FAR6M



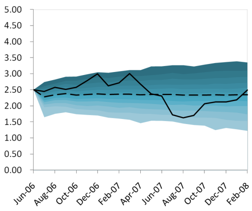
(d) FAR9M



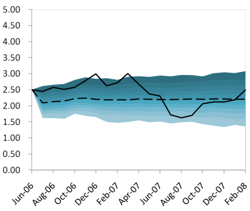
(e) FAR12M



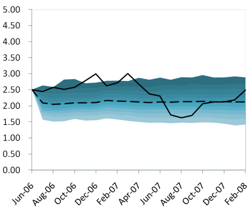
(f) AVE



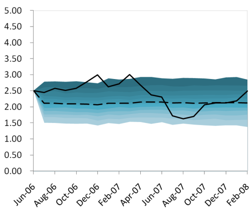
(g) AVE3M



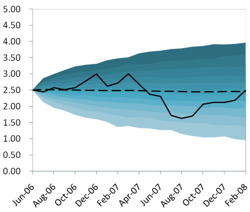
(h) AVE6M



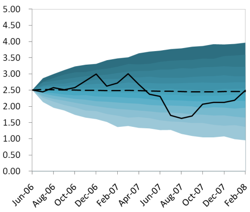
(i) AVE9M



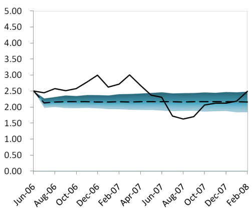
(j) AVE12M



(k) LAST

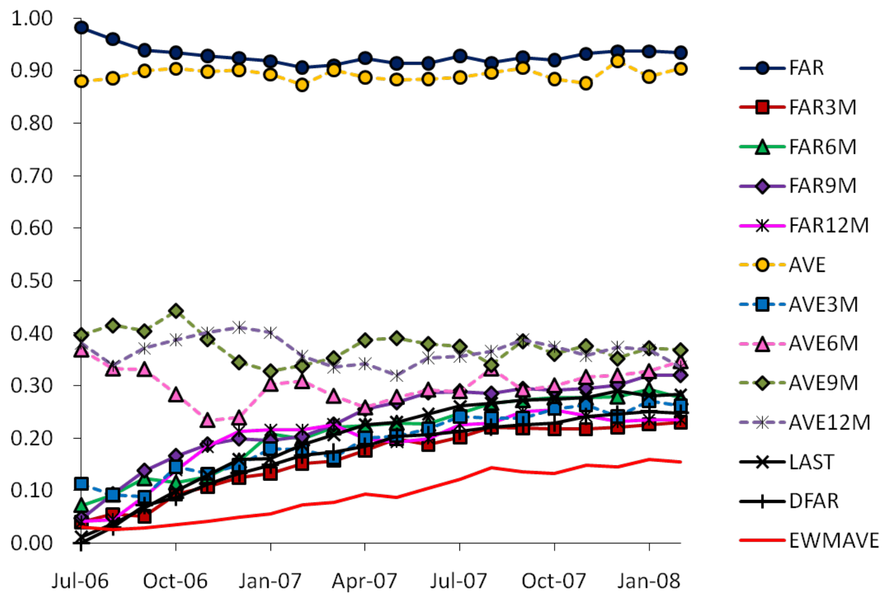


(l) DFAR

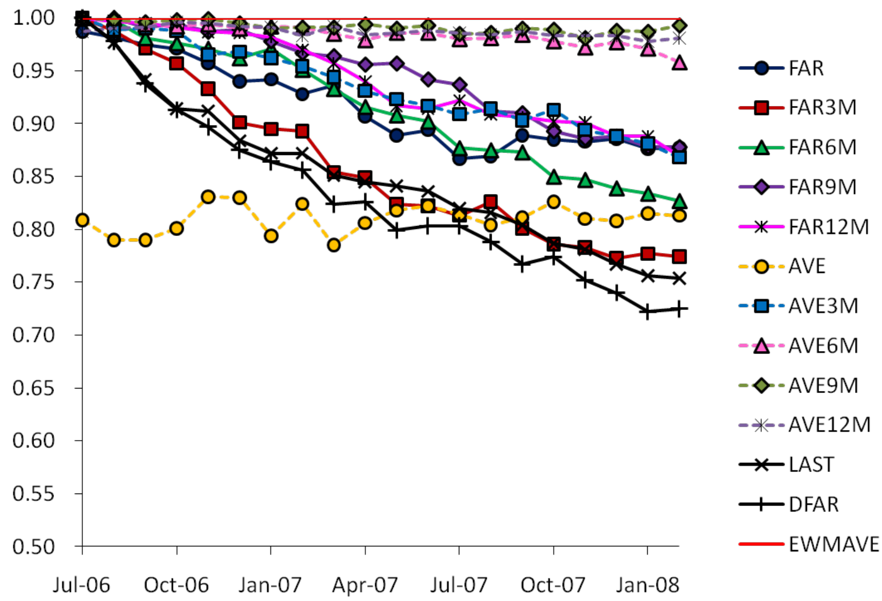


(m) EWMAVE

Figure 3.4: In-sample Pseudo Real Time Probability Event Forecasting.

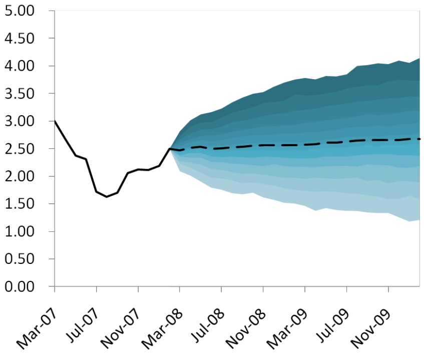


(a) $\mathbb{P}\{\pi < 2\%\}$

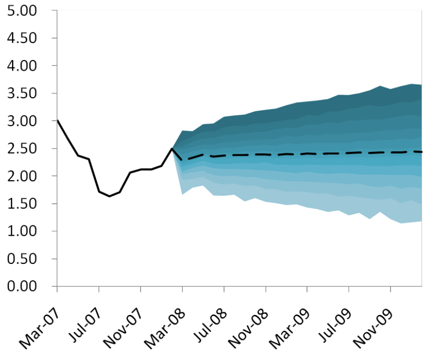


(b) $\mathbb{P}\{1\% < \pi < 3\%\}$

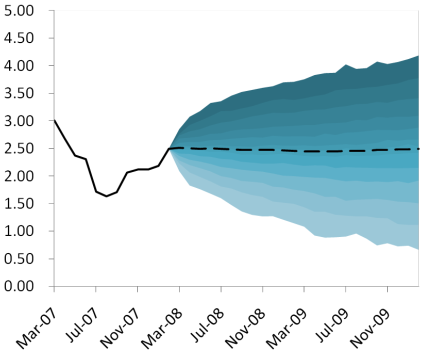
Figure 3.5: Out-of-sample Pseudo Real Time Forecasting of Mean Inflation Rate



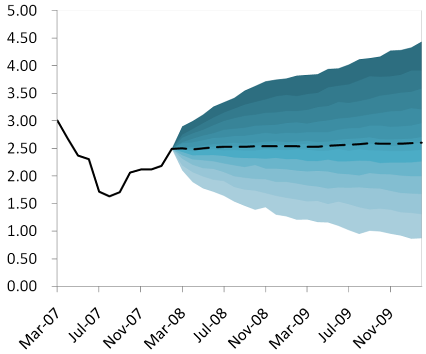
(a) FAR3M



(b)AVE3M



(e) LAST



(f) DFAR

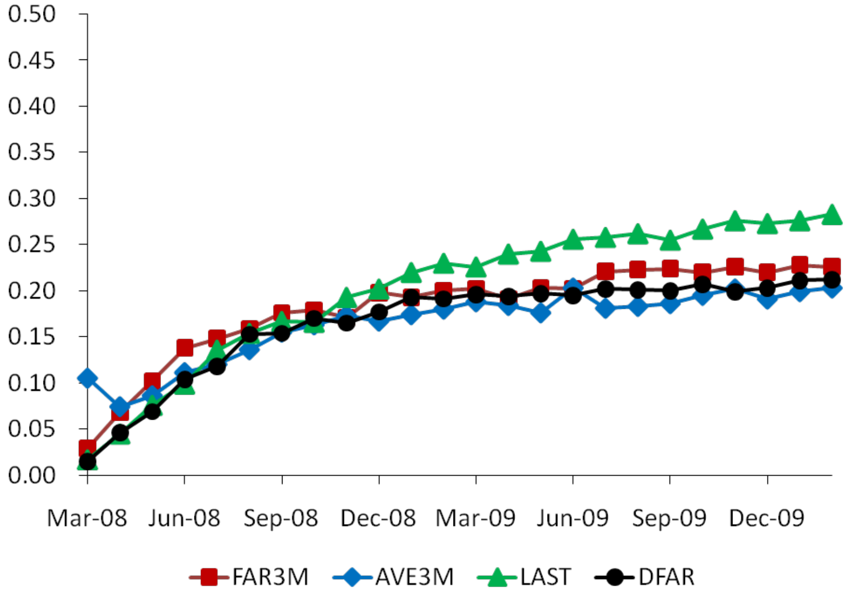
to February 2008. Also the inflation rate as announced in February 2010 by the BOE is 4.0%, lying well within our predicted band. Focusing on the better performing FAR3M and DFAR, the mean is projected to be stable around 2.4-2.6% over March 2008 - February 2010, but uncertainty bands quickly widen within 3-4 months and stay between 1.5% and 4.5%. On the other hand, we observe that the projection made by the AVE3M in general seem to somewhat underestimate the mean inflation and the associated uncertainty and in cases it is outside the announced inflation rates by the BOE.

Figure 3.6 shows out-of-sample probability forecasting results for the above four models. We find that the probabilities of achieving the first target ($\pi < 2\%$) are almost zero over the short-horizon and stay fairly low (around 20-25%) over the two-year horizon. Turning to the second target ($1\% < \pi < 3\%$), the probabilities of achieving the target are initially close to unity but keep decreasing as the horizon increases, reaching around 65-70% at the two-year horizon. The survey conducted by the BOE in February 2008 regarding public attitudes to inflation reveals that median expectations of the rate of inflation over the coming year is 3.3% and lies within our uncertainty band. Combining these results, we may conclude that the (recent) high inflationary pressure has been successfully predicted by our proposed approach.

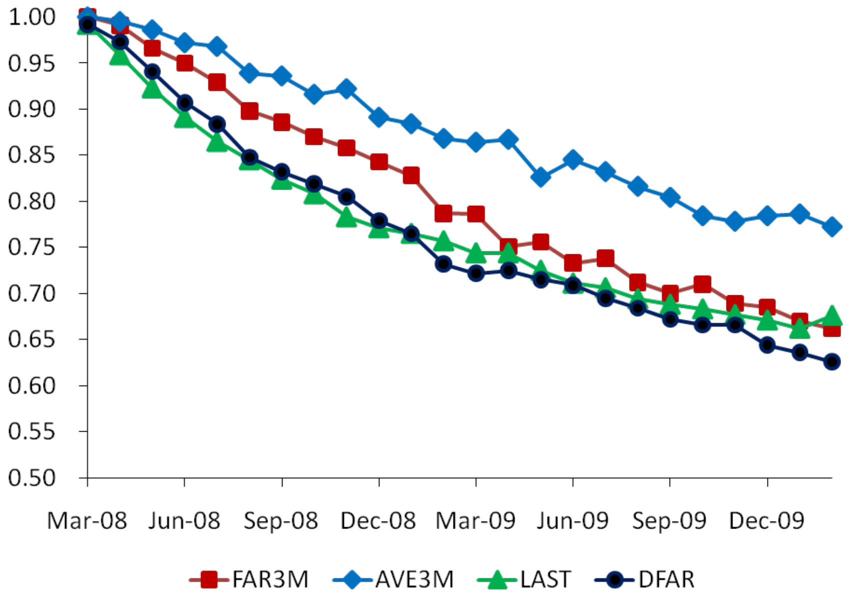
In sum, we conclude that a thorough examination of the cross-sectional distribution provides satisfactory performance based on both in-sample and out-of-sample forecasting exercises. Notice that Clements (2004) provides evidence that the MPC current and next quarter forecasts performs better based on mean square forecast error compared to no-change forecasts (based on Gaussian density with mean given by the actual rate of inflation in the last period), but not the year-ahead density forecasts. Based on our proposed model, we can offer more in this direction. Furthermore, our modelling is parsimonious in nature and provides additional information in understanding the national inflation rates and the associated uncertainty. Our approach is simpler compared to the system-based approach of Garratt, Lee, Pesaran, and Shin (2006), and is in the line with Stock and Watson (2003) who demonstrate the superior out-of-sample forecasting performance of univariate framework for inflation. Therefore, our proposed methodology can be fruitful for policy makers concerned with the complete density forecast for inflation.

Our framework provides an alternative for use by the independent researchers to the forecasts made by the central bank. Use of time-varying cross-sectional distribution of sectoral inflation rates via an autoregressive model in a functional space allows us to incorporate complex and time-varying responses by the policy makers to the disaggregate shock that could affect sectoral inflation rates. The results show that sectoral inflation rate is persistent not only in its first moment but also in higher-order moments both in low and high inflation regimes. The policy reaction of the central bank could differ depending on the nature of the shock and the regime. Assume that the economy consists of two sectors: one with low variance in inflation and the other with high variance. If the shock occurring in the high variance sector shifts the distribution of this sector rightwards, then the overall mean inflation increases. To compensate for this inflation, the low variance sector could follow the distribution of the high variance sector due to wage indexation. Here the central bank would follow an aggressive policy intervention in terms of a contradictory monetary policy by reducing money supply or an increase in policy rate to prevent the low inflation sector to move in the same direction and thereby causing a nation-wide inflation. On the other hand, a shock occurring in the high variance sector could plausibly shift the distribution of this sector leftwards thereby decreasing the overall mean inflation. In this situation, the central bank may be passive in terms of policy intervention as long as it does not create a deflationary environment. Therefore, matching of price changes across sectors is less in low inflation regime compared to that in high inflation regime and it is important to incorporate the degree of occurrence of such price changes for inflation forecasting (Taylor,

Figure 3.6: Out-of-sample Pseudo Real Time Probability Event Forecasting



(a) $\mathbb{P}\{\pi < 2\%\}$



(b) $\mathbb{P}\{1\% < \pi < 3\%\}$

2000). By modelling the time-varying cross-sectional distribution of sectoral inflation rates, our methodology accommodates this feature and thus provides an attractive alternative for inflation forecasting.

3.5 Conclusion

In this study, we use the semi-parametric functional autoregressive approach to model the time-varying density of the UK monthly inflation rates over the period January 1997 - February 2008 using sectoral inflation rates. Our approach exhibits several novel features. First, we allow for cross-sectional dependence and the impact of higher-order moments. Second, instead of using the SPF data, we can exploit the variation across sectors to construct our forecasts. Third, our approach can easily provide the time-varying moments obtained at each period which can help to analyse the descriptive features amongst all the moments. Fourth and importantly, our non-parametric modelling of the time-varying distribution of the inflation rates across sectors does not impose any assumptions on the class or structure of the distributions, or on the number of cross-section units in which distributions may vary over time. Additionally, our approach enables us to analyse the associated probability of inflation lying within a particular range, which is important for an optimal monetary policy making. There is broad agreement that high inflation is qualitatively different from low and stable inflation, while at the same time there are also concerns about excessively low inflation due to the implications of the zero lower bound on nominal interest rates for monetary policy (e.g. Greenwood-Nimmo, Nguyen, and Shin, 2011).

Although our work is limited to the UK data, we believe that this approach can be easily applied to various other countries. A few issues remain to be resolved for future research. First, in order to deal with the high and the low inflation regimes observed across most industrial countries since 1970's, one can construct a mixed model, where the time-varying mean inflation rates are modelled using the parametric AR type models (possibly with a structural break) in the first step, then applies the FAR modelling to the residuals, and forecast uncertainty by bootstraps. Second, our proposed methodology can easily be applied to the SPF data or for that matter any other survey data such as the Livingston Survey and the survey data published by Consensus Economics Incorporation. The SPF data exhibits strong persistence in heterogeneity in the predictions made by the panel members and suffers from the unbalanced nature due to the change in the composition of the panel members over time. The process through which the composition of the panel is being determined is also unknown, see Manski (2010). Such unbalanced and uncertain nature of the SPF data will not create any problem in our framework. Finally, one can extend our framework in a multi-dimensional context by adding some other macroeconomic and financial variables along with inflation. One feasible approach is to combine the marginal distributions of each series forecasted by the FAR model using the Copula approach. This would enable one to perform several multi-dimensional tasks, which would enhance our understanding of a range of macroeconomic stylised facts.

Chapter 4

Time Aggregated Utility from Emerging and Developed Stock Market Investments: Spatial Dominance Approach

4.1 Introduction

Evaluating the performance of alternative investment opportunities is fundamental to portfolio management. In this study, we propose a spatial dominance testing approach (hereafter SPD, Park, 2007) for the evaluation of alternative investments over the planned investment horizon. In particular, we apply our approach to evaluate the performance of emerging vs. developed equity investment, a topic of great interest to global investment managers.

Our study is motivated by the challenges in the recent developments in investment analysis. First, asset return distributions are found to be non-normal and time-varying (Bekaert and Harvey, 1997). Second, evaluating investment performance in different investment horizons is more challenging given the time-varying nature of return distributions (Johnson, 2004; Levy and Duchin, 2004). Given the challenges, an analysis based on the stochastic dominance (hereafter SD) has been shown to offer superior rules to compare the relative performance of investment opportunities. The SD analysis has less restrictive assumptions, incorporates information on the entire return distribution, and requires no specific form of the investor's utility function. It therefore provides rules for an unambiguous ranking of risky choices based on utility theory rather than narrowly specified asset pricing models (see, e.g., Levy, 1992 for a survey of the early literature for the use of the SD approach in analysing economic behaviour under uncertainty). However, the SD approach is likely to face difficulties when the return distributions are time-varying, e.g., an investor interested in evaluating two alternative investment opportunities over a holding period. The distributions of returns achieved at the end of the holding period are assumed to be the same whilst distributions of the interim returns over the period are not the same at any time point.[1] Specifically, one investment has a higher probability of a negative jump in mean or a positive jump in volatility in its return distribution. Without considering this interim information over the holding period, an investor's ranking is indifferent to the alternative investment opportunities. Otherwise, the investors prefer investing in the project with a lower probability of a jump in mean or volatility. Since the SD analysis uses the information of the returns at the end of the planned investment horizon, it cannot explicitly rank the two alternative opportunities of this example. If it tries to aggregate all interim information over the planned investment horizon, it would require very strong assumptions on both the distribution and the utility function. Hence, it has limitations in evaluating the investment opportunities incorporating interim information. Especially, such information is often considered in the emerging markets, e.g., currency crises, political crises, economic reforms, etc. We therefore propose using an SPD approach which has the distinct advantage over the SD test in that it allows explicit comparison of the time aggregated utilities obtained from alternative investment returns accumulated over the planned holding period instead of at the terminal time of the holding period. This is especially important given that asset returns distributions are non-normal and time-varying.

In addition, we contribute to the empirical literature as follows: First, this study aims at shedding light on the debate on the market integration and segmentation hypothesis, which is central in emerging market investment strategies. Bekaert and Harvey (2003) define markets to be integrated when assets with an identical risk require the same expected return irrespective of their domicile. Given this definition, there should be no dominance found between the emerging and developed market investments if the world markets are integrated. It is well documented that the correlation of the emerging market returns and the global market returns is low and shifting through time. Harvey (1995a) finds that the addition of the emerging markets to the portfolio basket significantly shifts the investment opportunity set. Furthermore, Bekaert, Harvey, Lundblad, and Siegel (2007) point out that a simple implementation of the Capital Asset Pricing Model (CAPM) is problematic in analysing emerging market investments. In emerging markets, there are more risks in addition to the

[1] Returns achieved at the end of the holding period is the annual return of investment. The interim returns are regarded as cumulative returns at any point over the holding period.

risk measured by the classical first two moments CAPM. If so, the risk measured in the higher order moments of the return distribution may be relevant, e.g., Harvey and Siddique (2000).

Second, this study examines the impact of currency risk on the relative performance between emerging and developed market investments. Whilst other risk factors such as political risk, failed economic reforms, and regulatory changes may affect both domestic and international investors significantly, currency risk is unique in the sense that it affects the return and volatility of an international investor's holding instantly. The need for currency risk management was underlined after the 1994 Mexico and the 1997 Asia currency crises. An interaction between the foreign exchange rate and the stock price can be seen as a measure of financial market development. Using an equilibrium model for the foreign exchange rate, the stock price and capital flows with incomplete foreign exchange risk trading, Hau and Rey (2005) show that higher returns in the home equity markets relative to foreign markets are associated with home currency depreciation. Generalising this finding, we expect that after combining stock returns and foreign exchange rate movements, all the investments from both markets should offer the same level of return, provided that they have the same level of risk. As Solnik and McLeavey (2003) point out, setting aside consideration of portfolio diversification, the reaction of asset prices to fluctuations in currency values is the matter of prime concern for international investors. The major question is whether stocks and bonds provide a hedge against foreign exchange rate movement. If not, then additional measures need to be taken to manage the foreign exchange rate risk for international investments.

Using the developed and the emerging market indices constructed by Morgan Stanley Capital International over 1988–2007, we evaluate the investment performance in both markets over different investment horizons, namely 3-month, 6-month, 1-year and 5-year. We find that the emerging market investments dominate the developed counterparts over a horizon longer than 6 months only when returns are denominated in the local currency, i.e., without taking currency risks into account. This is consistent with the nature of the emerging markets as they often have a relatively higher growth rate. On the other hand, once the returns are denominated in the US dollar and thus the currency risk is explicitly accommodated, there is no longer evidence in favour of any dominance between the two markets over any investment horizon. This finding clearly demonstrates that the currency risk should be an important factor in the global market environment, especially from the international investors' point of view. This finding supports the view that the emerging and the developed markets are likely to be integrated in both the short- and the long-term once the currency risk factor is fully taken into account.

Furthermore, investigating the effect of currency risk in emerging market investments, we find that investments denominated in the local currency dominate those in the US dollar only over the long-term investment horizon (one year plus) but not over the short-term (up to 6 months). This result suggests that the currency risk is only relevant for the long-term investments in the emerging markets. The indifference to short-term investments indicates that there is no need for hedging currency risk in the short-term investment horizon. However, the discrepancy in the time aggregated utilities obtained from the long-term investments indicates the need of long-term hedging instruments for the currency risk in emerging markets. On the other hand, we find that there is no dominance between the developed market investments denominated in the local currency and the US dollar. This implies that the currency movements between the local currency and the US dollar are fully integrated in the developed markets. Investments in the developed markets therefore prove the hedge against the currency movement between the local currency and the US dollar. In other words, hedging for the currency risk would not increase utility gains for the international investors whose fund is based on the US dollar.

We also conduct sub-period analyses over 1988–1996 and 1999–2007, taking out the 1997 Asian Currency Crisis period, and find that the emerging market investments dominate the developed counterparts in the long-term investment (5-year) even when they are denominated in the US dollar. Combining these results with the whole period results, we may reach the profound yet obvious conclusion that a huge financial crisis such as the Asian Currency Crisis is the key risk factor to weakening the attractiveness of the emerging market investments. If such a crisis could be avoided, emerging market investments could be preferable. However, there should be a non-negligible probability that the emerging markets undergo such a crisis, given the unstable and immature nature of the emerging markets. We argue that the investors should factor this risk into their emerging market investment decisions and the international financial markets are efficient only when such risk factors are priced appropriately. In this regard, such risk factors can be reflected only in the full-period analysis and thus the longer period analysis is preferable in order to better characterise the evolving dynamic nature of the emerging markets. Furthermore, we find evidence that the emerging markets become more integrated into the global market, a finding consistent with Bekaert and Harvey (1995).

The plan of this study is organised as follows. Section 2 introduces a spatial analysis of time series and provides a new testing procedure based on the SPD. Furthermore, it provides an illustration via a Monte Carlo simulation study. Section 3 briefly reviews studies related to emerging market investments. Section 4 describes the data and applies the SPD test to the emerging and the developed market returns accumulated over different investment horizons.

Section 5 provides concluding remarks.

4.2 Spatial Dominance

In this section, we briefly present the SD theory and develop the SPD testing procedure in details.

4.2.1 Spatial Dominance Theory

International investors maximise their expected time aggregated utilities obtained from their investments in international stock markets by continuously rebalancing their portfolios over their planned investment horizon. Hence, the rebalancing of investment strategies can be seen as the solution of a dynamic utility maximisation problem. Our study focusses on comparing investor's expected time aggregated utilities in the emerging and the developed stock markets over the planned investment holding period rather than at the terminal point. Notice that the SD approach can provide a valid ranking comparison of expected time aggregated utilities only when the underlying price (or cumulative investment return) is (strictly) stationary. Considering that the cumulative returns of these investments are likely to have a time-varying distribution and/or to be nonstationary, the SD approach based on a time invariant distribution may be inappropriate and likely to result in misleading results. In this regard, it is necessary to extend the SD approach to nonstationary time series. Here we employ "the spatial analysis of time series" advanced by Park (2007) and propose the SPD approach which can provide a more reliable investment ranking strategy.

We denote the cumulative returns of the emerging and developed market investments over the planned investment holding period, respectively, by X_t and Y_t. In order to evaluate their relative performance, we compare the expected time aggregated utilities obtained from the value of investments in each market, namely,

$$\mathbb{E}\left[\int_0^T u(X_t)\,dt\right] \lesseqqgtr \mathbb{E}\left[\int_0^T u(Y_t)\,dt\right], \tag{4.1}$$

where T is a planned investment horizon and $u(\cdot)$ is an admissible utility function $u \in U$, where U is the set of all non-decreasing functions. This is a very general issue involving expected returns and market-specific and global risk factors in a dynamic stochastic optimisation problem.

The SD theory provides a general framework for analysing economic and financial behaviour under uncertainty. Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970) and Whitmore (1970) advance the foundations of SD analysis. Since the SD makes decisions in terms of utility comparisons, it is an efficient analysis of choices involving risk. The SD is an optimal selection rule when all individuals' utility functions are assumed to be of a given general class of admissible functions. Hence, the SD rules are more general than the conventional mean-variance analysis and asset pricing models, the validity of which is confirmed only if asset returns follow a normal distribution and/or the utility functions are quadratic. Dominance and SD are formally defined as follows:

Dominance Given two risky assets X and Y with the respective cumulative distribution functions, Π_X and Π_Y, X (stochastically) dominates Y if $\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ for every admissible utility function, $u \in U$, where U is the class of all non-decreasing functions which are assumed to have finite values for any finite value of x.

First Order Stochastic Dominance (FSD) X first order stochastically dominates Y if and only if either

$$\Pi_X(x) \leq \Pi_Y(x), \ \forall x \in \mathbb{R},$$

or $\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ for all u, where u is a non-decreasing utility function $(u' \geq 0)$.

Second Order Stochastic Dominance (SSD) X second order stochastically dominates Y if and only if either

$$\int_{-\infty}^{x} \Pi_X(s)ds \leq \int_{-\infty}^{x} \Pi_Y(s)ds \ \forall x \in \mathbb{R},$$

or $\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ for all u, where u is non-decreasing $(u' \geq 0)$ and strictly concave $(u'' \leq 0)$.

Third Order Stochastic Dominance (TSD) X third order stochastically dominates Y if and only if either

$$\int_{-\infty}^{x} \int_{-\infty}^{t} \Pi_X(s)dsdt \leq \int_{-\infty}^{x} \int_{-\infty}^{t} \Pi_Y(s)dsdt \ \forall x \in \mathbb{R},$$

or $\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ for all u, where u is non-decreasing $(u' \geq 0)$, strictly concave $(u'' \leq 0)$, and preferable for positive skewness $(u''' \geq 0)$.

FSD imposes a highly stringent condition, and thus its applicability is limited in most financial applications involving risky asset choices. SSD allows for investors' risk aversion and TSD for investors' positive skewness preference. These general concepts provide better utility interpretations as these classes of agents increasingly prefer either less risky and/or positively skewed returns as they are prepared to trade-off lower average returns for the chance of obtaining extremely certain and/or positive returns.

Several testing procedures for the SD have been proposed in the literature. An early SD test was advanced by McFadden (1989), which is a Kolmogorov–Smirnov (KS) type test for the FSD and the SSD for independent samples with equal numbers of observations. Barrett and Donald (2003) generalise and develop these KS tests for the SD of any pre-specified order and of independent samples with different numbers of observations. Since the asymptotic null distribution of the KS tests depends on the unknown distributions, a Monte Carlo permutation procedure or simulation method is proposed to compute the associated critical values. Linton, Massoumi, and Whang (2005) propose a subsampling scheme for evaluating the critical values of the extended KS tests for the SD of an arbitrary order in the general K-prospect case, also allowing for the observations to be dependent and for the prospects ranked to be of general dependence using the full-sample.[2]

However, the SD approach is not appropriate in the case where the underlying stochastic process does not follow a time invariant distribution, since the SD only allows us to assess the expected utilities either at a given fixed time in a completely static setting or under the assumption of strict stationarity. To overcome this important issue we employ the concept of the spatial distribution of time series developed by Park (2007). The spatial distribution is developed primarily for time series that are nonstationary, i.e., the time series are generated with time-varying distributions. Unlike the time invariant distribution that exists only under stationarity, the spatial distribution is well defined for both nonstationary and stationary time series. In fact, a spatial distribution allows us to extend various models that have been developed under the presumption of stationarity, and make them applicable to nonstationary time series as well. For nonstationary time series, the spatial distribution may be interpreted as the time aggregation of its time-varying distributions over the given time period (e.g., $\int_0^T P_t(X_t)\,dt$). Park (2007) proves that the sum of the expected utilities generated by the underlying stochastic process is determined solely by its spatial distribution. Hence, the spatial distribution can provide an efficient ranking rule which compares the time aggregated utilities generated by nonstationary time series.

We briefly review the spatial analysis of time series[3] which is built upon an empirical assessment of the expected value of the local time of the underlying stochastic process that generates the observed sample path. Local time, denoted $\ell(T,x)$, represents the frequency at which the stochastic process $X = (X_t)$ visits the spatial point x up to time T. Hence, the local time itself is a stochastic process defined in the functional space. The expectation of local time then becomes the spatial density, $\lambda(T,x)$, which represents the expected frequency at which the underlying stochastic process visits each spatial point. Similarly, the spatial distribution $\Lambda(T,x)$ is given by the expectation of the integrated local time.[4]

[2] Alternatively, Davidson and Duclos (2000) propose a testing procedure which compares the cumulative distribution functions over an arbitrary grid of points so that their test can be applied to dependent samples.

[3] See Bosq (1998); Revuz and Yor (1994); Park (2007) for further details.

[4] The spatial density and the spatial distribution are defined by $\lambda(T,x) = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\int_0^T \mathbb{P}\{|X_t - x| < \varepsilon\}\,dt$ and $\Lambda(T,x) = \int_0^T \mathbb{P}\{X_t \leq x\}\,dt$. They will provide the whole distributional information for the underlying

The expected time aggregated utilities under the stochastic dynamic optimisation framework is hard to evaluate in the time domain due to several theoretical and empirical restrictions. To evaluate this, we need a joint distribution defined in infinite dimensions and the specific form of utility function. Surprisingly, this problem can be overcome in the spatial domain by the occupation time formula[5]

$$\mathbb{E}\left[\int_0^T u(X_t)dt\right] = \int_{-\infty}^{\infty} u(x)\lambda(T,x)dx, \tag{4.2}$$

which allows the SPD to be flexibly applied to both stationary and nonstationary time series. Eq. (4.2) explicitly shows that only the spatial distribution is required to compare the expected time aggregated utilities. Formally, we define the first, the second and the third order SPD as follows.

First Order Spatial Dominance (FSPD) X first order spatially dominates Y if and only if either

$$\Lambda_X^{(0)}(T,x) \leq \Lambda_Y^{(0)}(T,x) \; \forall x \in \mathbb{R}$$

or $\mathbb{E}\left[\int_0^T u(X_t)\,dt\right] \geq \mathbb{E}\left[\int_0^T u(Y_t)\,dt\right]$ for all u, where u is a non-decreasing utility function ($u' \geq 0$) and $\Lambda_\cdot^{(0)}(T,x) = \Lambda_\cdot(T,x)$ is a spatial distribution function.

Second Order Spatial Dominance (SSDP) X second order spatially dominates Y if and only if either

$$\Lambda_X^{(1)}(T,x) \leq \Lambda_Y^{(1)}(T,x) \; \forall x \in \mathbb{R},$$

where $\Lambda_\cdot^{(1)}(T,x) = \int_{-\infty}^{x} \Lambda_\cdot(T,s)\,ds$ is an integrated spatial distribution function, or $\mathbb{E}\left[\int_0^T u(X_t)dt\right] \geq \mathbb{E}\left[\int_0^T u(Y_t)dt\right]$ for all u, where u is non-decreasing ($u' \geq 0$) and strictly concave ($u'' \leq 0$).

Third Order Spatial Dominance (TSPD) X third order spatially dominates Y if and only if either

$$\Lambda_X^{(2)}(T,x) \leq \Lambda_Y^{(2)}(T,x) \; \forall x \in \mathbb{R},$$

where $\Lambda_\cdot^{(2)}(T,x) = \int_{-\infty}^{x}\int_{-\infty}^{t} \Lambda_\cdot(T,s)dsdt$ is a double-integrated spatial distribution function, or $\mathbb{E}\left[\int_0^T u(X_t)dt\right] \geq \mathbb{E}\left[\int_0^T u(Y_t)dt\right]$ for all u, where u is non-decreasing ($u' \geq 0$), strictly concave ($u'' \leq 0$), and preferable for positive skewness ($u''' \geq 0$).

stochastic process, which is either stationary or nonstationary. The spatial density and distribution will trivially become the time invariant density and distribution function under stationarity.

[5] See Lemma 2.1 in Park (2007).

4.2.2 Testing for Spatial Dominance

To test the validity of FSPD, SSPD or TSPD, we set the null hypothesis, X s-order spatially dominates Y, against the alternative hypothesis, X does not s-order spatially dominate Y:

$$H_0^{(s)} : \Lambda_X^{(s-1)}(T,x) \leq \Lambda_Y^{(s-1)}(T,x) \text{ for all } x \in \mathbb{R},\ s = 1,2,3, \tag{4.3}$$

$$H_1^{(s)} : \Lambda_X^{(s-1)}(T,x) > \Lambda_Y^{(s-1)}(T,x) \text{ for some } x \in \mathbb{R},\ s = 1,2,3, \tag{4.4}$$

where $\Lambda_{\cdot}^{(s-1)}(T,\cdot)$ is the $(s-1)$-order integrated spatial distribution function defined in the previous subsection. Then, the test statistics are derived by comparing the Kolmogorov–Smirnov uniform distance between the respective estimated $(s-1)$ integrated spatial distribution functions of X and Y as follows:

$$D_N^{(s)}(T) = \sqrt{N} \sup_{x\in\mathbb{R}} \left[\hat{\Lambda}_{N,X}^{(s-1)}(T,x) - \hat{\Lambda}_{N,Y}^{(s-1)}(T,x)\right],\ s = 1,2,3, \tag{4.5}$$

where $\hat{\Lambda}_{\cdot}^{(s)}(T,x)$ is the consistent estimator of $\Lambda_{\cdot}^{(s)}(T,x)$.[6]

However, the limiting distributions of the $D_N^{(s)}(T)$ statistics are dependent upon the (generally unknown) probability law of the underlying stochastic processes, X and Y, in a quite complicated manner. Hence, it is hard to apply the conventional bootstrap approach to obtain valid critical values for the SPD test. In general, a subsampling scheme, which we adopt in this study, appears to be most readily available to obtain the limiting distributions of $D_N^{(s)}(T)$, see Linton, Massoumi, and Whang (2005) for SD and Park (2007) for SPD.

We now describe in detail the SPD testing procedure to assess the performance of the emerging market investments relative to the developed counterparts. We assume that (X_t) and (Y_t) are stochastic processes with stationary independent increments, assumed to be strictly stationary and α-mixing, where (X_t) and (Y_t) are obtained from the emerging and the developed market investments. Under this assumption, we set the following sequence of the null hypotheses:

$$H_{0,XY}^{(s)} : X \nsim Y,\ H_{0,X}^{(s)} : X \succeq_{(s)} Y,\ H_{0,Y}^{(s)} : Y \succeq_{(s)} X,\ s = 1,2,3,$$

where $\nsim$ stands for the non-maximality (McFadden, 1989) and $\succeq_{(s)}$ for an s-th order spatial dominance in favour of the former stochastic process under the null.[7] The validity of $H_{0,X}^{(s)}$

[6] As $N \to \infty$, Park (2007) shows that $D_N^{(s)}(T) \xrightarrow{d} \sup_{x\in\mathbb{R}}\left[U_X^{(s-1)}(T,x) - U_Y^{(s-1)}(T,x)\right]$ for $s = 1,2,3$, where $U_{\cdot}^{(0)}(T,\cdot)$ is a zero-mean Gaussian process with finite covariance kernel, $U_{\cdot}^{(1)}(T,x) = \int_{-\infty}^{x} U_{\cdot}^{(0)}(T,s)ds$, and $U_{\cdot}^{(2)}(T,x) = \int_{-\infty}^{x}\int_{-\infty}^{t} U_{\cdot}^{(0)}(T,s)dsdt$.

[7] McFadden (1989) proposes to consider $H_{0,XY}^{(s)}$ prior to testing $H_{0,X}^{(s)}$ and $H_{0,Y}^{(s)}$ to avoid any ambiguity.

and $H_{0,Y}^{(s)}$ can be tested by comparing the Kolmogorov–Smirnov uniform distance between the estimated $(s-1)$-order integrated spatial distribution functions of X and Y respectively for $s = 1,2,3$:

$$D_{N,X}^{(s)} = \sqrt{N} \sup_{x \in \mathbb{R}} \left[\hat{\Lambda}_{N,X}^{(s-1)}(T,x) - \hat{\Lambda}_{N,Y}^{(s-1)}(T,x) \right], \tag{4.6}$$

$$D_{N,Y}^{(s)} = \sqrt{N} \sup_{x \in \mathbb{R}} \left[\hat{\Lambda}_{N,Y}^{(s-1)}(T,x) - \hat{\Lambda}_{N,X}^{(s-1)}(T,x) \right], \tag{4.7}$$

where $\hat{\Lambda}_{N,X}^{(s-1)}(T,x)$ and $\hat{\Lambda}_{N,Y}^{(s-1)}(T,x)$ are the estimates of the $(s-1)$-order integrated spatial distribution functions of X and Y, obtained by

$$\begin{aligned} \hat{\Lambda}_{N,X}^{(s-1)}(T,x) &= \frac{\delta}{N(s-1)!} \sum_{i=1}^{N} \sum_{j=1}^{p} \left(x - X_{i,j\delta} \right)^{s-1} 1\left\{ X_{i,j\delta} \leq x \right\}, \\ \hat{\Lambda}_{N,Y}^{(s-1)}(T,x) &= \frac{\delta}{N(s-1)!} \sum_{i=1}^{N} \sum_{j=1}^{p} \left(x - Y_{i,j\delta} \right)^{s-1} 1\left\{ Y_{i,j\delta} \leq x \right\}, \end{aligned} \tag{4.8}$$

where $p = T/\delta$ is the total number of observations over the planned holding period, the sample observation inside the ith holding period is denoted by $X_{i,\cdot} = \left\{ X_{i,1\delta}, \ldots, X_{i,p\delta} \right\}$, and $1\{\cdot\}$ is an indicator function.[8] Finally, the validity of $H_{0,XY}^{(s)}$ can be tested by the MacFadden test statistic:

$$M_N^{(s)} = \min\left(D_{N,X}^{(s)}, D_{N,Y}^{(s)} \right),\ s = 1,2,3. \tag{4.9}$$

In principle, the SPD test should be carried out sequentially as follows: First, we begin with testing the null hypotheses, $H_{0,XY}^{(1)}$, $H_{0,X}^{(1)}$ and $H_{0,Y}^{(1)}$. If $H_{0,XY}^{(1)}$ and $H_{0,X}^{(1)}$ are not rejected but $H_{0,Y}^{(1)}$ is rejected, we conclude that the emerging market investment first-order spatially dominates the developed one. On the other hand, if $H_{0,XY}^{(1)}$ and $H_{0,Y}^{(1)}$ are not rejected but $H_{0,X}^{(1)}$ is rejected, we conclude that the emerging market investment is first-order spatially dominated by the developed one. Otherwise, there is no FSPD and we continue to test the null hypotheses for SSPD: $H_{0,XY}^{(2)}$, $H_{0,X}^{(2)}$ and $H_{0,Y}^{(2)}$. If $H_{0,XY}^{(2)}$ and $H_{0,X}^{(2)}$ are not rejected but $H_{0,Y}^{(2)}$ is rejected, we conclude that the emerging market investment second-order spatially dominates the developed one. On the other hand, if $H_{0,XY}^{(2)}$ and $H_{0,Y}^{(2)}$ are not rejected but $H_{0,X}^{(2)}$ is rejected, we conclude that the emerging market investment is second-order spatially dominated by the developed one. Otherwise, there is no SSPD, and we continue to test the null hypotheses for TSPD, $H_{0,XY}^{(3)}$, $H_{0,X}^{(3)}$ and $H_{0,Y}^{(3)}$, and repeat a similar inference as above.

[8] Notice in empirical applications that we normalise $T = 1$, and set p to the number of days in the holding period, e.g., $p = 126$ for a 6-month holding period.

As discussed earlier, the asymptotic null distributions of the $D_{N,\cdot}^{(s)}(T)$ and $M_N^{(s)}$ test statistics depend on the unknown population distributions. We therefore follow Linton, Massoumi, and Whang (2005) and Park (2007) and employ a subsampling scheme for evaluating the critical values of the spatial dominance tests derived above.[9] We describe a practical subsample scheme in the Appendix.

4.2.3 Monte Carlo Demonstrations

In this subsection, we consider the case in which there is the SPD but not the SD between two investment prospects. In such a situation, we will demonstrate that the SPD test is sufficiently powerful to rank two alternative investment prospects by comparing the time aggregated utilities over the planned investment horizon. By contrast we will show that the SD test suffers from non-negligible size distortion, suggesting a non-negligible probability of making an incorrect decision about a (dynamic) investment choice.

We consider two investment projects over the investment horizon $[0,T]$, denoted X_t and Y_t, and assume that the increment of X_t (dX_t) is more volatile than that of Y_t (dY_t) before the (known) break point (τ),[10] and less volatile after the break. We also assume that the variances of $X_T - X_0 = \int_0^T dX_t$ and $Y_T - Y_0 = \int_0^T dY_t$ are the same at the terminal time, T.[11]

We now describe the two investment opportunities by the diffusion processes incorporating the structure break as follows:

$$dX_t = \begin{cases} \sigma_{X_a} dV_t, & t \in (0,\tau] \\ \sigma_{X_b} dV_t & t \in (\tau, T] \end{cases} \tag{4.10}$$

$$dY_t = \sigma_Y dW_t,\ t \in (0,T] \tag{4.11}$$

where V_t and W_t are independent standard Brownian motions with $\sigma_{X_b} < \sigma_Y < \sigma_{X_a}$. Then,

[9] Linton, Massoumi, and Whang (2005) propose non-parametric tests of stochastic dominance by extending the Kolmogorov–Smirnov type statistics developed by McFadden (1989) and suggest a subsampling method to evaluate the critical values and the associated p-values.

[10] In principle, we may consider a more general case with unknown break point, though the results will be similar to those obtained in this subsection.

[11] For example, suppose the increment of investment value in the emerging market is more volatile than that in the developed market for the first few years, but the volatility of the emerging market becomes smaller after capital liberalisation or successful economic reform. The reverse case can be considered with economic failure or currency/financial crisis.

the transition probability distributions of X_t and Y_t given $X_0 = 0$ and $Y_0 = 0$ are

$$X_t|X_0 \sim \begin{cases} N\left(0, \sigma_{X_a}^2 t\right), & t \in (0, \tau] \\ N\left(0, \sigma_{X_a}^2 \tau + \sigma_{X_b}^2 (t-\tau)\right), & t \in (\tau, T] \end{cases} \tag{4.12}$$

$$Y_t|Y_0 \sim N\left(0, \sigma_Y^2 t\right), \ t \in (0, T] \tag{4.13}$$

By construction, the variances of X and Y are the same at the terminal point of the investment horizon, i.e.,

$$\sigma_{X_a}^2 \tau + \sigma_{X_b}^2 (T-\tau) = \sigma_Y^2 T, \tag{4.14}$$

and thus there is no SD between X and Y. This implies that risk-averse investors expect the same level of utility from the increments of X and Y at the terminal point T.

We now examine whether there is an SPD between X and Y over the investment period $[0,T]$. The analytic form of the spatial density and distribution are hard to derive in this case, but it is clear that $Var(X_t|X_0 = 0) > Var(Y_t|Y_0 = 0)$ for any $t \in (0,T)$. Hence, risk-averse investors will prefer Y over X over the full investment horizon. This intuition can be easily demonstrated by the spatial distribution function simulated using sufficiently many sample paths of X and Y. We present the spatial distribution functions and the integrated spatial distribution functions of X and Y obtained using 1,000 sample paths in Figure 4.1. We generate X and Y such that (4.14) holds. Given $T = 5$, $\tau = 2.5$, $dt \approx 1/250$, $\sigma_Y = 1$ and $\sigma_{X_a} = 1.4$, we obtain $\sigma_{X_b}^2 = 0.2$. In this case, we find $\sup_{x \in R}\left[\Lambda_X^{(1)}(T,x) - \Lambda_Y^{(1)}(T,x)\right] \approx 1.993$ and $\sup_{x \in R}\left[\Lambda_Y^{(1)}(T,x) - \Lambda_X^{(1)}(T,y)\right] \approx 0$. Hence, it is clear that X is second-order spatially dominated by Y, supporting our intuition that the SPD provides a ranking of alternative dynamic investment projects over the investment period.

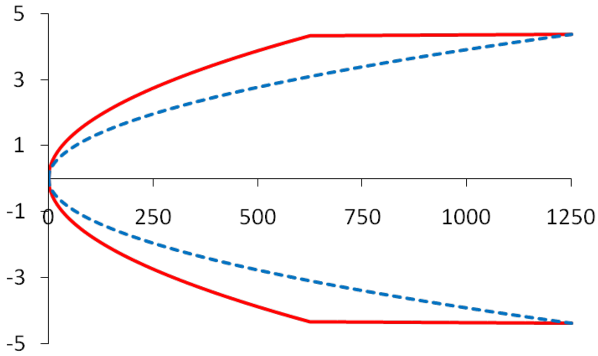
In sum, the SD and the SPD approaches provide the same ranking between two investment projects only when the distribution of increments on investment value is time-invariant.[12] On the other hand, when the distribution is time-varying, the SD test is restricted to evaluating the utility only at the terminal point while the SPD approach can be applied to evaluate the time aggregated utilities over the whole investment period.

To construct the (discrete) data generating process, we approximate the continuous time diffusion processes of X and Y in (4.10) and (4.11) with analytic transition probability distributions, (4.12) and (4.13) as follows: In a discrete time, we generate $N = 1,250$ daily observations by setting the interval $\Delta = 1/250$ for investment horizon T (e.g., $T = 5$ denotes

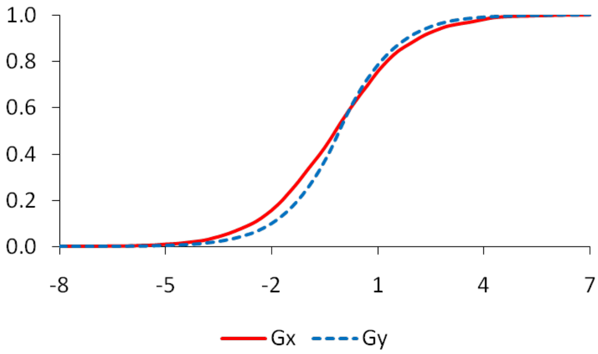
[12]Of course, their interpretations are different.

Figure 4.1: The Spatial Distribution Functions and Integrated Spatial Distribution Functions of Simulated Sample Paths

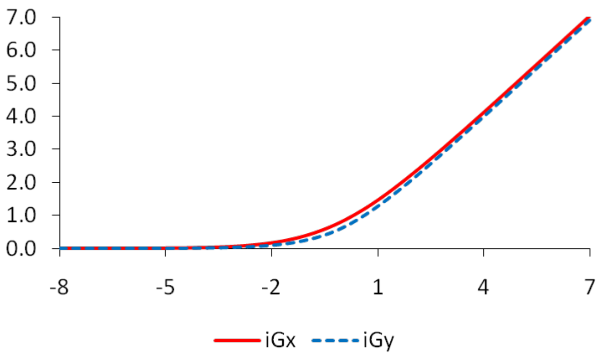
In Figure (a), the solid line denotes the 95% lower and upper bounds of sample path X and the dashed line denotes those of the sample path Y for the DGP specification of $\sigma_{X_a} = 1.4$, $\sigma_{X_b} = 0.2$, $\sigma_Y = 1$, $\tau = 2.5$, $T = 5$ and $\Delta = 1/250$ in (4.15) and (4.16). Gx and Gy denote the spatial distribution functions of X and Y in Figure (b) and iGx and iGy denote the integrated spatial distribution functions of X and Y in Figure (c). To obtain the spatial distributions and the integrated spatial distributions, we generate $N = 1,000$ sample paths of X and Y.



(a) 95% lower/upper bounds of sample paths



(b) Spatial distribution functions



(c) Integrated spatial distribution functions the 5-year investment horizon)[13] such that

$$X_{i\Delta} = X_{(i-1)\Delta} + \begin{cases} \sigma_{X_a}\sqrt{\Delta}\Phi, & i = 1,\ldots,N_\tau \\ \sigma_{X_b}\sqrt{\Delta}\Phi, & i = N_\tau + 1,\ldots,N \end{cases} \tag{4.15}$$

$$Y_{i\Delta} = Y_{(i-1)\Delta} + \sigma_Y\sqrt{\Delta}\Phi,\ i = 1,...,N \tag{4.16}$$

where $X_0 = 0$, $Y_0 = 0$, $N_\tau\Delta = \tau$, $N\Delta = T$ and Φ is a standard normal distribution function. As before, we set $\sigma_Y = 1$, $\sigma_{X_a} = 1.4$ and $\sigma_{X_b} = 0.2$.[14] Notice that there is no stochastic dominance between X and Y at the terminal point T, but X is second-order spatially dominated by Y over the period $[0,T]$.

We compare the finite sample performance of the spatial-dominance and the stochastic dominance tests at 5% significance level. We consider two sample observations of 250 and 500, and 1,000 iterations. Given the maximality at given significance level (McFadden, 1989), we test the joint null hypothesis,

$$H_{0,XY} : X \succeq_2 Y \ \&\ H_{0,XY} : Y \succeq_2 X$$

where $H_{0,XY}$ denotes "X is not second-order spatially (stochastically) dominated by Y" and $H_{0,YX}$ denotes "Y is not second-order spatially (stochastically) dominated by X". The DGP given by (4.15) and (4.16) is false under the joint null with respect to the SPD test,[15] so the rejection frequency of the test is its power. Specifically, we are interested in the frequency that the test rejects $H_{0,XY}$ but does not reject $H_{0,YX}$. On the other hand, the DGP is true under the joint null with respect to the SD test,[16] so the rejection frequency of the test is the size of test.

Panel A in Table 4.1 presents the rejection frequency of the SPD test for four cases; (1) non-rejection of both $H_{0,XY}$ and $H_{0,YX}$, (2) non-rejection of $H_{0,XY}$ and rejection of $H_{0,YX}$, (3) rejection of $H_{0,XY}$ and non-rejection of $H_{0,YX}$ and (4) rejection of both $H_{0,XY}$ and $H_{0,YX}$. The SPD test rejects the (false) joint null hypothesis by 100% (=(2)+(3)+(4)) at 5% significance level for both 250 and 500 observations. In particular, the rejection frequency for Case (3) is 99%, correctly indicating that X is second-order spatially dominated by Y. These results are almost the same when critical values are evaluated by mean and median. On the other hand, surprisingly, the SD test rejects the (true) joint null hypothesis too often as reported

[13] We follow a standard finance calender, so one year consists of 250 business days.

[14] Selecting the break point τ as the middle of the investment horizon, $N_\tau = 625$ for the 5-year investment horizon. Given $\sigma^2_{X_a}$ and σ^2_Y, $\sigma^2_{X_b}$ is obtained by (4.14).

[15] $H_{0,XY}$ is false but $H_{0,YX}$ is true under the DGP. Hence, the joint null is false.

[16] $H_{0,XY}$ and $H_{0,YX}$ are true under the DGP since neither X nor Y are stochastically dominated by the other.

Table 4.1: Monte Carlo Comparison of Spatial Dominance and Stochastic Dominance

The data generating process specifies $\sigma_{X_b} = 1.4$ and $\sigma_Y = 1$ in DGP (4.15) and (4.16). It generates two processes X and Y in daily frequency base. The two stochastic processes are not stochastically dominated each other at the terminal point T but X is second-order spatially dominated by Y over the period $[0,T]$. The terminal point T is set to 5 years. One year consists of 250 business days following the standard finance calender, so the number of observation is 1,250. Then σ_{X_a} is obtained by (4.14) and its value is 0.2. We compare the finite sample performance of the spatial-dominance and the stochastic-dominance test at 5% significance level with respect to $N = 250$ and $N = 500$ sample observations. The critical values at 5% significance level are provided by the median and mean value from the subsampling scheme. The figures in the table present the frequency for each event out of 1,000 experiments.

Number of observations	250		500	
Critical value by	Mean	Median	Mean	Median
Panel A: Spatial Dominance				
(1) Non-rejection of both $H_{0,XY}$ and $H_{0,YX}$	0.000	0.000	0.000	0.000
(2) Non-rejection of $H_{0,XY}$ and rejection of $H_{0,YX}$	0.000	0.000	0.000	0.000
(3) Rejection of $H_{0,XY}$ and non-rejection of $H_{0,YX}$	0.989	0.984	0.985	0.987
(4) Rejection of both $H_{0,XY}$ and $H_{0,YX}$	0.011	0.016	0.015	0.013
Panel B: Stochastic Dominance				
(1) Non-rejection of both $H_{0,XY}$ and $H_{0,YX}$	0.658	0.659	0.706	0.713
(2) Non-rejection of $H_{0,XY}$ and rejection of $H_{0,YX}$	0.178	0.177	0.147	0.143
(3) Rejection of $H_{0,XY}$ and non-rejection of $H_{0,YX}$	0.163	0.163	0.146	0.143
(4) Rejection of both $H_{0,XY}$ and $H_{0,YX}$	0.001	0.001	0.001	0.001

in Panel B. The rejection frequency is 34% for 250 observations, and it slightly decreases to 29% even for 500 observations, suggesting that the SD test may encounter a non-negligible size distortion in the presence of the time-varying nature of the underlying distribution even though the mean and variance are the same at the terminal period.

In sum, the simulation results clearly demonstrate the relative advantage of using SPD tests over SD tests in evaluating investment performance when the increment of investment value follows a time-varying distribution over the holding period. This may suggest that the SPD approach is more appropriate for evaluating the relative performance of the emerging market investment.[17]

In next section we will apply the SPD test to the emerging and developed stock market returns accumulated over different investment horizons and assess their relative performance. To explicitly deal with the general nature of the current dynamic decision making problem, we should employ the discounted spatial analysis of time series.[18]

[17] Although the market integration hypothesis in the long-run suggests that there is no SD between emerging and developed market at the end of the terminal point, we still observe that the risk factors are more or less prominent in the emerging market during the early stage.

[18] d-local time is defined as (see Park (2007) for more details) $\ell^r(T,x) = \int_0^T e^{-rt}\ell(dt,x)$ for some discount rate $r > 0$. Obviously, the d-spatial distribution function of X is defined accordingly by $\Lambda^r(T,x) = \int_0^T e^{-rt}\mathbb{P}\{X_t \leq x\}dt$. Hence, in what follows, we provide all the spatial dominance test results based on the d-spatial distribution function, $\Lambda^r(T,x)$.

4.3 Emerging Market Investment

4.3.1 Characteristics of Emerging Market Investments

A large and growing body of literature on emerging market investments has documented at least four distinguishing features of emerging market returns from the developed ones: (i) higher expected returns, (ii) low correlations with developed market returns, (iii) more predictable returns, and (iv) higher volatility. The higher expected returns and the higher volatility in the emerging markets are well supported by both theory and empirical findings. For example, Kohers, Kohers, and Kohers (2006) find that the risk, measured by the standard deviation of returns and the overall risk score taking account of country's economic, fiscal, legal, and governmental environment, are higher in emerging markets for most periods over 1998–2003, whilst their returns are also higher. The higher volatility in the emerging markets is supported by a number of other studies, e.g., Bekaert and Harvey (1997), Bekaert, Erb, Harvey, and Viskanta (1997), DeSantis and Imrohoroglu (1997), Kawakatsu and Morey (1999). This evidence may indicate that risk-averse investors have been compensated for the higher risks associated with the emerging markets investments.

Common risk factors, affecting the higher expected return and the higher volatility in the emerging markets, include currency risk, political risk, economic risk, financial risk and so on. Aggarwal, Inclan, and Leal (1999) argue that most events causing a large shift in the volatility tend to be local. Onder and Simga-Mugan (2006) report that the political and the economic news influence both the volatility of returns and the trading volume in the emerging stock markets. There are also a number of studies finding that the average return will decrease after financial liberalisations (Henry, 2000). Bekaert and Harvey (2000) argue that the cost of capital always decreases after a capital market liberalisation, which causes equity capital to flow to the emerging markets. We also expect that the gradual development and diversification of the markets should lead to lower volatility of returns in the long-run.

There have been two main streams of the literature in studying the emerging market investments. One stream focusses on examining the diversification benefits from exposure to the emerging markets. The most often used is a mean-variance test where the set of asset returns is said to provide the diversification benefits relative to the set of benchmark returns if adding these returns to the benchmark leads to a significant leftward shift in the mean-variance (standard deviation) frontier (Bekaert and Urias, 1996). Early studies by Harvey (1994a,b) found that adding emerging market investments to a well-diversified portfolio reduces the overall volatility even though the emerging-market equities are much more volatile than the developed-market equities, but critically argue that investors require information in addition to the mean and variance of the portfolio before making their decisions.

Another stream aims to examine the risk–return relationship in the context of the asset pricing model (Dumas and Solnik, 1995; Harvey, 1995a), though the success of such models in uncovering the risk–return relationship is contingent on the model specification. Levhari and Levy (1977) demonstrate that if the assumption of the holding period is different from the true horizon, then there will be a systematic bias of the performance measure as well as the beta estimate. Harvey (1995b) finds that there is little sensitivity between the emerging market returns and the traditional measures of risk such as the world stock return, exchange rate investment index, oil price, growth in OECD industrial production, and inflation. Furthermore, as pointed out by Roll (1997), testing the CAPM is equivalent to testing the mean-variance efficiency of the portfolio only. Therefore, any asset pricing model (that assumes complete integration of capital markets) will be unlikely to be able to fully account for the behaviour of the emerging markets. To this end an extended asset pricing model has also been proposed. Focussing on the time-varying nature of the returns in emerging markets, Harvey (1995a) proposes a conditional CAMP test, where conditioning is set by both common global variables and country-specific variables, and find that emerging market returns are strongly influenced by local information. Furthermore, when all the moments are allowed to change through time, there is some evidence of time-varying risk exposures. However, the conditional asset pricing model still fails to price the emerging market assets correctly and to account for the time variation in the expected return (Bekaert, Harvey, Lundblad, and Siegel, 2007).

Bekaert, Erb, Harvey, and Viskanta (1997) point out that most empirical findings on the relative performance based on the (conventional) mean-variance and the asset pricing model should be interpreted with extreme care, since the associated performance evaluation is inappropriate in the case where the return distribution is non-normal and time-varying and/or where investors' utility functions are not quadratic. Importantly, there is compelling evidence that almost all emerging market returns depart from normality (see Claessens, Dasgupta, and Glen, 1995; Harvey, 1995a). Salomonsa and Grootveldb (2003) find that an equity risk premium in the emerging markets is significantly higher than in the developed markets, suggesting that the investors focus more on the downside risk instead of the standard deviation, since the distribution of equity risk premium in the emerging markets is neither normal nor symmetric.

Given the non-normal and time-varying return distributions in the emerging markets, most prior studies employing either the mean-variance or the asset pricing models to assess the overall performance of the emerging market investments (Disyatat and Gelos, 2001; Bekaert and Harvey, 2003) are limited in terms of efficiency. Hence, we propose to follow an alternative approach based on the dominance test that offers superior criteria for comparing the relative performance of investments.

4.3.2 Market Integration and Effect of Currency Risk

Although several anomalous findings have been routinely reported in both developed and emerging market studies, most come to the conclusion that markets will be efficient once transaction costs are considered appropriately. In general, developed markets are found to be more efficient than emerging markets. On a global scale, international market efficiency is often addressed in terms of international market integration or segmentation, e.g., Solnik and McLeavey (2003). There are two different views. First, a research group with an ex ante operational view will highlight the importance of analysing the role of market openness and the impediments to capital mobility such as legal restrictions, transaction costs, political risk, and currency risk. Second, international asset pricing models investigate whether securities with similar levels of risk are priced in the same manner in different markets, e.g., Bekaert and Harvey (2003).

An efficient international equity market requires that the real prices of (consumption) goods should be equalised, namely, that purchasing power parity holds. If so, it does not matter whether one invests in the domestic or the foreign currency since real foreign exchange risk would not exist. When the market is less than perfectly integrated, however, the currency risk premium matters. Dumas and Solnik (1995) find evidence that a significant currency risk premium exists and reject a model that excludes currency risk factors. DeSantis and Imrohoroglu (1997) also find strong support for the ICAPM that includes both market and currency risks. Hence, we will examine the effect of currency risk in emerging market investments by comparing the time aggregated utilities obtained from investments denominated in the local currency and in the US dollar over the planned investment horizon. Effectively, this is tantamount to studying whether the additional risk and return accrued from the currency risk will increase or decease the utility gains of the investor investing in the emerging markets.

4.4 Empirical Evaluation of Time Aggregate Utility

In order to compare the expected time aggregated utilities obtained from investments in the emerging and the developed markets, we consider investments over 3 months, 6 months, 1 year and 5 years holding periods. The cumulative investment returns for emerging and developed markets, denoted by X_i and Y_i, are constructed by

$$X_i = \sum_{j=1}^{p} r^E_{j+(i-1)w},\ Y_i = \sum_{j=1}^{p} r^D_{j+(i-1)w},\ i = 1,2,...,N,$$

where $r^E_\cdot$ and $r^D_\cdot$ are the daily returns obtained by differencing the log prices of the emerging and the developed markets, p is the number of daily observations over the planned investment horizon, $N = \left[\frac{T^*-p}{w}\right] + 1$ is the total number of observations of the cumulative returns over the planned investment horizon, T^* is the total number of full sample observations, w is the size of the moving step allowing for overlapping, and $[\cdot]$ denotes the largest integer part of the argument. Thus we have $p = 63$, 126, 252, and 1260 for 3-month, 6-month, 1-year and 5-year investment horizons, respectively.

To examine the effects of currency risk on investment performance, we consider the investment returns denominated in the local currency and in the US dollar. Furthermore, we examine the effect of the 1997 Asian Currency Crisis on investment performance by splitting the sample into two sub-periods (1988–1998 and 1999–2007).

4.4.1 Data

This research uses the emerging and the developed market indices reported by Morgan Stanley Capital International (MSCI). They are based on individual MSCI country indices that are aggregated into the composite indices. As of June 2006, the MSCI Emerging Markets Index consists of the following 25 emerging market country indices: Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, Israel, Jordan, Korea, Malaysia, Mexico, Morocco, Pakistan, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand, and Turkey; and the MSCI Developed Markets Index consists of the following 23 developed country indices: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, The Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, and the United States. These MSCI indices are regarded as performance measurement benchmarks of global stock markets by global portfolio managers. Both indices are collected from 4/Jan/1988 to 7/May/2007 (5,046 observations). To further investigate the currency risk effect, we use indices denominated in the local currency and in the US dollar.

Tables 4.2, 4.3, and 4.4 present the descriptive statistics (the number of observations, mean, standard deviation, skewness, excess kurtosis, and the Jarque–Bera test statistic for normality) of daily returns and cumulative returns over 3-month (63 days), 6-month (126 days), 1-year (252 days) and 5-year (1260 days) holding periods denominated in the local currency and the US dollar over the whole sample period (1988–2007) and the sub-periods (1988–1996 and 1999–2007). Notice that the cumulative returns are constructed allowing an overlapping with a 5-day moving-step.[19]

We first consider the full sample period (see Table 4.2). Consistent with the previous literature, there are several stylised facts about the emerging market investments. Comparing returns denominated in the local currency, we find that the emerging markets have higher returns (about 4–5 times) and higher volatility (about 2–3 times) than the developed markets have. Similar patterns are found in comparing returns denominated in the US dollar but in much smaller magnitudes.

Comparing the same market returns denominated in the local currency and in the US dollar, the emerging market returns denominated in the local currency are higher than those denominated in the US dollar whilst the opposite is true for the developed market returns. This suggests that the emerging market currency has been depreciated against the US dollar whilst the developed market currency has been appreciated against the US dollar over the full sample period. Interestingly, we find that the volatility of the emerging market returns decrease rather than increase when taking the currency risk into consideration. Overall, the descriptive statistics suggest that there are higher average returns and risks for the emerging market investments.

Examining the higher order moments of the returns, a clear pattern emerges. Although normality of the return distribution is rejected for almost all the investment horizons, the extent of the violation decreases as the investment horizon increases. The kurtosis and the Jarque–Bera test statistic decrease as the investment horizon increases from 1 day to 5 year, which is consistent with a finding by Aggarwal and Aggarwal (1993).

Figure 4.2 provides the time series plots of cumulative returns over 3-month, 6-month, 1-year and 5-year holding periods. A careful investigation of the figures clearly suggests the different market movements for the different investment horizons. The discrepancy is clearer in the cumulative returns of the emerging markets. First, both the short- and the long-term emerging market returns are significantly higher than the developed counterparts but continue to decrease during the first sub-period whilst they tend to increase during the second sub-period but the average returns are smaller than the first sub-period ones. Second, there is a sizeable (positive) gap between the emerging market returns denominated in the local currency and in the US dollar in the first sub-period, whilst the gap gets narrower and then its sign reverses sometimes in the second sub-period. The emerging markets have

[19]We have also attempted to construct both cumulative returns using an overlapping with a 21-day moving-step. The results are qualitatively similar, and will be available upon request.

Table 4.2: Descriptive Statistics over 1988-2007

Table presents the number of observations (OBS), mean (MEAN), standard deviation (STDEV), skewness (SKEW), excess kurtosis (KURT) and the Jarque–Bera test statistic for normality (JB) for the emerging and the developed market daily stock returns, and cumulative returns over 3-month (63 days), 6-month (126 days), 1-year (252 days) and 5-year (1260 days) holding periods over the whole sample period (1988–2007). The standardised skewness and kurtosis, computed respectively by $\frac{\sum_{i=1}^{N}(r_i-\bar{r})^3/(N-1)}{\sigma^2}$ and $\frac{\sum_{i=1}^{N}(r_i-\bar{r})^4/(N-1)}{\sigma^4}$, respectively. JB stands for the Jarque–Berra statistic for the null hypothesis of normality of the respective holding period returns. The returns are denominated in both local currency and US dollar over the whole sample period. The daily returns are computed as the log difference of the respective stock price indexes, and the cumulative returns are constructed allowing an overlapping with a 5-day ahead moving window. D^L and $D^\$$ stand for the investment returns denominated in the local currency and the US dollar in the developed markets. Analogously E^L and $E^\$$ stand for the investment returns denominated in the local currency and the US dollar in the emerging markets. *, ** and *** indicate a rejection of the null at 10%, 5%, and 1%, respectively.

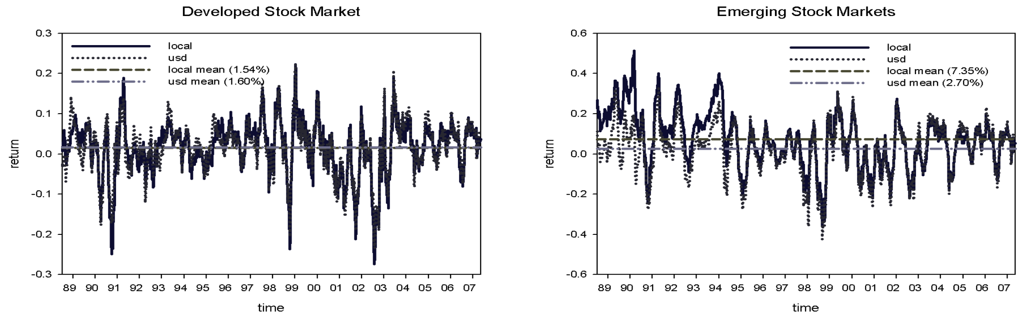
Market/Currency	OBS	MEAN	STDEV	SKEW	KURT	JB
A. Holding period: 1 day						
D^L	5046	0.03%	0.76%	-0.16	3.98	3354 ***
$D^\$$	5046	0.03%	0.78%	-0.14	3.38	2413***
E^L	5046	0.12%	0.92%	-0.53	4.19	3924***
$E^\$$	5046	0.05%	1.00%	-0.65	4.46	4538***
B. Holding period: 3 months						
D^L	997	1.54%	6.67%	-0.89	1.93	288***
$D^\$$	997	1.60%	6.47%	-0.53	0.95	84***
E^L	997	7.35%	13.94%	-0.01	0.02	0
$E^\$$	997	2.70%	12.55%	-0.49	0.03	40***
C. Holding period: 6 months						
D^L	985	3.03%	9.67%	-0.77	0.66	116***
$D^\$$	985	3.16%	8.95%	-0.65	0.46	78***
E^L	985	14.43%	22.81%	0.25	-0.18	12***
$E^\$$	985	5.14%	17.73%	-0.54	0.09	48***
D. Holding period: 1 year						
D^L	959	5.81%	14.64%	-0.93	0.13	138***
$D^\$$	959	6.05%	13.65%	-0.93	0.35	145***
E^L	959	28.32%	37.58%	0.58	-0.20	56***
$E^\$$	959	10.01%	25.04%	-0.52	-0.07	44***
E. Holding period: 5 years						
D^L	758	25.16%	33.99%	0.04	-1.02	33***
$D^\$$	758	26.42%	30.40%	-0.21	-1.03	39***
E^L	758	112.02%	117.35%	0.83	-0.71	103***
$E^\$$	758	31.05%	55.26%	-0.07	-1.31	55***

indeed undergone many political, economic and financial events locally and globally such as capital market liberalisations, currency and financial crises. In this regard we may consider that the emerging markets have been more or less segmented and unstable in the first sub-period and become more or less integrated into the global markets in the second sub-period. Interestingly, we also observe that most local currencies in the emerging markets have been depreciated or devalued against the US dollar during the first sub-period, whereas the currency movements have been reversed during the second sub-period.

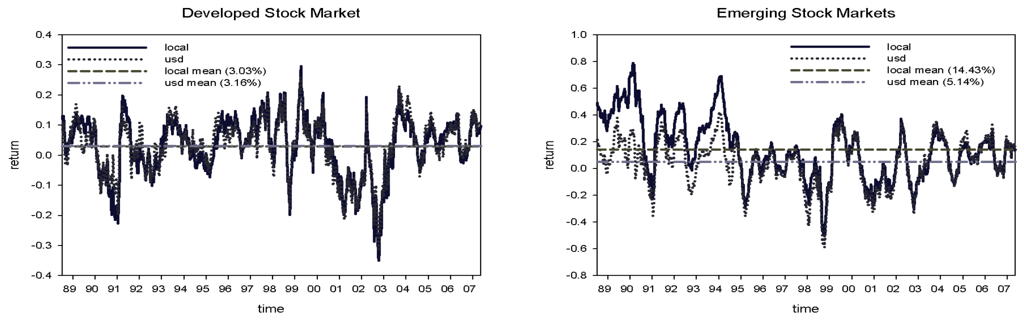
Next, we summarise the descriptive statistics for two sub-periods excluding the 1997 Asian Currency Crisis period during 1997–98 in Tables 4.3 and 4.4. Looking at the daily returns denominated in the local currency over the first sub-period, the emerging market returns are much higher on average, slightly more volatile, negatively skewed and fatter tailed as compared to their developed counterparts. As in the full sample period, there is a trade-off between higher average return (but the difference is far more significant) and more downside risk (negative skewness). On the other hand, the returns denominated in the local currency over the second sub-period show a similar patten as above but at a smaller scale. The return distributions over the sub-periods are somewhat similar to those observed over the full sample period, though the effect of currency fluctuation on emerging market returns seems to be much moderate in the post-crisis period.

Overall, the emerging market returns are higher on average and more volatile than the development counterparts both in the local currency and in the US dollar. The local currency of the developed markets is a bit appreciated against the US dollar. On the other hand, the local currency of the emerging markets is significantly depreciated against the US dollar so that the growth premium of the emerging market may be deteriorated by the currency devaluation. This finding is more clearly observed in the first sub-period, which is consistent with the nature of the early emerging market. After the Asian Currency Crisis, such early characteristics of the emerging market are more or less moderated. The average return and volatility are significantly reduced and also the currency movement is more integrated with the US dollar. Hence, the emerging market looks like being in the process of integration into the global market. Currency risk and a non-negligible likelihood of extreme events such as the Asian Currency Crisis are therefore most important factors for the international investors to take into account when building their investment strategies for the emerging markets.

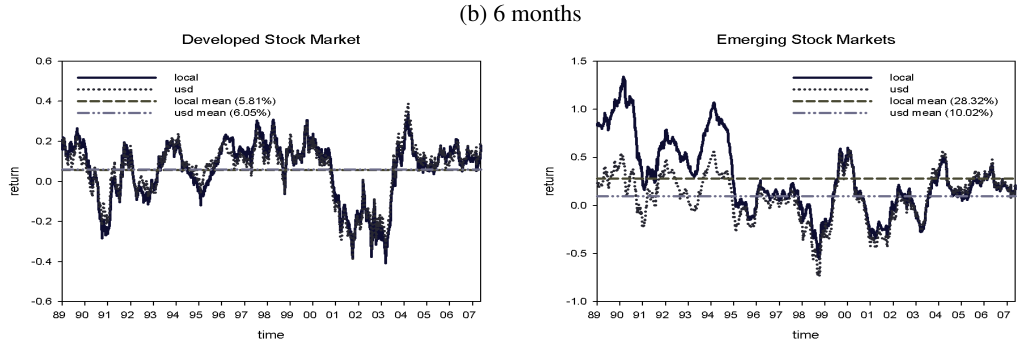
Figure 4.2: Time Series Plot of Emerging and Developed Equity Returns Accumulated over Different Holding Periods over 1988–2007



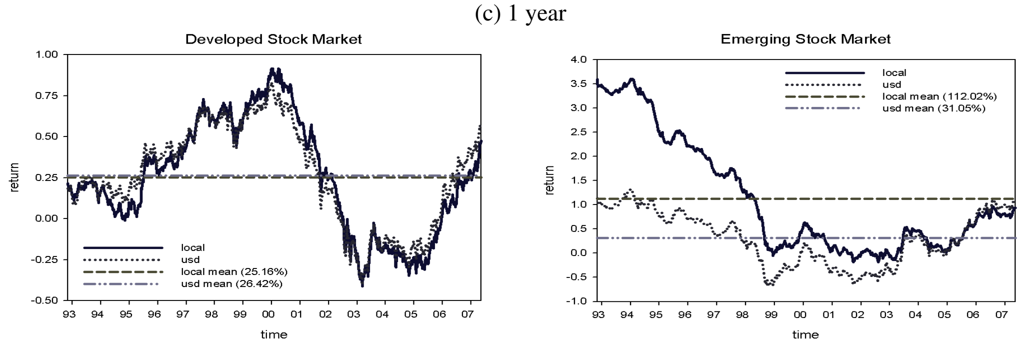
(a) 3 months



(b) 6 months



(c) 1 year



(d) 5 years

Table 4.3: Descriptive Statistics over 1988–1996 before the 1997 Asian Currency Crisis

The table presents the number of observations (OBS), mean (MEAN), standard deviation (STDEV), skewness (SKEW), excess kurtosis (KURT) and the Jarque–Bera test statistic for normality (JB) for the emerging and the developed market daily stock returns, and cumulative returns over 3-month (63 days), 6-month (126 days), 1-year (252 days) and 5-year (1260 days) holding periods over the sub-period sample (1988–1996). The standardised skewness and kurtosis, computed respectively by $\frac{\sum_{i=1}^{N}(r_i-\bar{r})^3/(N-1)}{\sigma^2}$ and $\frac{\sum_{i=1}^{N}(r_i-\bar{r})^4/(N-1)}{\sigma^4}$, respectively. JB stands for the Jarque–Berra statistic for the null hypothesis of normality of the respective holding period returns. The returns are denominated in both local currency and US dollar over the whole sample period. The daily returns are computed as the log difference of the respective stock price index, and the cumulative returns are constructed allowing an overlapping with a 5-day ahead moving window. D^L and $D^\$$ stand for the investment returns denominated in the local currency and the US dollar in the developed markets. Analogously E^L and $E^\$$ stand for the investment returns denominated in the local currency and the US dollar in the emerging markets. *, ** and *** indicate a rejection of the null at 10%, 5%, and 1%, respectively.

Market/Currency	OBS	MEAN	STDEV	SKEW	KURT	JB
A. Holding period: 1 day						
D^L	2347	0.03%	0.59%	-0.09	5.34	2791***
$D^\$$	2347	0.03%	0.66%	-0.08	5.05	2495***
E^L	2347	0.21%	0.86%	-0.28	5.65	3151***
$E^\$$	2347	0.07%	0.90%	-0.51	5.98	3599***
B. Holding period: 3 months						
D^L	457	1.64%	5.63%	-1.10	3.40	312***
$D^\$$	457	1.74%	5.37%	-0.63	1.18	57***
E^L	457	13.32%	14.09%	-0.11	-0.29	3
$E^\$$	457	4.04%	11.55%	-0.23	0.07	4
C. Holding period: 6 months						
D^L	445	3.10%	8.06%	-0.70	0.26	38***
$D^\$$	445	3.34%	6.63%	-0.42	-0.07	13***
E^L	445	26.49%	23.30%	-0.09	-0.76	11**
$E^\$$	445	7.75%	15.30%	-0.32	-0.08	8**
D. Holding period: 1 year						
D^L	419	6.00%	11.93%	-0.65	-0.07	30***
$D^\$$	419	6.56%	9.51%	-0.67	0.09	31***
E^L	419	54.04%	37.29%	0.00	-1.06	20***
$E^\$$	419	16.32%	19.75%	-0.15	-0.87	15***
E. Holding period: 5 years						
D^L	218	20.76%	12.33%	0.12	-1.15	13***
$D^\$$	218	26.88%	11.80%	0.31	-1.36	20***
E^L	218	277.35%	59.98%	-0.14	-1.51	21***
$E^\$$	218	87.22%	21.67%	-0.19	-0.55	4

Table 4.4: Descriptive Statistics over 1999–2007 after the 1997 Asian Currency Crisis

The table presents the number of observations (OBS), mean (MEAN), standard deviation (STDEV), skewness (SKEW), excess kurtosis (KURT) and the Jarque–Bera test statistic for normality (JB) for the emerging and the developed market daily stock returns, and cumulative returns over 3-month (63 days), 6-month (126 days), 1-year (252 days) and 5-year (1260 days) holding periods over the sub-period sample (1999–2007). The standardised skewness and kurtosis, computed respectively by $\frac{\sum_{i=1}^{N}(r_i-\bar{r})^3/(N-1)}{\sigma^2}$ and $\frac{\sum_{i=1}^{N}(r_i-\bar{r})^4/(N-1)}{\sigma^4}$, respectively. JB stands for the Jarque–Berra statistic for the null hypothesis of normality of the respective holding period returns. The returns are denominated in both local currency and US dollar over the whole sample period. The daily returns are computed as the log difference of the respective stock price index, and the cumulative returns are constructed allowing an overlapping with a 5-day ahead moving window. D^L and $D^{\$}$ stand for the investment returns denominated in the local currency and the US dollar in the developed markets. Analogously E^L and $E^{\$}$ stand for the investment returns denominated in the local currency and the US dollar in the emerging markets. *, ** and *** indicate a rejection of the null at 10%, 5%, and 1%, respectively.

Market	OBS	MEAN	STDEV	SKEW	KURT	JB
A. Holding period: 1 day						
D^L	2177	0.01%	0.87%	-0.01	2.69	659***
$D^{\$}$	2177	0.02%	0.86%	-0.03	2.37	510***
E^L	2177	0.06%	0.91%	-0.55	2.28	584***
$E^{\$}$	2177	0.06%	1.01%	-0.62	2.42	672***
B. Holding period: 3 months						
D^L	436	1.01%	7.24%	-0.81	1.23	76***
$D^{\$}$	436	1.15%	7.13%	-0.53	0.60	27***
E^L	436	3.64%	10.59%	-0.27	-0.54	11***
$E^{\$}$	436	3.48%	11.76%	-0.37	-0.61	16***
C. Holding period: 6 months						
D^L	436	1.80%	10.90%	-0.72	0.40	41***
$D^{\$}$	436	2.13%	10.80%	-0.55	-0.22	23***
E^L	436	6.99%	15.67%	-0.13	-0.92	16***
$E^{\$}$	436	6.65%	17.20%	-0.30	-1.01	25***
D. Holding period: 1 year						
D^L	436	2.92%	16.85%	-0.77	-0.64	50***
$D^{\$}$	436	3.62%	17.04%	-0.62	-0.79	39***
E^L	436	11.67%	21.97%	-0.19	-0.55	8***
$E^{\$}$	436	10.57%	25.22%	-0.39	-0.68	19***
E. Holding period: 5 years						
D^L	236	-8.59%	21.20%	0.81	-0.54	28***
$D^{\$}$	236	-1.94%	24.37%	0.79	-0.66	29***
E^L	236	42.37%	31.13%	0.03	-1.27	16***
$E^{\$}$	236	39.49%	43.51%	0.05	-1.10	12***

4.4.2 Emerging Market vs. Developed Market Investments

We study the stochastic and spatial dominance relations between the emerging and the developed market investments performance in this subsection. The SD and SPD test results can be interpreted from two different points of view. First, from an individual investor's point of view, that investment in asset A dominates investment in asset B, will indicate a (possible) utility gain if from switching investment from B to A. Second, from a market integration and efficiency point of view, such dominance can be regarded as an indication of market inefficiency under all regularity conditions. This is so because if random asset A dominates random asset B, rational investors will take make switch from B to A, and thus the less preferable investment in B would be either driven out of the market or its return will improve until the dominance disappears.

4.4.2.1 Full sample period (1988–2007)

Throughout the empirical evaluation of investment performance, we use a subsampling approach to obtain the critical values of both tests for the 31 subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$. We then select the empirical median as the critical values for inference. X and Y indicate the emerging and the developed market investments over the planned investment horizon. $X \nsim_s Y$ implies that there exists s-order (stochastic or spatial) dominance rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10%, 5% and 1%, respectively. X is said to dominate Y only if neither $X \nsim_s Y$ nor $X \succeq_s Y$ is rejected and vice verse.

Returns denominated in the local currency We first compare the cumulative returns denominated in the local currency from the emerging and the developed market investments over the planned investment horizons. Since the two market investments do not involve any currency risk, we take the position of domestic investors who do not diversify their investments internationally. This situation occurs due to several barriers to international investments and home biases constraints.

Table 4.5 reports the SD and SPD test results. Panel A reports the test results for the investment performance denominated in the local currency. Our SPD test results suggest that the domestic investors of the emerging markets obtain the higher level of time aggregated utilities from investments than the domestic investors of the developed markets over investment horizons greater than 6 months in terms of FSPD whilst there are no significant difference between time aggregated utilities in the emerging and developed market investments in the shorter-term (up to 6 month). This is consistent with the nature of emerging markets in that a market which is emerging will have a faster growth but will be eventually finish emerging, and come out as a developed market (Chapter 9 in Solnik and McLeavey, 2003). This finding suggests that the longer-term (one- to five-year horizon) investments should be held before the local investors will substantially benefit from the growth premium in the emerging markets.

As a robustness check for our proposed SPD tests, we have evaluated the p-value of the tests for a number of different subsample sizes. Figure 4.3 plots the p-values of the SPD tests across 31 different subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$. For the shorter-term holding periods (3- and 6-month), the null of the FSPD and SSPD are rejected for all subsample sizes used whilst there is very weak evidence that the null of the TSPD is not rejected marginally only over the 6-month investment horizon when the subsample size is less than 250. Turning to the long-term holding periods, the null of FSPD, SSPD and TSPD are not rejected over the 5-year investment horizon whilst there is weak evidence that the null of the FSPD and SSPD are rejected marginally for 1-year investment horizon when the subsample size is greater than 300. This is clearly consistent with our inference above based on the median as the critical values, as discussed in Section 3.

We intend to account for the specific risk factors of the emerging markets and evaluate the time aggregated utilities obtained over the planned investment horizons. As investigated in detail in subsection 4.2.3, the SPD test is more proper in this context. Hence, we focus our discussions on the SPD test results, though pointing out any inconsistency between SPD and SD results if necessary.[20]

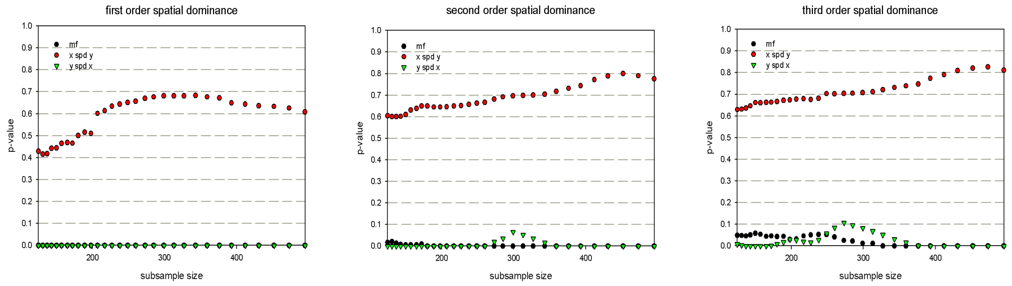
Returns denominated in the US Dollar Panel B of Table 4.5 reports the results for the cumulative investment returns denominated in the US dollar. Since the fluctuation of the currency movement is explicitly accommodated, we take the position of the international investors whose base currency is the US dollar. The results show that the SPD test results

[20]For comparison we also provide the SD test results. The results summarised in Panel A show that domestic investors obtain the higher level of utility from investment than the domestic investors of the developed markets at the terminal point of all the planned investment horizons in terms of FSD. There is a discrepancy between SD and SPD test results only for the 3- and 6-month horizons. Surprisingly, however, the SSD and TSD test results suggest that there is no dominance in both investment horizons, but this is in conflict with the first-order result. Theoretically, this should not be feasible. Indeed we find that the SD test results tend to provide inconsistent results especially for the short-term horizon investments. On the other hand, the SPD test results are mostly consistent across different dominance orders. This is not clearly demonstrated for FSD test in the Monte Carlo experiments but the size distortion problem is strongly observed in SSD test. Hence, we reason that the FSD may be observed in the case of non-FSD data generating process. Further complicated Monte Carlo experiments may be desirable but this is beyond our study.

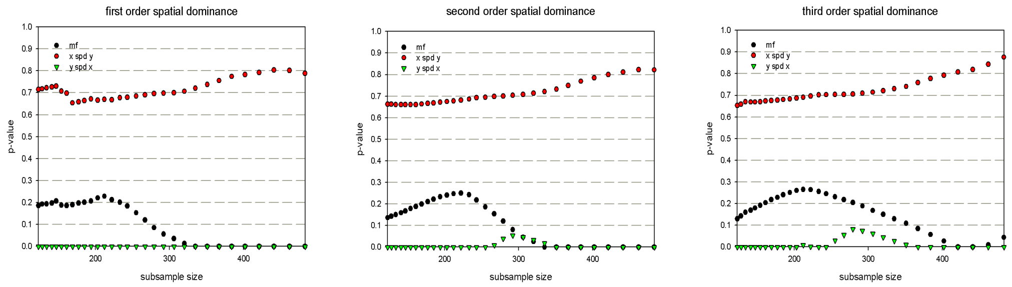
Figure 4.3: The p-values of Spatial Dominance Tests between Emerging and Developed Market Investment Returns Denominated in Local Currency across Different Sub-sample Sizes

The figures plot the p-values of FSPD, SSPD and TSPD test across 31 different subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ over 3-month, 6-month, 1-year and 5-year investment horizons. mf, x spd y and y spd x denote the p-value of SPD test for the null of $X \nsim_{(s)} Y$, $X \succeq_{(s)} Y$ and $Y \succeq_{(s)} X$ for $s = 1, 2, 3$, respectively.

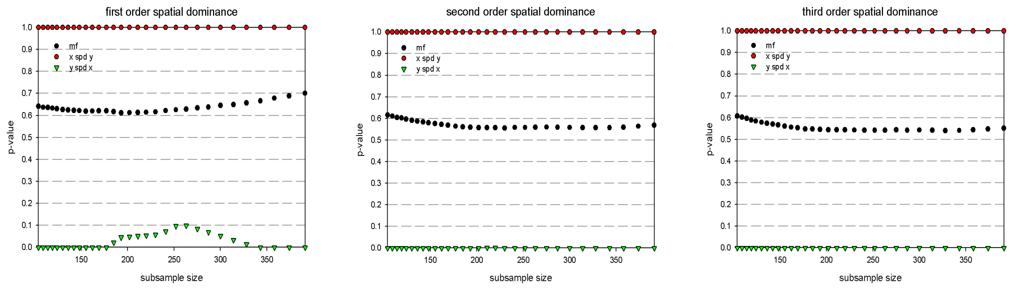
(a) 3 months



(b) 6 months



(c) 1 year



(d) 5 years

strongly indicate that there is no dominance at any order. This evidence suggests that the international investors are indifferent to the investments in the emerging and the developed markets, once the currency risk is fully taken into account.[21] This can be interpreted as that currency depreciation with a high probability of the currency crisis deteriorates the time aggregated utilities obtained from emerging market investments whilst the currency appreciation with a low probability of currency crisis enhances the time aggregated utilities obtained from the developed market investments. From the descriptive statistics, we note that this finding may reflect the volatile currency movement of the emerging market in the early period.

Overall, this finding supports the market efficiency and market integration hypotheses, since the same level of risk (including the higher order risks such as skewness risk than the second order risk) will reward the same level of return for both emerging and developed market investments. The exchange rate parity condition signifies that the currency risk allowance would control for other important macro-economic factors such as inflation and interest rates. Therefore, non-dominance between the emerging and the developed market investments may suggest that the global financial markets are efficient and integrated, once the currency fluctuation is explicitly accounted for.

Figure 4.4 shows the associated p-values of the SPD test across 31 different subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$. The null hypotheses of the FSDP and SSPD are rejected for all subsample sizes and for all investment horizons. There is weak evidence of the TSPD, but both hypotheses, $X \succeq_3 Y$ and $Y \succeq_3 X$, are not rejected either. We therefore conclude that there is no SPD between X and Y, which is also consistent with our formal test results reported in Table 4.5.

In summary, the performance of the emerging market investments strongly dominates the developed counterpart in the 1-year and 5-year investment horizon over the full sample period, 1988–2007, only when the cumulative returns of emerging and developed market investments are denominated in the local currency. In other words, local investors in emerging markets have enjoyed higher levels of utility than those in developed markets. However, once the currency risk factor is accounted for, such dominance disappears. Therefore, both emerging and developed markets are likely to be integrated both in the short- and long-term once the currency risk is fully taken into account.

[21] Consistent results are obtained with the SD tests.

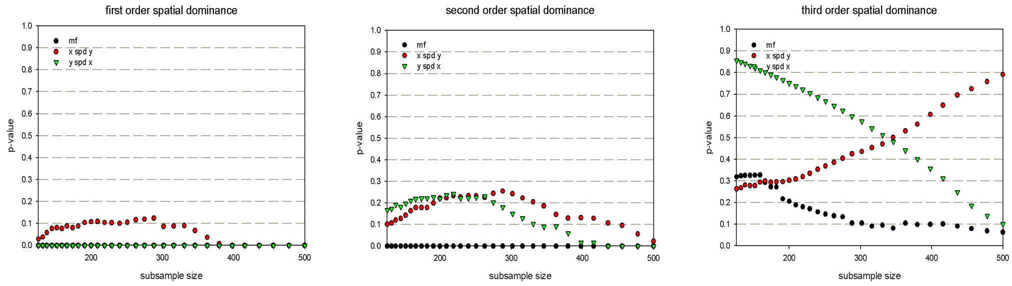
Table 4.5: Test Results for Stochastic and Spatial Dominance between Emerging and Developed Market Investments over 1988–2007

Table present the spatial dominance test (SPD) and stochastic dominance test (SD) results between emerging and developed market investments over 1988–2007. The returns are denominated in the local currency in Panel A and in the US dollar in Panel B. We use the subsampling scheme and obtain the critical values for subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the emerging and the developed market investments, respectively over the planned investment horizon. $X \not\sim_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10%, 5% and 1%, respectively.

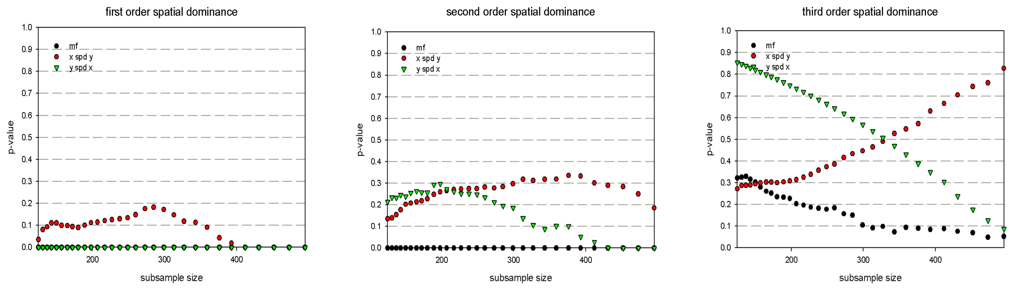
Holding Period		3 months		6 months		1 year		5 years	
Order	Null	SD	SPD	SD	SPD	SD	SPD	SD	SPD
Panel A: Local Currency									
1st	$X \not\sim_1 Y$	2.090***	1.373***	3.079	0.530***	0.452	0.970	0.000	0.000
	$X \succeq_1 Y$	2.090	1.373	3.079	0.530	0.452	0.970	0.000	0.000
	$Y \succeq_1 X$	12.415***	7.833***	28.311***	2.847***	12.594***	8.826***	24.180***	7.523***
2nd	$X \not\sim_2 Y$	0.269***	0.147***	0.513**	0.071***	0.060	0.161	0.000	0.000
	$X \succeq_2 Y$	0.269	0.147	0.513	0.071	0.060	0.161	0.000	0.000
	$Y \succeq_2 X$	1.834***	0.933***	8.062***	0.618***	6.969***	3.488***	53.429***	12.574***
3rd	$X \not\sim_3 Y$	0.039***	0.017***	0.113**	0.011**	0.016	0.030	0.000	0.000
	$X \succeq_3 Y$	0.039	0.017	0.113	0.011	0.016	0.030	0.000	0.000
	$Y \succeq_3 X$	0.624**	0.359**	4.237	0.341***	6.296***	3.543***	119.008	33.365***
Panel B: US dollar									
1st	$X \not\sim_1 Y$	3.769***	2.759***	7.855***	1.116***	2.422*	2.881***	13.081***	2.884***
	$X \succeq_1 Y$	3.769*	2.759*	7.855	1.116	2.422	2.881	13.081	2.884
	$Y \succeq_1 X$	7.379***	4.705***	18.318***	1.734***	9.429***	5.929***	16.364***	3.369
2nd	$X \not\sim_2 Y$	0.347***	0.179***	1.465***	0.107***	0.972***	0.601***	2.814***	0.826***
	$X \succeq_2 Y$	0.655	0.363	1.937	0.198	0.972	0.721	5.646	1.098
	$Y \succeq_2 X$	0.347	0.179	1.465	0.107	1.224	0.601	2.814	0.826
3rd	$X \not\sim_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_3 Y$	0.130	0.056	0.536	0.042	0.373	0.206	4.695	0.601
	$Y \succeq_3 X$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Figure 4.4: The p-values of Spatial Dominance Tests between Emerging and Developed Market Investment Returns Denominated in US Dollars across Different Sub-sample Sizes

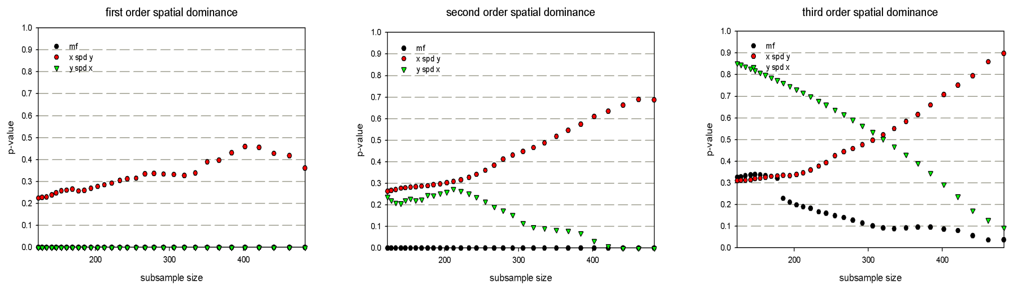
The figures plot the p-values of FSPD, SSPD and TSPD test across 31 different subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ over 3-month, 6-month, 1-year and 5-year investment horizons. mf, x spd y and y spd x denote the p-value of SPD test for the null of $X \sim_{(s)} Y$, $X \succeq_{(s)} Y$ and $Y \succeq_{(s)} X$ for $s = 1, 2, 3$, respectively.



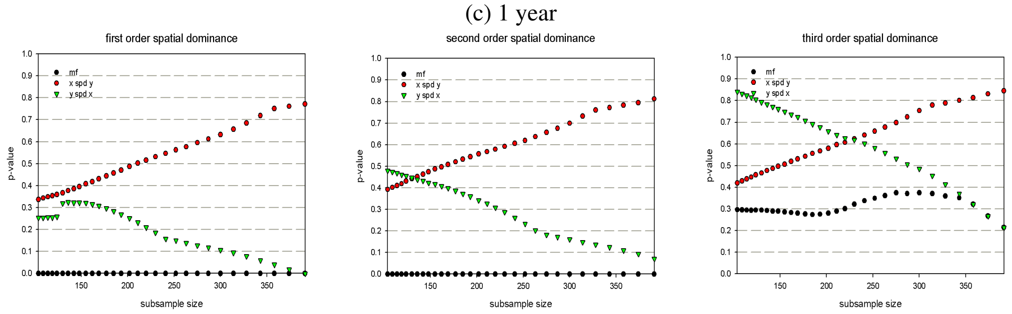
(a) 3 months



(b) 6 months



(c) 1 year



(d) 5 years

4.4.2.2 Sub-period analysis (1988–1996; 1999–2007)

Panel A of Table 4.6 shows that the sub-period results of the SPD test for the cumulative returns denominated in the local currency are generally consistent with the findings in the full sample analysis. The emerging market investments dominates the developed ones over investment horizons greater than or equal to one-year for the first sub-period (1988–1996) before the Asian Currency Crisis. The emerging market investments dominates developed ones for the 5-year investment horizon even when the investment returns are denominated in the US dollar in Panel B of 4.6. This pattern is also consistently observed over the second sub-period (1999–2007) after the crisis (see Table 4.7).

These findings provide an additional insight to the effect of the currency crisis in conjunction with the full sample period results. If the probability of a currency crisis occurring in the emerging markets is negligible, the international investors obtain the higher level of time aggregated utilities from the emerging market investments over the long-term investment horizon even though the currency fluctuation is taken into account. Of course, such excess utility gains disappear with a non-negligible probability of currency risk. Therefore, the probability of a currency crisis is a key risk factor which should be fully considered when considering a long-term investment strategy in emerging markets.

4.4.3 Effect of Currency Risk: The Local Currency vs. The US Dollar

International investors measure their total returns from investments as the sum of returns on the assets denominated in the local currency and any foreign exchange rate changes. Hence, the investors bear both market risks and currency risks. The empirical findings in the previous subsection clearly demonstrate the importance of the currency risk especially in emerging market investments. In this subsection we further study the direct effect of the currency risk on investments.

4.4.3.1 Full sample period (1988–2007)

We first compare the time aggregated utilities obtained from investment returns denominated in the local currency and the US dollar over the planned investment horizon in emerging markets (see Panel A of Table 4.8). Investors are indifferent for the short-term investment horizons (3 or 6 months) whilst they obtain higher time aggregated utilities from the investments denominated in the local currency over the longer-term horizons in terms of FSPD. From the international investors' point of view, this finding shows that the effect of currency risk will become more important for an investment horizon greater than or equal to

Table 4.6: Test Results for Stochastic and Spatial Dominance between Emerging and Developed Market Investments over 1988–1996 before the 1997 Asian Currency Crisis

Table present the spatial dominance test (SPD) and stochastic dominance test (SD) results between emerging and developed market investments over 1988–1996 before the 1997 Asian Currency Crisis. The returns are denominated in the local currency in Panel A and in the US dollar in Panel B. We use the subsampling scheme and obtain the critical values for subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the emerging and the developed market investments, respectively over the planned investment horizon. $X \approx_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10%, 5% and 1%, respectively.

Holding Period		3 months		6 months		1 year		5 years	
Order	Null	SD	SPD	SD	SPD	SD	SPD	SD	SPD
Panel A: Local Currency									
1st	$X \approx_1 Y$	0.561	0.368	0.284	0.297	0.000	0.000	3.007	0.000
	$X \succeq_1 Y$	0.561	0.368	0.284	0.297	0.000	0.000	5.168	0.000
	$Y \succeq_1 X$	12.443 ***	8.289 ***	13.463 ***	9.264 ***	14.785 ***	10.079 ***	5.177 ***	11.698 ***
2nd	$X \approx_2 Y$	0.042	0.026 **	0.022	0.030 *	0.000	0.000	1.072	0.000
	$X \succeq_2 Y$	0.042	0.026	0.022	0.030	0.000	0.000	6.554	0.000
	$Y \succeq_2 X$	2.467 ***	1.262 ***	4.947 ***	2.517 ***	9.804 ***	4.941 ***	5.756 ***	19.355 ***
3rd	$X \approx_3 Y$	0.004	0.002 **	0.003 **	0.003 **	0.000	0.000	0.557	0.000
	$X \succeq_3 Y$	0.004	0.002	0.003	0.003	0.000	0.000	8.791	0.000
	$Y \succeq_3 X$	0.909 ***	0.517 ***	2.712 ***	1.558 ***	8.836 ***	5.075 ***	6.623 ***	49.751 ***
Panel B: US dollar									
1st	$X \approx_1 Y$	2.853 ***	1.861 ***	1.896 ***	1.895 ***	1.025	1.669 ***	0.000	0.039
	$X \succeq_1 Y$	2.853 **	1.861	1.896	1.895	1.025	1.669	0.000	0.039
	$Y \succeq_1 X$	5.286 ***	3.590 ***	6.542	4.446 ***	8.442 ***	5.246 ***	14.088 ***	8.480 ***
2nd	$X \approx_2 Y$	0.335 ***	0.189 ***	0.353 ***	0.273 ***	0.163 ***	0.245 ***	0.000	0.003
	$X \succeq_2 Y$	0.335	0.189 *	0.353	0.273	0.163 *	0.245	0.000	0.003
	$Y \succeq_2 X$	0.454 **	0.242 ***	0.955 ***	0.510 ***	1.940 ***	0.988 ***	8.915 ***	5.214 ***
3rd	$X \approx_3 Y$	0.023 ***	0.023 ***	0.072 ***	0.043 ***	0.033 ***	0.041 ***	0.000	0.000
	$X \succeq_3 Y$	0.053	0.023	0.072	0.043	0.033	0.041	0.000	0.000
	$Y \succeq_3 X$	0.023	0.025	0.159	0.106	0.555	0.324 *	6.597 ***	4.992 ***

Table 4.7: Test Results for Stochastic and Spatial Dominance between Emerging and Developed Investments over 1999–2007 after the 1997 Asian Currency Crisis

Table present the spatial dominance test (SPD) and stochastic dominance test (SD) results between emerging and developed market investments over 1999–2007 after the 1997 Asian Currency Crisis. The returns are denominated in the local currency in Panel A and in the US dollar in Panel B. We use the subsampling scheme and obtain the critical values for subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the emerging and the developed market investments, respectively over the planned investment horizon. $X \approx_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10%, 5% and 1%, respectively.

Holding Period		3 months				6 months				1 year				5 years			
Order	Null	SD		SPD		SD		SPD		SD		SPD		SD		SPD	
Panel A: Local Currency																	
1st	$X \approx_1 Y$	1.437	**	0.862	***	0.718		1.054	***	0.000		0.000		4.671		0.000	
	$X \succeq_1 Y$	1.437		0.862	*	0.718		1.054		0.000		0.000		4.671		0.000	
	$Y \succeq_1 X$	6.178	***	3.493	***	6.745	***	3.873	***	6.178	***	4.162	***	5.186	***	4.617	***
2nd	$X \approx_2 Y$	0.121	***	0.099	***	0.033		0.103	***	0.000		0.000		1.147		0.000	
	$X \succeq_2 Y$	0.121		0.099	**	0.033		0.103		0.000		0.000		1.651		0.000	
	$Y \succeq_2 X$	0.548	***	0.255	***	1.077	***	0.523	***	1.826	***	1.043	***	1.304	***	3.417	***
3rd	$X \approx_3 Y$	0.013	***	0.013	***	0.000		0.011	***	0.000		0.000		0.209		0.000	
	$X \succeq_3 Y$	0.013		0.013		0.000		0.011		0.000		0.000		0.885		0.000	
	$Y \succeq_3 X$	0.091	***	0.042	*	0.258	***	0.134	***	0.754	*	0.489	***	0.702	***	3.513	***
Panel B: US dollar																	
1st	$X \approx_1 Y$	2.011	***	1.379	***	1.579		1.601	***	1.964		0.576		0.521		0.017	
	$X \succeq_1 Y$	2.011	**	1.379	***	1.579		1.601	***	1.964		0.576		0.521		0.017	
	$Y \succeq_1 X$	6.082	***	3.610	***	6.506	***	3.874	***	6.465	***	4.424	***	7.876	***	3.179	***
2nd	$X \approx_2 Y$	0.253	***	0.187	***	0.240	***	0.233	***	0.239	*	0.151	***	0.038		0.001	
	$X \succeq_2 Y$	0.253		0.187	***	0.240		0.233	*	0.239		0.151		0.038		0.001	
	$Y \succeq_2 X$	0.487	***	0.210	***	0.935	***	0.446	***	1.451	***	0.900	***	6.364	***	2.497	***
3rd	$X \approx_3 Y$	0.039	***	0.015	***	0.051	***	0.039	***	0.067	***	0.038	***	0.005		0.000	
	$X \succeq_3 Y$	0.039		0.028		0.051		0.039		0.067		0.038		0.005		0.000	
	$Y \succeq_3 X$	0.051		0.015		0.130	*	0.070		0.375		0.320	***	4.905	***	2.737	***

1-year. Hence, effective currency hedging is worthwhile for these holding periods, suggesting that international investors can enjoy the growth premium from the emerging market investments only if such currency risks can be hedged with relatively low costs.

Turning to the developed market investments, we find that investors are indifferent as to whether investments are denominated in the local currency or in the US dollar over all investment horizons (see Panel B of Table 4.8). This indicates that the current risk is more or less hedged in the developed market investments, suggesting that the developed markets are already integrated for the foreign exchange markets as well as the stock markets. Hence, international investors (e.g., any investors whose fund is based on the US dollar) can invest in the developed markets without bearing a currency risk or its hedge.

Therefore, international investors who expect a growth premium in emerging market investments, should take the currency risk into account for their investment strategies, and thus consider efficient instruments for hedging the currency risk.

4.4.3.2 Sub-period analysis (1988–1996 and 1999–2007)

A sub-period analysis provides further evidence that the emerging markets are as yet underdeveloped. In the first sub-period (before the Asian Currency Crisis), international investors obtain the higher time aggregated utilities from the investments denominated in the local currency than in the US dollar in terms of FSPD over all investment horizons in the emerging markets. This strongly suggests that the currency risk should be regarded as a negative factor in the emerging market investments during the 90s. Taking the currency risk would not be rewarded with the proper return in both short- and long-term investments. We also find that if the international investors could have an effective instrument for hedging the currency risk of their emerging market investments, they would gain a higher utility.

During the second sub-period, the FSPD disappears, which can be seen as a sign of improved market development and integration after the Aisan Currency Crisis. This finding is consistent with the earlier finding by Bekaert and Harvey (1995), who provide evidence that many emerging markets were segmented in their early years but have become more integrated in the global market. However, the SSPD of 5 year investments in the emerging markets suggests that international investors are still likely to enjoy potential utility gains if the currency risk is hedged properly. Finally, the results for the developed market investments are almost the same as those obtained for the full sample period.

There are some noticeable differences between the SD and the SPD tests in this section. The SD test appears to conclude a dominance relationship than the SPD test. For example, for the full sample period, the SD tests suggest that the investments denominated in the

Table 4.8: Test Results for Stochastic and Spatial Dominance between Investment Returns Denominated in the Local Currency and in the US dollar over 1988–2007

Table presents the spatial dominance test (SPD) and stochastic dominance test (SD) results between investment returns denominated in the local currency and the US dollar over 1988–2007. Results for emerging market investments are presented in Panel A and those for developed market investments are in Panel B. We use the subsampling scheme and obtain the critical values for subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the investment returns denominated in the local currency and in the US dollar, respectively over the planned investment horizon. $X \nsim_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10%, 5%, and 1%, respectively.

Holding Period		3 months				6 months				1 year				5 years			
Order	Null	SD		SPD		SD		SPD		SD		SPD		SD		SPD	
Panel A: Emerging Markets																	
1st	$X \nsim_1 Y$	0.000		1.373	***	0.000		1.582	***	0.000		0.970		0.000		0.000	
	$X \succeq_1 Y$	0.000		1.373		0.000		1.582		0.000		0.970		0.000		0.000	
	$Y \succeq_1 X$	4.149	*	7.833	***	6.022	**	8.458	***	7.104	***	8.826	***	9.625	***	7.523	***
2nd	$X \nsim_2 Y$	0.000		0.147	***	0.000		0.222	***	0.000		0.161		0.000		0.000	
	$X \succeq_2 Y$	0.000		0.147		0.000		0.222		0.000		0.161		0.000		0.000	
	$Y \succeq_2 X$	1.467	***	0.933	***	2.918	***	1.790	***	5.670	***	3.488	***	22.290	***	12.574	***
3rd	$X \nsim_3 Y$	0.000		0.017	***	0.000		0.034	**	0.000		0.030		0.000		0.000	
	$X \succeq_3 Y$	0.000		0.017		0.000		0.034		0.000		0.030		0.000		0.000	
	$Y \succeq_3 X$	0.622	***	0.359	**	1.698	***	1.043	***	5.297	***	3.543	***	49.670	***	33.365	***
Panel B: Developed Markets																	
1st	$X \nsim_1 Y$	1.140	**	2.759	***	1.370	***	3.206	***	1.647	***	2.881	***	2.070	***	2.884	***
	$X \succeq_1 Y$	1.140	*	2.759	*	1.370		3.206	*	1.647		2.881		2.288		2.884	
	$Y \succeq_1 X$	1.362		4.705	***	1.625	*	5.043	***	2.196	**	5.929	***	2.070		3.369	
2nd	$X \nsim_2 Y$	0.018		0.179	***	0.000		0.298	***	0.000		0.601	***	0.000		0.826	***
	$X \succeq_2 Y$	0.037		0.363		0.128		0.582		0.226		0.721		0.782	*	1.098	
	$Y \succeq_2 X$	0.018		0.179		0.000		0.298		0.000		0.601		0.000		0.826	
3rd	$X \nsim_3 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$X \succeq_3 Y$	0.008		0.056		0.031		0.126		0.068		0.206		0.546	**	0.601	
	$Y \succeq_3 X$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	

Table 4.9: Test Results for Stochastic and Spatial Dominance between Investment Returns Denominated in the Local Currency and in the US dollar over 1988–1996

Table presents the spatial dominance test (SPD) and stochastic dominance test (SD) results between investment returns denominated in the local currency and the US dollar over 1988–1996 before the 1997 Asian Currency Crisis. Results for emerging market investments are presented in Panel A and those for developed market investments are in Panel B. We use the subsampling scheme and obtain the critical values for the subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the investment returns denominated in the local currency and in the US dollar, respectively over the planned investment horizon. $X \nsim_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10, 5 and 1%, respectively.

Holding Period		3 months		6 months		1 year		5 years	
Order	Null	SD	SPD	SD	SPD	SD	SPD	SD	SPD
Panel A: Emerging Markets									
1st	$X \nsim_1 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_1 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_1 X$	6.923 ***	4.453 ***	9.102 ***	5.136 ***	10.832 ***	6.144 ***	14.765 ***	8.529 ***
2nd	$X \nsim_2 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_2 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_2 X$	1.987 ***	1.005 ***	3.962 ***	1.993 ***	7.748 ***	3.892 ***	28.050 ***	13.824 ***
3rd	$X \nsim_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_3 X$	0.782 ***	0.437 ***	2.157 ***	1.259 ***	6.578 ***	3.899 ***	48.118 ***	31.731 ***
Panel B: Developed Markets									
1st	$X \nsim_1 Y$	0.889 ***	0.262	1.517 ***	0.249	2.098	1.081 ***	0.000	0.000
	$X \succeq_1 Y$	0.889	0.471	1.517	0.253	2.098	1.081	3.454	0.898 *
	$Y \succeq_1 X$	0.982	0.262	2.275 ***	0.249	2.489	1.114 ***	0.000	0.000
2nd	$X \nsim_2 Y$	0.002	0.004	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_2 Y$	0.031	0.016	0.149	0.029	0.315	0.148 ***	0.906 ***	0.317 ***
	$Y \succeq_2 X$	0.002	0.004	0.000	0.000	0.000	0.000	0.000	0.000
3rd	$X \nsim_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_3 Y$	0.006	0.002	0.026	0.008	0.076	0.036 ***	0.230 ***	0.134 ***
	$Y \succeq_3 X$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.10: Test Results for Stochastic and Spatial Dominance between Investment Returns Denominated in the Local Currency and in the US dollar over 1999–2007

Table presents the spatial dominance test (SPD) and stochastic dominance test (SD) results between investment returns denominated in the local currency and the US dollar over 1999–2007 after the 1997 Asian Currency Crisis. Results for emerging market investments are presented in Panel A and those for developed market investments are in Panel B. We use the subsampling scheme and obtain the critical values for the subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the investment returns denominated in the local currency and in the US dollar, respectively over the planned investment horizon. $X \not\sim_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10, 5 and 1%, respectively.

Holding Period		3 months		6 months		1 year		5 years	
Order	Null	SD	SPD	SD	SPD	SD	SPD	SD	SPD
Panel A: Emerging Markets									
1st	$X \not\sim_1 Y$	1.006 ***	0.501 ***	1.292 ***	0.594 ***	1.437 ***	0.629 ***	1.953 ***	0.386 ***
	$X \succeq_1 Y$	1.341 *	0.677 ***	1.435 *	0.594 **	1.437	0.635	3.255 *	0.386
	$Y \succeq_1 X$	1.006 **	0.501 *	1.292	0.657 ***	1.964 **	0.629	1.953	1.353 ***
2nd	$X \not\sim_2 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_2 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_2 X$	0.130	0.079	0.184	0.115	0.429	0.178	1.132 *	0.787 ***
3rd	$X \not\sim_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_3 X$	0.036	0.021	0.076	0.041	0.273	0.102	1.023 **	0.983 ***
Panel B: Developed Markets									
1st	$X \not\sim_1 Y$	0.718 ***	0.293 ***	0.574	0.207	0.431	0.144	0.065	0.000
	$X \succeq_1 Y$	1.006	0.348	1.005	0.406	0.910	0.567	3.059 *	0.585
	$Y \succeq_1 X$	0.718	0.293	0.574	0.207	0.431	0.144	0.065	0.000
2nd	$X \not\sim_2 Y$	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_2 Y$	0.027	0.026	0.077	0.040	0.145	0.064	1.022 ***	0.302
	$Y \succeq_2 X$	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3rd	$X \not\sim_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_3 Y$	0.008	0.006	0.022	0.013	0.045	0.024	0.532 ***	0.186
	$Y \succeq_3 X$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

local currency dominate those denominated in the US dollar for both short- and long-term horizon whilst the SPD tests only conclude the latter. The SD tests also suggest second-order dominance of the US dollar over the local currency investments in the developed market, which is not found in SPD tests.[22]

In summary, the emerging market investments do not provide a hedge against currency risk. Hence, international investors need to take currency risk into account in making their global investment mixture. Without hedging the currency risk, they would receive less time aggregated utilities from their emerging market investments over the planned investment horizon. This is especially the case for an investment horizon longer than 1-year. On the other hand, the international investors who invest in the developed markets are less sensitive to currency risk, since the currency movement is well integrated with the US dollar except for the 5-year investment horizon before the Asian Currency Crisis.

4.4.4 Effect of Structural Change: Before and After the Asian Currency Crisis

Finally, we compare the investment performance before and after the Asian Currency Crisis and examine the possibility of the market timing effect. The SPD test results reported in Tables 4.11 and 4.12 clearly demonstrate that the emerging market investments before the Crisis dominate those after the Crisis over all investment horizons only when the currency risk is not considered whilst such a dominance is observed only over the 5-year holding period when the returns are denominated in the US dollar. On the other hand, the developed market investments before the Crisis dominate those after the Crisis only over the 5-year holding period irrespective of the currency denomination. This finding may suggest the possibility of market timing, indicating that emerging market investments were clearly a better alternative during most of the 1990s before the Crisis than after the Crisis. This could be the reflection of the prolonged effect of the currency crisis on the (Asian) emerging markets. Alternatively, it could imply that the high growth premium and the high average return of the emerging market investments seen in 1990s may no longer be sustainable.

Therefore, a very rare event such as the Asian Currency Crisis in the emerging markets is a crucial break point for international investors to construct their investment strategies in emerging markets. Especially if they plan a long-term investment strategy, the probability of extreme events must be the first risk factor considered. Except for the effect of the Asian Currency Crisis, the dominance results of the SD and the SPD tests provide consistent

[22] Not a direct but a similar finding is observed in the Monte Carlo simulation study in which the SD test is more likely to reject the null hypothesis of no dominance.

Table 4.11: Test Results for Stochastic and Spatial Dominance between Subperiods in Emerging Markets

Table presents the spatial dominance test (SPD) and stochastic dominance test (SD) results between investment returns between sub-periods, before/after the 1997 Asian Currency Crisis in the emerging markets. Results for investment returns denominated in the local currency are presented in Panel A and those for investment returns denominated in the US dollar are in Panel B. We use the subsampling scheme and obtain the critical values for subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the investment returns before the 1997 Asian Currency Crisis (1988–1996) and after the Asian Currency Crisis (1999–2007), respectively over the planned investment horizon. $X \nsim_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10%, 5%, and 1%, respectively.

Holding Period		3 months		6 months		1 year		5 years	
Order	Null	SD	SPD	SD	SPD	SD	SPD	SD	SPD
Panel A: Local Currency									
1st	$X \nsim_1 Y$	0.098	0.027	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_1 Y$	0.098	0.027	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_1 X$	5.356 ***	3.043 ***	6.356 ***	3.560 ***	7.725 ***	4.764 ***	10.645 ***	7.135 ***
2nd	$X \nsim_2 Y$	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_2 Y$	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_2 X$	1.427 ***	0.728 ***	2.921 ***	1.517 ***	6.176 ***	3.229 ***	25.000 ***	13.418 ***
3rd	$X \nsim_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$X \succeq_3 Y$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_3 X$	0.468 ***	0.311 ***	1.549 ***	0.968 ***	5.959 ***	3.397 ***	53.684 ***	34.259 ***
Panel B: US Dollar									
1st	$X \nsim_1 Y$	1.184 ***	0.588 ***	1.312 ***	0.388 **	0.382	0.054	0.000	0.000
	$X \succeq_1 Y$	1.403	0.669	1.312	0.388	0.382	0.054	0.000	0.000
	$Y \succeq_1 X$	1.184	0.588	1.535	1.086	2.144	1.625	6.184 ***	4.738 ***
2nd	$X \nsim_2 Y$	0.010 *	0.002	0.009	0.000	0.000	0.000	0.000	0.000
	$X \succeq_2 Y$	0.010	0.002	0.009	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_2 X$	0.093	0.056	0.293	0.156	0.841	0.555	5.083 ***	4.032 ***
3rd	$X \nsim_3 Y$	0.008 ***	0.000	0.001 **	0.000	0.000	0.000	0.000	0.000
	$X \succeq_3 Y$	0.010	0.000	0.001	0.000	0.000	0.000	0.000	0.000
	$Y \succeq_3 X$	0.008	0.014	0.090	0.073	0.620	0.389	4.811 ***	4.423 ***

Table 4.12: Test Results for Stochastic and Spatial Dominance between Sub-periods in Developed Markets

Table presents the spatial dominance test (SPD) and stochastic dominance test (SD) results between investment returns between sub-periods, before/after the 1997 Asian Currency Crisis in the developed markets. Results for investment returns denominated in the local currency are presented in Panel A and those for investment returns denominated in the US dollar are in Panel B. We use the subsampling scheme and obtain the critical values for subsample sizes ranging between $N^{0.7}$ and $N^{0.9}$ with a total of 31 grids. We then select the empirical median values as the critical values for inference. X and Y indicate the investment returns before the 1997 Asian Currency Crisis (1988–1996) and after the Asian Currency Crisis (1999–2007), respectively over the planned investment horizon. $X \nsim_s Y$ implies that there exists s-order SD or SPD rank between X and Y for $s = 1,2,3$. $X \succeq_s Y$ implies that X s-order (stochastically or spatially) dominates Y for $s = 1,2,3$. $Y \succeq_s X$ implies that Y s-order (stochastically or spatially) dominates X for $s = 1,2,3$. *, ** and *** indicate a rejection of the null at 10%, 5%, and 1%, respectively.

Holding Period		3 months				6 months				1 year				5 years			
Order	Null	SD		SPD		SD		SPD		SD		SPD		SD		SPD	
Panel A: Local Currency																	
1st	$X \nsim_1 Y$	1.076	***	0.586	***	0.720		0.488	*	1.617	***	0.314		0.135		0.032	
	$X \succeq_1 Y$	1.076		0.586		0.720		0.488		1.617		0.314		0.135		0.032	
	$Y \succeq_1 X$	1.655		1.130	*	1.723		1.401	*	2.849		1.736	***	7.398	***	3.954	***
2nd	$X \nsim_2 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$X \succeq_2 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$Y \succeq_2 X$	0.155		0.087		0.255		0.163		0.538		0.276	***	3.125	***	1.654	***
3rd	$X \nsim_3 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$X \succeq_3 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$Y \succeq_3 X$	0.020		0.017		0.084		0.052		0.270		0.133	***	1.605	***	1.006	***
Panel B: US Dollar																	
1st	$X \nsim_1 Y$	1.572	***	0.712	***	1.669	***	0.769	***	1.690	***	0.656		0.226		0.012	
	$X \succeq_1 Y$	1.572		0.712		1.669	**	0.769		1.690		0.656		0.226		0.012	
	$Y \succeq_1 X$	1.694		1.058		2.359		1.493		3.318		2.100	***	7.398	***	4.271	***
2nd	$X \nsim_2 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$X \succeq_2 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$Y \succeq_2 X$	0.158		0.085		0.336		0.173		0.676		0.354	***	3.082	***	1.702	***
3rd	$X \nsim_3 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$X \succeq_3 Y$	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
	$Y \succeq_3 X$	0.020		0.017		0.085		0.046		0.338		0.147	***	1.810	***	1.149	***

results over the longer-term investment horizon. However, the SD tests provide incorrect results over a longer-term investment horizon when the period of the Crisis is considered in Table 4.5.

4.5 Conclusions

When the investment returns accumulated over the specified holding period are likely to have a time-varying distribution, we propose using the SPD test, and we study international investor's expected time aggregated utilities obtained from investments in emerging and developed markets over different investment horizons: 3-month, 6-month, 1-year, and 5-year holding periods. The SPD test is a more general approach for analysing nonstationary time series than the SD approach, which requires the assumption of stationary distributions. We show the advantage of the SPD approach over the SD approach through a Monte Carlo simulation study. Applying our tests to the MSCI emerging and developed market indexes, we find that the emerging market investments strongly dominate their developed counterparts over long-term (1 and 5-year) investment horizons for the full sample period (1988–2007) only when the respective investment returns are denominated in the local currency. Such dominance disappears once the currency risk factor is explicitly taken into account. Our findings suggest that both emerging and developed markets are more likely to be integrated into the global market only after the currency risk is appropriately factored into. Therefore, international investors need to explicitly take into account currency risk in their investment. Without hedging the currency risk, international investors would receive less utility from their investments in emerging markets. On the other hand, for the international investors who invest in the developed markets, our results suggest that there is no difference between investing in the local currency and the US dollar.

Moreover, we find that the Asian Currency Crisis affects the international investors' decision on emerging market investments. First, the growth premium of the emerging markets disappears and the currency movement of the emerging markets becomes integrated with the US dollar after the Crisis. Second, the non-negligible probability of a currency crisis deteriorates the time aggregated utilities from a emerging market investments. Hence, international investors should take this probability into account for their investment strategies in emerging market investments even though it is a very rare event. On the other hand, developed market investments are robust to the Crisis. Consistent with the previous literature, the SPD test presents the importance of currency risk in emerging market investments. In addition, our study indicates that currency risk is a key factor for investors to keep in mind in emerging market investments. Without currency risk, the emerging market investments dominate developed market investments over the long-term investment horizon before/after the Crisis. Hence, the emerging markets are not yet completely integrated into the global market. In the short-term, the emerging markets look to be integrated with the developed markets. However, over a longer investment horizon, the emerging markets are segmented from the developed market due to an increased currency risk. This finding is not easily observed by the SD test whilst the SPD test properly detects this by taking into account the interim risk factors of a long-term investment.

Appendix

To evaluate the critical values of the spatial dominance test, we first need to draw a subsample of size M (without shuffling the time order).[23] This results in $N-M+1$ subsamples of size M for X and Y: $X_{(h)} = (X_h, X_{h+1}, \ldots, X_{h+M-1})$ and $Y_{(h)} = (Y_h, Y_{h+1}, \ldots, Y_{h+M-1})$ for $h = 1, \ldots, N-M+1$. From each subsample we evaluate the (spatial dominance) test statistics and obtain the estimated (subsample) $(s-1)$ integrated spatial distributions of X and Y for $s = 1,2,3$ and $h = 1, \ldots, N-M+1$ by

$$\hat{\Lambda}^{(s-1)}_{M,X_{(h)}}(T,x) = \frac{1}{M}\frac{\delta}{(s-1)!}\sum_{i=1}^{M}\sum_{j=1}^{p}\left(x - X_{i+h-1,j\delta}\right)^{s-1} 1\left\{X_{i+h-1,j\delta} \leq x\right\},$$
$$\hat{\Lambda}^{(s-1)}_{M,Y_{(h)}}(T,x) = \frac{1}{M}\frac{\delta}{(s-1)!}\sum_{i=1}^{M}\sum_{j=1}^{p}\left(x - Y_{i+h-1,j\delta}\right)^{s-1} 1\left\{Y_{i+h-1,j\delta} \leq x\right\}. \quad (4.17)$$

Then the subsample test statistics are obtained for $s = 1,2,3$, and $h = 1, \ldots, N-M+1$ by

$$D^{(s)}_{M,X_{(h)}} = \sup_{x\in\mathbb{R}}\sqrt{M}\left(\left[\hat{\Lambda}^{(s-1)}_{M,X_{(h)}}(T,x) - \hat{\Lambda}^{(s-1)}_{N,X}(T,x)\right] - \left[\hat{\Lambda}^{(s-1)}_{M,Y_{(h)}}(T,x) - \hat{\Lambda}^{(s-1)}_{N,Y}(T,x)\right]\right),$$
$$D^{(s)}_{M,Y_{(h)}} = \sup_{x\in\mathbb{R}}\sqrt{M}\left(\left[\hat{\Lambda}^{(s-1)}_{M,Y_{(h)}}(T,x) - \hat{\Lambda}^{(s-1)}_{N,Y}(T,x)\right] - \left[\hat{\Lambda}^{(s-1)}_{M,X_{(h)}}(T,x) - \hat{\Lambda}^{(s-1)}_{N,X}(T,x)\right]\right),$$
$$M^{(s)}_{M,(h)} = \min\left(D^{(s)}_{M,X_{(h)}}, D^{(s)}_{M,Y_{(h)}}\right), \quad (4.18)$$

where $\hat{\Lambda}^{(s-1)}_{N,X}(T,x)$ and $\hat{\Lambda}^{(s-1)}_{N,X}(T,x)$ are defined in (4.8). The subsampling scheme can be successfully justified to show that the sampling distribution of $D^{(s)}_{N,X}$, $D^{(s)}_{N,Y}$ and $M^{(s)}_{N}$ can be approximated by the sampling distribution of $D^{(s)}_{M,X_{(h)}}$, $D^{(s)}_{M,Y_{(h)}}$, and $M^{(s)}_{M,(h)}$ over $N-M+1$ different subsamples of size M. Specifically, we can approximate these sampling distributions by

$$G^{(s)}_{M,X}(\omega) = \frac{\sum_{j=1}^{N-M+1} 1\left\{D^{(s)}_{M,X_{(h)}} \leq \omega\right\}}{N-M+1},\ G^{(s)}_{M,Y}(\omega) = \frac{\sum_{j=1}^{N-M+1} 1\left\{D^{(s)}_{M,Y_{(h)}} \leq \omega\right\}}{N-M+1},$$
$$G^{(s)}_{M}(\omega) = \frac{\sum_{j=1}^{N-M+1} 1\left\{M^{(s)}_{M,(h)} \leq \omega\right\}}{N-M+1}. \quad (4.19)$$

Let $x^{(s)}_{M,(1-\alpha)}$, $y^{(s)}_{M,(1-\alpha)}$ and $m^{(s)}_{M,(1-\alpha)}$ denote the $(1-\alpha)$th sample quantile of $G^{(s)}_{M,X}$, $G^{(s)}_{M,Y}$ and $G^{(s)}_{M}$, i.e., $x^{(s)}_{M,(1-\alpha)} = \inf\left\{\omega : G^{(s)}_{M,X}(\omega) \geq 1-\alpha\right\}$. Then these will become the

[23]In small samples, we need to allow overlapping, to increase the number of subsampling observations.

subsample critical values of significance level α. To make the test procedure robust to the selected subsample size, Linton, Massoumi, and Whang (2005) suggest selecting a sequence of subsample values from $\{M_1, M_2, \ldots, M_K\}$, where M_1 and M_K are the lower and upper bounds, and select the median value as the critical value.[24] For significance level α, we then obtain the estimated critical values, denoted $x^{(s)}_{M_k,(1-\alpha)}$, $y^{(s)}_{M_k,(1-\alpha)}$ and $m^{(s)}_{M_k,(1-\alpha)}$, $k = 1, ..., K$ and $s = 1, 2, 3$, and select the median as the critical value, denoted $\tilde{x}^{(s)}_{M_k,(1-\alpha)}$, $\tilde{y}^{(s)}_{M_k,(1-\alpha)}$ and $\tilde{m}^{(s)}_{M_k,(1-\alpha)}$ for inference.

[24]Linton, Massoumi, and Whang (2005) provide simulation evidence in favour of this approach.

Chapter 5

Concluding Remarks

This study has taken up the semi-parametric modelling of the time-varying non-normal distribution in finance and monetary economics. In Chapter 2 we apply FAR (Park and Qian, 2007, 2011) to modelling the dynamics of the non-parametric asset return density function to improve a VaR analysis. This takes into account advantages of parametric and non-parametric approaches in a hybrid manner to minimise the economic cost and maximise the coverage ability of the VaR model. The superiority of the proposed methodology over the existing ones is verified via the Monte Carlo simulation, as well as several backtestings associated with thirty components of the Dow Jones Industrial Average and their equal-weighted portfolio. Overall, our findings support the fact that the FAR approach improves the performance of the internal VaR model for both regulators and banks in a fair and satisfactory manner.

In Chapter 3 we deal with the problem of modelling and forecasting the time-varying distribution of the UK monthly inflation rate using sectoral inflation rates. To this end, FAR is applied to the cross-sectional distribution of sectoral inflation rates. This framework provides for easy modelling of the cross-sectional variation across sectors, and its dynamics, in a semi-parametric manner. The in-sample forecasting evaluation clearly demonstrates that using the cross-sectional variation improves the performance of forecasting the aggregated inflation rate compared with benchmark autoregressive models which use only the time variation of the aggregated inflation rate. In addition, the out-of-sample mean inflation forecasting, with an uncertainty band, by the bootstrap scheme also accommodates recently realised inflation rates in a satisfactory manner.

In Chapter 4 we rank investment opportunities in the emerging and the developed stock markets through time aggregated utilities obtained from the investment values accumulated over the planned investment horizon, instead of those evaluated only at the terminal point of the horizon. The time aggregated utility is evaluated by the spatial dominance approach (Park, 2007), which is distribution-free and can be applied to both stationary and non-stationary time series. In other words, it solves the standard stochastic dynamic utility optimisation problem regardless of the stationarity of time-series in a non-parametric manner. The proposed test results indicate that the investment in emerging markets is indifferent to that in developed markets over all the investment horizons ranging from 3 months to 5 years, but only if the currency risk is explicitly taken into account. This may suggest that the two markets have become integrated according to the definition of Bekaert and Harvey (2003). We also find that accumulated investment values of emerging markets, denominated in the local currency, dominate those in US dollar related markets over 1- and 5-year investment horizons, implying that there is still an insufficient interaction between equity prices and foreign exchange rates in emerging markets over the longer-term period. As expected, the currency risk is found to be mostly irrelevant for developed market investments.

A natural extension of this study is modelling multivariate time-varying distributions for the financial risk management. Firstly, a functional autoregressive CoVaR (Adrian and Brunnermeier, 2010) will be considered for forecasting a systemic risk. Given the samples of two random variables, return of financial system and return of institution, the conditional quantile function of the return of financial system, given the return of institution, can be estimated in a non-parametric manner (Li and Racine, 2008; Takeuchi, Le, Sears, and Smola, 2006). Since it is the function of the return of financial institution, given α-quantile, the sequence of the quantile function can be modelled by FAR. Once we obtain the forecast of the conditional quantile function, CoVaR is easily obtained. The second extension will be an integrated functional autoregressive VaR. It suggests combining marginal density functions, which are forecasted by FAR, using a Copula approach. This will make it feasible to perform a multi-dimensional risk analysis: if two marginal density functions are generated from different markets (stock market and foreign exchange market) or assets (stock and bond) or regions (US and China), we could access the integrated risk analysis as well as the interdependence risk analysis via the joint density function constructed by the proposed methodology. Thirdly, we will explore the theory of the time aggregated VaR from the spatial distribution and extend it into the time aggregated CoVaR and the integrated time aggregated VaR, analogous to the FAR extension. Since the time aggregated VaR explicitly includes all risks over the investment horizon, it can be interpreted as the maximum loss of asset values incurred over the horizon whilst the standard VaR is simply the maximum loss of asset values occurred at the terminal point of the investment horizon.

All of the proposed extensions will be promising and useful in financial risk management applications. An abundant financial data source makes extensions feasible in practice. Since financial risk management is the most urgent demand at both micro- and macro-level, the accurate and robust modelling of time-varying multivariate asset return distributions must be a fruitful research area in the near future.

Bibliography

Adrian, T., and M. K. Brunnermeier, 2010, Covar, Working Paper 348 Federal Reserve Bank of New York.

Adrian, T., and J. Rosenberg, 2008, Stock returns and volatility: pricing the short-run and long-run components of market risk, *Journal of Finance* 63, 2997–3030.

Aggarwal, R., and R. Aggarwal, 1993, Security return distribution and market structure: evidence from nyse/amex and the nasdaq market, *Journal of Financial Research* 16, 209–220.

Aggarwal, R., C. Inclan, and R. Leal, 1999, Volatility in emerging stock markets, *Journal of financial and Quantitative Analysis* 34, 33–55.

Andersen, T. G., and T. Bollerslev, 1998, Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies, *Journal of Finance* 53, 219–265.

———, F. X. Diebold, and H. Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43–76.

Antoniadis, A., and T. Sapatinas, 2003, Wavelet methods for continuous-time prediction using hilbert-valued autoregressive processes, *Journal of Multivariate Analysis* 87, 133–158.

Bachelier, L, 1900, Théorie de la spéculation, *Annales Scientifiques de l'École Normale Supérieure* 3, 21–86.

Bakshi, G., N. Kapadia, and D. Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bali, G. T., H. Mo, and Y. Tang, 2008, The role of autoregressive conditional skewness and kurtosis in the estimation of conditional var, *Journal of Banking & Finance* 32, 269–282.

Balke, N. S., and M. A. Wynne, 2000, An equilibrium analysis of relative price change and aggregate inflation, *Journal of Monetary Economics* 45, 269–292.

Ball, L., and N. G. Mankiw, 1995, Relative-price changes as aggregate supply shocks, *Quarterly Journal of Economics* 1, 161–193.

Bandorff-Nielsen, O. E., 1997, Normal inverse gaussian distribution and stochastic volatility modelling, *Scandinavian Journal of Statistics* 24, 1–13.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper, 1999, Var without correlation for portfolios of derivative securities, *Journal of Future Markets* 19, 583–602.

——— , 2002, Backtesting derivative portfolios with filtered historical simulation (fhs), *European Financial Management* 8, 31–58.

Barrett, G. F., and S. G. Donald, 2003, Consistent test for stochastic dominance, *Econometrica* 71, 71–104.

BCBS, 1996, Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements, Discussion paper, Bank for International Settlements Basel, Switzerland.

——— , 2004, International convergence of capital measurement and capital standards: A revised framework, Discussion paper, Bank for International Settlements Basel, Switzerland.

Bekaert, G., C. B. Erb, C. R. Harvey, and T. E. Viskanta, 1997, What matters for emerging equity market investments, *Emerging Markets Quarterly* 1, 101–116.

Bekaert, G., and C. R. Harvey, 1995, Time-varying conditional world market integration, *Journal of Finance* 50, 403–445.

——— , 1997, Emerging equity market volatility, *Journal of Financial Economics* 43, 29–77.

——— , 2000, Foreign speculators and emerging equity markets, *Journal of Finance* 55, 565–613.

——— , 2003, Emerging markets finance, *Journal of Empirical Finance* 10, 3–55.

——— , C. Lundblad, and S. Siegel, 2007, Liquidity and expected returns: Lessons from emerging markets, *Review of Finanical Studies* 20, 1783–1831.

Bekaert, G., and M. S. Urias, 1996, Diversification, integration and emerging market closed-end funds, *Journal of Finance* 51, 835–869.

Berkowitz, J., P. Christofferse, and D. Pelletier, 2011, Evaluating value-at-risk models with desk-level data, *Management Science* Forthcomming.

Bertail, P., C. Haefke, D. N. Politis, and H. White, 2004, A subsampling approach to estimating the distribution of diverging statistics with applications to assessing financial market risks, *Journal of Econometrics* 120, 295–326.

Besse, P., H. Cardot, and D. Stephenson, 2000, Autoregressive forecasting of some functional climatic variations, *Scandinavian Journal of Statistics* 27, 673–688.

Black, F., 1976, Studies of stock price volatility change, in *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section*.

——— , M. C. Jensen, and M. Scholes, 1972, *Studies in the theory of capital markets* . chap. The capital asset pricing model: some empirical tests, pp. 79–121 (Praeger Publishers: New York).

Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Blattberg, R. C., and N. J. Gonedes, 1974, A comparison of the stable and student distribution as statistical models for stock prices, *Journal of Business* 47, 244–280.

Bollen, B., and B. Inder, 2002, Estimating daily volatility in financial markets utilizing intraday data, *Journal of Empirical Finance* 9, 551–562.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics* 31, 307–327.

——— , and H. O. Mikkelsen, 1996, Modelling and pricing long memory in stock market volatility, *Journal of Econometrics* 73, 151–184.

Bollerslev, T., and J. M. Wooldridge, 1992, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariance, *Econometric Review* 11, 143–172.

Bosq, D., 1998, *Nonparametric statistics for stochastic processes* (Springer-Verlag: New York).

———, 2000, *Linear Processes in Function Spaces* (Springer: New York).

Bowsher, C. G., and R. Meeks, 2008, The dynamics of economic functions: modeling and forecasting the yield curve, *Journal of the American Statistical Association* 103, 1419–1437.

Brooks, C., S. P. Burke, S. Hervi, and G. Persand, 2005, Autoregressive conditional kurtosis, *Journal of Financial Econometrics* 3, 399–421.

Bryan, M. F., S. G. Cecchetti, and R. L. Wiggins, 1997, Efficient inflation estimation, Working Paper Series 5793 NBER.

Calvo, G., 1983, Staggered prices in a utility-maximizing framework, *Journal of Monetary Economics* 12, 383–398.

Cardot, H., F. Ferraty, A. Mas, and P. Sarda, 2003, Testing hypotheses in the functional linear model, *Scandinavian Journal of Statistics* 30, 241–255.

Cardot, H., A. Mas, and P. Sarda, 1999, Functional linear model, *Statistics & Probability Letters* 45, 11–22.

———, 2007, Clt in functional linear regression models, *Probability Theory and Related Fields* 138, 325–361.

Casillas-Olvera, G., and D. A. Bessler, 2006, Probability forecasting and central bank accountability, *Journal of Policy Modeling* 28, 223–234.

Chen, S-W., C-H. Shen, and Z. Xie, 2008, Evidence of a nonlinear relationship between inflation and inflation uncertainty: the case of the four little dragons, *Journal of Policy Modeling* 30, 363–376.

Choi, C. Y., 2010, Reconsidering the relationship between inflation and relative price variability, *Journal of Money, Credit and Banking* 42, 769–798.

Christoffersen, P., 1998, Evaluating interval forecast, *International Economic Review* 39, 841–861.

———, and D. Pelletier, 2004, Backtesting value-at-risk: a duration-based approach, *Journal of Financial Econometrics* 2, 84–108.

Chung, Y. P., H. Johnson, and M. J. Schill, 2006, Asset pricing when returns are nonnormal: Fama-french factors versus higher-order systematic comoments, *Journal of Business* 79, 923–940.

Claessens, S, S. Dasgupta, and J. Glen, 1995, The cross-section of stock returns : evidence from emerging markets, Policy Research Working Paper Series 1505 The World Bank.

Clark, P. K., 1973, A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* 41, 135–155.

Clements, M. P., 2004, Evaluating the bank of england density forecasts of inflation, *The Economic Journal* 114, 844–866.

——— , 2010, Explanations of the inconsistencies in survey respondents' forecasts, *European Economic Review* 54, 536–549.

——— , and D. F. Hendry, 2002, Pooling of forecast, *Econometrics Journal* 5, 1–26.

——— , 2006, *Handbook of Economic Forecasting*vol. 1 . chap. Forecasting with structural breaks, pp. 605–658 (Elsevier: Amsterdam).

Cukierman, A., and A. H. Meltzer, 1986, A theory of ambiguity, credibility, and inflation under discretion and asymmetric information, *Econometrica* 54, 1099–1128.

Davidson, J., 2004, Moment and memory properties of linear conditional heteroscedasticity models, and a new model, *Journal of Business & Economic Statistics* 22, 16–29.

Davidson, R., and J. Y. Duclos, 2000, Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica* 68, 1435–1464.

DeSantis, G., and S. Imrohoroglu, 1997, Stock returns and volatility in emerging financial markets, *Journal of International Money and Finance* 16, 561–579.

Diebold, F. X., T. A. Gunter, and A. S. Tay, 1998, Evaluating density forecasts with applications to financial risk management, *International Economic Review* 39, 863–883.

Ding, Z., and C. W. J. Granger, 1996, Modeling volatility persistence of speculative returns: a new approach, *Journal of Econometrics* 73, 185–215.

——— , and R. F. Engle, 1993, A long memory property of stock market returns and a new model, *Journal of Empirical Finance* 1, 83–106.

Disyatat, P., and R. G. Gelos, 2001, The asset allocation of emerging market mutual funds, Working Paper 01/111 IMF.

Dumas, B., and B. Solnik, 1995, The world price of foreign exchange risk, *Journal of Finance* 50, 445–479.

Eberlin, E., and U. Keller, 1995, Hyperbolic distributions in finance, *Bernoulli* 1, 289–299.

Elliot, G., and A. Timmermann, 2008, Economic forecasting, *Journal of Economic Literature* 46, 3–56.

Embrechts, P., C. Klüppelberg, and T. Mikosch, 1997, *Modelling Extreme Events for Insurance and Finance* (Springer: Berlin).

Engelberg, J., C. F. Manski, and J. Williams, 2009, Comparing the point predictions and subjective probability distributions of professional forecasters, *Journal of Business and Economic Statistics* 27, 30–41.

Engle, R. F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, *Econometrica* 50, 987–1007.

——— , and T. Bollerslev, 1986, Modeling the persistence of conditional variance, *Econometric Review* 5, 1–50.

Engle, R. F., and G. G. L. Lee, 1999, *Cointegration, Causality, and Forecasting* . chap. A long-run and short-run component model of stock return volatility (Oxford University Press).

Engle, R. F., D. M. Lilien, and R. P. Robins, 1987, Estimating time varying risk premium in the term structure: The arch-m model, *Econometrica* 55, 391–407.

Engle, R. F., and S. Manganelli, 2004, Caviar: conditional autoregressive value at risk by regression quantiles, *Journal of Business and Economic Statistics* 22, 367–381.

Engle, R. F., and V. K. Ng, 1993, Measuring and testing the impact of news on volatility, *Journal of Finance* 48, 1749–1778.

Fama, E. F., 1963, Mandelbrot and the stable paretian hypothesis, *Journal of Business* 36, 420–429.

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

Fernández, C., and M. Steel, 1998, On bayesian modelling of fat tails and skewness., *Journal of the American Statistical Association* 93, 359–371.

Friedman, M., 1977, Nobel lecture: inflation and unemployment, *Journal of Political Economy* 85, 451–472.

Gabor, D., 1946, Theory of communication, *J. IEE* 93, 429–457.

Gaglianone, W. P., R. Lima, O. Linton, and D. Smith, 2009, Evaluating value-at-risk models via quantile regression, Working Paper 09-4, Universidad Carlos III de Madrid.

Garratt, A., K. Lee, M. H. Pesaran, and Y. Shin, 2006, *Global and National Macroeconometric Modelling: A Long-Run Structural Approach* (Oxford University Press: Oxford).

Giordani, P., and P. Söderlind, 2003, Inflation forecast uncertainty, *European Economic Review* 47, 1037–1059.

Giot, P., and S. Laurent, 2003, Value at risk for long and short position, *Journal of Applied Econometrics* 18, 641–663.

Glosten, L., R. Jagannathan, and D. Runkle, 1993, Relationship between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

González-Rivera, G., 1998, Smooth-transition garch models, *Studies in Nonlinear Dynamics and Econometrics* 3, 61–78.

Gourieroux, C., and J. Jasiak, 2001, Local likelihood density estimation and value at risk, Unpublished Manuscript, York University.

Granger, C. W. J., and M. H. Pesaran, 2000, Economic and statistical measure of forecast accuracy, *Journal of Forecasting* 19, 53–70.

Greenwood-Nimmo, M. J., V. H. Nguyen, and Y. Shin, forthcoming, Probabilistic forecasting of output growth, inflation and the balance of trade in a gvar framework, *Journal of Applied Econometrics*.

Guidolin, M., and A. Timmermann, 2008, International asset allocation under regime switching, skew, and kurtosis preferences, *Review of Finance* 21, 889–935.

Hadar, J., and W. R. Russell, 1969, Rules for ordering uncertain prospects, *American Economic Review* 59, 25–34.

Hanoch, G., and H. Levy, 1969, The efficiency analysis of choices involving risk, *The Review of Economic Studies* 36, 335–346.

Hansen, B. E., 1994, Autoregressive conditional density estimation, *International Economic Review* 35, 705–730.

Harvey, C. R., 1994a, Emerging markets, predictability and active investment strategies, in *Asset Management Investment Symposium*.

——— , 1994b, International investment opportunities which capture the predictability of asset returns, in *Asset Management Investment Symposium*.

——— , 1995a, Predictable risk and returns in emerging markets, *Review of Financial Studies* 8, 773–816.

——— , 1995b, The risk exposure of emerging equity markets, *World Bank Economic Review* 9, 19–50.

——— , and A. Siddique, 1999, Autoregressive conditional skewness, *Journal of Financial and Quantitative Analysis* 34, 465–487.

——— , 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263–1295.

Hasbrouck, J., 2007, *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading* (Oxford University Press: New York).

Hau, H., and H. Rey, 2005, Exchange rates, equity prices, and capital flows, *Review of Financial Studies* 19, 273–317.

Hazelton, M. L., and B. A. Turlach, 2009, Nonparametric density deconvolution by weighted kernel estimators, *Statistics and Computing* 19, 217–228.

Henry, P. B., 2000, Stock market liberalization, economic reform and emerging market equity prices, *Journal of Finance* 55, 529–564.

Holland, A. S., 1995, Inflation and uncertainty: tests for temporal ordering, *Journal of Monetary, Credit and Banking* 27, 827–837.

Hsiao, C., Y. Shen, and H. Fujiki, 2005, Aggregate vs. disaggregate data analysis: a paradox in the estimation of a money demand function of japan under the low interest rate policy, *Journal of Applied Econometrics* 20, 579–602.

Hvidkjaer, S., 2006, A trade-based analysis of momentum, *Review of Financial Studies* 19, 457–491.

Johansson, F., M. Seiler, and J. Michael, 1999, Measuring downside portfolio risk,, *Journal of Portfolio Management* 26, 96–107.

Johnson, W. T., 2004, Predictable investment horizons and wealth transfers among mutual fund shareholders, *Journal of Finance* 59, 1979–2012.

Jondeau, E., and M. Rockinger, 2003, Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements, *Journal of Economic Dynamics & Control* 27, 1699–1737.

Jorion, P., 2006, *Value-at-Risk: The New Benchmark for Managing Financial Risk* (McGraw-Hill: Chicago).

Kapetanios, G., V. Labhard, and S. Price, 2007, Forecast combination and the bank of england's suite of statistical forecasting models, Working Paper 323 Bank of England.

Kargin, V., and A. Onatski, 2008, Curve forecasting by functional autoregression, *Journal of Multivariate Analysis* 99, 2508–2526.

Kawakatsu, H., and M. R. Morey, 1999, An empirical examination of financial liberalization and the efficiency of emerging market stock prices, *Journal of Financial Research* 4, 385–411.

Koenker, R., and G. Bassett, 1978, Regression quantiles, *Econometrica* 46, 33–50.

Kohers, G., N. Kohers, and T. Kohers, 2006, The risk and return characteristics of developed and emerging stock markets: the recent evidence, *Applied Economic Letters* 13, 737–743.

Komunjer, I., 2004, Quasi-maximum likelihood estimation for conditional quantile, *Journal of Econometrics* 128, 137–168.

Kuester, K., S. Mittnik, and M. S. Paolella, 2006, Value-at-risk prediction: a comparison of alternative strategies, *Journal of Financial Econometrics* 4, 53–89.

Kupiec, P. H., 1995, Techniques for verifying the accuracy of risk management models, *Journal of Derivatives* 3, 73–84.

Lahiri, K., and F. Liu, 2006, Modeling multi-period inflation uncertainty using a panel of density forecasts, *Journal of Applied Econometrics* 21, 1199–1219.

Lahiri, K., and X. Sheng, 2008, Evolution of forecast disagreement in a bayesian learning model, *Journal of Econometrics* 144, 325–340.

Lee, C. G., and M. J. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733–746.

Lee, S. W., and B. E. Hansen, 1994, Asymptotic theory for the garch(1,1) quasi-maximum likelihood estimator, *Econometric Theory* 10, 29–52.

Levhari, D., and H. Levy, 1977, The capital asset pricing model and the investment horizon, *Review of Economics and Statistics* 59, 92–104.

Levy, H., 1992, Stochastic dominance and expected etility : survey and analysis, *Management Science* 38, 555–593.

———, and R. Duchin, 2004, Asset return distributions and the investment horizon, *Journal of Portfolio Management* 30, 47–62.

Li, Q., and J. S. Racine, 2008, Nonparametric estimation of conditional cdf and quantile function with mixed categorical and continuous data, *Journal of Business and Economic Statistics* 26, 423–434.

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.

Linton, O., E. Massoumi, and Y. J. Whang, 2005, Consistent testing for stochastic dominance under general sampling schemes, *Review of Economic Studies* 72, 735–765.

Lumsdaine, R. L., 1996, Consistency and asymptotic normality of the quasimaximum likelihood estimator in igarch(1,1) and covariance stationary garch(1,1) models, *Econometrica* 64, 575–596.

Mandelbrot, B., 1963, The variation of certain speculative prices, *Journal of Business* 36, 394–419.

Manski, C. F., 2010, *Oxford Handbook on Economic Forecasting* . chap. Interpreting and combining heterogeneous survey forecasts (Oxford University Press).

Markowitz, H. M., 1952, Portfolio selection, *Journal of Finance* 7, 77–91.

Marzio, M. D., and C. C. Taylor, 2004, Boosting kernel density estimates: a bias reduction technique?, *Biometrika* 91, 226–233.

Mas, A., 2007, Weak convergence in the functional autoregressive model, *Journal of Multivariate Analysis* 98, 1231–1261.

McDonald, J. B., and W. Newey, 1988, Partially adaptive estimation of regression model via the generalized t distribution, *Econometric Theory* 4, 428–457.

McDonald, J. B., and Y. J. Xu, 1995, A generalization of the beta distribution with application, *Journal of Econometrics* 66, 133–152.

McFadden, D., 1989, *Studies in the economics of uncertainty (Part II), (in honor of J. Hadar)* . chap. Testing for stochastic dominance (Springer-Verlag: New York).

McNeil, A., and F. Rudiger, 2000, Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach, *Journal of Empirical Finance* 7, 271–300.

Merton, R. C., 1973, Theory of rational option pricing, *Bell Journal of Economics and Management Science* 4, 141–183.

Mitchell, J., and S. G. Hall, 2005, Evaluating, comparing and combining density forecasts using the klic with an application to the bank of england and niesr fan charts of inflation, *Oxford Bulletin of Economics and Statistics* 67, 995–1033.

Mittnik, S., and M. S. Paolella, 2000, Conditional density and value-at-risk prediction of asian currency exchange rates, *Journal of Forecasting* 19, 313–333.

Mitton, T., and K. Vorkink, 2007, Equilibrium underdiversitication and the preference for skewness, *Review of Financial Studies* 20, 1255–1288.

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

Nelson, D. B., 1991, Conditional heteroscedasticity in asset returns: A new approach, *Econometrica* 59, 347–370.

Netftci, S., 2000, Value at risk calculations, extreme events, and tail estimation, *Journal of Derivatives* 7, 23–38.

Onder, Z., and C. Simga-Mugan, 2006, How do political and economic news affect emerging markets? evidence from argentina and turkey, *Emerging Markets Finance and Trade* 42, 50–77.

Osborne, M. F. M., 1959, Brownian motion in the stock market, *Operationas Research* 7, 145–173.

Panorska, A. K., S Mittnik, and S. T. Rachev, 1995, Stable garch models for financial time series, *Applied Mathematics Letters* 8, 33–37.

Park, J. Y., 2007, The spatial analysis of time series, Unpublished Manuscript, Texas A&M University.

———, and J. Qian, 2007, Autoregressive modeling of time-varying densities in functional space, Unpublished Manuscript, Texas A&M University.

———, 2011, Functional regression of continuous state distributions, *Journal of Econometrics* Forthcomming.

Perignon, C., Z. Y. Deng, and Z. J. Wang, 2008, Do bank overstate their value-at-risk?, *Journal of Banking & Finance* 32, 783–794.

Perignon, C., and D. R. Smith, 2009, The level and quality of value-at-risk disclosure by commercial banks, *Journal of Banking & Finance* 34, 362–377.

Pesaran, M. H., and R. Smith, 1995, Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics* 68, 79–113.

Post, T., P. Van Vliet, and H Levy, 2008, Risk aversion and skewness preference, *Journal of Banking & Finance* 32, 1178–1187.

Pritsker, M., 2001, The hidden dangers of historical simulation, Finance and Economics Discussion Series 27 Board of Governors of the Federal Reserve System Washington, D. C.

Revuz, D., and M. Yor, 1994, *Continuous martingale and brownian motion* (Springer-Verlag: New York).

Rich, R., and J. Tracy, 2010, The relationships among expected inflation, disagreement, and uncertainty: evidence from matched point and density forecasts, *Review of Economics and Statistics* 92, 200–207.

RiskMetrics, 1996, Riskmetrics technical document, Discussion paper, J. P. Morgan Fourth Edition.

Roger, S., 2000, Relative prices, inflation and core inflation, Working Paper 58 IMF.

Roll, R., 1997, A critique of the asset pricing theory's tests; part i: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129–176.

Ross, S., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Rothschild, M., and J. E. Stiglitz, 1970, Incresing risk: I. a definition, *Journal of Economic Theory* 2, 225–243.

Salomonsa, R., and H. Grootveldb, 2003, The equity risk premium: emerging vs. developed markets, *Emerging Markets Review* 4, 121–144.

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

Silverman, B. W., 1986, *Density estimation for statistical data analysis* (Chapman & Hall).

Solnik, B., and D. McLeavey, 2003, *International Investments* (Pearson Education).

Stock, J. H., and M. W. Watson, 2003, Forecasting output and inflation: the role of asset prices, *Journal of Economic Literature* 41, 788–829.

Subbotin, M. T., 1923, On the law of frequency of error, *Mathematicheskii Sbornik* 31, 296–301.

Takeuchi, I., Q. V. Le, T. D. Sears, and A. J. Smola, 2006, Nonparametric quantile estimation, *Journal of Machine Learning Research* 7, 1231–1264.

Taylor, J., 2000, Low inflation, pass-through, and the pricing power of firms, *European Economic Review* 44, 1389–1408.

Theodossiou, P., 1998, Financial data and skewed generalized t distribution, *Management Science* 44, 1650–1661.

———, 2000, Skewed generalized error distribution of financial assets and option pricing, Unpublished Manuscript, Rutgers Business School.

Tobin, J., 1972, Inflation and unemployment, *American Economic Review* 62, 1–18.

Ullah, A., 1996, Entropy, divergence and distance measures with econometric applications, *Journal of Statistical Planning and Inference* 49, 137–162.

Vanden, J. M., 2006, Option coskewness and capital asset pricing, *Review of Financial Studies* 19, 1278–1320.

Venkataraman, S., 1997, Value at risk for a mixture of normal distributions: the use of quasi-bayesian estimation techniques, *Economic Perspectives* 21, 2–13.

Whitmore, G. A., 1970, Third degree stochastic dominance, *American Economic Review* 60, 457–459.

Zakoian, J., 1994, Threshold heteroskedastic models, *Journal of Economic Dynamics and Control* 18, 931–955.

Zarnowitz, V., and L. A. Llambros, 1987, Consensus and uncertainty in economic prediction, *Journal of Political Economy* 95, 591–621.