The Effectiveness of Margin-Setting with Historical, Implied, or Realized Volatility

** Chor-yiu SIN
* Kin LAM

** Associate Professor, School of Business, Hong Kong Baptist University, Kowloon Tong, Hong Kong, Tel: (852) 3411 5200, Fax: (852) 3411 5580, e-mail: cysin@hkbu.edu.hk
* Chair Professor & Head, School of Business, Hong Kong Baptist University, Kowloon Tong, Hong Kong, Tel: (852) 3411 5007, Fax: (852) 3411 5585, e-mail: lamkin@hkbu.edu.hk

** Corresponding author

This version: June 5, 2006

Abstract

In this paper, we investigate the effectiveness of volatility forecast in a risk management context. We look at a clearinghouse's margin-setting system, which is primarily designed to control the risk resulting from members' defaults. Once the default risk is judged to be prudential enough, the clearinghouse's remaining concern is the opportunity cost of the investors. Such a framework is applied to evaluate the effectiveness of volatility forecasts based on historical, implied and realized volatility using HSI (Hang Seng Index) futures and options data. Our results generally support the conclusion that IV (implied volatility) outperforms the RV (realized volatility) model, which in turn also outperforms the HV (historical volatility) model.

JEL Classification: G14, G15

Keywords: Historical volatility, implied volatility, information content, margin-setting, realized volatility

Acknowledgments: This is a substantial revision of the paper entitled, "The Regulatory Value of Volatility Forecast: A Comparison between GARCH, Implied Volatility and Realized Volatility". We thank comments form the editor and two anonymous referees, which substantially improve the paper. Comments of the participants at the Inaugural Conference of AAD 2004 (especially those of Jay M. Chung) and those of the participants at the 2004 Annual Meeting of the Financial Management Association International are much appreciated. Kin Lam would like to thank the financial support by the Research Grants Council of the Hong Kong Special Administrative Region Government, in the form of an earmarked grant (HKBU2069/00H), while Chor-yiu Sin would like to thank those of two earmarked grants (HKBU2014/02H and HKBU2472/05H). The usual disclaimers apply.

1. Introduction

What if the margin committee of a clearinghouse uses IV (implied volatility) or RV (realized volatility) to set margin for the futures contracts, instead of using HV (historical volatility)? Can they introduce more savings for the investors without putting the clearinghouses under extra risk? In this paper, we assess the usefulness of forecasts based on (1) HV, (2) IV, and (3) RV, in terms of the effectiveness of margin-setting.

A clearinghouse for futures contracts collects margin money from its clearing members with non-zero open position in the contract. In Hong Kong, when an investor starts a new position, an amount called the initial margin will be collected by the brokerage firm. To maintain their positions, investors have to meet the maintenance margin requirement in the sense that the trader must replenish the account balance up to the maintenance margin level. In the clearing member level, a clearing member is required by the clearinghouse to maintain a sum of money equal to the maintenance margin in its account to satisfy the margin requirement. The margin accounts are then adjusted for gains and losses at the end of each trading day. The maintenance margin at the clearinghouse level is the same as the maintenance margin at the brokerage level.

The maintenance margin is usually set to be 80% of the initial margin applicable to the brokerage level only. Since the clearinghouse deals directly with its clearing members, it is the maintenance margin which is more important for risk management purposes. Throughout this paper, margin refers to the maintenance margin in the Hong Kong index futures market. The purpose of a margin system is to protect the clearinghouse from members' defaults resulting from big losses due to adverse movements of futures prices. Hence, a clearinghouse should impose a margin level such that margin money collected can cover the future price swing with a large probability. To achieve this goal, Duffie (1989) suggested statistical methods to determine a margin level to guard against a default. In the earlier literatures, normality assumption was usually adopted for calculating an appropriate margin. Later, Warshawsky (1989) showed that the usual normality assumption is inappropriate. In lieu of the normal assumption, Kofman (1993), Longin (2000) and Cotter (2001) used extreme value theory to determine an appropriate margin level. Booth, Broussard, Martikainen and Puttonen (1997) and Booth and Broussard (1998) documented that the use of extreme-value statistical techniques to various futures contracts may be beneficial to the margin-setting committee which holds the final authority in margin determination.

While a margin committee of a clearinghouse would not follow a mechanical formula in setting the margin, they do use some benchmark formula as a reference in their margin decision process. This applies not only to the maintenance margin but also to the initial margin. By now it is well-known that asset return volatilities are typically serially dependent (and thus forecastable), and there are fat tails in the asset return. [1] All in all, the benchmark margin is often set to be equal to a constant multiple of a volatility forecast, in which the constant multiple captures the non-normality of the return on futures. See for instance the report by Lam, Lee, Cox, Leung and Zhou (1999). Different clearinghouses would adopt different constant multipliers, depending on how much coverage probability they want to achieve. Needless to say, the larger the multiplier is,

[1] See, for instance, Christoffersen and Diebold (2005).

the larger the coverage probability will be. In Table 4.1 of Lam, Sin and Leung (2004), they tabulated the multipliers adopted by clearinghouses in various markets in the world.

To compare various volatility forecasts, Lam et al. (2004) introduced another dimension of margin effectiveness. They argued that while margin determination should be prudential, an increase in margin level drives up the investors' cost. The opportunity cost imposed on investors who pay up the margin should also be taken into consideration. It is this extra dimension that can differentiate between the effectiveness of various volatility forecasts. They then formulated a measure of opportunity cost and compare volatility forecasts in terms of such opportunity costs.

Parallel to the previous development, the recent literatures show that volatility forecasts based on implied volatility (IV) or realized volatility (RV) sometimes can give more effective volatility forecasts, by and large with a traditional metric for comparing effectiveness of the volatility forecasts, namely goodness-of-fit. As a forecast of the subsequent market volatility, IV is widely believed to be informationally superior to other alternatives, because it is the "market's" forecast of future volatility.[2] Nonetheless, the literature documented equivocal results, when different frequencies of data and/or different definitions of future volatility are used. For instance, although the in-sample results in Day and Lewis (1992) suggested that the IVs may contain incremental information relative to the conditional volatility from generalized autoregressive conditional heteroskedasticity (GARCH) and exponential GARCH (EGARCH) models, they failed to draw any strong conclusions concerning the relative post-sample

[2] For brevity, when no ambiguity arises, we simply call "future volatility" as "volatility".

performance. On the other hand, Canina and Figlewski (1993) found IV to be a poor forecast of the subsequent future volatility. The in-sample test in Lamoureux and Lastrapes (1993) found that IV explains the major part of the conditional variance of daily return, though other terms in a GARCH(1,1) model are also significant.[3] In their post-sample encompassing test, similar results were obtained. Similar results were also obtained when Amin and Ng (1997) compared the IV with various forms of GARCH models with the conditional variance of daily or monthly interest rate in the Eurodollar options market. Fleming (1998) found bias of IV in forecasting the stock market volatility. The degree of bias does not seem large enough to signal the existence of abnormal trading profits though. On the other hand, using monthly time series of non-overlapping data, Christensen and Prabhala (1998) found that IV outperforms the historical volatility and the latter has essentially no explanatory power. Moreover, they also found that a regime shift around the October 1987 crash explains why IV is more biased in the previous literature. Their findings are qualitatively similar to those in Szakmary, Ors, Kim and Davidson (2003), who used daily data from 35 futures options markets from eight separate exchanges.

Using intradaily returns to construct estimates of volatility of various horizons can be traced back to Müller, Dacorogna, Olsen, Pictet, Schwarz and Morgenegg (1990) and Dacorogna, Müller, Nagler, Olsen and Pictet (1993). The literature on this realized volatility has been voluminous over the last 15 years. For instance, using five-minute returns, Taylor and Xu (1997) compared the volatility information found in high-frequency exchange rate quotations and in implied volatility. Blair, Poon and Taylor

[3] Contrast to other papers in the literature, the IV used in that paper is the implied volatility from the Hull-

(2001) compared the information content of implied volatilities and intradaily returns, in the context of forecasting index volatility over horizons from 1 to 20 days. Fleming, Kirby and Ostdiek (2003) found that switching from daily to intradaily returns to estimate the *daily* conditional covariance matrix can increase substantially the economic value of volatility timing. Developing the link between realized volatility and the conditional covariance matrix, Andersen, Bollerslev, Diebold and Labys (2003) used five-minute, fifteen-minute and thirteen-minute returns to estimate the daily volatility. Both Fleming et al. (2003) and Andersen et al. (2003) concluded, respectively with different loss criteria, that an exponentially weighted moving averages (EWMA), rather than the raw realized volatility, performs better.

This paper is organized as follows. In Section 2, we start off with a discussion on the raw data. Following a description in Section 2.1 on the Hang Seng Index Futures (HSIF) data and the corresponding implied volatility data used, Section 2.2 is devoted to the three approaches in volatility forecasts. We first consider in Section 2.2.1 the GARCH-GJR model, which has been found to perform well in margin determination over other statistical historical approaches such as simple moving averages or exponentially weighted moving averages. The IV approach and the RV approach are introduced respectively in Section 2.2.2 and 2.2.3. Section 3 contains a review on the theoretical framework to compare the effectiveness of various volatility forecasts. Section 4 reports the empirical results of the three margin-setting approaches with the HSIF data. Discussions and conclusions can be found in Section 5.

White model, which explicitly models the time-varying volatility.

2. Raw Data and Various Volatility Forecasts

2.1 Raw Data

The Hang Seng Index Futures (HSIF) market is chosen for comparing the effectiveness of the three approaches of volatility forecasts when they are used for margin determination. Data used for computing the HV (historical volatility), IV (implied volatility) and RV (realized volatility) are obtained for the period from July 31, 1996 to September 5, 2003. Daily closing prices for futures contracts are collected to compute the daily returns. In order to mitigate the expiration effects, we follow the suggestion of Puttonen (1993) and shift over to the next nearest futures contract one day before the expiration of the nearest futures. The IV data are provided by the courtesy of the Hong Kong Exchanges and Clearing (HKEx). They are based on prices of near-the-money options at the market close. Table 2.1 below gives some summary statistics of these raw daily data.

Table 2.1 Summary statistics of the original daily data

	Mean	*Median*	*SD*	*Kurtosis*	*Skewness*	*Min.*	*Lower Quartile*	*Upper Quartile*	*Max.*
IV [a]	29.086	27.000	11.580	6.412	1.680	9.500	20.600	34.800	139.000
HSIF [b]	12121	11593	2571	-0.779	0.352	6610	10040	13890	18390
HSIF Return [c]	0.023	-0.027	2.234	17.087	0.400	-22.509	-1.087	1.105	24.799

[a] The implied volatility is in percentage. It runs from July 31, 1996 to September 5, 2003. That amounts to 1710 observations.

[b] The HSIF is the daily close, in index point. It also runs from July 31, 1996 to September 5, 2003. That amounts to 1710 observations.

[c] The HSIF Return is the close-to-close daily return, in percentage. Computed from the HSIF series, it runs from August 1, 1996 to September 5, 2003. That amounts to 1709 observations.

The mean of the IV is around 29%, while the median is 27%. Its minimum is 9.5%, which happened on August 26, 1996. Its maximum is 139%, which happened on October 27, 1997, the date when the HSIF dropped 22.5%. On the other hand, October 27, 1997 is also the date when the return registered a minimum. The maximum return is 25%, which happened on October 29, 1997 when the market rebounded after it had slumped for a few days amidst the Asian financial crisis. On the same day, IV went down from 139% to 69%. These and other extreme values render a huge kurtosis of about 17 for HSIF return. The mean HSIF daily return is 0.023%, which is close to 0%, as expected for close-to-close daily returns. The median is -0.027%. There seems to be no evidence of asymmetry as the skewness stands at 0.4, quite close to 0.

Apart from the daily data, we also use the intradaily tick-by-tick data to construct the realized volatility (RV) for each trading day. Following Taylor and Xu (1997) and Fleming, Kirby and Ostdiek (2003) (see also Müller et al., 1990 and Dacorogna et al., 1993), we compute the five-minute RV from the tick-by-tick data provided by the HKEx. More precisely, for each trading day, the *raw* RV is simply the *average* of the square of five-minute returns of HSIF. [4] Summary statistics of the 5-minute returns can be found in Table 2.2 below.

Table 2.2 Summary statistics of the five-minute data

	Nobs.	*Mean*	*Median*	*SD*	*Min.*	*Lower Quartile*	*Upper Quartile*	*Max.*
5-minute return	81619	0.001	0.002	0.191	-3.412	-0.087	0.089	3.482
5-minute return Squared	81619	0.037	0.008	0.136	0.000	0.001	0.030	12.122

[4] After careful investigation of the tick-by-tick data set, we find that there is only one observation that contains less than 16 5-minute returns in a day, which occurs on May 4, 2000. We delete that observation in the entire sample and that results in 1709 observations.

2.2 Volatility Forecasts Based on HV, IV and RV

As discussed in Section 1, there are numerous ways to forecast the future volatility one-day ahead using the HV. Due to their simplicity or otherwise, the simple moving average or the exponentially weighted moving average volatility forecast are often used by many clearinghouses as inputs to its benchmark margin formula. As documented in the literature, these two forecasts were out-performed by the GARCH forecast, which is the focus of our paper, as far as HV is concerned. This HV forecast is compared with other forecasts such as variants of IV and variants of RV.

2.2.1 Generalized autoregressive conditional heteroskedasticity-GJR (HV)

We consider the GARCH-GJR model first proposed by Glosten, Jagannathan and Runkle (1993). For each trading day t-1, we use 400 past daily returns to estimate the following quasi-log-likelihood function: [5] [6]

$$\sum_{s=t-400}^{t-1} l_{s,t-1}\ , \quad \text{where } l_{s,t-1} = -\frac{1}{2}\ln(2p) - \frac{1}{2}\ln(h_{s-1,t-1}) - \frac{(R_s - m_{t-1})^2}{2h_{s-1,t-1}}, \tag{2.1}$$

$$h_{s-1,t-1} = a_{0,t-1} + a_{1,t-1}(R_{s-1} - m_{t-1})^2 + a_{2,t-1}(R_{s-1} - m_{t-1})^2 I_{s-1,t-1} + a_{3,t-1}h_{s-2,t-1}, \tag{2.2}$$

[5] The conditional distribution of return is assumed to be normal. This is a convenient assumption rather than a realistic one. Statistical justification on the consistency of the estimates, regardless of the validity of this assumption, can be found in Bollerslev and Wooldridge (1992).

[6] For the first forecast where t is August 1, 1996, the variance forecast h_{t-1} is based on the previous 400 daily returns, that is, the daily returns from December 9, 1994 to July 31, 1996.

where the indicator function $I_{s-1,t-1} = 1$ if $R_{s-1} - \mu_{t-1} < 0$ and $I_{s-1,t-1} = 0$ if $R_{s-1} - \mu_{t-1} \geq 0$. μ_{t-1}, $\alpha_{0,t-1}$, $\alpha_{1,t-1}$, $\alpha_{2,t-1}$ and $\alpha_{3,t-1}$ are the model parameters at the end of day t-1.

On the other hand, the one-day-ahead forecast for the mean is simply μ_{t-1}, the one-day-ahead forecast at time t for the volatility (variance) is: [7]

$$h_{t-1} \equiv \alpha_{0,t-1} + \alpha_{1,t-1}(R_{t-1} - \mu_{t-1})^2 + \alpha_{2,t-1}(R_{t-1} - \mu_{t-1})^2 I_{t-1,t-1} + \alpha_{3,t-1} h_{t-2,t-1}. \qquad (2.3)$$

For notational convenience, in the remaining of the paper, we denote the variance forecast of this model as HV-GARCH.

2.2.2 <u>Implied volatility (IV)</u>

Under an IV approach, the one-day-ahead forecast for the volatility (variance) is simply the square of implied volatility calculated at the close of trading day t-1. For sake of comparison, it is the daily variance of the return, in percentage. The lag of IV is denoted as IV-LAG, while its simple averages and its exponentially weighted moving averages are denoted as IV-SMA and IV-EWMA, respectively.

2.2.3 <u>Realized volatility (RV)</u>

Under the RV approach, the realized volatility is defined as the average value of the squared five-minute (or other time interval) intraday return. [8] The lag of RV is

[7] Unlike the usual applications of GARCH or GARCH-GJR model, we allow not only time-varying heteroskedasticity but also time-varying parameters. Further, unlike the usual applications, on the right-

denoted as RV-LAG, while its simple averages and its exponentially weighted moving averages are denoted as RV-SMA and RV-EWMA, respectively.

All in all, seven models are considered. Referring to Sub-sections 2.2.1-2.2.3, they are HV-GARCH, IV-LAG, IV-SMA, IV-EWMA, RV-LAG, RV-SMA and RV-EWMA. For IV-SMA and RV-SMA, 60-day moving averages are used while for IV-EWMA and RV-EWMA: [9]

$$\textit{IV-EWMA} = 0.06\ \textit{IV-LAG} + 0.94\ \textit{IV-EWMA-LAG}. \tag{2.4}$$

$$\textit{RV-EWMA} = 0.06\ \textit{RV-LAG} + 0.94\ \textit{RV-EWMA-LAG}. \tag{2.5}$$

Table 2.3 contains the summary statistics of the daily volatility forecasts. It should be noted that the RV's reported in Table 2.3 is normalized such that its mean is the same as that of HV-GARCH. It is clear that the normalization does not alter the evaluation of the performance of various models on the one hand, and facilitates the interpretation of the multiplier k on the other.

hand side of (2.15), the lagged variance is $h_{t-2,t-1}$ rather than simply h_{t-2}. The latter depends on data up to the end of day *t-2* while the former depends on data up to the end of day *t-1*.

[8] This is in contrast with the usual definition (summation of squared intraday return). We adopt this new definition because of different numbers of 5-minute returns in some trading days when the market was open only for the morning session.

[9] We also tried different parameter values for the SMA and the EWMA. The results are qualitatively the same and thus they are not reported.

Table 2.3 Summary statistics of daily volatility forecast [a]

	Mean	*Median*	*SD*	*Kurtosis*	*Skewness*	*Min.*	*Lower Quartile*	*Upper Quartile*	*Max.*
HV-GARCH	3.792	1.510	8.834	156.491	10.737	0.966	1.166	2.904	173.255
IV-LAG	3.891	2.893	3.784	85.084	5.959	0.358	1.684	4.806	76.671
IV-SMA	3.879	2.990	2.849	1.662	1.472	0.602	1.746	4.695	14.541
IV-EWMA	3.885	3.043	2.967	2.947	1.726	0.620	1.769	4.942	16.195
RV-LAG	3.792	2.292	6.872	390.139	15.633	0.167	1.201	4.270	197.864
RV-SMA	3.776	2.628	3.490	5.211	2.186	0.543	1.493	4.631	19.378
RV-EWMA	3.782	2.699	3.771	12.251	3.001	0.485	1.430	4.566	27.793

[a] All the data run from August 1, 1996 to September 5, 2003, in daily percentage. That amounts to 1708 observations.

3 A Theoretical Framework to Evaluate the Effectiveness of Volatility Forecast: A Review

In this section, we evaluate the performance of volatility forecasts in the context of its usefulness in a clearinghouse’s risk management process. From a clearinghouse's point of view, the utmost concern is prudentiality. That is why, regardless of volatility forecasts used, the constant multiplier should be adjusted to achieve a pre-determined coverage probability. However, if the margin level is set too high, prudentiality is achieved at the expense of the investors’ opportunity cost and defeats the purpose of efficient design of futures contracts. In this section, a brief review on the theoretical framework developed in Lam et al. (2004) is given.

3.1 Measuring prudentiality: coverage probability or expected shortfall

A common measure of the degree of prudentiality resulted from a prescribed margin level is its coverage probability (CP), which is defined as the probability that the margin collected is sufficient to cover the losses arising from the actual price changes in the market. The margin is usually set at a level so that CP is higher than 95%. For a more prudent approach, CP can be set at 99%, or even higher.

In risk management, we often have to decide if we are interested in the conditional or the unconditional losses. In our context, other than using coverage probability as an unconditional measure of prudentiality, we also consider the margin shortfall, which is a conditional measure defined as the loss beyond what could be covered by the margin money. Since the clearinghouse is exposed to default risk once the margin shortfall becomes positive, it is natural to use the expected shortfall (ESF) as a measurement of risk, conditional on a loss beyond the protected level. For detailed discussions on ESF, see Bates and Craine (1999), Acerbi and Tasche (2002) and Tasche (2002). Specifically, we define the shortfall (SF) as:

$$SF = \begin{cases} 0 & if \quad M \geq L \\ L - M & if \quad M < L \end{cases} \tag{2.6}$$

where M represents margin money collected for a position and L is the loss resulting from taking the position. The expected shortfall ESF = E[SF] defined as the expectation of the shortfall can be estimated with historical data once the margin level has been well specified.

3.2 <u>Measuring opportunity cost: margin collected or expected overcharge</u>

In Hong Kong, an individual investor on HSI futures must pay cash to the brokerage firm as margin deposits that will not earn any interest. His or her opportunity cost is equal to the cost of raising the margin money and is hence directly proportional to the margin money. In the clearing members' level, margin requirements imposed on members can be met using Treasury Bills or cash that can earn interest linked to the inter-bank rate in Hong Kong. Under both methods of payment, interest is earned at a rate lower than the clearing member's borrowing rate. Thus a clearing member also has an opportunity cost directly proportional to the margin money although the proportionality constant will depend on the interest spread obtainable by a clearing member. Based on the above argument, we can assume that the opportunity cost for an investor or a clearing member is proportional to the margin money paid up for the futures contract. However, one may argue that as the investor has to pay the investment loss anyway, it is the part of fund exceeding the loss that should constitute the opportunity cost. In other words, from an investor's viewpoint, paying a margin money M is envisaged to be excessive, if the actual loss of her/his position amounts to L, which is smaller than M. In other words, what is relevant is the overcharge, defined as follows:

$$OC = \begin{cases} M - L & if \quad M \geq L \\ 0 & if \quad M < L \end{cases} \tag{2.7}$$

Since futures contracts are marked to market everyday, investors have to pay up the loss (=L) anyway, OC seems to be a more reasonable measure of opportunity cost. Thus, the expected overcharge, EOC = E[OC] is also a sensible measure of opportunity cost.

Similar to ESF, EOC can be estimated by historical data once the margin level has been specified.

3.3 A benchmark formula

It is quite common for a clearinghouse to adopt a formula to determine a benchmark margin level. In most situations, such a formula dictates a margin level which directly varies with the standard deviation of daily returns. [10] More precisely, at the end of each trading day t-1, the margin for the next trading day t is set at a level $P_{t-1}|m_{t-1}+k\sqrt{h_{t-1}}|$, where P_{t-1} is the close of futures price in t-1, k is a prescribed constant, and m_{t-1} and h_{t-1} are the one-day ahead mean forecast and the volatility (variance) forecast, respectively. [11] Often the mean forecast m_{t-1} is close to 0 and hence margin level is directly proportional to volatility forecast. Since m_t is usually of very small magnitude and our purpose is to compare the effectiveness of various volatility forecasts, we assume m_{t-1}=0 in all approaches. It should be emphasized that we do not assume that the daily return is normally distributed and any targeted coverage probability (such as 95% or 98%) can be achieved by varying the constant that takes care of the tails of the distribution.

Since there is more than one measure of a clearinghouse's level of prudentiality (that is, coverage probability or expected shortfall), and there is more than one measure

[10] Some clearinghouses charge a flat margin level independent of market volatility. For example, the Korea Stock Exchanges charges flat 10% margin money for its index futures contract.

[11] Unlike some other papers such as Booth et al. (1997), we assume here that equal margin requirement is applied to both short and long position. This is a common practice for many, if not all, clearinghouses, as they are not supposed to hold a viewpoint on the direction of market movement.

of the investor's opportunity cost (i.e. margin level or expected overcharge), effectiveness of a benchmark formula can be compared using one of the following methods:

(1) Comparing the expected overcharge while holding coverage probability fixed;

(2) Comparing the margin level while holding coverage probability fixed;

(3) Comparing the expected overcharge while holding expected shortfall fixed;

(4) Comparing the margin level while holding expected shortfall fixed.

Empirical results using these fours methods of comparison will be presented in Section 4 below.

4 Empirical Results

As discussed in Section 3, we use two different measures of prudentiality, namely coverage probability and expected shortfall; and two different measures of opportunity cost, namely overcharge and margin level. This results in four different methodologies for comparing the relation between prudentiality and opportunity cost. The results turn out to vary little across these methodologies. See Tables 4.2(i)-4.2(iv) below. Therefore, our discussions concentrate on comparing coverage probability with overcharge.

The targetted coverage probability varies from one clearinghouse to another, thus we consider a spectrum of coverage probabilities ranging from 95% to 99.8%. Specifically, coverage probabilities equal to 95%, 96%, 97%, 99%, 99.5% and 99.8% are considered in this paper. For each coverage probability and for each volatility forecast, we determine a constant multiplier (denoted by k) so that the empirical coverage probability (denoted by ECP) is as targetted. Naturally such multiplier increases as the targetted coverage probability increases. On the other hand, it is also interesting to note that the multiplier varies substantially across different volatility forecasts. For instance, the k value for IV-LAG is substantially smaller than that for HV-GARCH. Nevertheless, it should be pointed out that despite the varying multipliers, the same ECP is achieved across all volatility forecasts.

4.1 Comparing with IV-LAG

For sake of exposition, we first concentrate on the comparisons among HV-GARCH, IV-LAG, and RV-EWMA, which are followed by brief comparisons with their variants. That latter two rather than their variants are chosen because they are by and large performing the best, as well as more commonly used. We also perform a classical z-test [12] on the difference in the overcharges or the margin levels. For this, it is our hypothesis that IV-LAG is informationally most efficient and thus every model is compared with IV-LAG.

Table 4.1(a) summarizes the average overcharge (AOC) while holding empirical coverage probability (ECP) fixed. In this table, the ECP, the multiplier k in the margin setting formula, and AOC are presented for each volatility forecasts. Here, AOC is the total overcharge divided by the number of observations, rather than divided by the

[12] In the classical z-test, we assume that the *difference* of the overcharge is i.i.d. and under the null hypothesis, the population mean equals to 0. The test statistic is $\sqrt{(n-1)}\,\overline{X}/S$, where $\overline{X}$ is the sample mean, S is the sample standard deviation and n is the sample size. See, for instance, Section 4.8, pp.214-217 in Hogg and Craig (1995).

number of failure days or by the number of non-failure days. In Table 4.1, we first align the models so that they have the same prudentiality value, i.e., they have the same empirical coverage probability. This is consistent with the usual practice of clearinghouses to choose a suitable k value so as to achieve a pre-determined prudential level.

Refer to Table 4.1(a). If the return is normally distributed, to achieve an ECP = 95%, k should be 1.960; for ECP = 98%, k should be 2.326; while for ECP = 99.8%, k should be 3.090, etc. However, in Table 4.1, we observe, as expected, that the actual k value needed to achieve the required ECP is larger than the theoretical k value suggested by normality. In general, HV-GARCH requires the highest k while IV-LAG (or its variants) requires the lowest k, but all models suggest a distribution with a fatter tail than normality.

It is quite clear from Table 4.1(a) that IV-LAG performs the best, followed by RV-EWMA and then HV-GARCH. The smoothed version of implied volatility, IV-SMA, performs worse than IV-LAG, though IV-EWMA performs a bit better, when ECP = 96%. However, the improvement is not statistically significant, as one can see from Column (4) of Table 4.1(b). In contrast, the performance of the realized volatility is improved if we consider its smoothed versions RV-SMA and RV-EWMA. [13] This is in line with the results found in Fleming, Kirby and Ostdiek (2003), in which the economic value of volatility timing is considered. All in all, judging from Table 4.1(b), IV-LAG

[13] This is also the case for HV-GARCH. However, the improvement is not substantial and thus the results are not reported.

by and large outperforms other models, at significance level of 1%, for almost all ECP considered.

Tables 4.1(a)-(b) are here

In Figure 4.1, we consider more empirical coverage probabilities and produce an AOC versus ECP plot. One can see from Figure 4.1 that for all ECPs, the AOC of IV-LAG is smaller than the other model.

We close this sub-section by reporting the summary statistics of various versions of opportunity costs (average overage and average margin level), while holding various versions of the prudentialities (estimated coverage probability and average shortfall) fixed. Table 4.2(i) presents the summary statistics of overcharge while Table 4.2(ii) presents those of margin level, with ECP = 98%. On the other hand, Table 4.2(iii) presents the summary statistics overcharge while Table 4.2(iv) presents those of margin level, with ASF = 2.0. The corresponding time plot of the recursive average overcharge, with ECP = 98%, is depicted in Figure 4.2.

Tables 4.2(i)-(iv) are here

Figures 4.1-4.2 are here

4.2 <u>Comparisons among IV-LAG, IV-SMA and IV-EWMA</u>

In this sub-section, we look closely at the performance of IV-LAG and its smoothed versions IV-SMA and IV-EWMA. Similar to Figure 4.1, Figure 4.3 plots AOC versus ECP of these three volatility forecasts. One can see from Figure 4.3 that for all ECPs, the AOC of IV-LAG is smaller than the others, or at least more or less the same.

The corresponding time plot of the recursive average overcharge is presented in Figure 4.4.

Figures 4.3-4.4 are here

All in all, we find overwhelming evidence that IV-LAG outperforms its smoothed versions. This supports the idea that IV-LAG is the markets' expectation on the future volatility and there may not be any improvement when statistical smoothing methods are used. Moreover, as one can see from Table 4.2(i) (see also Tables 4.2(ii)-(iv)), IV-LAG is a good volatility model for setting margins not only because it results in a lower average margin for a given empirical coverage probability, but also because the required fluctuations in margins using this method are lower than for the HV-GARCH and (generally) the RV approaches. [14]

5. Discussions and Conclusions

[14] We are indebted to a referee who makes this careful observation.

The information content of a particular volatility forecast differs, when a different loss function is used. Following the lines in Christoffersen and Jacobs (2004) who studied different option valuation, this paper investigates the effectiveness of volatility forecast in a risk management context. We look at a clearinghouse's margin-setting system, which is primarily designed to control the risk resulting from members' defaults. Once the default risk is judged to be prudential enough, the clearinghouse's remaining concern is the opportunity cost of the investors. Such a framework is applied to evaluate the effectiveness of volatility forecasts based on historical, implied and realized volatility using HSI (Hang Seng Index) futures and options data. Our results generally support the conclusion that IV (implied volatility) outperforms the RV (realized volatility) model, which in turn also outperforms the HV (historical volatility) model. In other words, the IV is found to have the highest information content, in the sense of striking a balance between prudentiality and opportunity cost. In sum, the empirical results in this paper suggest that both RV and IV outperform HV, which is commonly used by the clearinghouses. On the other hand, as in Andersen et al. (2003) and Fleming et al. (2003), the smoothed versions of RV outperform the original one. Though in contrast, we find overwhelming evidence that the original IV outperforms its smoothed versions. This supports the idea that IV is the markets' expectation on the future volatility and there may not be any improvement when statistical smoothing methods are used. In sum, our results throw light on using volatility forecast for not only margin-setting but also for risk management in general, as well as the information contents of different volatility forecasts.

References

Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall. Journal of Banking and Finance 26, 1487—1503.

Amin, K.I., Ng, V.K., 1997. Inferring future volatility from the information in implied volatility in Eurodollar options: a new approach. The Review of Financial Studies 10, 333—367.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579—625.

Bates, D., Craine R., 1999. Valuating the futures market clearinghouse's default exposure during the 1987 crash. Journal of Money, Credit and Banking 31, 248—272.

Blair, B.J., Poon, S.-H., Taylor, S.J., 2001. Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index return. Journal of Econometrics 105, 5—26.

Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11, 143—172.

Booth, G.G., Broussard, J.P., 1998. Setting NYSE circuit breaker triggers. Journal of Financial Services Research 13, 187—204.

Booth, G.G., Broussard, J.P., Martikainen, T., Puttonen, V., 1997. Prudent margin levels in the Finnish Stock Index futures market. Management Science 43, 1177—1188.

Canina, L., Figlewski, S., 1993. The informational content of implied volatility. The Review of Financial Studies 6, 659—681.

Christensen, B.J., Prabhala, N.R., 1998. The relation between implied and realized volatility 50, 125—150.

Christoffersen, P.F., Diebold, F.X., 2005. Financial asset returns, direction-of-change forecasting, and volatility dynamics. Manuscript, downloadable from http://www.ssc.upenn.edu/%7Efdiebold/papers/paper47/tempcd9.pdf.

Christoffersen, P.F., Jacobs, K., 2004. The importance of the loss function in option valuation. Journal of Financial Economics 72, 291—318.

Cotter, J., 2001. Margin exceedences for European stock index futures using extreme value theory. Journal of Banking and Finance 25, 1475—1502.

Dacorogna, M.M., Müller, U.A., Nagler, R.J., Olsen, R.B., Pictet, O.V., 1993. A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. Journal of International Money and Finance 12, 413—438.

Day, T.E., Lewis, C.M., 1992. Stock market volatility and the information content of stock index options. Journal of Econometrics 52, 267—287.

Duffie, D., 1989. Futures Markets, New Jersey: Prentice-Hall.

Fleming, J., 1998. The quality of market volatility forecasts implied by S&P 100 index option prices. Journal of Empirical Finance 5, 317--345.

Fleming, J., Kirby, C., Ostdiek, B., 2003. The economic value of volatility timing using "realized" volatility. Journal of Financial Economics 67, 473—509.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779—1802.

Kofman, P., 1993. Optimizing futures margins with distribution tails. Advances in futures and options research, 6, 263-278.

Lam, K., Lee, L., Cox, R., Leung, R., Zhou, Y., 1999. A report from the ad-hoc working group on risk management. Internal Document of the Hong Kong Futures Exchange Clearinghouse.

Lam, K., Sin, C.-y., Leung, R., 2004. A theoretical framework to evaluate different margin-setting methodologies. Journal of Futures Markets 24, 117—145.

Lamoureux, C.G., Lastrapes, W.D., 1993. Forecasting stock-return variance: toward an understanding of stochastic implied volatilities. The Review of Financial Studies 6, 293—326.

Longin, F., 2000. From value at risk to stress testing: The extreme value approach. Journal of Banking and Finance, 24, 1097—1130.

Müller, U.A., Dacorogna, M.M., Olsen, R.B., Pictet, O.V., Schwarz, M., Morgenegg, C., 1990. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. Journal of Banking and Finance 14, 1189—1208.

Puttonen, V., 1993. Short sale restrictions and the temporal relationship between stock index cash and derivatives markets. Journal of Futures Markets 13, 645—664.

Szakmary, A., Ors, E., Kim, J.K., Davidson, W.N. III, 2003. The predictive power of implied volatility: Evidence from 35 futures markets, Journal of Banking and Finance 27, 2151—2175.

Tasche, D., 2002. Expected shortfall and beyond. Journal of Banking and Finance 26, 1519—1533.

Taylor, S.J., Xu, X., 1997. The incremental volatility information in one million foreign exchange quotations. Journal of empirical finance 4, 317—340.

Warshawsky, M.J., 1989. The adequacy and consistency of margin requirements in the markets for stocks and derivative products. Board of Governors Staff Study, No.15.

Table 4.1(a) Empirical coverage probability (ECP) versus average overcharge (AOC)

	(i) HV-GARCH		(ii) IV-LAG		(iii) IV-SMA		(iv) IV-EWMA		(v) RV-LAG		(vi) RV-SMA		(vii) RV-EWMA	
ECP (%)	K	AOC	K	AOC	k	AOC	K	AOC	k	AOC	K	AOC	k	AOC
95	2.645	331.50 (6.34)	2.110	288.58 (4.24)	2.177	312.55 (4.49)	2.107	295.49 (4.28)	2.516	341.37 (5.95)	2.307	322.86 (4.96)	2.214	300.94 (4.76)
96	2.848	368.60 (6.82)	2.247	317.17 (4.47)	2.273	332.97 (4.65)	2.203	315.85 (4.44)	2.689	375.67 (6.36)	2.438	349.82 (5.21)	2.355	329.80 (5.03)
97	3.092	413.62 (7.40)	2.438	357.41 (4.78)	2.449	370.77 (4.95)	2.409	359.97 (4.77)	2.897	417.22 (6.85)	2.600	383.48 (5.52)	2.557	371.52 (5.42)
98	3.427	475.91 (8.19)	2.668	406.39 (5.16)	2.859	459.71 (5.63)	2.725	428.33 (5.28)	3.166	471.33 (7.49)	2.901	446.55 (6.10)	2.886	440.16 (6.06)
99	3.894	563.50 (9.29)	2.981	473.61 (5.68)	3.159	525.44 (6.13)	3.108	511.97 (5.90)	3.815	603.03 (9.04)	3.517	576.96 (7.29)	3.269	520.87 (6.80)
99.5	4.402	659.59 (10.49)	3.407	565.90 (6.39)	3.891	687.01 (7.37)	3.621	624.82 (6.75)	4.485	739.47 (10.69)	3.771	631.04 (7.78)	3.813	636.24 (7.86)
99.8	5.376	844.52 (12.83)	4.124	721.95 (7.64)	5.669	1080.96 (10.56)	5.215	976.81 (9.54)	5.484	943.73 (13.18)	5.268	951.27 (10.78)	4.624	808.83 (9.50)

Figures in brackets are the standard errors of AOC.

Table 4.1(b) Difference in average overcharge (AOC) (compared with IV-LAG)

	(i) HV-GARCH	(ii) IV-LAG	(iii) IV-SMA	(iv) IV-EWMA	(v) RV-LAG	(vi) RV-SMA	(vii) RV-EWMA
ECP (%)	Difference in AOC	Difference in AOC	Difference in AOC	Difference in AOC	Difference in AOC	Difference in AOC	Difference in AOC
95	42.92** (4.88)	NA	23.97** (2.32)	6.91** (1.65)	52.78** (3.90)	34.27** (2.44)	12.35** (1.95)
96	51.44** (5.28)	NA	15.80** (2.46)	-1.31 (1.76)	58.50** (4.20)	32.66** (2.59)	12.63** (2.08)
97	56.20** (5.76)	NA	13.35** (2.68)	2.56 (1.94)	59.81** (4.56)	26.07** (2.77)	14.11** (2.27)
98	69.52** (6.42)	NA	53.32** (3.08)	21.94** (2.16)	64.93** (5.02)	40.16** (3.10)	33.77** (2.60)
99	89.89** (7.33)	NA	51.84** (3.43)	38.37** (2.45)	129.43** (6.25)	103.35** (3.84)	47.27** (2.98)
99.5	93.69** (8.29)	NA	121.11** (4.16)	58.92** (2.85)	173.57** (7.47)	65.14** (4.08)	70.34** (3.53)
99.8	122.57** (10.14)	NA	359.00** (6.18)	254.86** (4.27)	221.78** (9.23)	229.32** (5.97)	86.88** (4.32)

Figures in brackets are the standard errors of differences in AOC.
†, * and ** denote significance at 10%, 5% and 1% respectively. Assuming a 2-tailed z-test under the usual assumptions and estimation procedure, the critical values used are 1.645, 1.960 and 2.576 respectively.

Table 4.2(i) Summary statistics of the overcharge of different volatility models (ECP = 98%)

	Mean	Median	SD	Kurtosis	Skewness	Minimum	Lower Quartile	Upper Quartile	Maximum
HV-GARCH	475.910	397.128	338.290	23.075	3.347	0	298.742	562.844	4279.194
IV-LAG	406.392	379.521	213.361	0.269	0.548	0	255.156	545.663	1254.231
IV-SMA	459.707	428.383	232.741	-0.349	0.361	0	288.532	618.629	1221.147
IV-EWMA	428.328	399.939	218.218	-0.109	0.431	0	268.056	580.301	1143.734
RV-LAG	471.326	421.050	309.372	12.485	1.925	0	255.123	631.237	4080.576
RV-SMA	446.549	412.543	251.966	0.100	0.590	0	254.667	614.928	1327.625
RV-EWMA	440.160	403.783	250.241	0.737	0.759	0	248.435	594.346	1547.891

Table 4.2(ii) Summary statistics of the margin level of different volatility models (ECP = 98%)

	Mean	Median	SD	Kurtosis	Skewness	Minimum	Lower Quartile	Upper Quartile	Maximum
HV-GARCH	653.145	567.864	360.275	28.048	3.962	307.544	432.166	720.221	4903.194
IV-LAG	583.503	569.419	224.777	0.814	0.690	180.182	395.523	725.888	2034.784
IV-SMA	635.455	636.931	231.541	-0.897	0.290	235.449	438.598	799.982	1231.151
IV-EWMA	604.125	609.787	219.477	-0.675	0.357	227.812	408.405	753.385	1153.738
RV-LAG	648.827	594.616	344.767	16.347	2.395	119.999	396.309	828.775	4720.576
RV-SMA	622.725	608.071	261.615	-0.458	0.503	234.944	412.403	784.279	1374.584
RV-EWMA	617.190	605.539	265.389	0.488	0.750	232.253	397.187	768.456	1702.170

Table 4.2(iii) Summary statistics of the overcharge of different volatility models (ASF = 2.0)

	Mean	Median	SD	Kurtosis	Skewness	Minimum	Lower Quartile	Upper Quartile	Maximum
HV-GARCH	536.793	447.837	370.060	23.961	3.442	0	340.655	626.804	4744.164
IV-LAG	469.073	439.678	233.379	0.274	0.565	0	298.822	619.178	1396.958
IV-SMA	793.876	739.454	339.636	-0.513	0.404	0	519.414	1045.512	1873.497
IV-EWMA	687.200	641.115	298.783	-0.214	0.475	0	441.670	901.668	1642.876
RV-LAG	541.443	484.266	343.442	13.196	1.993	0	301.972	717.186	4596.438
RV-SMA	610.786	567.793	313.930	0.020	0.600	0	362.213	824.874	1690.297
RV-EWMA	527.860	489.163	283.453	0.738	0.781	0	308.018	700.693	1778.197

Table 4.2(iv) Summary statistics of the margin level of different volatility models (ASF = 2.0)

	Mean	Median	SD	Kurtosis	Skewness	Minimum	Lower Quartile	Upper Quartile	Maximum
HV-GARCH	715.083	621.715	394.439	28.048	3.962	336.708	473.148	788.519	5368.164
IV-LAG	647.361	631.735	249.377	0.814	0.690	199.900	438.808	805.328	2257.465
IV-SMA	972.163	974.422	354.227	-0.897	0.290	360.206	670.997	1223.868	1883.502
IV-EWMA	865.487	873.599	314.430	-0.675	0.357	326.370	585.093	1079.322	1652.880
RV-LAG	719.730	659.596	382.443	16.347	2.395	133.113	439.617	919.343	5236.438
RV-SMA	789.075	770.506	331.501	-0.458	0.503	297.705	522.569	993.785	1741.779
RV-EWMA	706.149	692.818	303.641	0.488	0.750	265.728	454.435	879.218	1947.511

Figure 4.1 Average Overcharge vs Empirical Coverage Probability

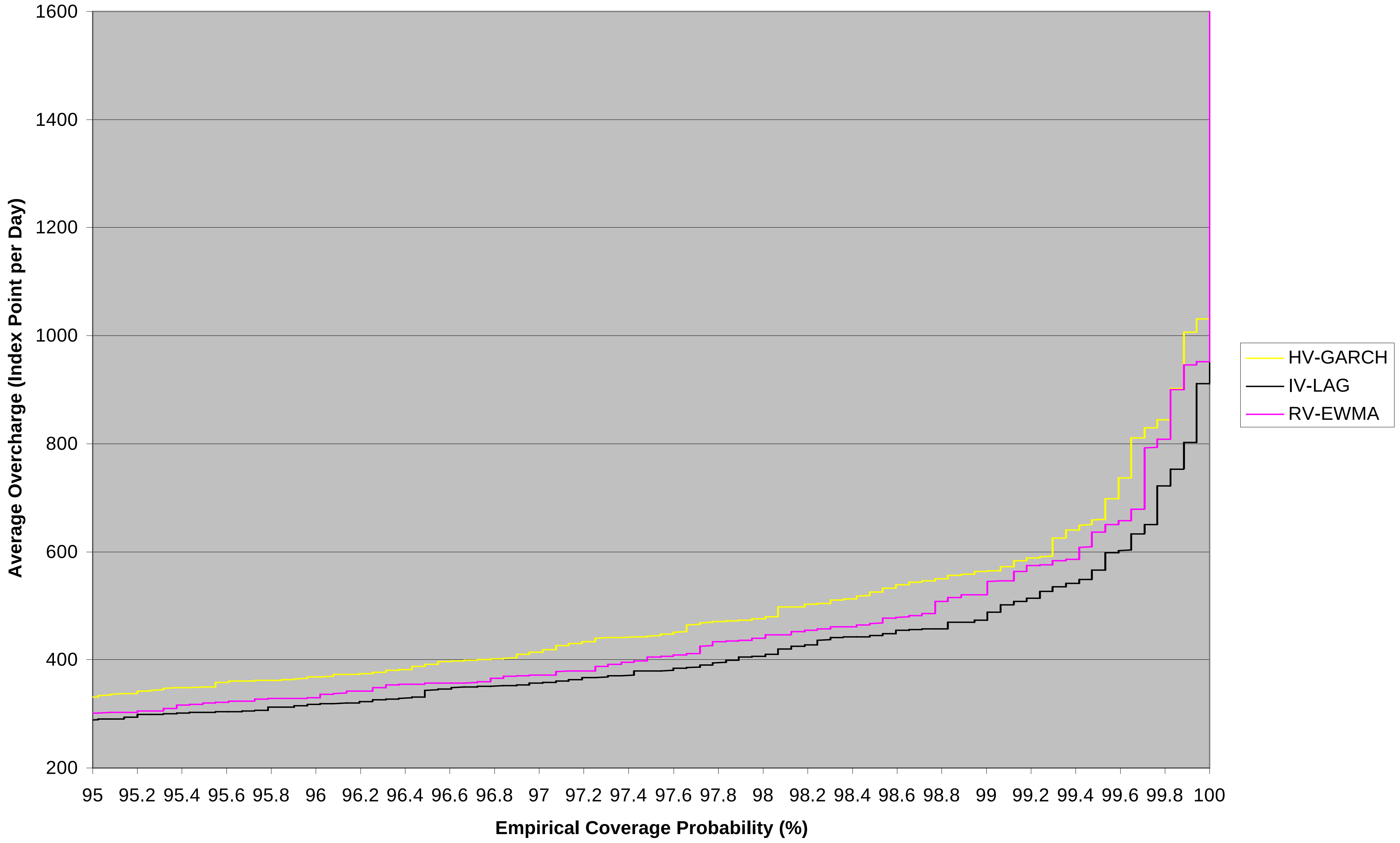

Figure 4.2 Recursive Average Overcharge of Different Models (ECP=98%)

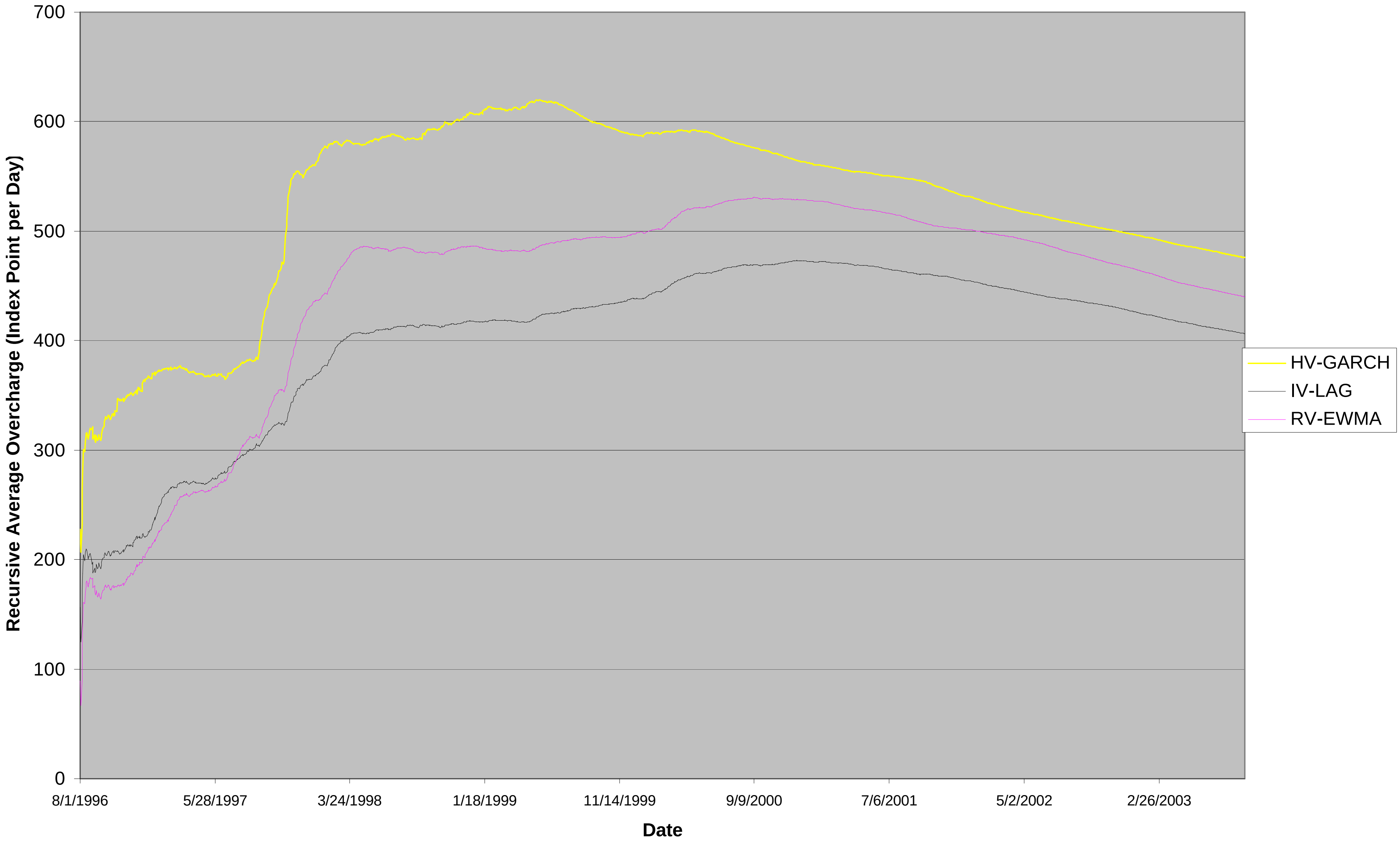

Figure 4.3 Average Overcharge vs Empirical Coverage Probability

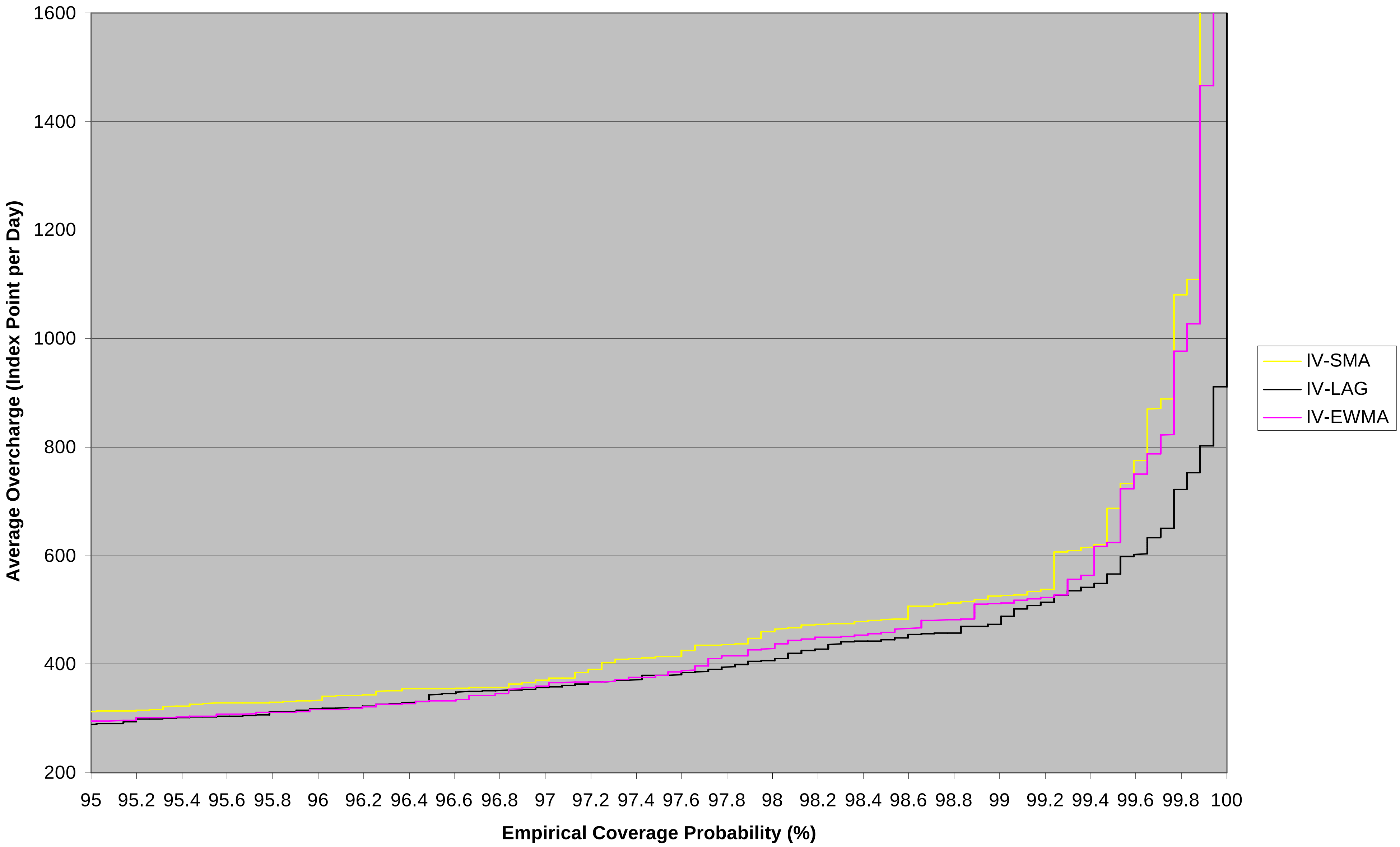

Figure 4.4 Recursive Average Overcharge of Different Models (ECP=98%)