Aalto University
School of Business

# From Smile To Smirk: The Relevance Of Implied Volatility Skew Changes In Swaption VaR Estimation

Finance

Master's thesis

Ilkka Juutilainen

2013

Department of Finance
Aalto University
School of Business

**Aalto University**
**School of Business**

Aalto University, P.O. BOX 11000, 00076
AALTO
www.aalto.fi

**Abstract of master's thesis**

**Author** Ilkka Juutilainen

**Title of thesis** From Smile To Smirk: The Relevance Of Implied Volatility Skew Changes In Swaption VaR Estimation

**Degree** Master of Science (M.Sc)

**Degree programme** Finance

**Thesis advisor(s)** Matti Suominen

**Year of approval** 2013 **Number of pages** 97 **Language** English

**Abstract**

This thesis aims to provide contribution to further development of Value at Risk (VaR) models utilized in the risk measurement and management of financial instruments. More specifically, this study concentrates on VaR measurement of swaptions by employing Historical Simulation and its variations. Furthermore, the objective is to find out whether or not it is worth the added measurement system complexity to incorporate fluctuations in the observed shape of swaption volatility smile as a risk factor into VaR estimation process.

A set of interest rate and swaption implied volatility data from the period between March 8, 2011 and February 1, 2013 is used in this study to generate VaR estimates, the validity of which are evaluated using a variety of backtesting methods that compare the estimates with actual profit and loss figures. The VaR estimates are computed for swaption contracts including maturity-tenor -pairs of 1x2, 1x5, 5x2, 5x5, 10x2 and 10x5. Moreover, the considered contract strike rates in addition to at-the-money (ATM) level comprise the following: +25 bps, +50 bps, +100 bps and +200 bps with respect to the ATM levels. The different main VaR models employed are Historical Simulation (HS), Filtered Historical Simulation (FHS) and Time-Weighted Historical Simulation (TW). These models are applied with modifications regarding the method used for incorporating different implied volatility fluctuations into the simulation. The VaR estimates are generated using a historical observation period of 250 trading days, which leaves 228 trading as the backtesting period.

The backtesting results show partial support with mixed evidence for the validity of the considered VaR models. None of them is able to pass each of the backtests with all tested swaption contracts, but some models could be considered sufficiently accurate in terms of regulatory boundaries. Overall, TW models seem to yield best results, but the estimates suffer from clustering of VaR exceptions, which leads to rejection by Christoffersen's (1998) test of conditional coverage. All of the considered models perform adequately well when tested with short positions, but the number of VaR breaches for long positions is in most cases clearly above the acceptable region.

**Keywords** Value at Risk, swaption, risk management

**Aalto-yliopisto**
**Kauppakorkeakoulu**

**Aalto-yliopisto, PL 11000, 00076 AALTO**
**www.aalto.fi**

**Maisterintutkinnon tutkielman tiivistelmä**

**Tekijä** Ilkka Juutilainen

**Työn nimi** From Smile To Smirk: The Relevance Of Implied Volatility Skew Changes In Swaption VaR Estimation

**Tutkinto** Kauppatieteiden maisteri

**Koulutusohjelma** Rahoitus

**Työn ohjaaja(t)** Matti Suominen

**Hyväksymisvuosi** 2013 **Sivumäärä** 97 **Kieli** Englanti

## Tiivistelmä

Tutkimuksen tavoitteena on kehittää Value at Risk (VaR) -perusteista markkinariskien mittausta ja hallintaa tukevia malleja. Tutkimus kohdistuu erityisesti swap-optioiden markkinariskin mittaamiseen soveltuvien historialliseen simulaatioon perustuvien VaR -mallien kehittämiseen. Tarkoituksena on lisäksi arvioida ns. at-the-money (ATM) -tasolta laskettujen historiallisten implisiittisten volatiliteettimuutosten riittävyyttä riskilaskennassa sellaisille optioille, joiden toteutushinta poikkeaa ATM -tasosta. Tällöin mallin oletuksena on, että implisiittisissä volatiliteeteissa havaittavissa oleva vinouma (nk. smile-efekti) pysyy muodoltaan vakiona. Tutkimuksen pitäisi siis vastata kysymykseen, voiko kyseistä oletusta pitää oikeellisena, vai olisiko mahdolliset muutokset smile-efektissä syytä huomioida riskilaskennassa.

Tutkimuksen aineisto koostuu pääasiassa koronvaihtosopimusten korkodatan sekä swap-optioiden historiallisten implisiittisten volatiliteettien avulla lasketuista swap-optioiden päivittäisistä hintamuutoksista. Hintamuutosten aikasarjoja käytetään riskimallien validoinnissa eli ns. back-testauksessa. Tutkimus perustuu aikaväliltä 8.3.2011 – 1.2.2013 poimittuun dataan. VaR -laskenta on toteutettu yhteensä kuudelle eri maturiteetti-tenori -parille ja kullekin näistä viidelle eri toteutushinnalle. Maturiteetti-tenori -parit ovat 1x2, 1x5, 5x2, 5x5, 10x2 ja 10x5, sekä toteutushinnat suhteessa ATM -tasoon ovat +0, +25 bps, +50 bps, +100 bps ja +200 bps. Riskilaskenta kyseisille sopimuksille toteutetaan erilaisilla historialliseen simulointiin perustuvilla VaR malleilla. Käsiteltävät mallit ovat painottamaton, volatiliteettipainotettu sekä aikapainotettu historiallinen simulointi. Malleja sovelletaan vaihtoehtoisilla metodeilla smile-efektin huomioimisen suhteen.

Mallien validoinnin perusteella tutkimuksessa käsitellyt mallit eivät tuota kaikkien määrittelyiden mukaan hyväksyttäviä riskiarvioita. Yksikään malleista ei läpäise kaikkia testejä kaikilla sovelletuilla optiosopimuksilla. Sen sijaan osa malleista läpäisee kaikilla sopimuksilla sääntelyyn perustuvan testin, joten näiden käyttö riskilaskennassa olisi periaatteessa mahdollista säännösten puitteissa. Aikapainotettujen mallien tulokset ovat suurimmaksi osaksi hyväksyttäviä, mutta myös näiden arviot tuottavat tietyillä sopimuksilla liian ryhmittyneitä VaR-ylityksiä. Kaikki tutkittavat mallit tuottavat hyväksyttäviä arvioita myydyille optiopositiolle, mutta suurin osa malleista tuottaa selvästi liian paljon VaR-ylityksiä ostetuille optioille.

**Avainsanat** Value at Risk, swap-optiot, riskienhallinta

Acknowledgements

I would like to express my sincere gratitude towards the sponsoring institution, Pohjola Bank Plc., and I would especially like to thank Juha Soini, Head of Market and Insurance Risk Management, for providing me the opportunity to write this thesis at Pohjola Bank and for giving valuable insight regarding the theory and implementation of the thesis.

Helsinki, May 2013

Ilkka Juutilainen

**Table of Contents**

## List of Tables

**List of Figures**

# 1 INTRODUCTION

## 1.1 Background

Risk management is an essential concern to any participant who operates in financial markets. The need for practical yet accurate risk management methods increases as both the external requirements arising from increased regulation as well as internal demand for more timely reports continue to create challenges to risk managers. Moreover, simultaneously increasing complexity of financial instruments and positions generates additional challenges to risk management in financial institutions. As a result, the combination of these factors set partially conflicting demands for risk management systems, and avoiding trade-offs is difficult.

A case of such trade-off arises when balancing between the performance and practicality of a risk measurement system: whilst accounting for a higher number of risk factors in market risk computation enhances the accuracy of the estimates by default, it correspondingly adds to the complexity of the system. Increased complexity in turn leads to slower performance despite the constant increase in computational power. Moreover, further investments into the system are needed, and both of these consequences are undesired effects from the managerial viewpoint.

Option contracts in general present a more specified example of the complexities associated with market risk measurement. Option value is determined by a function of several variables, which accordingly translates to several sources of risk. The risks are often quantified in terms of option price sensitivities with respect to changes in the different variables, the most important of which are usually considered to be the sensitivities to changes in the price of the underlying asset and the level of implied volatility. However, in addition to describing the market risk of a given instrument using differential calculus, a common metric that is used for both regulatory and internal reporting of portfolio risk is Value at Risk (VaR).

VaR is designed to provide an aggregated estimate of a portfolios risk by combining the effects of different risk factors into a single figure. It represents the lower percentile of an assumed profit and loss distribution that is based on the movements of an appropriate set of market risk factors over a given time horizon. Nevertheless, depending on the instruments in a portfolio, it is not always straightforward to determine the relevant risk factors to be used. A special case in point is swaption risk measurement. Obvious risk factors for swaptions include the underlying swap rates and the general level of implied volatilities, which are often referred to as delta and vega risks respectively. However, whilst the level of the implied volatility represents a first order source of risk, the volatility of the implied volatility presents a risk source of a second order. Although Malz (2001) introduces a general means for incorporating vega risk into VaR framework by presenting an example for FX options, the importance of the risk related to the volatility in implied volatilities for swaptions has so far received no attention in the literature. More specifically, Malz (2001) illustrates two methods that can be applied with Monte Carlo simulation. The first method takes into account only the general level of the implied volatilities whereas in the second method also the volatility of implied volatility is accounted for. The fundamental idea of these two methods is translated into Historical Simulation based VaR estimation in this study and then employed to provide insight whether or not it would be worth the extra efforts to include the so called smile risk into the swaption VaR framework.

## 1.2 The research problem

The purpose of this study is to first examine the suitability of Historical Simulation (HS), Filtered Historical Simulation (FHS) and time-weighted (TW) methods for swaption VaR estimation with different combinations of risk factors that are accounted for. The risk factors include the following:

- interest rate curves: Euribor 6M for projection and EONIA for discounting
- implied volatilities for swaptions

The second and perhaps more interesting question to be answered is if the addition of risk factor depth through accounting for historical changes in out-of-the-money volatilities significantly improves the swaption VaR estimation accuracy or not compared to the base case of using merely changes in at-the-money volatilities. Particularly, the objective is to examine if the risk arising from conceivable changes in the shape of the volatility smile is significant enough to be included in risk measurement computations for swaptions. This second question embodies also practical importance, since despite the advances in IT and computing power, the data management still requires substantial amount of time in many financial institutions.

## 1.3 Motivation and contribution to the existing literature

Despite the fact that different Value at Risk methods have been widely covered in literature, relatively few papers have been written about VaR application for interest rate instruments, and even fewer for interest rate options. However, the volume of interest rate products traded annually clearly surpasses for instance the volume of equity trading. In this light, it is rather surprising that bulk of the previous VaR articles focus on estimating equity and currency risks.

Moreover, one line of earlier VaR literature focuses more on how to estimate changes in the risk factors, while another major line of work concentrates more on how to translate the estimated changes in risk factors to changes in the portfolio value. The former includes for instance comparisons of different volatility forecasting methods and distributional assumptions that can be integrated into VaR computation. As an example, ARCH family models are well covered. Examples of the former line of work are Eberlein et al. (1998), Billo and Pelizzon (2000), Giot and Laurent (2004) and Shao et al. (2009). In contrast, the alternative line of study pays more attention to relative performance of different VaR models, such as Monte Carlo simulation, Historical Simulation and so called “delta-gamma” method, as well as their practical implementation including different interpolation methods. Examples of these include Britten-Jones and Schaeffer (1999), Jamishidian and Zhu (1996), Barone-Adesi et al. (2002).

This study contributes to both of the aforementioned lines of exploration by providing an empirical investigation of VaR estimation for interest rate options through first introducing a set of alternative methods for incorporating volatility smile changes into swaption risk estimation and then evaluating the differences in VaR backtesting results depending on how the changes in risk factors are estimated.

## 1.4 Limitations of the study

When reading this study, there are three key limitations that should be kept in mind. The first and most important is the fact that the number of backtesting days is quite limited, which means that the statistical power of the implemented backtests is restricted. Moreover, the testing period covers a time period during which the interest rates are at historically low levels and decline gradually throughout most of the observation period. While VaR is assumed to provide an estimation of the risk during so called "normal market conditions", it is questionable whether interest rates very close to or even below zero can be considered "normal".

The second limitation is related to the coverage of different swaption positions. Although this study utilizes swaptions with six different maturity-tenor -pairs and five different strike levels, the moneyness levels are restricted to strike levels above at-the-money level and many possible maturity-tenor -pairs are left out of scope. As a result, the general applicability of the results for swaption positions not covered in this study is uncertain.

The third limitation is associated with the implementation of the VaR estimation for a swaption position: the estimates are not computed for actual positions but for imaginary contracts opened and closed on a daily interval, which is a drastic simplification compared to real life positions. Furthermore, the position pricing accuracy is imperfect as the decline in time until maturity is ignored and also the calibration of the SABR model is not thoroughly optimized. However, since each swaption is held only for one day, the price effect stemming from a one day decline in time to maturity is miniscule as the minimum maturity used in this study is one year. Also, whereas the SABR calibration applied in this study might be insufficient for swaption trading purposes, it should still provide accurate enough values to be used in VaR backtesting as the same parameters are used in both VaR estimation and profit and loss computations.

## 1.5 Main findings

The main finding of this study is that while VaR models based on historical simulation are able to generate sufficiently accurate market risk estimates for most swaption contracts investigated in the empirical part of this paper, the accuracy depends on position attributes and model specifications. As a result, this study is unable to present strong support for applicability of historical simulation based VaR estimation in swaption risk measurement. Even so, the results indicate that the models considered in this study are fundamentally sound but are swayed by the shortness of historical observation period.

Moreover, the results do not provide evidence that would support the use of moneyness-dependent implied volatility changes over the changes observed on at-the-money level in VaR estimation for out-of-the-money swaptions. Instead, the results suggest that the use of at-the-money level changes in swaption implied volatilities is generally sufficient in VaR estimation for also out-of-the-money swaption positions. Furthermore, the results suggest that it is adequate to use implied volatility changes derived from the at-the-money level without considering the simultaneous change in interest rates on successive days and changes in actual moneyness resulting thereof.

### 1.6 Structure of the study

The rest of the paper is organized as follows. The second section combines the theoretical base on which the study is based on with literature review about the topic. Section 3 introduces the hypotheses. Section 4 explains the data, VaR methods and VaR backtesting methods used. Section 5 presents the empirical results from VaR estimation. Section 6 concludes the study and offers suggestions for further research. A list of frequently used abbreviations and model names introduced in this paper is presented in Appendix A.

## 2 LITERATURE REVIEW AND THEORETICAL BACKGROUND

This section is divided into three main parts that illustrate the cornerstones of the study. The first part covers VaR in risk management framework and the second introduces interest rate swaptions. The section concludes with a combination of the first two parts through introducing a selection of alternative methods for swaption VaR estimation that are compared in the empirical part of the study.

### 2.1 Risk management and capital requirements

The purpose of a risk management function in any enterprise is to manage the risks that are necessary for conducting the core business of the enterprise. Hence, the purpose is not to minimize the risks as such, but to control the level of risks so that they are in line with the strategic objectives of the entity. While it can be argued, based on the corporate finance theory (Modigliani and Miller, 1958), that well-diversified investors do not gain any value from company-level hedging, it is still undertaken by most entities due to practical reasons such as minimizing taxes and costs of financial distress or due to managerial risk aversion (Smith and Stulz, 1985).

While non-financial companies engage in risk management mainly in order to maximize their enterprise value, financial companies are also subject to regulatory requirements to do so. The main reason behind such rules and regulations are the adverse effects that banking crises may have on the economy as a whole. The regulatory requirements have been developed gradually, and the Basel Committee on Bank Supervision published *The Accord* in 1988 that was an agreement between bank regulators on how much a bank is required to hold capital against credit losses. In 1996 the Basel Committee on Bank Supervision published an amended to *The Accord,* called *The Amendment,* to include additional minimum capital reserves for covering market risks as well. Accordingly, The Amendment distinguished between a bank's trading book and its banking book so the market risk was defined as the risk arising from fluctuations in the market prices of trading positions (Basel Committee on Banking Supervision, 1996a)

Furthermore, the 1996 Amendment included an Internal Models Approach (IMA) according to which a financial institution's capital requirements are based on the institution's internal risk measurement systems. In order to be allowed to use IMA as basis for capital requirement calculations, a bank is expected to fulfill certain qualitative as well as quantitative requirements. The qualitative requirements dictate that a bank should be able to demonstrate that it has a sound and sophisticated risk management system and an independent risk-control unit. Furthermore, the bank must also conduct regular stress test and external audits.

The IMA is based on a bank's Value at Risk (VaR) figure computed using the following inputs:

- a 99% confidence interval
- an observation period of at least one year

- a horizon of 10 trading days that can be derived from daily VaR by scaling it up using by the square root of ten (Basel Committee on Banking Supervision, 1996b)

The market risk charge is then set at the average VaR over the last 60 trading days times a multiplier *k*. Naturally, this reliance on the bank's self-reported VaR to determine capital requirements leads to an adverse selection problem as it creates an incentive to report unrealistically low VaR figures in order to minimize its capital requirements. To address this issue, the banking regulators evaluate the quality of a bank's VaR measurements by observing the frequency of its VaR exceptions based on backtest reports, and adjust the multiplier *k* accordingly to penalize banks with inferior measurement accuracy (Basel Committee on Banking Supervision, 1996c). Therefore, the financial institutions have an incentive to report their VaR figures truthfully due to threat of increased future capital requirements.

## 2.2 Introduction to Value at Risk

VaR was first developed by major financial institutions in the late 1980s in order to measure the market risk of their trading portfolios caused by fluctuations in asset values (Linsmeier & Pearson, 1996). Different VaR models have gained a central role in the risk management of financial institutions after JP Morgan publicly introduced their internal VaR model in 1994. The initial model was based on the variance-covariance of past security returns and the method can be traced back to the early days of Markowitz's Modern Portfolio Theory.

Taking risks is an inevitable part of conducting business, and firms face a variety of risks arising from their operations, financing activities and general economic, legal and regulatory environments. Moreover, risks related in finance can be further divided into liquidity risks, credit risks and market risks. Risk as a term is generally understood as "a threat of loss", whereas in financial theory it is defined as the dispersion of returns. As such, the measure of risk is the standard deviation of unexpected outcomes of financial assets that is also called *volatility,* or *sigma* ($\sigma$). Jorion (2007) separates the main sources of market risk broadly as interest rate risk, foreign exchange risk, equity price risk and commodity price risk. Losses that a firm may face depend both on its exposure to these different sources of market risk and on the level of the underlying volatility of the financial assets that the firm is holding. VaR was originally designed for measuring the market risks originating from the fluctuations in asset prices in a way that captures both a firm's exposure to the risks and the underlying volatility.

The main purpose of a VaR model is to measure the size of possible future losses of a portfolio at a given probability. Stated in a more formal way, VaR is defined as *the worst expected loss that a portfolio may suffer during a specified period under normal market conditions with a specified level of confidence*. For instance, if an institution's daily VaR is stated as 12 m€ with a 99% level of confidence, the probability of facing a loss that exceeds 12 m€ the next day should be 1%. Consequently, VaR provides a simple and easy to understand measure of a portfolio's downside risk.

Following Jorion (2007), in mathematical terms, for a portfolio whose value at the end of a period is given by

$$W = W_0(1+R) \qquad (1)$$

where $W_0$ is the initial portfolio value and $R$ the portfolio's rate of return, there is a distribution of future portfolio value $f(W)$ and the VaR for the portfolio is defined as

$$1 - c = \int_{-\infty}^{W*} f(W)dw \qquad (2)$$

where $c$ is the specified confidence level and $W*$ is the end of period portfolio value when the worst portfolio return with the given confidence level is realized.

Consequently, assuming that asset returns are normally distributed (ie. $(W) \sim N(0,1)$ ) VaR is calculated as follows:

$$VaR_c = \alpha * \sigma * W_0 \qquad (3)$$

where $\alpha$ is the normal deviate associated with the confidence level $(1-c)$ , $\sigma$ is the portfolio volatility and as before, $W_0$ is the initial value of the portfolio. It is important to remember to use consistent time horizons in estimating the figures for return and its volatility with respect to the VaR time period. For example, if a portfolio's value is €50 million and the annual volatility of its returns is 15%, a 1-day VaR at 95% confidence level would be as follows:

$$VaR_{95\%} = -1.65 * \sqrt{\frac{1}{250}} * 15\% * €50m = -€0.78m \tag{4}$$

The value of $\alpha$ can be read off from standard normal distribution tables and the annual volatility is converted into daily volatility by $\sigma_T = \sigma\sqrt{T}$, where $T$ is time horizon expressed in years.

As already stated above, VaR is a fairly simple and intuitive concept in theory, and that is one of the main reasons for its popularity among financial practitioners. However, its implementation in practice is hardly a straightforward process - at least not for portfolios that contain a large number and different kinds of securities - and measuring VaR is actually a demanding statistical problem. Following Dowd (2002), the different models used for estimating VaR can be divided into groups based on their approach to risk exposure and on how they define the distribution of risk factors. With respect to risk exposure, the models can be further divided into two groups: *local-valuation* methods and *full-valuation* methods. In local-valuation methods the risk is modeled using local derivatives to infer price movements, whereas in full-valuation methods the portfolio is fully repriced over a variety of scenarios. Local-valuation methods include delta- and delta-gamma-approximation methods. In the former only the first derivative is used and in the latter also the second order derivative is taken into account. Furthermore, different models for risk factors can be divided into *parametric* and *nonparametric* methods (Jorion, 2007). It should also be pointed out that the different methods yield somewhat differing results. Further, using same model gives naturally differing measures depending on the practical implementation of the model. The differences between the models arise from their approach to estimating the changes in value of the portfolio. Nevertheless, what the models have in common is that they all try to account for the empirical findings about financial markets that were first documented already a half a century ago by Mandelbrot (1963) and Fama (1965). These can broadly be summarized as follows:

- Financial return distributions are leptokurtic (ie. the returns have fatter tails than in the normal distribution)

- Equity returns are typically negatively skewed (ie. the left side of the distribution is longer than the right side)
- Volatility is typically clustered in time so that large changes in asset values are followed by large changes and vice versa.

An overview of the different VaR models is presented in the subsequent chapters of this section and the models employed in this study are covered more comprehensively in Section 4.

## 2.3 Parametric methods

As the name suggests, parametric models are based on parameterization of the behavior of financial instruments' price changes. Put more explicitly, these models require making an assumption about the statistical distribution of asset returns from which the data is drawn. Parametric approach can be perceived as fitting curve across the data and then reading off the VaR measure from that fitted curve. This is also the primary advantage of parametric models: computing requires relatively little information and so the practical implementation is less burdensome than with the other models. Furthermore, since parametric VaR figure is simply a multiple of the standard deviation of the distribution multiplied by an adjustment factor that depends on the confidence level and holding period length, normality enables simple rescaling of VaR figures for differing confidence levels and holding periods through changing the adjustment factor accordingly (Dowd, 1998). However, the problem with parametric models is that the chosen statistical distribution may not reflect accurately the actual distribution, which leads to either under- or overestimation of the actual risk. This is especially problematic for portfolios that contain options or other instruments whose pay-off is highly asymmetric as this adds to the skewness and the kurtosis of the distributions which again leads to more extreme price variations and, consequently, to increased probability of more extreme losses (Jorion, 2007).

### 2.3.1 Variance-covariance method

Variance-covariance approach[1] is one of the basic VaR computing methodologies in the class of parametric models. The key step in variance-covariance VaR method is the computation of the standard deviation of changes in portfolio value. The portfolio VaR is obtained multiplying the standard deviation by the normal deviate and risk factor weights as shown in the previous chapter. However, even if the basic idea is very simple, the practical implementation can become challenging as the standard deviation of portfolio depends both on the standard deviations of the portfolio's individual instruments and on the correlation between them. As a result, the total number of required parameters grows rapidly as the number of instruments increases (Linsmeier & Pearson, 1996).

## 2.4 Non-parametric methods

Even though parametric methods are attractive because of their theoretical simplicity, Barone-Adesi and Giannopoulos (2000) point out that the parametric methods have materially underestimated the size and frequency of substantial losses due to the fact that normal distribution fails to accurately describe the actual distribution of portfolio returns. Unlike the parametric models discussed above, non-parametric methods do not make any distributional assumption about portfolio returns.

### 2.4.1 Historical Simulation

Historical VaR is one of the most used and perhaps the easiest to apply within the class of non-parametric methods. It is computed using past returns of the portfolio’s present assets so that one obtains a distribution of price changes that would have realized had the current portfolio been held throughout the observation period. The most important advantage of this model is that it accounts for also the fat tails and skewness observed in return distributions, as Angelidis and Benos (2006) points out. Moreover, it can be applied for basically all types of financial instruments (Jorion, 2007).

[1] Also frequently referred to as the *delta-normal method*.

Nevertheless, even though historical VaR does not make explicit distributional assumptions, it still contains an implicit assumption that the distribution of returns stays unchanged within the historical estimation time window (Engle and Manganelli, 2001). This assumption leads to a few problems. First, if returns within the estimation window are assumed to have the same distribution, it means that all the returns of different time series have to be independent and identically distributed. The assumption of independency of returns implies that the magnitude of price movement in one period of time would not influence the price fluctuations that occur during subsequent time periods. Further, if the returns were identically distributed, or stationary, through time, it would imply that that the probability of a given loss was the same for each day. However, as already pointed out, this is empirically not true as volatility has a tendency to cluster so that large price fluctuations are followed by further large changes. In practice this entails that during periods with higher volatility one would also expect losses that exceed the usual level. Consequently, using a constant volatility model such as basic historical simulation could be misleading as it underestimates risk during highly volatile market conditions, which is documented by van den Gloorbergh and Vlaar (1999) and Vlaar (2000). Second, choosing a proper length for the time window is not a trivial task: if it’s too short, it is not possible to obtain statistically significant figures, and if it’s too long, the market fundamentals may have changed since the beginning of the period and observations from the past – with either too low or too high volatility – may dominate the VaR estimation yielding either excessively low or high VaR figures (Dowd, 2002). For instance, Hendricks (1995) finds that longer historical sample periods result in less variability in VaR estimates, but that they also result in absolutely larger VaR estimates.

### 2.4.2 Time-weighted Historical Simulation

One solution to the jumping volatility arising from historical observations being dropped out of the estimation window is to assign heavier weights to more recent observations as suggested by Boudoukh et al. (1998). This way, VaR estimates are more responsive to present market conditions. Moreover, as the impact of distant observations declines in time, there is no need to drop old data out of the estimation window, which tackles the problem of jumps resulting from old observations falling outside the estimation period (Dowd, 2002). A more detailed description of the time-weighted historical simulation is presented in 4.2.3.

### 2.4.3 Volatility Adjustment and Filtered Historical Simulation

Nonetheless, time weighing suffers from a few shortcomings as well. For instance, Hull and White (1998b) show that a sequence of large gains or losses can create substantial distortions in the risk profile of the sample. Instead of weighing returns based on when they occurred, Barone-Adesi et al. (1998) and Barone-Adesi et al. (1999) suggest assigning different weights to observations based on their volatility, which is known as *filtered historical simulation* (FHS). Hence, the idea behind FHS is to adopt the historical simulation method to the prevailing level of volatility observed in the market. In this method, actual returns within the historical dataset are replaced with returns adjusted by forecast of volatility. As a result from incorporating the information from volatility forecast, the model generates estimates that are sensitive to current volatility and better captures the nature of current market conditions (Dowd, 2002). Accordingly, one of the most important advantages of FHS over basic HS is that the volatility filtering process increases the range of possible risk factor outcomes beyond the unadjusted historical record through change of scale. Thus, FHS effectively supplements the tails of the return distribution through generating extreme events that are not present in the historical record. This shortens the length of the historical period required for collecting return observations used in the simulation process compared to HS method. Furthermore, the capability of FHS to better adjust to the prevailing market conditions is also supported by Angelidis and Benos (2006) who propose that the FHS outperforms parametric and other non-parametric methods at higher confidence levels. The practical implementation of FHS in this study is further clarified in 4.2.2

### 2.4.4 Monte Carlo simulation

Monte Carlo (MC) simulation is a highly flexible method for computing VaR as it is capable of simultaneously accounting for various risk sources and it can deal with time variation in volatility and nonlinear price exposure arising from complex pricing models (Jorion, 2007). Consequently, as pointed out by Ammann and Reich (2001), combination of MC simulation and full valuation yields most accurate VaR results for portfolios with substantial option positions.

Implementation of the MC method begins with identifying the important market factors and assigning suitable stochastic processes for these factors. Then a future distribution of portfolio returns is created through simulation of price paths for the instruments, and the different confidence level VaR figures are drawn from this distribution (Wiener, 1999).

Even though MC model uses parametric inputs, such as volatility in Geometric Brownian Motion that is used for describing the dynamics of stochastic price process, the future distribution cannot be described by an analytical function and thus the model can be interpreted as a non-parametric method. Hence, for instance Dowd (2002) categorizes MC simulation as a "semi-parametric" method[2].

Despite its virtues, MC simulation has also attracted criticism. For example, Barone-Adesi et al. (2002) point out that the model's multivariate properties of the risk factors are based on historical correlations and the correlations tend to increase rapidly during crises, which may lead to underestimation of risk. Moreover, simulation methods require substantial computing capacity and are hence time consuming. This problem, however, is gradually mitigated as the computing capacity as well as the efficiency of simulation methods is evolving constantly.

## 2.5 VaR criticism

Despite all the positive attributes of VaR measure that explain its popularity, the model also has its weaknesses. For the sake of providing a comprehensive perspective on VaR, also some of the model’s shortcomings are discussed in the following chapter.

Artzner et al. (1998) have proposed a list of desirable properties that measures of risk should have in order to be considered as “coherent” risk measures. These include

- Monotonicity: if a portfolio A yields in every possible situation better returns than portfolio B, then portfolio B should be assigned with a higher risk
- Sub-additivity: the combined risk of two portfolios cannot be higher than the sum of the separate risks
- Positive homogeneity: if the size of the portfolio is doubled, the risk should double as well
- Relevance: the risk of holding no assets is zero

[2] Dowd also classifies weighted historical simulation methods under the term semi-parametric methods, as for instance, GARCH model that is utilized in the filtered historical method, is parametric.

It is a known fact that VaR fails to meet the requirement of sub-additivity, which means that using VaR might discourage diversification. Moreover, it could possibly lead to regulatory arbitrage in the sense that if the capital requirements of an institution depend on its VaR figure, by splitting its assets into separate subsidiaries a company would be able to appear less risky than it actually is.

Furthermore, while the conceptual simplicity is perhaps the main reason why VaR has become such a widespread method for risk measurement, it is also one of its fundamental shortcomings. As all available information is condensed into a single easy-to-digest figure it is evident that some relevant information will be lost. For instance, two positions with different risk characteristics beyond the VaR confidence level can still have the same VaR figure. This is due to the property of VaR that it provides no information regarding the losses that exceed the VaR estimate, and why it is often said that VaR fails to account for the "tail risk". Consequently, VaR figures solely do not provide sufficient estimate of the risks that an entity faces. A point in case is the $2 billion mark-to-market loss suffered by JPMorgan Chase's Chief Investment Office in May 2012, while its daily average VaR in the first quarter of 2012 was reported to be $67 million. Indeed, this incident has fuelled the debate about the reliability of VaR as a risk measure especially when it was JPMorgan that originally developed the measurement concept.

However, it is possible to partially overcome this shortcoming by using a so called "conditional VaR" method that measures the expected loss given that the VaR is exceeded. This method is also known as "expected shortfall" (ES) or "expected tail loss" (ETL) and it is gradually gaining more popularity. It should be pointed out that while ES is based on value at risk method, it is a coherent measure of risk while VaR is not. Also, it is expected that ES will take VaR's position as a regulatory measure in the future as the Basel Committee on Banking Suprevision has recently stated that under the prospective Basel III the market risk capital requirements would be based on ES rather than VaR measurements. However, even though ES is a coherent risk measure and hence theoretically better than VaR, it still has its limitations. For instance, the challenges in its implementation exceed those of VaR, and if the calculation method used in ES is the same as in VaR, i.e. based on boot-strapping data from the past 250 days, it does not make a significant difference which method is used.

One further point of general VaR criticism has been the model's intrinsic feature of merely considering the loss at the end of the estimation period, which, as for example Boudoukh et al. (2004) and Kritzman and Rich (2002) point out, becomes a problem with longer estimation horizons. For instance, certain investors, such as insurance companies and money managers, are interested not only in their long-horizon VaR but also in what happens in the interim: deterioration in asset prices might force them to unwind their positions already before the VaR horizon and hence the actual losses could become substantially worse than predicted by VaR. Kritzman and Rich (2002) propose using "continuous VaR" in which the normal end-of-horizon probability of loss is transformed into intra-horizon path-dependent loss.

Furthermore, all VaR methods are at least partially dependent on historical data, and as is well known, history does not predict future very well. All in all, regardless of how VaR is computed, it is far from being a perfect tool for risk measurement. As a result, other risk management techniques are required in addition to VaR estimates. These include stress tests and scenario analysis together with various sensitivity analyses with respect to different risk factors.

## 2.6 VaR backtesting

The usefulness of a VaR model for generating risk estimates is heavily dependent on the model's ability to accurately predict future losses. The precision of a VaR model can and should be backtested by comparing actual losses to corresponding VaR estimates. However, there exist a few different viewpoints that can be taken into account when evaluating goodness of a VaR model. When determining whether the model in question is accurate or not, some kind of a definition for accuracy is needed. For instance, accuracy could refer to the ability of the model to measure a particular percentile of the profit and loss distribution, or it could mean the model's capability of predicting the size and frequency of portfolio losses. For that reason there is no one single test that provides a correct answer. The purpose of this chapter is to provide a brief overall view over commonly used backtesting methods, while a more detailed description of the statistical framework of backtesting and the specific backtests applied in this study is in chapter 4.4.

As Christoffersen (1998) points out, the evaluation of a VaR model's accuracy can be reduced into studying the *unconditional* and *conditional* coverage properties of the exception sequence generated by the model. Hence, most backtesting methods can be divided into tests of unconditional coverage and into tests of conditional coverage.

Tests of unconditional coverage measure the frequency of VaR exceptions over a specified time period. In short, these tests compare the actual failure rate with the model's theoretical failure rate. For instance, when using a 95% confidence level for daily VaR computing, one should expect to face losses greater than the model has predicted five times during 100 trading days on average. Hence, even the estimates generated by a sound VaR model are breached occasionally but it is the number of those exceptions, or violations, that counts. Consequently, the most obvious determinant of a model's validity is the number of occasions when the actual loss for the observed period exceeds the model's respective forecast.

While tests of unconditional coverage mainly focus on the number of VaR violations, the tests of conditional coverage account also for the *time variation of the occurred exceptions*. The reason behind this is that a sound VaR model is expected to generate an acceptable number of exceptions that are also evenly distributed in time. If a model generates an acceptable number of exceptions during a given backtesting period, the model could still be deemed deficient in case the exceptions suffer from clustering, which could be a sign of the model's poor ability to capture changes in market volatility and correlations.

In addition to conditional and unconditional tests, it is also possible to utilize the information provided by the size of the exception through applying a loss function that penalizes a model that has provided a worse estimate of the loss given that the VaR figure is estimated. Consequently, an expected shortfall figure is needed to use loss function based evaluation, and hence, the test provides indirect insight about the quality of a VaR model through studying the tail of the distribution used in the given model rather than the hit sequence the model generates.

### 2.6.1 Backtesting in the regulatory framework

Since the market risks of banks' trading books are subject to minimum regulatory capital requirements, supervisors are also interested in the risk figures reported by the banks. Under certain conditions, it is possible that the banks' internal risk measurement models are accepted as a tool for measuring the capital requirements. Naturally, the soundness of the internal model suggested by a given institution has to be validated by supervisors before it can be applied as the basis for capital requirements. For this purpose, the Basel Committee has chosen a fairly simple test of unconditional coverage based only on the number of exceptions during the last 250 trading days as the official method for validating banks' internal VaR models.

The regulatory backtesting procedure is aligned with the capital requirement ratio calculation so that it is also implemented using 99% VaR confidence level. However, the Committee has allowed the use of 1-day estimation horizon in backtesting although the capital requirements are based on 10-day VaR horizon (Basel Committee on Banking Supervision, 1996b). Also, the result of the backtest has an effect on the Internal Models Approach based capital charge $c$ that is calculated according to the following formula:

$$c = \max\left\{VaR_{t-1}; m_c * VaR_{avg}\right\} + max\left\{sVaR_{t-1}; m_s * sVaR_{avg}\right\} \qquad (5)$$

where $VaR_{t-1}$ is a bank's previous day's VaR figure measured according to the parameters specified in above, $VaR_{avg}$ is the average of daily VaR figures on each of the preceding 60 trading days, and $sVaR_{t-1}$ and $sVaR_{avg}$ are stressed VaR measures for the previous day and average of the previous 60 business days respectively. The stressed VaR metric is similar to the normal VaR, but it is simulated with using risk factor changes that occurred during a continuous 12-month period of significant financial stress relevant to the institution's portfolio. The multiplication factors $m_c$ and $m_s$ have a minimum value of 3 to which a "plus" factor is added. The plus values range from 0 to 1 and they are set by supervisory authorities based on their assessment of the quality of the bank's risk management. Moreover, the values of the plus factors are linked to the backtesting results, which creates an incentive for banks to develop the quality of their models (Basel Committee on Banking Supervision, 2011).

The regulatory backtesting results are described by a three-zone approach that is also known as the "traffic-light" approach as the test results are classified into *green, yellow* and *red zones* based on the number of VaR exceptions that the model generates during a backtesting period of one year. If the backtesting results fall into the green zone, the results are not deemed to suggest that there were problems with the quality or accuracy of the given model. The yellow zone indicates that the model's quality or accuracy could be questioned but no definitive conclusions can be directly drawn from the results alone. For instance, a model's results could gain a yellow classification based on bad luck even if the model was "fundamentally sound". Hence, while the yellow zone generally results into heightened capital charge ratio, the Committee points out that the supervisor may consider revising the requirement based on the bank's further demonstrations about the model's quality. Finally, a backtesting result that falls into the red zone can be interpreted as sign that there are severe problems with the model: as Table 1 shows, the probability of an accurate model producing ten or more exceptions at 99% confidence level is microscopic. Consequently, red zone classification should lead to an almost automatic rejection of the model (Basel Committee on Banking Supervision, 2006).

**Table 1:**
**Three zone approach (Basel Committee on Banking Supervision)**

The backtesting result categories with 250 observations and 99% VaR confidence level. The cumulative probability in the right shows the probability of obtaining a given number or fewer exceptions when the model is accurate.

| Zone | Number of exceptions | Increase in scaling factor | Cumulative probability |
|---|---|---|---|
| Green | 0 | 0.00 | 8.11 % |
| | 1 | 0.00 | 28.58 % |
| | 2 | 0.00 | 54.32 % |
| | 3 | 0.00 | 75.81 % |
| | 4 | 0.00 | 89.22 % |
| Yellow | 5 | 0.40 | 95.88 % |
| | 6 | 0.50 | 98.63 % |
| | 7 | 0.65 | 99.60 % |
| | 8 | 0.75 | 99.89 % |
| | 9 | 0.85 | 99.97 % |
| Red | 10 ≤ | 1.00 | 99.99 % |

## 2.7 Interest Rate Swaps and Swaptions

This chapter lays down the foundations on which another central part of the thesis is built on through first introducing interest rate basics and then continuing with plain vanilla interest rate swaps and swaptions. The interest rate swaps and swaptions are covered with a practical viewpoint together with a short introduction on how they are treated in reality. The valuation of interest rate swaps and swaptions is covered subsequently.

### 2.7.1 Interest rates and interest rate swap essentials

Plain vanilla interest rate swaps (IRS) is one of the most actively traded interest rate instruments and thus also one of the most commonly used financial instruments in general. An interest rate swap is a contractual agreement between two parties to exchange fixed interest rate payments for floating rate payments on a specified notional during a defined time interval. Consequently, an IRS consists of two components: a *floating leg* and a *fixed leg*. Also, each party's position in the swap contract is named relative to the fixed leg so that the party paying the fixed rate has entered into a *payer swap* and the party that pays the floating rate had entered into a *receiver swap.* In addition, terms *buyer* and *seller* are also used. The swap buyer buys the floating leg for a fixed price and is thus the fixed leg payer.

The floating leg payments are tied to a reference rate, usually to an Ibor (InterBank Offered Rate) rate. There are a few different Ibor rates that are fixed by different entities and the fixing entity is differentiated by the prefix. For instance, Libor refers to *London InterBank Offered* rate that is fixed by the *British Banker's Association* and Euribor fixings are determined by the *European Banking Federation.*

Ibor rates are quoted using the money market convention, which means that the interest paid is calculated as $\delta L N$ where $\delta$ is the interest rate accrual period year fraction, or *coverage*, $L$ is the reference Ibor rate and $N$ refers to the loan notional. Moreover, the loan coverage depends on the market's day count convention that determines the exact length of the accrual period. Market conventions, and hence also day count conventions are slightly different in different currencies. The market conventions for plain vanilla interest rates in some of the main currencies are provided in the Table 1. The day count conventions in the table are as follows:

- Actual/360 (ACT/360). With this convention the length of a year is 360 days. Hence, the year fraction between two dates is the actual number of days between them divided by 360:
$$\frac{D_2 - D_1}{360}.$$
- Actual/365 (ACT/365). The same as ACT/360 but a year is assumed 365 days long.
- 30/360. In this convention each month is assumed to be 30 days long leading to a 360-day long year. Thus, the year fraction between dates $d_1$ and $d_2$ is computed in the following way:

$$\frac{\max(30 - d_1, 0) + \min(d_2, 30) + 360 * (y_2 - y_1) + 30 * (m_2 - m_1 - 1)}{360}$$

where $m_i$ refers to month and $y_i$ to year of $d_i$.

Furthermore, there exist different adjustments to the day count conventions regarding the treatment of holidays.

The floating leg interest rate is typically reset and paid semi-annually or four times a year depending on the currency in which the contract is denominated. The fixed rate (or swap rate) is usually paid on an annual or semi-annual basis and it is set to a level that initially makes the value of the swap worth zero so that the present value of the fixed leg equals the present value of the floating leg. One should note, however, that although the present value (PV) of the swap is zero it does not mean that it did not have any value: as soon as the interest rates change, one of the legs becomes more valuable than the other, which leads to either mark-to-market gain or loss. Moreover, the floating rate is typically set in advance (i.e. a few days before each accrual period) and paid in arrears (at the end of each accrual period). The number of days between the interest rate swap trade date and the first fixing period start date is called the spot lag. Index spot lag also determines the lag between interest rate reset date and accrual period starting date for the floating leg rates. The length of spot lag also depends on the market convention.

**Table 2:**
**Conventions for plain vanilla IRS contracts**

| | | Fixed leg | | Floating leg | | |
|---|---|---|---|---|---|---|
| Currency | Spot Lag | Period | Convention | Reference | Period | Convention |
| USD | 2 | 6M | 30/360 | Libor | 3M | ACT/360 |
| EUR: 1Y | 2 | 1Y | 30/360 | Euribor | 3M | ACT/360 |
| EUR: >1Y | 2 | 1Y | 30/360 | Euribor | 6M | ACT/360 |
| GBP: 1Y | 0 | 1Y | ACT/365 | Libor | 3M | ACT/365 |
| GBP: >1Y | 0 | 6M | ACT/365 | Libor | 6M | ACT/365 |
| JPY | 2 | 6M | ACT/365 | Tibor | 3M | ACT/365 |
| JPY | 2 | 6M | ACT/365 | Libor | 6M | ACT/360 |
| CHF: 1Y | 2 | 1Y | 30/360 | Libor | 3M | ACT/360 |
| CHF: >1Y | 2 | 1Y | 30/360 | Libor | 6M | ACT/360 |

In addition to plain vanilla IRS contracts, there are also other types of swaps as well. For instance, a swap whose fixing date for floating rate payment is the index spot lag before the period end date is called an in-arrears swap. Additionally, the fixed rate does not have to be the same for each coupon: step-up swaps have an increasing rate and step-down decreasing fixed rates. Also the notional of the swap can vary between coupons. In case the notional is decreasing through time, the swap is called amortized swap, whereas a swap with increasing notional is referred to as an accruing swap. Furthermore, if the notional first increases and then decreases towards the end of the contract, it is called a roller coaster.

Even though swap principal is not usually exchanged at the end of the swap, for the sake of intuition it often helps to think that there is a mutual exchange of one euro at the end of the swap. As Longstaff et al. (2000) point out, from this perspective, the cash flows from the fixed leg equal to cash flows from a bond whose coupon rate is the swap rate, whereas the floating leg cash flows equal to those of a floating rate note. Accordingly, a swap can be thought as exchanging a fixed rate coupon to a floating rate note. Also, one can think of a swap to be a series of consecutive forward rate agreements over the swap period.

### 2.7.2 Swaptions

As the name suggests, a swaption, is an option that grants its owner the right but not the obligation to enter into an underlying swap at a specified future time, that is at the swaption maturity. Normally the swaption maturity matches with the first reset date of the underlying swap. The length of the underlying swap contract is called the tenor of the swaption. While it is possible to trade options on a variety of swaps, within the context of this study swaptions refer to options on interest rate swaps. Swaption contracts are divided into two types subject to the direction of the fixed leg of the contract: *payer swaptions* and *receiver swaptions*. A payer swaption grants the owner of the option a right to enter into a swap where she pays the fixed leg and receives the floating leg. Respectively, a receiver option gives its owner a right to enter into a swap where she receives the fixed leg and pays the floating leg. Analogous to other option positions, the party who has bought an option is said to have a long position, while the seller has a short position. For instance, a 4% 5x10 ("5 into 10") receiver swaption gives the holder the right to receive 4% on a 10 year swap starting in 5 years. Accordingly, a payer swaption can be perceived as a call on paying fixed swap and a receiver swaption as a call on receiving fixed swap. Swaption thus allows its holder to benefit from favorable interest rate development while providing protection against unfavorable movements. Moreover, swaptions, like other options, can either be European, American or Bermudan with respect to their exercise style. Swaptions can also be grouped into *physically settled* and *cash-settled* contracts. In the former, the contract parties enter into the underlying swap if the swaption is in the money from the buyer's viewpoint at the maturity, while in the latter, merely the present value of the underlying swap is exchanged at the maturity if it is positive for the buyer. In the European market, the most actively traded swaptions are cash-settled (Mercurio, 2008).

As swaps and swaptions resemble more contractual obligations than securities, a substantial legal infrastructure is required for functional markets. Hence, primary and secondary markets for swaps and swaptions are made by a network of swap dealers and most of them are members of the International Swaps and Derivatives Association (ISDA). ISDA is an independent organization that has developed standards and contractual terms for swap markets. These standards dictate, for instance, what happens in the event of default by either side. Market for swaptions has grown together with interest rate swap markets: due to the enormous growth and size of the swap market – approximately $300 trillion as of April 20, 2012 according to the Interest Rate Trade Repository Report published by ISDA – swaptions have become one of the most important fixed income derivative instruments together with interest rate caps and floors. Typical participants in the swaption market are banks and financial institutions, corporations and funds that wish either to hedge their interest rate exposure or to speculate on interest rate fluctuations.

As a summary, in order to specify a swaption contract, the following properties have to be indicated:

- the maturity of the option
- the strike rate (or swap rate, i.e. the fixed rate on the underlying swap)
- the tenor of the underlying swap
- notional amount
- option exercise style (European, American or Bermudan)
- settlement style (physical or cash)

### 2.7.3 Swap pricing

After introducing the basics of swap and swaption contracts, it is possible to advance to their valuation. Furthermore, in order to value a swaption contract one must first determine an appropriate forward rate for the underlying swap contract to be used as the strike rate for the swaption, which again is based on basic IRS valuation explained in this chapter.

Traditionally swap rates were valued using a single forward curve for both the projection of the future interest rates and for discounting the interest payments. However, after the credit crisis this is not a valid method anymore as the swap curves with longer tenors have now an implicit risk element built into them arising from the incorporation of counterparty credit risk, whereas the discount factor in derivatives pricing is based on risk free rates. Consequently, it has become a market standard to apply so called *dual curve* method in pricing interest rate instruments where the reference rate for floating leg payments is still projected from Ibor curve and another interest rate curve is used for determining appropriate discount factors. Typically, overnight swap-rates, such as EONIA (Euro OverNight Index Average), are used for discounting (Bianchetti, 2009). The use of overnight rates as risk-free rate can be justified by the fact that typically interbank operations are collateralized and the collateral is assumed to be revalued daily. Hence, overnight rates can be deemed as close to risk-free as possible, whereas the forward rates derived from interbank reference rates cannot. Indeed, Pallavicini and Tarenghi (2010) show that market quotes of interest rate swaps are coherent with the use of EONIA-based discounting curve. However, the authors also find that swaptions, on the other hand, were still consistent with the traditional "text-book" type of pricing in which the same curve is used for both projection and discounting.

Since a swap can be characterized as a portfolio of forward rate agreements, it is possible to value one with using the no-arbitrage principle by assuming that market-observable forward interest rates are realized (Hull, 2008). Thus, the valuation procedure is as follows:

- Use the Euribor or swap zero curve to calculate forward rates for each of the Euribor rates that will determine the cash flows of the swap's floating leg
- Calculate swap cash flows assuming that the Euribor rates will equal the forward rates calculated at present time
- Discount the swap cash flows using EONIA zero curve rates to obtain the present value of the swap

The fixed rate is chosen so that the value of the swap is initially zero. This further implies that the sum of the underlying forward rate agreements is zero. Nevertheless, it does not mean that the value of each individual FRA was zero: some of the FRAs will have positive values and others have negative values and the values of the FRAs at different future points depend on the prevailing term structure of interest rates.

As an illustration how the fixed rate is determined, the value of an IRS to the fixed rate payer at time *t* is determined in the following way:

$$IRS(t,K) = \sum_{k=a+1}^{b} \tau_k P_D(t,T_k) L_k(t) - K \sum_{j=c+1}^{d} \tau_j^S P_D\left(t,T_j^S\right) \qquad (6)$$

where

$\tau_k$ denotes the year fraction between $T_{k-1}$ and $T_k$

$P_D(t,T_k)$ is the value of a zero-coupon bond at *t* with maturity $T_k$, i.e. the discount factor for $T_k$ starting at *t*

$L_k(t)$ is the floating rate at *k* observed at *t*, that is, the forward rate between $T_{k-1}$ and $T_k$

*K* is the fixed rate

$a$ and $b$ are starting time instant and number of floating rate payments for the floating rate leg of the contract and

$c$ and $d$ are the starting time instant and number of payments for the fixed rate leg.

Therefore, the first term is the present value of the floating rate payments and the second term is the present value of the fixed rate payments. As a result, the corresponding forward swap rate that makes the swap value equal to zero at time *t* can be calculated, and for time *t=0,* it is:

$$S_{0,b,0,d} = \frac{\sum_{k=1}^{b} \tau_k P_D(0,T_k) L_k(0)}{\sum_{j=1}^{d} \tau_j^S P_D\left(0,T_j^S\right)} \qquad (7)$$

where $L_1(0)$ is the first floating payment known at time *t = 0.*

### 2.7.4 Swaption pricing

As generally in option markets, swaptions are quoted in terms of implied volatility relative to a standard pricing model. For European swaptions, the quoted implied volatilities are relative to the Black (1976) model. The Black's model was originally developed for valuing options on commodities based on the idea that it would be reasonable to model the forward prices with a geometric Brownian motion even if it would not be a viable method for modeling the spot prices. Similarly, the model can be applied to forward swap rates when adjusted with an annuity factor that incorporates the time structure of the underlying swap into the formula.

The Black-like formula states that the price of a European receiver swaption (RS) starting at $t$=0 with payment dates T, i.e. with maturity $a$ and tenor $b$, is calculated as follows:

$$RS\big(0,\mathrm{T},\tau,N,K,\sigma_{a,b}\big) = N\mathrm{BL}\big(K,S_{a,b}(0),\sigma_{a,b}\sqrt{T_a},-1\big)\sum_{i=a+1}^{\beta}\tau_i P(0,T_i) \qquad (8)$$

Where

$$\mathrm{BL}\big(K,S_{a,b}(0),v,\omega\big) = S_{a,b}(0)\omega\Phi\left(\omega d_1\big(K,S_{a,b}(0),v\big)\right) - K\omega\Phi\left(\omega d_2\big(K,S_{a,b}(0),v\big)\right) \qquad (9)$$

N = swap notional

$$d_1(K,S_{a,b}(0),v) = \frac{\ln\big(S_{a,b}(0)/K\big)+v^2/2}{v} \qquad (10)$$

$$d_2\big(K,S_{a,b}(0),v\big) = \frac{\ln\big(S_{a,b}(0)/K\big)-v^2/2}{v} \qquad (11)$$

$$v = \sigma_{a,b}\sqrt{T_{i-1}} \qquad (12)$$

$S_{a,b}(0)$=current forward swap rate
t = option starting date
T = option expiry date
$\sigma_{a,b}$= volatility for F that is retrieved from market quotes
$\Phi(.)$ = the cumulative normal distribution function

Similarly, the formula for pricing a European payer swaption (PS) at $t$=0 is

$$PS\big(0,\mathrm{T},\tau,N,K,\sigma_{a,b}\big) = N\mathrm{BL}\big(K,S_{a,b}(0),\sigma_{a,b}\sqrt{T_a},1\big)\sum_{i=a+1}^{\beta}\tau_i P(0,T_i) \qquad (13)$$

The formulae above are used for physically settled swaptions. For cash-settled swaptions, the formula is the same otherwise except for the annuity term that is slightly different. The reason behind the adjusted annuity term is that the counterparties could end up with different prices for the underlying swap based on how they have bootstrapped their discounting curves. To prevent this, the swap rate is used for discounting as well. This way the amount due at swaption maturity will be unambiguously determined.

### 2.7.5 Implied volatilities and volatility models

When pricing any option, an important piece of information is some kind of measure of the uncertainty related to the future price of the underlying asset. Thus, a pricing method must include some a priori expectation regarding the asset price performance as well as an approximation of to what extent the price process will fluctuate. This expectation is denoted by the volatility factor used in the formula. However, it is not an easy task to determine an estimate of volatility as the results tend to depend heavily on the length of the time period from which the historical volatility is observed. Moreover, the magnitudes of asset price fluctuations have a tendency to change over time, which means that the volatility has a stochastic nature itself, as for instance Ball and Torous (1999) point out in their study about short-term interest rates. In spite of the fact that volatility is not constant, it is still assumed to be so in the Black-Scholes model. Despite this shortcoming, the Black-Scholes model has remained as a standard model in option pricing as it has the positive feature of providing a one-to-one relationship between volatility and monetary price of an option. As a result, it is possible to invert the Black-Scholes formula for the theoretical value of an option against the observed market price of that option, which leads to so called implied volatility[3]. For the vis-à-vis link between option price and implied volatility, it has become a market practice to quote option prices in terms of implied volatility.

Analogously to other options, also the implied volatilities of swaptions exhibit a so called "smile" or "skew" with respect to their strike rate. The reason behind the volatility smile arises from the above stated fact that the volatility of the underlying security is greater than assumed by the Black-Scholes or Black-like pricing models that are based on the assumption

[3] The terms volatility and implied volatility are used interchangeably in this paper when associated with the option price.

that the underlying instruments follow a log-normal process and that the returns are normally distributed. Hence, the actual return distributions have fatter tails, and consequently, the real probability of an option that is far out of the money to end up being in the money is higher than the probability predicted by the pricing model. As a result, options that are far out of the money are more valuable than they should be if the model's assumptions about return distributions were true. As the prices are quoted in implied volatilities, the seemingly higher implied volatilities for out of the money options reflect the relatively higher value of those options. Moreover, as the swaptions are quoted using implied Black volatilities that are based on a log-normal distribution of underlying interest rates, the level of implied volatilities tends to be inversely related to the level of interest rates if it is assumed that the changes in interest rates are more normally than log-normally distributed in reality. Hence, if the magnitude of moves in interest rates does not depend on their absolute level, the changes increase in relative terms when interest rates decrease. According to Corb (2012), this provides an explanation to higher observed implied volatilities for swaptions with lower strikes in addition to possible imbalances in supply and demand.

In addition to volatility skew in strike dimension, the implied volatility depends also on the maturity of the option. This is called the term structure of volatility and it is driven by the market's expectation about implied impact of upcoming events. When the implied volatility is plotted as a function of both strike price and time to maturity, it is possible to graph the volatility surface for a given underlying instrument.

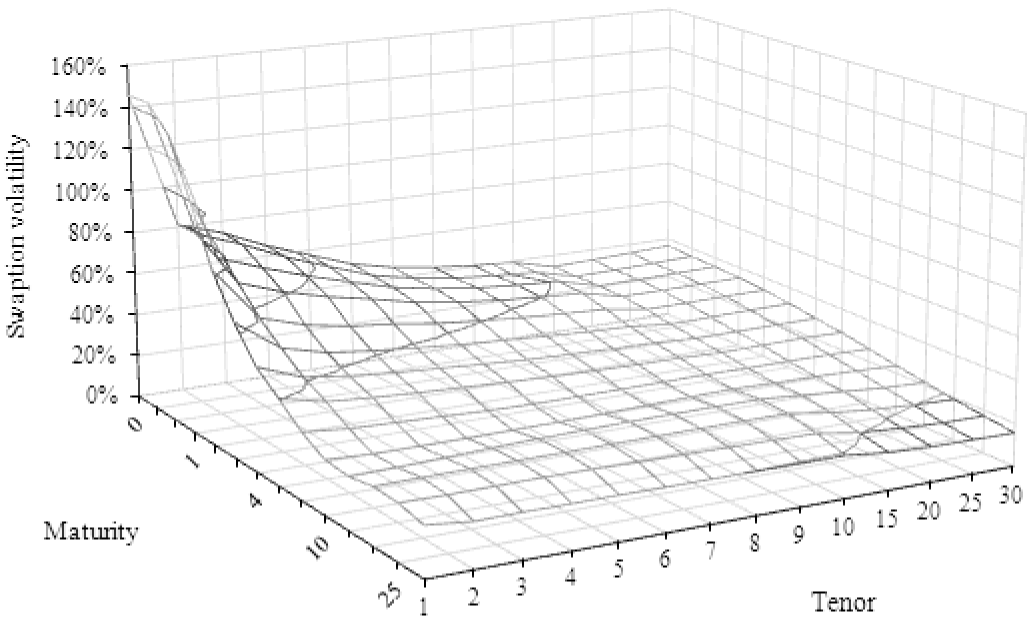
Brokers provide tables of implied volatilities for European swaptions with different maturity-tenor pairs. Swaption volatilities are quoted by the different strikes $K$ as a difference with respect to the at-the-money level. That is,

$$\Delta\sigma_{a,b}(\Delta K) = \sigma_{a,b}(K^{ATM} + \Delta K) - \sigma_{a,b}^{ATM} \tag{14}$$

and $\Delta K \in \{\pm 200, \pm 100, \pm 50, \pm 25, 0\}$, where $\Delta K$ are stated in basis points.

**Figure 1:**
**Swaption volatility surface**

The figure illustrates at-the-money Euro swaption volatility surface on September 27, 2012




Regardless of these well-known issues associated with the basic assumptions of Black-Scholes and other pricing models based on it, the models can be used for option and thus for swaption pricing with the idea of "plugging a wrong number into a wrong formula to get a correct price." As a result, Black-Scholes and its versions can be argued to be merely sophisticated interpolation tools used by traders so that an option is priced consistently with the market prices of actively traded options (Hull, 2008). Moreover, implied volatilities of interest rates are often interpolated using sophisticated modeling techniques and the Black-model is used only for translating the implied volatility figure into a monetary value. Hence, the models used for modeling the volatility play an important role: in addition to accurate pricing of options, it is also critical for hedging to have a model that correctly handles the market skews and smiles of the implied volatilities.

The models used for describing volatility smiles and skews can be divided into local volatility and stochastic volatility models. In the former, volatility is treated as a function of the current underlying asset level and time, whereas in the latter also the volatility process itself has a volatility of its own. The local volatility models, most of which are based on Dupire's (1994) work, are self-consistent, arbitrage free and can be calibrated to fit market skews and smiles exactly. However, as observed by Hagan et al. (2002), the local volatility models fail to capture the actual dynamics of the implied volatility curve: local volatility models predict that the market skew moves in the opposite direction as the underlying asset, which is contrary to the actual behavior observed in the markets, in which smiles and skews tend to move in the same direction with the underlying. This leads to poor hedging performance, and as a solution, the authors introduce a "stochastic αβρ" model, or *the SABR model*. The SABR model can be used to accurately fit the implied volatility curves observed in the marketplace, and as it also predicts the correct dynamics of the implied volatility curves, the SABR model is also an efficient means to manage smile risk. As a result, it has become widely used in the swaption and caplet/floorlet markets. For that reason, it is also employed in this study. Hagan et al. (2002) derive the following approximation for the implied volatility of the swaption with maturity $T_a$, strike $K$ and underlying forward swap rate $S_{a,b}(t)$ at time $t$:

$$\sigma\left(K, S_{a,b}(0)\right) \approx \frac{\alpha}{(S_{a,b}(0))K^{\frac{1-\beta}{2}}[1 + \frac{(1-\beta)^2}{24}\ln^2(\frac{S_{a,b}(0)}{K}) + \frac{(1-\beta)^4}{1920}\ln^4(\frac{S_{a,b}(0)}{K})]}\frac{z}{x(z)}$$

$$\cdot\left\{1 + \left[\frac{(1-\beta)^2\alpha^2}{24\left(S_{a,b}(0)K\right)^{1-\beta}} + \frac{\rho\beta\nu\alpha}{4\left(S_{a,b}(0)K\right)^{\frac{1-\beta}{2}}} + \nu^2\frac{2-3\rho^2}{24}\right]T_a\right\}, \qquad (15)$$

where

$$z = \frac{\nu}{\alpha}\left(S_{a,b}(0)K\right)^{\frac{1-\beta}{2}}\ln(\frac{S_{a,b}(0)}{K}) \qquad (16)$$

and

$$x(z) = ln\left\{\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}\right\} \qquad (17)$$

## 2.8 Swaption risk measurement

Swaptions, as any other traded instruments, have exposure to a range of risk factors that constitute the total market risk of the instrument in question. Generally, an option's sensitivity with respect to different risk factors is described by partial derivatives of the option price with respect to those different factors. The different derivatives are often called option *Greeks*.

*Delta* and *gamma* risks describe an option's exposure to the changes in the price of the underlying asset: delta is the first order derivative and gamma is of the second order. Specifically, in the case of a swaption, delta describes a swaption's sensitivity to the changes in the corresponding swap rate. Further, in addition to delta and gamma risks that arise from the interest rate fluctuations, the price of a swaption is subject to changes in implied volatilities as well. The volatility of an option's implied volatility is called the *vol-of-vol*, and the sensitivity of an option price to fluctuations in implied volatility is called the *vega*. The vega exposure of a swaption position can be defined as the change in the swaption price resulting from an increase in the swaption implied volatility by one percentage point. Accordingly, the risk with respect to changes in option price due to changes in implied volatility is called *vega risk.* Vega risk is an important risk component in an option portfolio and it should be of interest for a risk manager overseeing a position that includes options: neglecting the effect of changes in implied volatilities entirely would provide an exceptionally optimistic picture about the risk level of a given swaption position, which is something any risk manager tries to avoid. However, the treatment of vega risk in portfolios is often impeded due to practical reasons. First, the availability of option implied volatilities is somewhat limited, and second, the prevalence of volatility smiles and term structures add to the complicatedness of nesting vega risk into a risk management framework. Yet, in certain option positions, the level of vega risk may well exceed the level of delta risk, which underlines the need for an accurate risk management model that is able to incorporate also the vega risk. For instance, in the following example swaption trading positions and strategies the proportion of vega risk is relatively high compared to the level of delta risk:

- a *delta-hedged position* in which a swaption is hedged with the underlying forward swap contract so that the book is insulated against short-term interest rate changes but the option position's sensitivity to general level of option prices remains, which acts as the source of the vega risk of the book
- a *straddle* that consists of a long (short) call and a long (short) put with equal exercise rates and
- a *strangle* that consists of a long (short) call and a long (short) put with different exercise rates.

### 2.8.1 Swaption VaR

When computing VaR for swaptions, the relevant risk factors are interest rates and implied volatilities. Hence, in the Historical VaR framework, the changes in both of those risk factors within a chosen historical period are used for simulating possible outcomes for the position. Then, depending on the selected confidence interval and number of simulations, the n:th worst simulation outcome is the VaR figure.

It is up to the risk manager to decide whether to use absolute or relative changes in the given risk factors when creating a suitable distribution for VaR computations. On the one hand, when using relative changes in interest rates or volatilities, the possibility of generating negative outcomes is avoided, which is a desired trait in order to remain within the range of realistic scenarios. On the other hand, applying relative changes leads to implicit assumption that the magnitude of changes in those risk factors would depend on their absolute levels. The studies in that subject have not provided solid empirical evidence to support either assumption over the other when evaluating the dynamics of interest rates (e.g. Corb, 2012).

While accounting for the interest rate risk in swaption VaR estimation is quite straightforward, enclosing the vega risk is slightly more intricate. A rather simple and tractable means of integrating the vega risk into VaR calculation is to treat implied volatility analogously to other market risk factors to which a portfolio is exposed to as Malz (2001) proposes. Despite the theoretical simplicity of treating the volatility of volatility in a similar fashion as any other risk factor, the key difficulty involved in modeling the vega risk is associated with the choice of strike level to be used. Generally speaking, the use of at-the-money (ATM) level implied volatilities provides insight on the fluctuations of option prices on a large scale, but due to the existence of volatility smile, option portfolios are also exposed the changes in implied volatility *along the smile*, and, moreover, due to non-parallel shifts of the volatility smile options are also exposed to *changes in the shape of the smile*. When the shape of the volatility smile changes, the change in the price of an out-of-the money (OTM) option is relatively higher or lower than the simultaneous price change of an ATM option. Thus, in such situation a VaR model that only incorporates ATM implied volatility changes generates either too optimistic or too pessimistic estimates for options whose strike prices are not close to the ATM level. Furthermore, even if the shape of the volatility smile remained constant, an option is still exposed to changes along the smile when the price of the underlying asset moves away from the option's strike level. For instance, as time passes and interest rates change, the strike rate of a swaption that was originally issued at ATM level subsequently becomes an OTM swaption, which means that the changes in ATM implied volatilities may no longer yield accurate predictions about the changes in its present implied volatility. Intuitively, it would be reasonable to take into account both exposures mentioned above, but in reality, the restricted access to historical implied volatilities for OTM strike levels together with increased computational complicatedness makes it difficult for institutions with limited resources to incorporate the volatility smile changes into their VaR models. Fortunately, the changes along the volatility smile are more straightforward to incorporate in VaR estimation as, after all, the present smile data are required for pricing and computing prevailing mark-to-market values regardless of the risk management models.

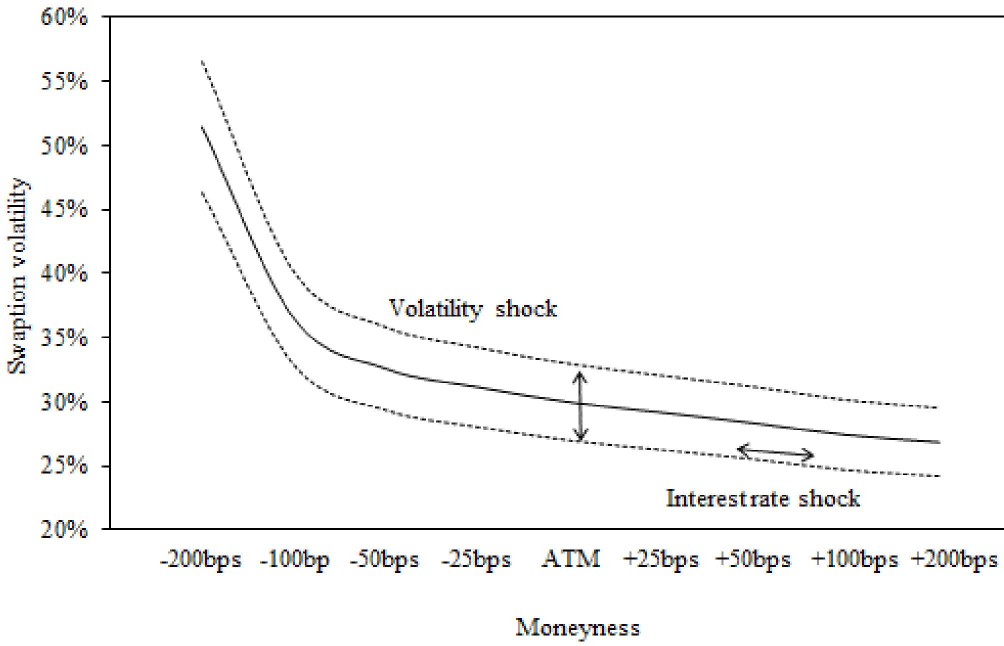
### 2.8.2 Fixed smile method

The simplest approach for incorporating vega risk into VaR computation is to ignore the changes in the implied volatility smile and to take into account merely the changes in the volatility smile as a whole, thus only allowing for parallel shifts in the smile. This is also the first of the models Malz (2001) presents for including vega risk into FX-option VaR estimation. In this approach it is assumed that the smile moves within the exercise rate – vol space but not in the delta – vol space. Consequently, this model assumes that the evolution of the implied volatility surface follows a common heuristic called *sticky delta approach* as Derman (1999) has labeled it. What this approach suggests is that if the underlying swap rate changes, the implied volatility of a swaption with a given moneyness does not change. Thus, in the fixed smile method, the changes in the implied volatility along the smile are taken into consideration through assigning correct implied volatilities to swaptions when their underlying swap rate moves away from the ATM level as a result from applying different historical interest rate changes as shocks to the prevailing interest rates. As a summary, in the fixed smile method, a number of swaption's mark-to-market value changes are simulated by shocking both the swap rates and the implied volatilities, but with the limitation that the same volatility shocks are applied to all swaptions with the same maturity regardless of their respective moneyness. Moreover, the volatility shocks are derived from a history of daily changes in implied volatilities of ATM swaptions, which means that no history of OTM implied volatilities is required in estimation process.

However, when the volatility fluctuations are modeled using relative changes, the shifts are not strictly parallel, as when there is a skew in the volatilities, the absolute changes are not uniform across the different strikes. Nevertheless, the method assumes equivalent relative changes, which seriously limits the shapes that the smile may adopt. Hence, the term "fixed smile" is still used in this paper even if slightly stretching the actual denotation of "fixed".

**Figure 2:**
**The fixed smile method**

In the fixed smile method, implied volatilities are shocked along the prevailing smile and/or in a parallel fashion. The solid line represents volatility smile on May 10, 2012 for a 5x5 swaption.

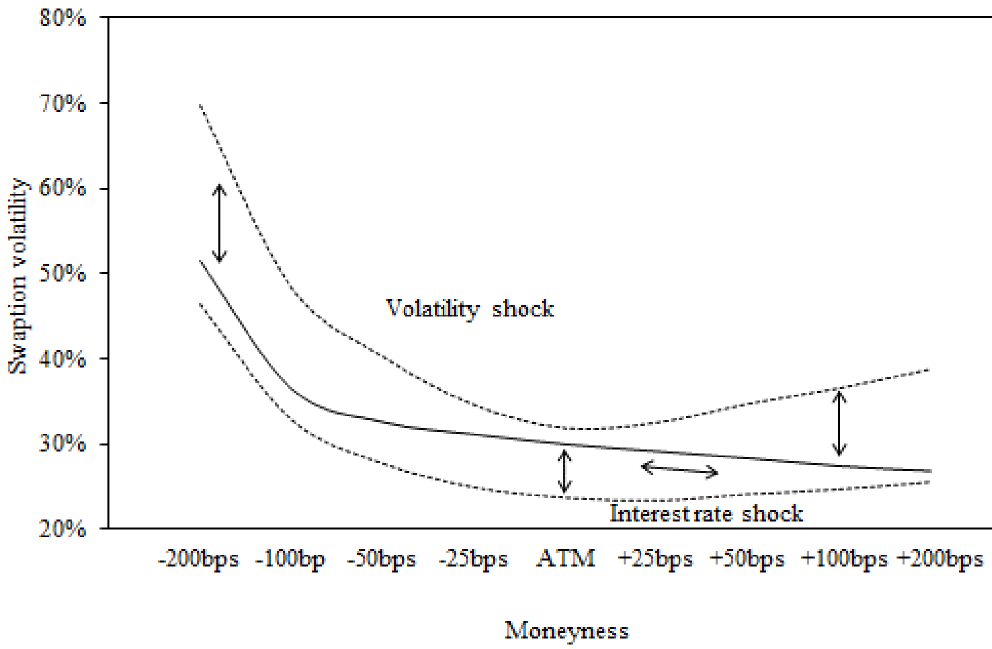



### 2.8.3 Random smile method

A more sophisticated method to incorporate the different dimensions of vega risk into VaR estimation is to allow for changes in the curvature and skewness of the implied volatility smile in addition to the parallel changes that are included in the fixed smile approach. Hence, the shape of the smile is no longer assumed to be fixed but rather allowed to vary more freely than in the previous model. This model is principally the second proposed in Malz (2001). The key difference between the fixed and the random smile methods is that the moneyness of a given swaption is accounted for when applying different volatilities to generate required simulations for its price changes. Thus, for instance, the historical implied volatility changes used for generating prices for a swaption whose moneyness is ATM+50 bps are chosen from changes that have occurred for swaptions with that specific moneyness in the past, while in the fixed smile method the changes in the ATM volatilities would be used instead.

**Figure 3:**
**The random smile method**

In the random smile method the shape of the smile is permitted to vary through applying volatility shocks that depend on the moneyness of a swaption.




### 2.8.4 Fixed and random smile methods with observed volatility changes

While the above presented fixed and random smile methods utilize observed simultaneous changes in interest rates and volatilities in scenario creation in a tractable way, it is assumed in both of the models that the implied volatility scenarios gained from using differences between implied volatilities with equal strike levels on successive days would provide a sufficiently accurate distribution of daily volatility changes. However, for a given swaption traded on day *t* the correct implied volatility on *t+1* is most likely different from the implied volatility of a swaption made on *t+1* with equal moneyness. Hence, the both models can be modified so that instead of using changes between *t+1* and *t* for distinct swaptions with same strike moneyness, the distribution of volatility moves is compiled using changes between actual observed implied volatilities for a given swaption on the consecutive days. The implied volatility on *t+1* is obtained using the SABR model, which means that a history of OTM implied volatilities is required even if using only ATM level changes since the historical

SABR parameters are needed to determine the volatility changes. More specifically, the observed difference is defined in the following way using *(15)*:

$$\Delta\sigma = \sigma\left(S_{a,b}(t) + \Delta K, S_{a,b}(t+1)\right) - \sigma\left(S_{a,b}(t) + \Delta K, S_{a,b}(t)\right) \qquad (18)$$

On the contrary, the method used in 2.8.2 and 2.8.3 can be formalized as follows:

$$\Delta\sigma = \sigma\left(S_{\alpha,\beta}(t+1) + \Delta K, S_{\alpha,\beta}(t+1)\right) - \sigma\left(S_{\alpha,\beta}(t) + \Delta K, S_{\alpha,\beta}(t)\right) \qquad (19)$$

where $S_{a,b}(t)$ is the ATM forward rate on *t* and $\Delta K$ is the difference between ATM level and the strike rate of the specific swaption.

To distinguish the models with implied volatility estimation described in 2.8.2 and 2.8.3 from estimation method defined in this chapter, the models that employ *(19)* are labeled with a prefix “Proxy” and the models that utilize *(18)* are titled with “Direct” in the following sections of this paper.

### 2.8.5 Skew dependent model

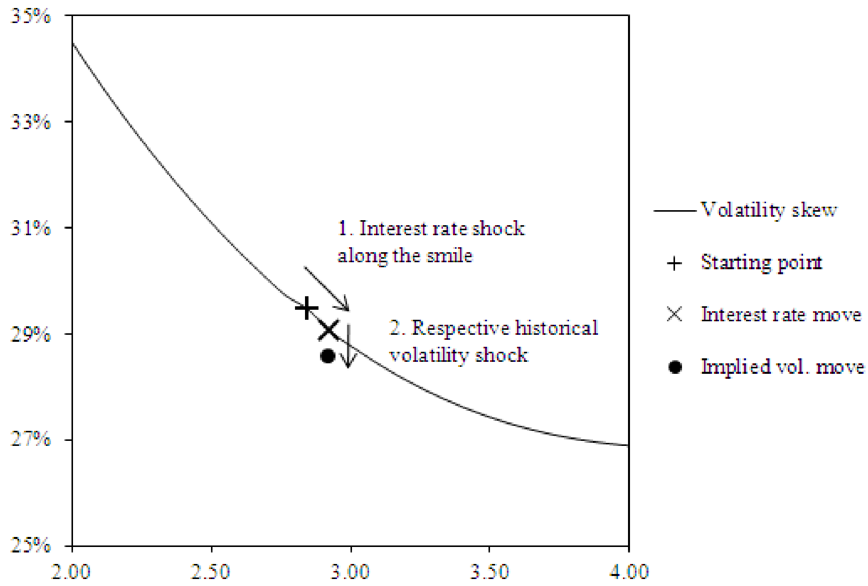
An additional potential source of imprecision in each of the above presented models arises from ignoring the prevailing shape of the volatility smile. For instance, it the skew is steep, a change in the underlying swap rate leads to a higher change in implied volatility than would occur as a result from a similar rate change when the smile is flatter. One conceivable method of considering the present shape is based on separating the observed implied volatility changes into two components stemming from either a change in interest rates or from a change in general level of implied volatilities. However, the problem is that there is no objectively specified method for determining which proportion of the observed change stems from interest rate changes or from overall changes in volatilities. Nevertheless, an approximating segregation can be implemented using the SABR model so that the change in implied volatility that follows from an interest rate move between days *t* and *t+1* is defined as the difference between the observed implied volatility on *t* and the SABR model output for the new interest rate level using model parameters from day *t.* Then, the residual difference between the actual implied volatility on *t+1* and the figure estimated using the SABR model for day *t* represents the move due to general change in volatilities.

As a result, the scenarios are obtained by first shocking the interest rate with historical rate changes for which the theoretical corresponding implied volatilities are computed with *(15)*. Next, each of these scenarios are further shocked with the previously obtained theoretical changes in the general level of implied volatilities to achieve the final swap rate-implied volatility -pairs. Figure 4 illustrates how a single scenario point is achieved under the skew dependent model.

The skew dependent model is referred to with a prefix "Component" in the empirical part of this study based on the technique how the total VaR estimate is obtained. However, it should not be fixed with the risk components of VaR figures.

**Figure 4:**
**Skew dependent model**

The figure illustrates the scenario creation method in skew dependent model. First, the prevailing point is moved from starting position (+) along the observed volatility skew to the point that matches the interest rate scenario (X). Then, the point is moved vertically by an amount that matches the corresponding change in implied volatilities that occurred together with the observed interest rate change (dot). As a result from repeating the described steps 250 times and valuating the contract in each of them, a distribution of mark-to-market price scenarios is achieved. When the prevailing price of the swaption is reduced from the price scenarios, a distribution of 250 profit and loss figures is achieved, and VaR estimate is then the percentile of the distribution that matches the chosen VaR confidence level (e.g. 5% percentile for 95% confidence level VaR).




# 3 HYPOTHESES

This section presents the hypotheses that are tested in the empirical part of this study. The hypotheses are structured in a way that facilitates finding answers to the research problems stated in chapter 1.2.

As this thesis aims to examine the importance of accounting for swaption moneyness in creating a distribution of historical changes in implied volatilities used in swaption VaR estimation based on variants of historical simulation, a relevant hypothesis to test is whether or not the alternative VaR methods introduced in the previous section generate differing risk estimates. However, before investigating differences between the alternative models, it is worth testing whether the VaR models are able to provide a sufficiently accurate estimates of swaption market risk to begin with. Therefore, the first hypothesis is aligned with the statistical test used for determining validity of a given VaR model and is as follows:

> Hypothesis 1: VaR models based on historical simulation generate acceptable estimates of swaption's market risk.

The second hypothesis is related to the differences between the previously mentioned swaption VaR models. If the changes in the shape of the implied volatility smile induce a relevant risk source, models with random smile method should provide better results than the fixed smile models. Nevertheless, due to the lack of previous empirical findings on the specific issue, the formulation of the second hypothesis is partially arbitrary: while Malz (2001) does not find evidence that the model in which smile changes are recognized as a risk factor would be more preferable than the model with fixed smile, the paper concentrates on options on FX-rates instead of swap rates. Moreover, the shape of the volatility smile is not fixed in reality, which supports the notion of accounting for the changes also in VaR estimation. Notwithstanding, the preliminary evaluation of the sample data suggests that the changes in OTM levels are highly correlated with the changes in ATM level, which indicates that the added information from considering OTM volatility changes may not improve the VaR precision in a notable fashion. Hence,

> Hypothesis 2: A model that employs OTM level implied volatility changes in swaption VaR estimation does not generate more accurate estimates than a model with fixed smile approach.

# 4 DATA AND METHODOLOGY

This section starts with a presentation of the data employed in this study, and then continues with a description of the VaR methods utilized in creating the historical daily VaR estimates.

The section concludes with an illustration of the methodology utilized in testing the hypotheses.

## 4.1 Data

The data required in this study comprises of the elements needed for swaption pricing, and hence, the following data are needed:

- EONIA zero rates for discounting
- 6 month Euribor zero rates for projection
- Implied swaption volatilities

The interest rate data are provided by Pohjola Bank Plc. As zero rates are not traded as such, the rates are bootstrapped from different market traded interest rate instruments, and the interest rate curve is smoothed via interpolation in order to cover also the maturities for which there are no market quotes.

The implied swaption volatilities for different maturity-tenor-strike -combinations are also provided by Pohjola Bank Plc. However, due to the low absolute level of interest rates, quotes are not available for all of the moneyness levels for which there usually exist quotes as highest negative relative strike levels would imply negative strikes in absolute terms. Hence, the VaR estimation in this study concentrates on moneyness levels above the at-the-money level for the data below the ATM level is incomplete. Furthermore, the reliability of historical implied volatilities for out-of-the-money strikes is also limited as the trading is concentrated around at-the-money level. Nevertheless, as there is no conclusive method for separating erroneous observations from valid figures, merely those observations that are definitely incorrect, such as negative volatilities, are disregarded. However, the plausible incorrectness of observations for strikes far OTM should be taken into account when interpreting the results. Moreover, the dataset covers dates only from March 8, 2011 onwards. Consequently, changes in implied volatilities needed in historical VaR estimation are available from March 9, 2011, which truncates the length of the backtesting period shorter than desired.

In total, the data consists of 480 daily observations of swap rates with maturities ranging from one month up to 60 years, and of equal number of daily swaption volatilities. Using a historical observation period of 250 trading days in VaR estimation, there is 228 days left for backtesting. However, due to practical reasons in spite of the abundant availability of data for different maturities, the number of different swaptions employed in this study is limited to the following set of maturity-tenor pairs: 1x2, 1x5, 5x2, 5x5, 10x2 and 10x5 with strike rates ranging from ATM to ATM +200 bps. Furthermore, the VaR estimates and backtesting results shown in this study are computed for long and short physically settled European receiver swaption positions with notional of 10 million EUR. To confirm the validity of the estimates, I also run analogous analyses for payer swaption positions.

### 4.1.1 Descriptive statistics

Swap rates during the overlapping estimation and backtesting periods are at historically low levels: for instance, five year rate starts at 2.98% and ends at 1.13% while maximum and minimum values are 3.20% and 0.74% respectively. Moreover, as Figure 5 shows, the swap rates decline until the end of 2012 quite steadily and then begin to climb again. Additionally, the shape of the yield curve stays relatively unchanged for the short end, although a minor twist can be observed as the long end ascends in relation to the short end during the latter part of the investigation period. The yield curve is presented for a few selected days from the beginning, middle and end of the observation period in Figure 6.

**Figure 5:**
**Euribor swap rates**

The figure shows the development of Euribor swap rates for maturities of 1, 5, 10 and 15 years under the VaR estimation period between March 8, 2011 and February 1, 2013.

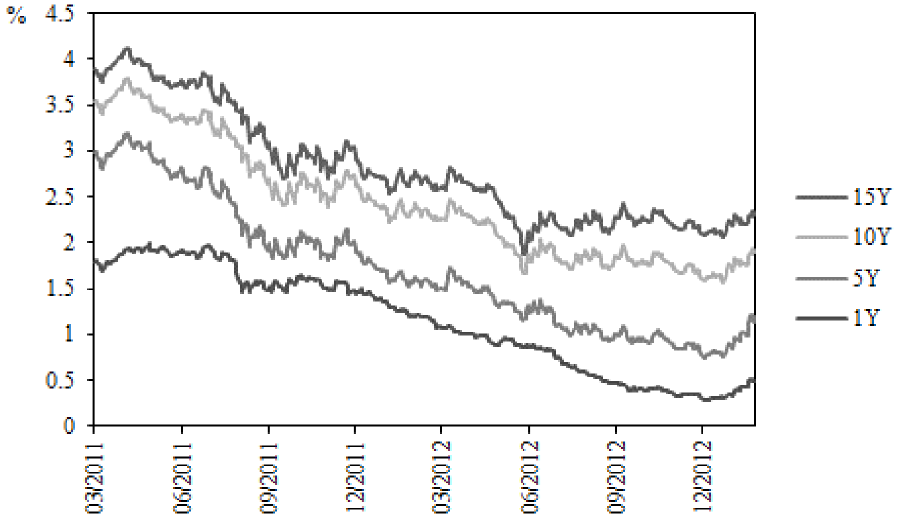



**Figure 6:**
**Yield curves**

Figure 6 illustrates yield curves for selected dates within the sample data. The curve ranges from one month to 60 years. The primary change during the period concerns the overall level of the curve, but also the shape of the yield curve has altered slightly so that the long end has gradually shifted upwards in relation to the short end.

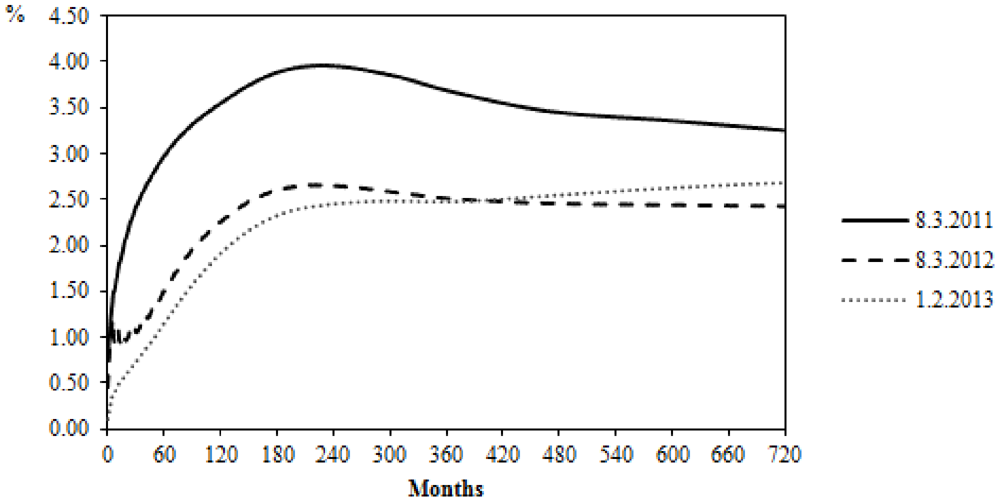



Table 4 collects the summary statistics of daily changes in swaption implied volatilities and underlying forward swap rates during the backtesting period, and the analogous statistics from the whole observation period are shown in Table 16 in Appendix B.

The steady decline in swap rates illustrated in Figure 5 is corroborated by the average daily change of approximately -0.4 basis points for each swaption maturity-tenor pairs investigated during the whole sample period. In contrast, implied volatilities tend to increase during the sample period and daily changes average between 0.06% and 0.14% for ATM options. Slightly surprisingly, however, the average daily log-returns for implied volatilities of 10x2 and 10x5 swaptions are negative during the backtesting period.

Table 3 shows the correlations between the risk factors during the whole observing period and correlations sampled from the backtesting time period are presented in Table 17. As expected on the basis of estimates presented in Andersen and Lund (1997), who study the dynamics of interest rates, the correlations between interest rates and lognormal implied volatilities are negative. This negative link between the risk factors indicates that neglecting the existence of the volatility smile by using an assumption of constant volatility in swaption VaR computations would lead to erroneous risk estimates as Malz (2001) puts forward. However, correlations between changes in implied volatilities of different strike levels are fairly high, which gives a reason to expect that using merely ATM volatility changes might be sufficient for VaR estimation also for swaptions with OTM strikes. Yet, as Figure 7 shows, the spreads between ATM and OTM implied volatilities are not constant, which indicates that the shape of the volatility skew fluctuates at least to some extent during the observation period. Nonetheless, as Figure 13 in Appendix B illustrates, changes in the shape of the volatility skew during the observation period are minimal.

Table 5 shows summary statistics of the swaption price fluctuations for the contracts examined in this study. Explicitly, the statistics are computed from one day changes for the contracts opened and closed on a daily basis. The return distributions exhibit positive skewness and excess kurtosis, and according to the Lilliefors test of normality most of the price fluctuations do not come from a normally distributed population. On the other hand, Jarque-Bera test of normality cannot reject the null hypothesis of normality for 1x2, 1x5 and 5x2 swaptions with strikes ranging from ATM to ATM +25bps. Nevertheless, these distributional characteristics refute the use of a parametric VaR method as it would likely underestimate the probability of the tail events. Consequently, the use of Historical Simulation is a preferable approach.

Interestingly, the average price changes for many of the long receiver swaption positions considered in the study are negative although the interest rates decline on average during the observation period. At first this remark appears counterintuitive, as a receiver swaption has a negative delta and a positive vega, which means that for the long position holder, a decline in the interest rates and an increase in the volatility both increase the value of the contract. Moreover, since the underlying forward rates and implied volatilities are negatively correlated, as presented in Table 3, a decline in interest rates should on average be accompanied by an increase in volatility, which should then deliver a double-boost to the long position holder. However, a closer examination of the dynamics of the risk factors reveals that even though the interest rates decline and implied volatilities increase in more than half of the days during the backtesting period, the average increase in volatility is nevertheless smaller than the average decrease. This asymmetry is also reflected on the negative skewness of the implied volatility fluctuations during the backtesting period. Furthermore, as the swaption sensitivities with respect to changes in interest rates and volatilities depend on the moneyness and maturity as well as on the tenor of the option, price changes arising from fluctuations in the risk factors are not equal for ATM and OTM options or for options whose time to maturity and tenor are not equal. In general, both delta and vega are at maximum for ATM options when other factors are held constant. However, their relative significance depends on the other factors as well, and hence pointing out which factor dominates the price movements requires simultaneous computations of the sensitivities.

Nevertheless, to summarize the behavior of the considered positions, the longer the time to maturity, the more important the vega becomes, whereas the further away from ATM the strike moves, the more central the delta becomes. That is, the impact of the delta tends to decrease at a slower pace with respect to changes in strike dimension than that of the vega. Consequently, while the volatility changes dominate the price fluctuations for ATM to ATM +100 bps swaptions, the trend of declining interest rates is better reflected on prices of ATM +200 bps swaptions. Numerical[4] delta and vega figures for long swaption positions are shown in Figure 12 (Appendix B) to illustrate the swaption sensitivities during the backtesting period. As a whole, the prices of the considered swaption positions do not move in unison as can be observed from less than perfect correlations between implied volatility fluctuations of different strike levels and from divergent sensitivities with respect to the risk factors.

---

[4]The sensitivities could be computed also analytically, i.e. they could be derived from the option pricing formula, but using numerical estimates often leads to more rational figures (e.g. Taleb, 1997).

**Table 3:**
**Risk factor correlations**

The table shows correlation coefficients for the swaption risk factors during the observation period from March 8, 2011 to February 1, 2013. F stands for the forward starting swap rate underlying the specified swaption contracts. The correlations are measured between absolute changes in F and log-changes in implied volatilities.

| | F | ATM | +25bps | +50bps | +100bps | +200bps |
|---|---|---|---|---|---|---|
| Panel A: 1x2 | | | | | | |
| F | 1.000 | -0.651 | -0.611 | -0.561 | -0.481 | -0.397 |
| ATM | -0.651 | 1.000 | 0.979 | 0.875 | 0.767 | 0.655 |
| +25bps | -0.611 | 0.979 | 1.000 | 0.861 | 0.758 | 0.648 |
| +50bps | -0.561 | 0.875 | 0.861 | 1.000 | 0.912 | 0.847 |
| +100bps | -0.481 | 0.767 | 0.758 | 0.912 | 1.000 | 0.897 |
| +200bps | -0.397 | 0.655 | 0.648 | 0.847 | 0.897 | 1.000 |
| Panel B: 1x5 | | | | | | |
| F | 1.000 | -0.728 | -0.718 | -0.659 | -0.529 | -0.375 |
| ATM | -0.728 | 1.000 | 0.993 | 0.889 | 0.705 | 0.476 |
| +25bps | -0.718 | 0.993 | 1.000 | 0.887 | 0.709 | 0.484 |
| +50bps | -0.659 | 0.889 | 0.887 | 1.000 | 0.905 | 0.740 |
| +100bps | -0.529 | 0.705 | 0.709 | 0.905 | 1.000 | 0.878 |
| +200bps | -0.375 | 0.476 | 0.484 | 0.740 | 0.878 | 1.000 |
| Panel C: 5x2 | | | | | | |
| F | 1.000 | -0.748 | -0.738 | -0.724 | -0.689 | -0.607 |
| ATM | -0.748 | 1.000 | 0.980 | 0.987 | 0.953 | 0.858 |
| +25bps | -0.738 | 0.980 | 1.000 | 0.968 | 0.939 | 0.845 |
| +50bps | -0.724 | 0.987 | 0.968 | 1.000 | 0.985 | 0.919 |
| +100bps | -0.689 | 0.953 | 0.939 | 0.985 | 1.000 | 0.966 |
| +200bps | -0.607 | 0.858 | 0.845 | 0.919 | 0.966 | 1.000 |
| Panel D: 5x5 | | | | | | |
| F | 1.000 | -0.811 | -0.800 | -0.806 | -0.792 | -0.729 |
| ATM | -0.811 | 1.000 | 0.979 | 0.989 | 0.959 | 0.880 |
| +25bps | -0.800 | 0.979 | 1.000 | 0.971 | 0.942 | 0.863 |
| +50bps | -0.806 | 0.989 | 0.971 | 1.000 | 0.985 | 0.918 |
| +100bps | -0.792 | 0.959 | 0.942 | 0.985 | 1.000 | 0.945 |
| +200bps | -0.729 | 0.880 | 0.863 | 0.918 | 0.945 | 1.000 |
| Panel E: 10x2 | | | | | | |
| F | 1.000 | -0.759 | -0.745 | -0.735 | -0.711 | -0.666 |
| ATM | -0.759 | 1.000 | 0.990 | 0.988 | 0.969 | 0.925 |
| +25bps | -0.745 | 0.990 | 1.000 | 0.980 | 0.961 | 0.915 |
| +50bps | -0.735 | 0.988 | 0.980 | 1.000 | 0.987 | 0.952 |
| +100bps | -0.711 | 0.969 | 0.961 | 0.987 | 1.000 | 0.971 |
| +200bps | -0.666 | 0.925 | 0.915 | 0.952 | 0.971 | 1.000 |
| Panel F: 10x5 | | | | | | |
| F | 1.000 | -0.798 | -0.784 | -0.802 | -0.781 | -0.760 |
| ATM | -0.798 | 1.000 | 0.972 | 0.982 | 0.960 | 0.918 |
| +25bps | -0.784 | 0.972 | 1.000 | 0.976 | 0.955 | 0.914 |
| +50bps | -0.802 | 0.982 | 0.976 | 1.000 | 0.986 | 0.956 |
| +100bps | -0.781 | 0.960 | 0.955 | 0.986 | 1.000 | 0.975 |
| +200bps | -0.760 | 0.918 | 0.914 | 0.956 | 0.975 | 1.000 |

**Table 4:**
**Risk factor summary statistics**

The table presents summary statistics about risk factor returns during the backtesting period between March 5, 2012 and February 1, 2013. Changes in forward swap rates (F) are measured in absolute terms and presented in basis points. Changes in implied volatilities are measured as log-returns and presented as percentage changes.

| | F | ATM | +25bps | +50bps | +100bps | +200bps |
|---|---|---|---|---|---|---|
| Panel A: 1x2 | | | | | | |
| Average | -0.09 | 0.12 | 0.12 | 0.10 | 0.08 | 0.07 |
| Median | -0.05 | 0.26 | 0.25 | 0.20 | 0.24 | 0.33 |
| Std dev | 3.63 | 4.52 | 4.35 | 4.70 | 5.28 | 5.70 |
| Skewness | 0.56 | -0.06 | -0.05 | -0.10 | -0.20 | -0.69 |
| Kurtosis | 4.24 | 4.09 | 4.35 | 4.01 | 5.91 | 8.79 |
| Panel B: 1x5 | | | | | | |
| Average | -0.15 | 0.05 | 0.05 | 0.05 | 0.06 | 0.05 |
| Median | -0.18 | 0.21 | 0.11 | -0.01 | 0.02 | 0.12 |
| Std dev | 4.19 | 3.10 | 3.02 | 3.37 | 4.15 | 5.79 |
| Skewness | 0.53 | -0.44 | -0.41 | -0.26 | -0.22 | -0.04 |
| Kurtosis | 3.31 | 4.71 | 4.75 | 4.80 | 6.99 | 7.75 |
| Panel C: 5x2 | | | | | | |
| Average | -0.21 | 0.03 | 0.03 | 0.03 | 0.03 | 0.04 |
| Median | -0.66 | 0.20 | 0.20 | 0.15 | 0.09 | 0.06 |
| Std dev | 4.90 | 2.36 | 2.33 | 2.30 | 2.28 | 2.43 |
| Skewness | 0.50 | -0.28 | -0.23 | -0.35 | -0.35 | -0.20 |
| Kurtosis | 3.46 | 4.78 | 4.66 | 4.60 | 4.48 | 4.44 |
| Panel D: 5x5 | | | | | | |
| Average | -0.15 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| Median | -0.48 | 0.12 | 0.12 | 0.07 | 0.02 | 0.06 |
| Std dev | 5.10 | 2.19 | 2.12 | 2.17 | 2.18 | 2.17 |
| Skewness | 0.43 | -0.34 | -0.18 | -0.26 | -0.17 | 0.06 |
| Kurtosis | 3.55 | 5.09 | 4.90 | 4.76 | 4.87 | 6.90 |
| Panel E: 10x2 | | | | | | |
| Average | -0.08 | -0.02 | -0.02 | -0.03 | -0.03 | -0.03 |
| Median | -0.06 | 0.16 | 0.15 | 0.11 | 0.07 | 0.11 |
| Std dev | 5.35 | 2.03 | 2.00 | 2.01 | 2.00 | 1.97 |
| Skewness | 0.42 | -0.12 | -0.10 | 0.04 | 0.14 | 0.19 |
| Kurtosis | 4.30 | 4.78 | 4.86 | 4.99 | 5.21 | 5.32 |
| Panel F: 10x5 | | | | | | |
| Average | -0.05 | -0.02 | -0.02 | -0.02 | -0.02 | -0.02 |
| Median | -0.09 | 0.07 | 0.06 | 0.05 | 0.03 | 0.07 |
| Std dev | 5.46 | 2.06 | 2.00 | 2.05 | 2.07 | 2.03 |
| Skewness | 0.40 | -0.39 | -0.19 | -0.25 | -0.16 | 0.15 |
| Kurtosis | 4.90 | 5.46 | 4.84 | 5.07 | 5.02 | 6.10 |

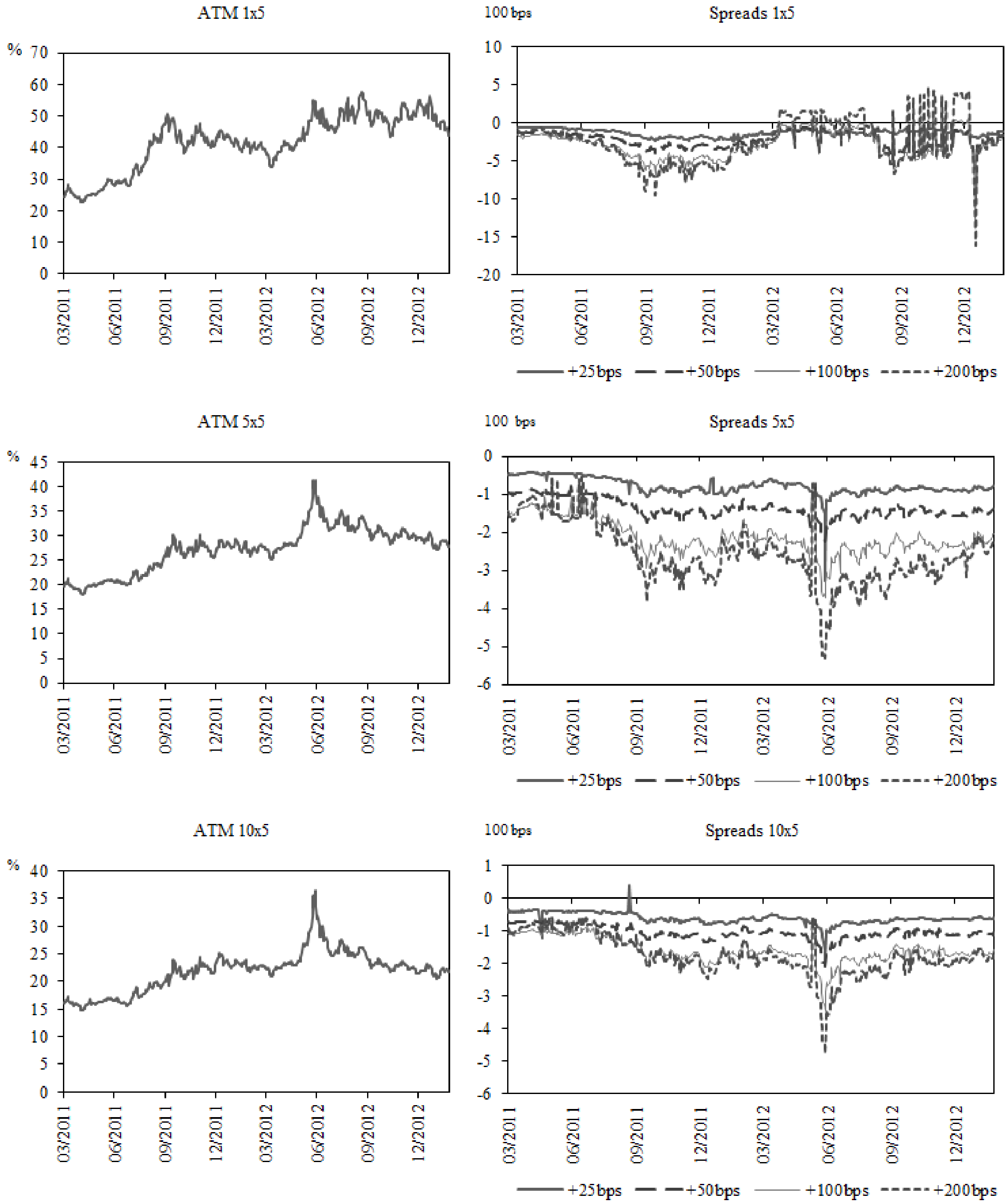
**Table 5:**
**Swaption return statistics**

This table summarizes statistical properties of the daily price fluctuations of the long receiver swaption positions during the observation period of March 5, 2012 to February 1, 2013. The statistics are based on absolute returns and the number of observations is 228.

| | ATM | +25 bps | +50 bps | +100 bps | +200 bps |
|---|---|---|---|---|---|
| Panel A: 1x2 | | | | | |
| Average | 490 | -368 | -251 | -53 | 351 |
| Median | 620 | -293 | -13 | 416 | 658 |
| Std dev | 3 926 | 5 333 | 6 148 | 7 021 | 7 529 |
| Skewness | -0.122 | -0.139 | -0.307 | -0.428 | -0.528 |
| Kurtosis | 3.204 | 3.288 | 3.444 | 3.812 | 4.215 |
| Panel B: 1x5 | | | | | |
| Average | 744 | -956 | -1 063 | -226 | 1 421 |
| Median | 607 | -918 | -727 | -146 | 2 004 |
| Std dev | 10 899 | 14 244 | 16 517 | 19 632 | 21 675 |
| Skewness | -0.291 | -0.268 | -0.405 | -0.477 | -0.522 |
| Kurtosis | 3.076 | 3.011 | 3.031 | 3.097 | 3.225 |
| Panel C: 5x2 | | | | | |
| Average | -599 | -1 910 | -1 863 | -1 238 | 3 000 |
| Median | -224 | -1 500 | -1 313 | -857 | 3 341 |
| Std dev | 5 244 | 5 936 | 6 676 | 7 876 | 9 646 |
| Skewness | -0.079 | -0.107 | -0.168 | -0.293 | -0.349 |
| Kurtosis | 3.820 | 3.643 | 3.515 | 3.317 | 3.039 |
| Panel D: 5x5 | | | | | |
| Average | -815 | -1 965 | -3 204 | -2 524 | 4 605 |
| Median | -77 | -861 | -2 316 | -1 898 | 5 963 |
| Std dev | 13 455 | 15 056 | 17 125 | 20 357 | 24 846 |
| Skewness | -0.347 | -0.269 | -0.355 | -0.367 | -0.353 |
| Kurtosis | 3.821 | 3.579 | 3.528 | 3.444 | 3.400 |
| Panel E: 10x2 | | | | | |
| Average | -257 | -472 | -1 230 | -1 150 | 1 634 |
| Median | -83 | -210 | -969 | -854 | 1 890 |
| Std dev | 5 426 | 6 064 | 6 776 | 8 045 | 10 071 |
| Skewness | -0.180 | -0.209 | -0.225 | -0.240 | -0.199 |
| Kurtosis | 4.111 | 4.157 | 4.160 | 4.163 | 4.007 |
| Panel F: 10x5 | | | | | |
| Average | -603 | -1 190 | -2 936 | -2 779 | 3 872 |
| Median | 452 | 417 | -1 196 | -981 | 5 165 |
| Std dev | 13 361 | 14 862 | 16 798 | 20 086 | 25 221 |
| Skewness | -0.369 | -0.323 | -0.404 | -0.380 | -0.261 |
| Kurtosis | 4.615 | 4.442 | 4.592 | 4.578 | 4.434 |

## Figure 7:
## ATM volatilities and OTM spreads

The figures below present ATM implied volatility development during the sample period for 1x5, 5x5 and 10x5 swaptions together with spreads for their OTM strikes above the ATM level.



## 4.2 VaR estimation

This study employs historical VaR estimation for generating a backtesting sample that is used for testing the hypotheses. The choice of the VaR method is based on its suitability for option portfolios and on the overall incentive to avoid having to make assumptions regarding the parameters of return distributions.

In VaR computation the following three methods are used:

- basic historical simulation (HS)
- filtered historical simulation (FHS)
- time-weighted historical simulation (TW)

The historical observation period is 250 trading days, which is approximately one calendar year. Hence, the VaR figure is computed for every day starting from March 5, 2012 until the end of the data set. The VaR figures are estimated using 95% confidence level and a risk horizon of 1 day. In addition, the figures for expected shortfall (ES) with confidence level of 95% are computed in order to obtain data for loss function -based backtesting method. For computational reasons, it is assumed that a swaption position is opened and closed on a daily basis. Otherwise one should also take into account the changes in time value of the option, which would introduce another level of complexity that, however, would not contribute additional relevant information to the original research question of this study. Moreover, for a swaption whose maturity is either two, five or ten years, the change in its mark-to-market value resulting from a one day reduction in its time until maturity is infinitesimal compared to changes stemming from fluctuations in interest rates and implied volatilities.

The historical changes in interest rates are measured in absolute arithmetic terms[5] in all models. Therefore changes in implied volatilities are modeled using relative changes to maintain consistency between interest rate and implied volatility changes. In all of the methods the VaR calculation is implemented using full-valuation method, which means that the position is revalued under each of the risk factor scenarios that are created through applying historical swap rate and implied volatility changes to the respective prevailing

[5] There is no clear evidence whether interest rate volatility depends on interest rate levels or not (eg. van Deventer et al., 2005). Consequently, arithmetic returns are used as they are more intuitively appealing when dealing with interest rate instruments.

levels. In FHS models the changes are adjusted using exponentially weighted moving average (EWMA) method with the lambda parameter set at 0.94 following RiskMetrics convention, while in models employing TW approach the lambda is set to be 0.98 as in Boudoukh et al. (1998). In HS method the changes are applied without adjustments.

### 4.2.1 Historical simulation

Historical simulation is implemented by generating different scenarios using the observed historical changes in interest rates and implied volatilities. Each of the historical daily changes in swap rates and implied volatilities are added one at a time to the prevailing interest rate and implied volatility curves, and as a result, a distribution of 250 possible scenarios is obtained. Then the scenarios are used for generating respective number of new mark-to-market scenarios for the position from which the present value of the position is subtracted. Accordingly, a distribution of 250 profit and loss figures is achieved from which the 5% quantile (for VaR at 95% confidence level) is drawn. In the fixed smile method, each historical swap rate change for a given date is accompanied by the respective change in at-the-money implied volatility on that same date, whereas in the random smile method the changes in implied volatility are chosen based on the moneyness of the swaption.

### 4.2.2 Filtered historical simulation

Filtered historical simulation is employed in a similar fashion as basic HS, but the actual returns within the historical dataset are replaced with returns adjusted by forecast of volatility for a variable *i* in the following way:

$$r_{t,i}^{*} = \frac{\sigma_{T,i}}{\sigma_{t,i}} r_{t,i} \tag{20}$$

where

$\sigma_{T,i}$ is the most recent forecast for the volatility for *i*

$r_{t,i}$ is the actual historical return in *i* on day *t*

$\sigma_{t,i}$ is the historical forecast of the volatility for changes in $i$ made on *t*

The volatility forecast can be gained by for instance using the generalized autoregressive conditional heteroskedastic (GARCH) model originally developed by Engle (1982) and

Bollerslev (1986). However, as this model requires a vast number of parameters to be estimated, its practical implementation is usually not viable in large scale risk measurement process carried out on a daily basis. Hence, more commonly used approach for modeling volatility is exponentially weighted moving average (EWMA) forecast[6] (Jorion, 2007). Stated in a formal way, the variance forecast for time *t* is

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_t^2 \qquad (21)$$

where the parameter $\lambda$ is again the decay factor that defines the weights for previous forecast and latest observation. Also, $\lambda$ is the only parameter that needs to be estimated in the EWMA approach. The value of $\lambda$ defines the reactiveness of the forecast to market events: the lower the figure, the more reactive the forecast becomes. Following J.P. Morgan's RiskMetric (1996) it is frequently set at 0.94 for daily observations.

### 4.2.3 Time-weighted historical simulation

Following Boudoukh et al. (1998), the time-weighted approach is implemented in three steps:

1. Realized return from *t-1* to *t* is denoted by *R(t)*. To each of the *H* most recent returns $(R(t), R(t-1), \ldots, R(t-(H+1))$ is assigned a weight $\frac{1-\lambda}{1-\lambda^H}, \frac{1-\lambda}{1-\lambda^H}\lambda, \ldots,$ $\frac{1-\lambda}{1-\lambda^H}\lambda^{H-1}$ respectively, where the parameter $\lambda$ is the decay factor.
2. The returns are sorted in ascending order
3. To find the VaR figure of the portfolio, the weights are accumulated starting from the lowest return until the desired quantile is reached (e.g. 5% when using 95% confidence level).

### 4.2.4 Quantile estimation

In addition, since the quantile estimation might be rather inaccurate due to rather short observation period, two different percentile estimation methods are utilized for determining the VaR figure from the distribution. The first alternative is to use the percentile -function in Matlab without any additional adjustments. The second alternative is to apply a distribution

[6] GARCH models have time-varying conditional volatility whereas EWMA models give time-varying estimates of the unconditional volatility (Alexander, 2008).

fitting method to the tail of the distribution, and to choose the value from the fitted distribution. The fitting is implemented using Epanechnikov kernel function following suggestion of Butler and Schachter (1998). Also the fitting is implemented in Matlab using ProbDistUnivKernel constructor.

### 4.2.5 VaR components

Moreover, the VaR estimation is divided into sub-steps in order to allow for separation of the total VaR figure into interest rate risk and volatility risk components. The interest rate component is obtained by generating scenarios using the historical changes in interest rates in a similar fashion as in total VaR computation, but using only the implied volatility that has prevailed on the given estimation date. Depending on the model in question, the historical changes are adjusted similarly as in total VaR computation. However, it should be pointed out that when computing the interest rate risk component of a swaption VaR estimate, one still needs to account for the changes also in the implied volatility used in Black model due to the existence of the volatility smile. Hence, for each interest rate scenario that is generated by shocking the swap curve with historical changes, the corresponding implied volatility for each swap rate scenario must be chosen accordingly from the volatility cube prevailing on the estimation day. The correct volatility is obtained with the SABR model.

The volatility component is calculated in a similar way as the interest rate component, but now the swap rate on the estimation date is kept constant while historical implied volatility changes are applied to the volatility prevailing on the estimation day. Similar adjustments are applied to implied volatility changes in FHS and TW methods that are used for adjusting the interest rates in the respective models.

### 4.2.6 Summary of VaR estimation

The VaR figures are estimated for the different combinations of the following parameters:

- Use of ATM or moneyness dependent historical changes in implied volatilities
- Implied volatility changes derived from daily differences between swaptions with equal moneyness on successive days

("Proxy") or from actual differences in implied volatilities for a specific contract ("Direct")

- Swaption moneyness relative to ATM: 0, +25bps, +50bps, +100bps or +200bps
- VaR confidence level: 95% and 90%
- Quantile estimation: interpolation from the observed distribution or from fitted distribution
- Volatility updating: lambda 0.94 or no weighing
- Time weighing: 0.98 or no weighing
- Maturity of the swaption: 2, 5 or 10 years
- Tenor of the underlying swap: 2 or 5 years

The ATM-moneyness is included for comparing the effect of lambda weighing. Naturally, the results for ATM-moneyness swaptions are the same regardless of whether the fixed smile or the random smile method is used.

Different combinations of the above mentioned dimensions of models and model configurations would enable computation of almost 4 000 VaR series when combined also with swaption type being receiver or payer. However, not all possible variations are considered in this study. For instance, distribution fitting is not evaluated for each of the combinations, and VaR estimates with 90% confidence level are computed only for a limited number of contracts. Moreover, time weighing is not combined with volatility updating method utilized in FHS models. Finally, the backtesting sample consists of 1 080 different VaR estimate series.

## 4.3 SABR -parameter estimation

In order to compute the 1-day change in a swaption's mark-to-market value, i.e. the profit or loss figure, or the interest rate risk component of the daily VaR estimate, one has to apply a volatility model to obtain the implied volatility needed in Black model that corresponds to the strike rate of the given contract. Likewise, the model is required in estimating the observed historical implied volatility changes to be used in models that employ "Direct" changes instead of "Proxy" changes. As the SABR model is able to fit the observed smile almost perfectly (e.g. Hagan, 2002) and since it can be regarded as the standard method in interest rate option markets for modeling implied volatilities, it is also applied in this study.

The SABR model's parameters are calibrated for every date included in the sample data and for each of the maturity-tenor -pairs. In general, model calibration means finding parameters that minimize the error between the observed implied volatilities and the points that are fitted by the model. More specifically, the calibration is conducted by using Levenberg-Marquardt algorithm to find the parameter values that minimize the mean squared error between observed and fitted volatilities. Thus, the problem can be formulated as follows:

$$\min_{\nu,\alpha_0,\rho,\beta} \sum_i ((\bar{\sigma}_i - \sigma_B(\nu, \alpha_0, \rho, \beta; K_i, S_{a,b}(0)))^2 \qquad (22)$$

where $\bar{\sigma}_i$ are the market observed implied volatilities and $\sigma_B\ (\nu, \alpha_0, \rho, \beta; K_i, S_{a,b}(0))$ are the fitted implied volatilities as a function of the SABR parameters, strike rates $K_i$ and ATM forward rate $S_{a,b}(0)$. The calibration is implemented in Matlab.

## 4.4 Backtesting

The backtesting process is implemented in two stages following the strategy of Angelidis et al. (2007). First, the different models are evaluated with conditional and unconditional tests based on the information provided by the number and frequency of VaR exceptions. Also the Basel traffic light test is implemented in this stage. Moreover, the first stage of backtesting process should provide an answer to the Hypothesis 1 by providing insight whether the used models are suitable for measuring swaption VaR.

Next, the second stage of testing is implemented to the models that have passed the first stage. In the second stage, the risk measurement accuracy of the models is compared using the loss function based backtesting method described in paragraph 4.7. More specifically, models with different methods with respect to smile modelling but otherwise similar configurations are compared together by gathering data from the series that passed the first stage. For instance, a TW model with fixed smile method is compared against TW model with random smile method with the same underlying swaption contract. Nevertheless, as the loss function based backtest would define a model with zero exceptions as the superior model, this test will be implemented by taking into account only the exceptions that have occurred on those days when both of the models have generated an exception.

### 4.4.1 The statistical framework of backtesting

The most straightforward method for assessing the quality of a given VaR model is to count the number of times when the actual portfolio losses exceed the model's respective estimates: if the number of exceptions goes above the limit indicated by the used confidence level, the model could be too optimistic in the sense it might underestimate the actual risk. In addition, if the number of exceptions is less than predicted by the confidence level, the model may overestimate the risk, which also indicates that the quality and hence the estimates of the model might be questionable. Obviously, the number of exceptions is a random variable, which means that the amount of exceptions rarely equals the number suggested by the confidence interval. Consequently, the decision whether the number of exceptions is acceptable or not should be based on study of appropriate statistical analyses.

Statistical tests provide valuable insight into VaR model quality estimation and, more importantly, a systematic approach to decision making when assessing the validity of a VaR model. In the tests of unconditional coverage a VaR model's *failure rate* is used as a basis for statistical analyses when assessing the quality of the model. The failure rate is based on the "hit" sequence of historical losses that have exceeded the respective VaR estimates over a given observation period. Following Campbell (2005), when the daily profit and loss figure of the portfolio is denoted as $x_{t+1}$, the hit function can be presented as follows:

$$I_{t+1}(\alpha) = \begin{cases} 1, x_{t,t+1} \leq \ -VaR_t\ (\alpha) \\ 0, x_{t,t+1} > \ -VaR_t\ (\alpha) \end{cases} \tag{23}$$

The failure rate is defined as the number of violations divided by the total number of observations *T.* Hence, the hit ratio is an unbiased estimator of the probability of observing a violation so that

$$\frac{1}{T} I(\alpha) = \hat{\alpha} \tag{24}$$

where the number of exceptions is

$$I(\alpha) = \sum_{t=1}^{T} I_t(\alpha) \tag{25}$$

When the sample size increases so that $\hat{\alpha}$ converges to $\alpha$, the following relation should hold for an accurate model:

$$\alpha = 1 - c \tag{26}$$

where $c$ denotes the chosen confidence level. Thus, for example when using 95% confidence level, $\alpha$ should equal to 5%. Hence, the backtesting procedure resembles a Bernoulli trial in which an action with two possible outcomes is repeated numerous times and in which each outcome is independent from the prior outcomes. Therefore, the number of violations follows binomial probability distribution as follows:

$$f\big(I(\alpha)\big) = \binom{T}{I(\alpha)} \alpha^{I(\alpha)} (1-\alpha)^{T-I(\alpha)} \tag{27}$$

With sufficiently large sample size, the binomial distribution can be approximated with the normal distribution so that

$$z = \frac{\sqrt{T}(\hat{\alpha} - \alpha)}{\sqrt{\alpha(1-\alpha)}} \approx N(0,1) \tag{28}$$

The hypothesis tests could then be conducted based on the known sample distribution of $z$.

However, when conducting statistical analysis in either accepting or rejecting a null hypothesis, there is always a tradeoff between type I and type II errors. When validating soundness of a given VaR model, the null hypothesis refers to the goodness of the VaR model, type I error stands for a rejection of a sound model, and type II error, respectively, refers to not rejecting a deficient model. In the field of risk management, incurring type II errors can be very costly and therefore a high threshold should be applied when accepting validity of a VaR model.

## 4.5 Tests of unconditional coverage

The purpose of tests of unconditional coverage is to determine whether the hit sequence generated by a VaR model satisfies the unconditional property i.e. the aim is to study if the sequence contains a tolerable amount of exceptions or not. The tests employed in this study that belong to the category of unconditional tests include the proportion of failures tests and the regulatory Basel traffic light test.

### 4.5.1 Proportion of failures test

The leading idea behind tests of unconditional coverage is to test whether the observed failure rate is consistent with the expected failure rate indicated by the confidence level. A commonly used test based solely on the failure rate and confidence interval is a *proportion of failures (POF) test* proposed by Kupiec (1995). In the POF-test it is assumed that the number of violations follows the binomial distribution, and the null hypothesis for a correct model is

$$H_0: \alpha = \hat{\alpha} = \frac{I(\alpha)}{T} \qquad (29)$$

Respectively, the null hypothesis is tested against an alternative hypothesis $H_A$:

$$H_A: \alpha \neq \hat{\alpha} \qquad (30)$$

Consequently, the test aims to provide an answer to the question whether the observed failure rate significantly differs from the expected rate and it can be performed as a likelihood-ratio test that expresses how many times more likely the observed data are under the null model compared to the alternative model. More specifically, the ratio to be investigated is the maximum probability of the observed result under the null hypothesis divided by the maximum probability of the observed result under the alternative hypothesis. The logarithm of the computed ratio is assumed to be asymptotically chi-square ($\chi^2$) distributed with one degree of freedom and thus the obtained test statistic is compared to a critical value obtained from $\chi^2$ distribution. The smaller the ratio is, the higher the value of the test statistic becomes, which leads to rejection of the null hypothesis if the critical value is exceeded.

The POF statistic is of the following form:

$$LR_{POF} = 2\log\left(\frac{(1-\hat{\alpha})^{T-I(\alpha)}\hat{\alpha}^{I(\alpha)}}{[1-(\alpha)]^{T-I(\alpha)}(\alpha)^{I(\alpha)}}\right) \qquad (31)$$

While calculating the log-likelihood ratio is a purely quantitative exercise, choosing which confidence level to use in rejecting the model resembles more art than science. Therefore, even though the chosen level should balance the probability of committing type I and type II errors, the decision is often an arbitrary one. For instance, 95% test confidence level implies that the model will be rejected only if the evidence against is fairly strong, and with 99% confidence level the evidence against a given model should be very strong before it is rejected.

**Table 6:**
**Non-rejection ranges for Proportion of Failures –test**

This table shows non-rejection ranges for a VaR model with different chosen VaR confidence and test confidence levels with samples sizes of 250 and 1 000. Probability level $\alpha$ is the expected proportion of failures, or exceptions, under a given VaR confidence level.

| | | Non-rejection range for number of exceptions y | | | |
|---|---|---|---|---|---|
| | Test confidence level | 95 % | | 99 % | |
| VaR Confidence Level | Probability Level p | T = 250 | T= 1 000 | T = 250 | T= 1 000 |
| 99 % | 1 % | $0 \leq y \leq 6$ | $5 \leq y \leq 16$ | $0 \leq y \leq 7$ | $4 \leq y \leq 19$ |
| 95 % | 5 % | $7 \leq y \leq 19$ | $38 \leq y \leq 64$ | $5 \leq y \leq 22$ | $34 \leq y \leq 68$ |
| 90 % | 10 % | $17 \leq y \leq 34$ | $82 \leq y \leq 119$ | $14 \leq y \leq 38$ | $77 \leq y \leq 125$ |

However, while POF test is relatively simple to implement, it suffers from two rather major shortcomings. First, the test is not statistically powerful with small sample sizes. For example, with 250 observations the acceptance percentage at 95% confidence level for a VaR model using 99% confidence level is

$$\left[\frac{0}{250}, \frac{6}{250}\right] = [0\%, 2.4\%]$$

whereas with sample size of 1000 the respective region is

$$\left[\frac{5}{1000}, \frac{16}{1000}\right] = [0.5\%, 1.6\%]$$

Hence, the relative acceptance region for the smaller sample size is substantially wider than in the latter case, which means that rejecting an inaccurate model becomes harder as sample size decreases.

The second shortcoming is that the test does not account for time variation in the observed exceptions, which implies that a supposedly accurate model that generates an acceptable amount of exceptions could still fail in capturing market volatility and correlations. For that reason Christoffersen (1998) proposes complementing unconditional coverage tests with independence tests that take into account possible clustering of the observed exceptions.

### 4.5.2 Regulatory test

The Basel test is simple to implement as it is based only on the number of exceptions during the preceding 250 business days. However, since the backtesting period is shorter than 250 days and the VaR confidence level is 95% instead of 99%, the number of exceptions that define the color of the category in which the model belongs are different from shown in Table 1. As a general rule for the test, the yellow and red zones begin at the points where the cumulative probability of obtaining a given number or fewer exceptions when the model is correct exceeds 95% or 99.99% respectively.

However, the three zone approach adopted by the Committee has the same limitations as the POF test by Kupiec, namely the statistical weakness as well as ignorance of possible time dependence of the observed exceptions. This implies that the test can hardly be used for comparing alternative models while it may on some level serve its purpose of providing a straightforward framework for model validation.

## 4.6 Tests of conditional coverage

While the POF test and the Basel three zone tests use the ratio of observed exceptions as the only input, tests of unconditional coverage are designed to account also for the time variation of the exceptions. Despite the fact that the clustering of VaR violations is not considered in the regulatory backtesting procedure, Christoffersen and Pelletier (2004) further emphasize that it should actually receive more attention since successive large losses are more likely to lead to a bankruptcy. Moreover, if the VaR violations are clustered in time and also across different banks, as Berkowitz and O'Brien (2002) find, it may be a significant source of systemic risk. Therefore, it can be argued that the clustering of VaR exceptions could also be used as basis for rejecting a given VaR model.

Chirstoffersen's (1998) Markov test is one of the first tests of conditional coverage. It is designed to examine whether or not the probability of a VaR exception is dependent on whether or not a VaR exception has occurred on the previous day. For a sound model, the probability of a VaR exception on a given day should be independent of whether or not an exception has taken place on the preceding day.

The test employs a two-state Markov process, and it is conducted through creating a 2 by 2 contingency matrix recording portfolio's VaR exceptions on successive days as shown below in Table 7. If the exceptions are independently distributed in time, the proportion of exceptions subsequent to a day when no exception has occurred $\pi_1 = \frac{N_3}{N_1 + N_3}$ should be the same as proportion of exceptions following a day when an exception has occurred $\pi_2 = \frac{N_4}{N_2 + N_4}$. Moreover, in order to satisfy the unconditional coverage property, the ratio of total number of exceptions should equal to the ratio indicated by the VaR level so that $\Pi = \frac{N_3 + N_4}{N} = \alpha$. Consequently, the following null hypothesis can be evaluated:

$$H_0: \pi_1 = \pi_2 \qquad (32)$$

**Table 7:**
**Contingency matrix for Markov independence test**

$N_1$ is the number of observations when VaR estimate is not exceeded subsequent to a day when an exception has not occurred and $N_2$ is the number of no-exception observations subsequent to days when exceptions have occurred. Correspondingly, $N_3$ and $N_4$ are the number of exceptions on days following no-exception and exception days.

| | $I_{t-1} = 0$ | $I_{t-1} = 1$ | |
|---|---|---|---|
| $I_t = 0$ | $N_1$ | $N_2$ | $N_1 + N_2$ |
| $I_t = 1$ | $N_3$ | $N_4$ | $N_3 + N_4$ |
| | $N_1 + N_3$ | $N_2 + N_4$ | $N$ |

The test statistic is again a log-likelihood ratio, and it is of the following form:

$$LR_{ind} = 2\log\left(\frac{(1-\pi_1)^{N_1+N_2}\pi_1^{N_3}(1-\pi_2)^{N_2}\pi_2^{N_4}}{(1-\Pi)^{N_1+N_2}\Pi^{(N_3+N_4)}}\right) \qquad (33)$$

$LR_{ind}$ is also chi-square distributed with one degree of freedom. Moreover, it can be combined with the POF test to obtain a joint test for studying simultaneously both the unconditional coverage and independence properties of the exception series. The combined statistic is simply the following:

$$LR_{cc} = LR_{POF} + LR_{ind} \tag{34}$$

and it is also chi-square distributed but with two degrees of freedom (Christoffersen, 1998). However, in order to gain more insight about the backtesting results, the tests are utilized separately in this study.

## 4.7 Loss function based backtesting

Using merely the information derived from the hit ratio and sequence of VaR series disregards substantial amount of data that could be used for evaluating the precision of a given model. Moreover, although testing unconditional and conditional properties assists in ruling out deficient models, the test results cannot be used for ranking the models. Lopez (1999) proposes a loss function based forecast evaluation framework in order to overcome this shortcoming of previous backtesting methods. The proposed loss function is as follows:

$$\psi_{t+1} = \begin{cases} 1 + (VaR_t(\alpha) - x_{t+1})^2 & , I_{t+1}(\alpha) = 1 \\ 0 & , I_{t+1}(\alpha) = 0 \end{cases} \tag{35}$$

The loss function is designed to account for the magnitude of the tail losses and a score of one is added when an exception is observed. Consequently, the model with the lowest total loss, $\sum_{t=1}^{T} \psi_t$, is preferred. However, as the author suggests, the loss function result cannot be used for separating accurate models from inaccurate since a model that generates zero exceptions would be deemed as the most accurate. Hence, the loss function result should be used for comparing the relative accuracy of a given number of alternative models. Nevertheless, as Angelidis et al. (2007) point out, the loss $x_{t+1}$ exceeding $VaR_t(\alpha)$ should actually be compared to expected shortfall $ES(\alpha)$measure and not to VaR estimate since the latter does not give any indication about the magnitude of the expected loss. For that reason, the authors suggest the following adjustments to the loss function:

$$\psi_{1,t+1} = \begin{cases} \left|x_{t+1} - ES(\alpha)^{(i)}{}_t\right| , & I_{t+1}(\alpha) = 1 \\ 0 \quad , & I_{t+1}(\alpha) = 0 \end{cases} \tag{36}$$

and

$$\psi_{2,t+1} = \begin{cases} \left(x_{t+1} - ES(\alpha)^{(i)}{}_t\right)^2, & I_{t+1}(\alpha) = 1 \\ 0 \qquad , & I_{t+1}(\alpha) = 0 \end{cases} \tag{37}$$

so that for each model $i$ the mean absolute deviation (MAD), $\frac{1}{T}\sum_{t=1}^{T}\psi_1$, and mean squared error (MSE), $\frac{1}{T}\sum_{t=1}^{T}\psi_2$, are computed. Moreover, the total loss value given a VaR exception is

$$L_i = \sum_{t=1}^{T}\psi_{l,t}^{(i)} \tag{38}$$

The authors further propose applying Hansen's (2005) test for superior predictive ability (SPA) for studying the statistical significance of the differences between a benchmark model and an alternative model. The null hypothesis that the benchmark model $i*$ is not outperformed by alternative models *i,* for *i*=1,...,*M,* is tested with the following statistic:

$$T_l^{SPA} = \max_{i=1,\dots,M}\frac{\sqrt{M}\bar{X}_{l,i}}{\sqrt{Var(\sqrt{M}\,\bar{X}_{l,i})}} \tag{39}$$

where

$$T_l^{SPA} = \max_{i=1,\dots,M}\frac{\sqrt{M}\bar{X}_{l,i}}{\sqrt{Var(\sqrt{M}\,\bar{X}_{l,i})}} \tag{40}$$

$$\bar{X}_{l,i} = \frac{1}{T}\sum_{t=1}^{T} L_{i,t} - L_{i^*,t} \tag{41}$$

Following Angelidis et al. (2007), the estimation of $\sqrt{Var(\sqrt{M}\,\bar{X}_{l,i})}$ and *p*-values for the test statistic are obtained by utilizing the stationary bootstrap of Politis and Romano (1994). Moreover, the block-length used in the bootstrap is obtained using the automatic block-length selection algorithm presented by Politis et al. (2009).

# 5 RESULTS

This section presents the empirical findings of my thesis. The section starts with an overview of the VaR estimation results and then proceeds to more detailed assessment of the differences between the tested models. A list of abbreviations used in the tables as well as short descriptions of the models introduced in section 2 are presented in Appendix A for a quick review.

## 5.1 VaR results

Table 8 and Table 9 show the average VaR figures at 95% confidence level for each of the swaption contracts during the backtesting period. In the following sections, models in which the daily volatility moves are measured from a given strike level are labeled with a prefix "Proxy". Correspondingly, the models in which the implied volatility changes are determined from the daily differences for a specific contract are tagged with term "Direct". Additionally, "ATM" and "Smile" in the column headers refer to the moneyness level from which the implied volatility changes are observed. Hence, "ATM" refers to the Fixed smile method and "Smile" refers to the Random smile method. The former table contains the estimates for long positions and the latter for short positions. Results for 90% confidence level are left unreported as they provide no additional information regarding the performance of the models. Furthermore, since no visible differences arise between estimates drawn from a fitted distribution versus those obtained by interpolating from the discrete scenarios, the results presented in the following chapters are acquired without the fitting method described in 4.2.4.

While investigating the average VaR estimates tells little from the actual performance of the models, it reveals some information about the differing risk levels that the models report. On average, FHS models appear to generate the lowest estimates while HS and TW models seem to provide rather similar figures. The lower level of FHS estimates reflects the fact that changes in the risk factors tend to be quite modest for most part of the backtesting period. Furthermore, the estimates for short positions are universally higher than for the long positions. This phenomenon is explained by the skewed risk factor returns during both the observation and the backtesting period as a whole. The numbers could be different at least for HS models in case the historical observation period was longer and contained a broader scale of different scenarios.

However, average numbers provide only one perspective of the differences among the considered models and do not tell much about the models' reactiveness to market fluctuations, for instance.

A graphical illustration in Figure 8 provides an example of the differences between HS, TW and FHS models for long and short 5x2 ATM receiver positions. The figure shows how the estimates obtained with TW and FHS models follow more closely the realized profit and loss numbers, whereas the level of HS estimates is more static.

While Figure 8 shows that the differences concerning the choice of VaR method are clearly visible, the differences between fixed and random models are less pronounced and depend on other model specifications. Figure 9 shows an instance where the difference between smile models is noticeable for the long position, while the VaR figures for the short position are nearly equal for the short position. However, when using HS model with "Proxy" method for implied volatility fluctuations, the difference between using ATM level or moneyness-dependent changes is negligible. One reason behind this lies in the VaR components: the risk arising from interest rate fluctuations is clearly higher than that of stemming from changes in implied volatilities. The relative differences of the components are illustrated in Figure 10 and average VaR components for Proxy_HS are presented in Table 19 (Appendix C).

Nevertheless, a more detailed assessment is still needed to systematically evaluate how the models succeed, which is the main theme of the subsequent chapters.

## Table 8:
## Average long receiver swaption VaR estimates

The table presents average VaR estimates over the backtesting period for long receiver swaption contracts and for the different models studied. VaR figures are absolute values in euros for swaption contracts with 10 MEUR notional. The table columns are sorted by VaR method (Historical Simulation, Time Weighted and Filtered Historical Simulation), by volatility change measurement (proxy, direct or component) and by used moneyness (ATM or Smile). The number of daily observations from which the averages are computed is 228.

| | Proxy_HS | | Proxy_TW | | Direct_HS | | Direct_TW | | Component_HS | | Proxy_FHS | | Direct_FHS | | Component_FHS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | | | | | | | | | | | | |
| ATM | 7 157 | | 6 540 | | 6 577 | | 5 970 | | 5 794 | | 5 383 | | 4 881 | | 4 103 | |
| ATM +25 bps | 9 907 | 9 832 | 8 832 | 8 693 | 9 234 | 9 621 | 8 194 | 9 114 | 9 240 | 9 357 | 7 248 | 7 036 | 6 590 | 7 192 | 6 648 | 7 066 |
| ATM +50 bps | 11 950 | 11 529 | 10 300 | 10 114 | 11 274 | 11 638 | 9 742 | 10 447 | 11 223 | 11 488 | 8 518 | 8 411 | 7 879 | 8 551 | 7 911 | 8 462 |
| ATM +100 bps | 13 933 | 13 521 | 11 621 | 11 416 | 13 568 | 13 693 | 11 309 | 11 694 | 13 638 | 13 732 | 9 708 | 9 693 | 9 261 | 9 747 | 9 233 | 9 725 |
| ATM +200 bps | 15 261 | 14 988 | 12 367 | 12 230 | 15 077 | 14 912 | 12 234 | 12 115 | 14 840 | 14 961 | 10 259 | 10 216 | 10 045 | 10 093 | 9 757 | 10 123 |
| Panel B: 1x5 | | | | | | | | | | | | | | | | |
| ATM | 20 523 | | 20 133 | | 18 924 | | 18 233 | | 16 173 | | 16 277 | | 14 334 | | 12 802 | |
| ATM +25 bps | 26 397 | 26 359 | 25 704 | 25 555 | 24 651 | 25 836 | 23 638 | 25 637 | 23 683 | 24 670 | 20 267 | 19 985 | 18 324 | 19 659 | 18 263 | 19 007 |
| ATM +50 bps | 31 647 | 31 393 | 30 044 | 29 681 | 29 836 | 32 496 | 28 203 | 30 705 | 28 992 | 31 217 | 23 383 | 23 367 | 21 509 | 23 966 | 21 633 | 23 188 |
| ATM +100 bps | 38 348 | 38 239 | 35 467 | 35 171 | 36 895 | 39 513 | 34 125 | 36 388 | 36 370 | 39 299 | 27 036 | 27 098 | 25 760 | 27 874 | 25 489 | 27 482 |
| ATM +200 bps | 43 235 | 43 233 | 39 471 | 39 193 | 42 618 | 42 532 | 38 907 | 38 621 | 41 358 | 43 163 | 29 697 | 29 845 | 29 176 | 29 062 | 28 011 | 29 179 |
| Panel C: 5x2 | | | | | | | | | | | | | | | | |
| ATM | 10 560 | | 10 731 | | 9 216 | | 9 429 | | 8 387 | | 8 801 | | 7 648 | | 6 939 | |
| ATM +25 bps | 11 671 | 11 685 | 11 885 | 11 843 | 10 222 | 11 194 | 10 486 | 11 726 | 10 506 | 10 998 | 9 639 | 9 740 | 8 484 | 9 429 | 8 932 | 9 542 |
| ATM +50 bps | 12 707 | 12 697 | 12 978 | 12 814 | 11 275 | 12 773 | 11 541 | 13 203 | 11 305 | 12 589 | 10 397 | 10 274 | 9 297 | 10 781 | 9 608 | 10 802 |
| ATM +100 bps | 14 553 | 14 469 | 14 810 | 14 538 | 13 162 | 15 090 | 13 425 | 15 404 | 12 703 | 15 052 | 11 730 | 11 633 | 10 709 | 12 493 | 10 475 | 12 696 |
| ATM +200 bps | 17 421 | 17 060 | 17 583 | 17 101 | 16 208 | 16 087 | 16 339 | 15 602 | 11 944 | 15 290 | 13 638 | 13 508 | 12 772 | 12 393 | 8 843 | 11 946 |
| Panel D: 5x5 | | | | | | | | | | | | | | | | |
| ATM | 27 704 | | 28 756 | | 24 418 | | 25 503 | | 21 059 | | 22 649 | | 20 316 | | 17 734 | |
| ATM +25 bps | 30 341 | 30 109 | 31 612 | 30 904 | 27 041 | 28 059 | 28 267 | 28 927 | 24 595 | 25 939 | 24 797 | 24 528 | 22 507 | 23 925 | 20 844 | 22 156 |
| ATM +50 bps | 33 114 | 32 839 | 34 300 | 34 200 | 29 468 | 32 813 | 30 998 | 34 377 | 28 304 | 30 923 | 26 765 | 26 707 | 24 662 | 27 938 | 24 089 | 27 010 |
| ATM +100 bps | 38 167 | 37 753 | 39 216 | 38 896 | 34 376 | 39 211 | 35 920 | 40 315 | 32 404 | 38 218 | 30 237 | 30 434 | 28 175 | 32 438 | 27 534 | 33 177 |
| ATM +200 bps | 45 823 | 44 696 | 46 359 | 45 504 | 42 637 | 42 303 | 43 405 | 42 758 | 34 709 | 42 922 | 35 295 | 34 902 | 33 416 | 33 608 | 26 576 | 34 922 |
| Panel E: 10x2 | | | | | | | | | | | | | | | | |
| ATM | 11 554 | | 11 730 | | 9 950 | | 10 114 | | 8 248 | | 8 954 | | 8 071 | | 7 001 | |
| ATM +25 bps | 12 543 | 12 623 | 12 698 | 12 609 | 10 815 | 11 343 | 11 057 | 11 414 | 9 279 | 10 035 | 9 712 | 9 634 | 8 712 | 8 949 | 7 816 | 8 371 |
| ATM +50 bps | 13 504 | 13 628 | 13 673 | 13 606 | 11 781 | 13 733 | 12 002 | 13 581 | 10 861 | 12 487 | 10 454 | 10 272 | 9 333 | 10 704 | 9 196 | 10 362 |
| ATM +100 bps | 15 297 | 15 107 | 15 471 | 15 207 | 13 699 | 15 856 | 13 791 | 15 889 | 12 629 | 15 611 | 11 740 | 11 567 | 10 673 | 12 411 | 10 471 | 12 744 |
| ATM +200 bps | 17 983 | 18 245 | 18 431 | 18 173 | 16 689 | 17 384 | 16 883 | 17 222 | 13 520 | 17 630 | 13 919 | 13 752 | 13 096 | 13 445 | 10 376 | 14 258 |
| Panel F: 10x5 | | | | | | | | | | | | | | | | |
| ATM | 28 744 | | 29 437 | | 24 221 | | 25 517 | | 20 114 | | 22 743 | | 19 184 | | 16 145 | |
| ATM +25 bps | 31 101 | 30 901 | 31 880 | 31 601 | 26 491 | 28 020 | 27 903 | 28 908 | 22 517 | 24 044 | 24 666 | 24 580 | 20 947 | 22 485 | 18 204 | 19 828 |
| ATM +50 bps | 33 352 | 33 005 | 34 213 | 33 886 | 28 871 | 32 932 | 30 177 | 34 052 | 26 269 | 29 792 | 26 472 | 26 414 | 22 730 | 26 511 | 21 550 | 25 273 |
| ATM +100 bps | 37 579 | 37 168 | 38 650 | 38 545 | 33 302 | 38 503 | 34 604 | 39 971 | 30 773 | 37 513 | 29 577 | 29 300 | 26 303 | 32 131 | 24 930 | 32 245 |
| ATM +200 bps | 45 360 | 44 099 | 46 134 | 45 289 | 40 938 | 42 150 | 42 240 | 42 840 | 32 996 | 43 549 | 35 044 | 34 221 | 32 630 | 33 174 | 25 416 | 35 466 |

## Table 9:
## Average short receiver swaption VaR estimates

The table presents average VaR estimates over the backtesting period for short receiver swaption contracts and for the different models studied. VaR figures are absolute values in euros for swaption contracts with 10 MEUR notional. The table columns are sorted by VaR method (Historical Simulation, Time Weighted and Filtered Historical Simulation), by volatility change measurement (proxy, direct or component) and by used moneyness (ATM or Smile). The number of daily observations from which the averages are computed is 228.

| | Proxy_HS | | Proxy_TW | | Direct_HS | | Direct_TW | | Component_HS | | Proxy_FHS | | Direct_FHS | | Component_FHS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | | | | | | | | | | | | |
| ATM | 8 794 | | 6 694 | | 8 247 | | 6 774 | | 7 791 | | 6 751 | | 5 897 | | 5 643 | |
| ATM +25 bps | 11 669 | 11 464 | 8 870 | 8 641 | 11 151 | 10 872 | 8 932 | 8 570 | 9 808 | 10 120 | 8 567 | 8 332 | 7 946 | 7 661 | 6 736 | 7 048 |
| ATM +50 bps | 13 652 | 13 573 | 10 310 | 9 886 | 13 201 | 12 973 | 10 369 | 9 772 | 12 230 | 12 678 | 9 767 | 9 849 | 9 209 | 9 040 | 8 326 | 8 768 |
| ATM +100 bps | 15 692 | 15 637 | 11 653 | 11 489 | 15 413 | 15 392 | 11 683 | 11 445 | 15 004 | 15 403 | 10 812 | 10 907 | 10 402 | 10 444 | 10 020 | 10 396 |
| ATM +200 bps | 16 945 | 16 911 | 12 378 | 12 539 | 16 839 | 17 033 | 12 386 | 12 570 | 16 933 | 17 186 | 11 264 | 11 269 | 11 096 | 11 372 | 11 210 | 11 391 |
| Panel B: 1x5 | | | | | | | | | | | | | | | | |
| ATM | 24 136 | | 17 930 | | 21 164 | | 18 427 | | 19 594 | | 18 126 | | 16 249 | | 15 018 | |
| ATM +25 bps | 29 933 | 29 894 | 22 691 | 21 959 | 26 998 | 26 732 | 23 125 | 22 201 | 23 662 | 24 440 | 22 094 | 21 967 | 20 127 | 19 942 | 17 139 | 18 349 |
| ATM +50 bps | 34 405 | 34 231 | 26 555 | 25 439 | 31 776 | 31 045 | 26 863 | 25 412 | 28 640 | 30 058 | 25 150 | 25 489 | 23 510 | 23 103 | 20 564 | 22 359 |
| ATM +100 bps | 39 867 | 39 589 | 31 196 | 30 908 | 38 040 | 38 484 | 31 466 | 31 200 | 36 450 | 38 504 | 29 206 | 29 517 | 28 087 | 28 263 | 26 498 | 28 512 |
| ATM +200 bps | 44 096 | 43 941 | 34 844 | 35 706 | 43 136 | 45 075 | 34 900 | 35 676 | 43 549 | 45 696 | 32 411 | 32 810 | 31 951 | 33 505 | 32 348 | 33 576 |
| Panel C: 5x2 | | | | | | | | | | | | | | | | |
| ATM | 11 939 | | 8 229 | | 9 882 | | 8 044 | | 7 863 | | 9 841 | | 8 067 | | 6 312 | |
| ATM +25 bps | 13 096 | 12 978 | 9 113 | 8 709 | 11 026 | 10 785 | 8 963 | 8 151 | 7 311 | 7 804 | 10 718 | 10 562 | 8 894 | 8 490 | 5 696 | 6 190 |
| ATM +50 bps | 14 177 | 14 049 | 9 994 | 9 480 | 12 127 | 11 601 | 9 790 | 9 164 | 8 354 | 9 458 | 11 499 | 11 833 | 9 622 | 9 404 | 6 473 | 7 581 |
| ATM +100 bps | 15 973 | 15 793 | 11 479 | 11 604 | 13 943 | 14 083 | 11 294 | 11 518 | 11 030 | 13 184 | 12 884 | 13 373 | 10 963 | 11 570 | 8 512 | 10 806 |
| ATM +200 bps | 18 624 | 18 425 | 13 755 | 17 544 | 16 996 | 20 501 | 13 632 | 18 058 | 18 552 | 22 075 | 14 908 | 15 565 | 13 142 | 17 485 | 15 073 | 18 534 |
| Panel D: 5x5 | | | | | | | | | | | | | | | | |
| ATM | 30 158 | | 20 334 | | 24 800 | | 20 375 | | 20 125 | | 25 026 | | 20 427 | | 17 005 | |
| ATM +25 bps | 32 969 | 32 792 | 22 718 | 22 601 | 27 668 | 27 764 | 22 692 | 22 400 | 21 246 | 22 799 | 27 384 | 27 220 | 22 607 | 22 316 | 17 652 | 19 027 |
| ATM +50 bps | 35 666 | 35 477 | 24 891 | 24 328 | 30 317 | 29 577 | 24 885 | 24 156 | 22 555 | 25 479 | 29 595 | 30 112 | 24 764 | 24 426 | 18 344 | 21 247 |
| ATM +100 bps | 40 185 | 40 491 | 28 878 | 29 658 | 35 398 | 36 494 | 28 870 | 29 716 | 28 300 | 34 240 | 33 358 | 34 442 | 28 772 | 30 361 | 22 901 | 28 841 |
| ATM +200 bps | 47 740 | 48 000 | 35 372 | 43 318 | 43 671 | 51 561 | 35 393 | 43 519 | 45 073 | 54 128 | 38 717 | 39 065 | 34 722 | 42 566 | 36 498 | 44 792 |
| Panel E: 10x2 | | | | | | | | | | | | | | | | |
| ATM | 12 945 | | 8 611 | | 10 450 | | 8 524 | | 8 251 | | 9 912 | | 8 182 | | 6 665 | |
| ATM +25 bps | 13 969 | 13 851 | 9 422 | 9 624 | 11 467 | 11 706 | 9 301 | 9 351 | 8 896 | 9 664 | 10 638 | 10 539 | 8 905 | 9 037 | 6 979 | 7 700 |
| ATM +50 bps | 14 952 | 14 898 | 10 220 | 9 698 | 12 458 | 12 263 | 10 043 | 9 568 | 9 048 | 10 682 | 11 370 | 11 511 | 9 619 | 9 508 | 6 836 | 8 279 |
| ATM +100 bps | 16 733 | 16 709 | 11 679 | 11 870 | 14 376 | 15 017 | 11 540 | 11 654 | 11 113 | 14 321 | 12 774 | 12 713 | 10 944 | 11 466 | 8 292 | 11 265 |
| ATM +200 bps | 19 733 | 19 723 | 14 155 | 17 404 | 17 651 | 21 759 | 14 022 | 17 480 | 17 659 | 22 516 | 15 174 | 14 882 | 13 346 | 16 701 | 13 860 | 18 169 |
| Panel F: 10x5 | | | | | | | | | | | | | | | | |
| ATM | 33 274 | | 20 646 | | 27 202 | | 20 575 | | 21 126 | | 25 498 | | 20 724 | | 16 085 | |
| ATM +25 bps | 35 856 | 35 562 | 22 614 | 22 616 | 29 789 | 29 496 | 22 520 | 22 620 | 22 779 | 24 692 | 27 398 | 26 869 | 22 512 | 22 195 | 16 944 | 18 761 |
| ATM +50 bps | 38 410 | 38 358 | 24 556 | 23 936 | 32 388 | 31 952 | 24 374 | 23 714 | 23 355 | 27 825 | 29 200 | 29 621 | 24 283 | 24 263 | 16 914 | 20 917 |
| ATM +100 bps | 43 210 | 42 870 | 28 228 | 29 169 | 37 170 | 38 695 | 28 154 | 29 087 | 28 757 | 37 010 | 32 514 | 33 512 | 27 714 | 29 557 | 20 862 | 28 485 |
| ATM +200 bps | 51 145 | 51 065 | 35 004 | 44 188 | 45 338 | 56 223 | 34 949 | 44 123 | 45 137 | 58 079 | 38 307 | 38 962 | 34 025 | 43 710 | 34 723 | 46 624 |

**Figure 8:**
**VaR estimates and actual Profit and Loss figures**

The estimates given by the HS, TW and FHS VaR models are shown with the actual daily Profit and Loss (PL) during the backtesting period. The long position estimates are shown below as negative numbers. The contract in question is a short position in 5x2 swaption with ATM strike.

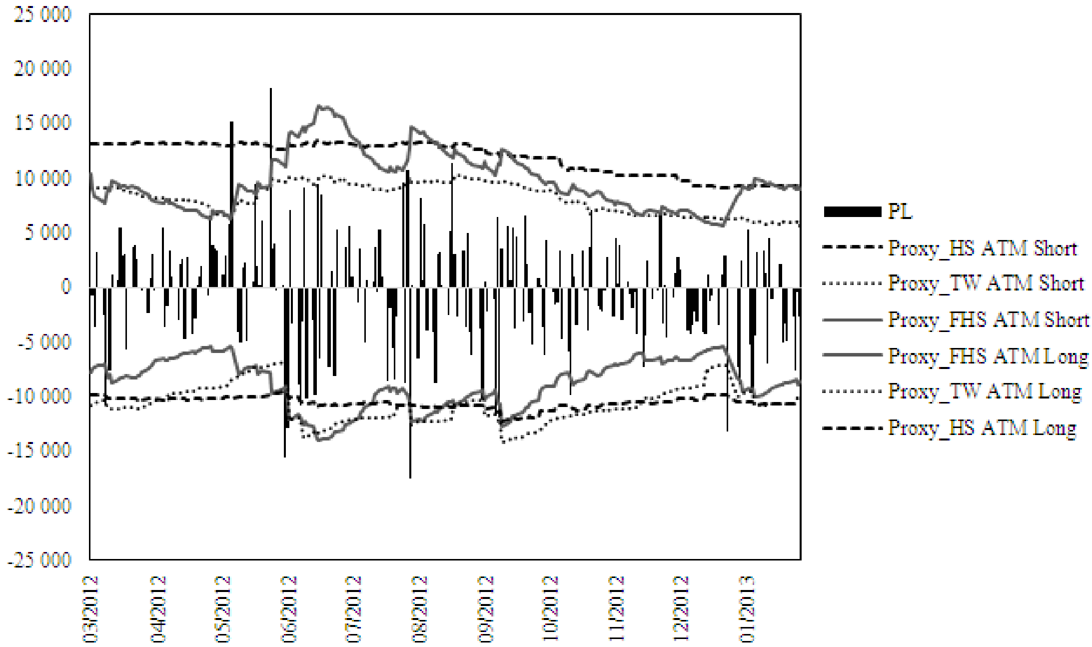



**Figure 9:**
**The difference between smile methods**

The graph depicts Direct_TW VaR estimtes for a 10x2 swaption with strike ATM +50 bps using fixed ("ATM") and random smile ("Smile") methods. Estimates for the long position are in the lower part of the figure and estimates for the short position are in the upper part. While the difference is not visually significant, the respective number of exceptions are 14 and 7 for long position in the favour of random smile method. For the short position, the respective counts of hits are 6 and 8.

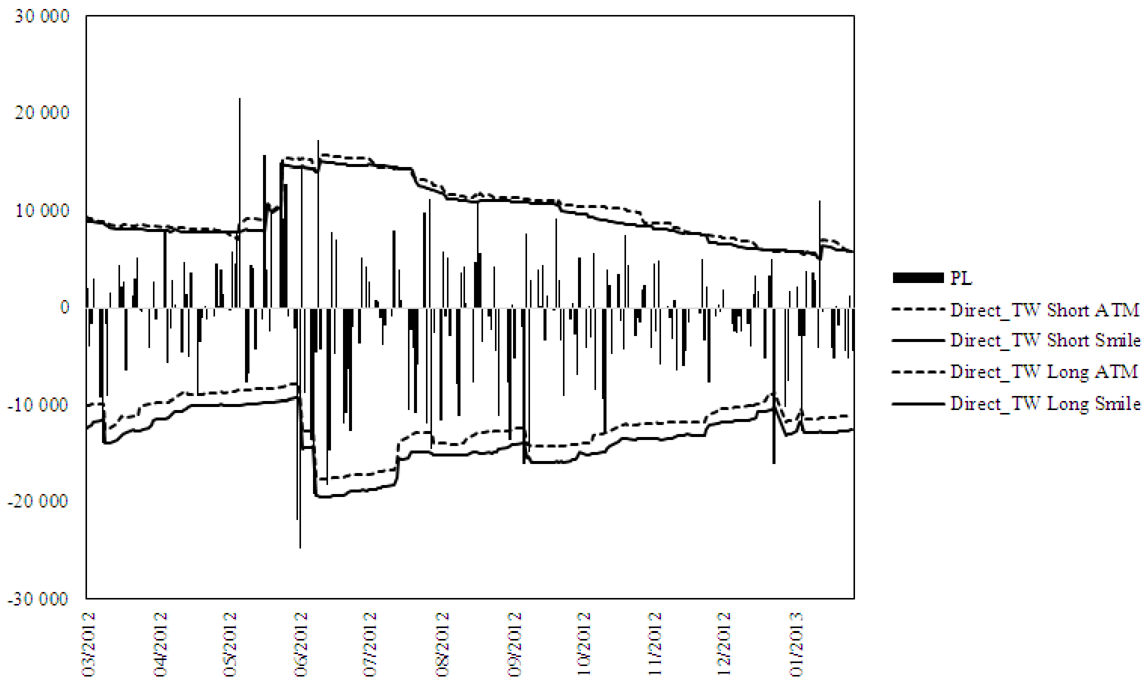
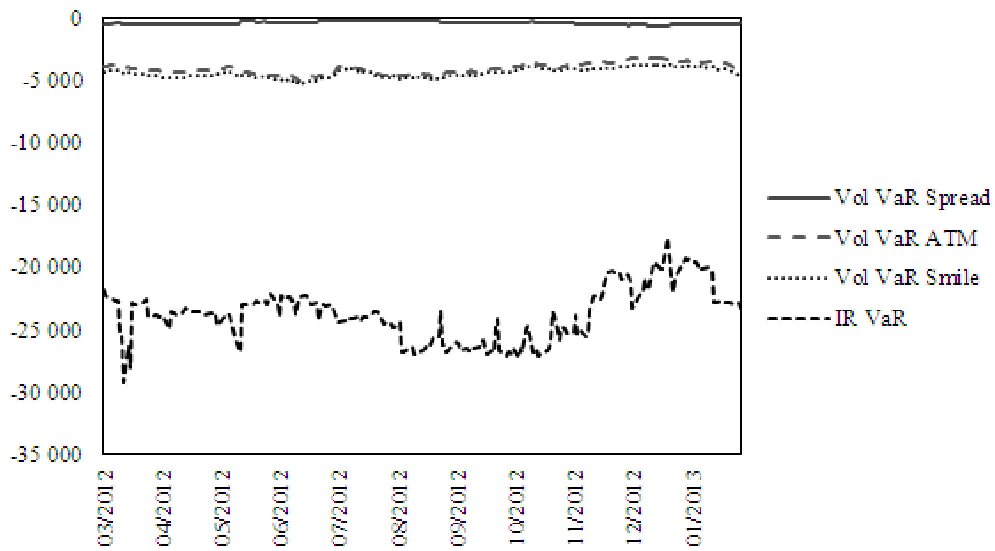


**Figure 10:**
**VaR components**

The figure illustrates the VaR risk components for long 1x5 ATM + 25 bps receiver position using HS_Proxy with fixed and random smile methods. IR stands for interest rate risk and ATM and Smile Vols are volatility components of the respective smile methods. Vol VaR Spread shows the spread between the smile methods. Interest rate risk is naturally same in the both methods. The average risk components over the backtesting period are shown in Table 19 (Appendix C).




## 5.2 Stage one backtesting results

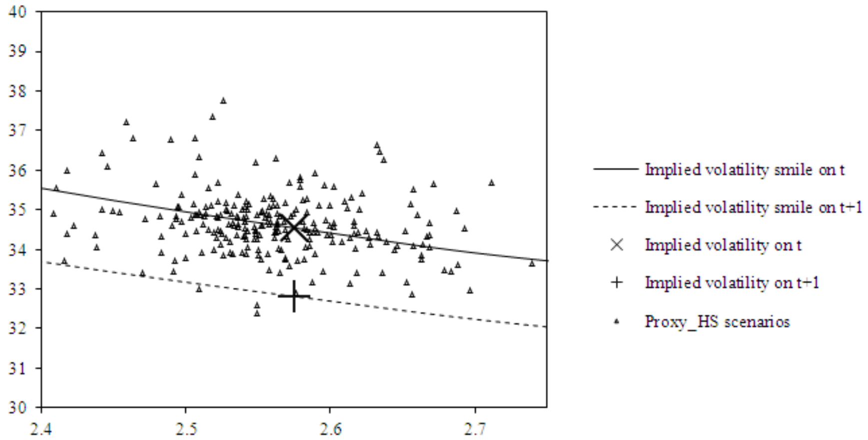
Table 10 presents the number of exceptions observed in each of the backtesting series estimated for long positions using 95% VaR confidence level. As the number of backtesting days is 228, the expected number of hits is approximately 11. Moreover, in order to obtain "Green" flag as a result from the regulatory three zone test a model is allowed to generate 16 exceptions with the parameters used. The results from the regulatory test are presented in Table 11. Additionally, to pass Kupiec's (1995) proportion of failures test, the acceptable region for number of exceptions is from 6 to 18 using 95% test confidence level. In this light, the VaR estimates tend to provide rather optimistic estimates: essentially each of the FHS models should be rejected on the basis of regulatory backtest and many of the HS and TW are either in the "Yellow" -zone or at the upper end of the acceptable region for "Green" test outcome. On the contrary, results for the short positions are on the other end of the scale, thus generating rather conservative VaR estimates. The results for short positions are presented in Appendix C. Consequently, based on the results for short positions, each of the models would be acceptable in terms of the regulatory test. However, one should not feel comfortable using a model whose performance depends on the direction of the position. Hence, FHS models are deemed inaccurate, and the rest of the analysis concentrates on the results for long positions models that have gained at least a “Yellow” outcome from the Basel test.

The HS and TW models generate the highest VaR estimates on average during the backtesting period, as shown in Table 8, which also results into least number of exceptions among the different models. The low number of hits for short receiver positions, and even zero in the case of 1x5 swaption contracts, is due to relatively sharp interest rate moves during the fall of 2011: the fluctuations remain in the historical distribution throughout most of the backtesting period for each of the considered models, but in the HS model, there is no weighing that would diminish the effect of older observations. However, due to the skewed distribution of historical returns, the estimates for long positions are correspondingly rather optimistic. On the contrary, TW models generate markedly more dynamic estimates that reflect and follow market movements more closely. Despite the noticeable rise in VaR figures during the early summer 2012 resulting from increased fear for further escalation of the Euro crisis, the numbers quickly return to lower levels, which results into lower overall VaR estimates for contracts with short maturities. Nevertheless, for swaptions with longer maturities there is no significant difference between average VaR estimates of HS and TW models. From the viewpoint of minimizing capital charges, it would be optimal to find a model with lowest average VaR estimates combined with the least number of exceptions. However, the differences between the models are relatively small and no such chance for optimization is available. Furthermore, on the one hand, VaR estimates that reflect the market fluctuations more closely are appropriate for monitoring the prevailing risk status, but on the other hand, using such vigorous numbers for risk limitation purposes is not very practical, and would most likely be objected by those subject to the VaR limits.

While the skew dependent model is perhaps the most theoretically appealing of the alternative models covered in this study, it nevertheless fails to provide sufficiently accurate estimates as shown by the backtesting results. The problem is again the distributional narrowness of historical observations that does not contain sharp moves in volatilities. Moreover, as the interest rates decline and implied volatilities increase on average during the observation period, most scenarios are positioned below the prevailing forward rate and somewhat above the prevailing volatility level if no other adjustments are applied. However, as the negative correlation between the interest rates and implied volatilities breaks for contracts with longer maturities more often during the backtesting period than within the whole observation period on average, this leads to profit and loss figures not covered by estimation scenarios. Hence, the models tend to generate excessive number of VaR breaches. Figure 11 provides an illustration of the scenarios formed using HS and skew dependent models with random smile method.

**Figure 11:**
**Model scenarios**

The graphs below illustrate scenarios generated by Proxy_HS and Component_HS models on June 5, 2012. The cross (X) shows the present forward rate and implied volatility pair for a 10x5 +25 bps receiver swaption and the plus (+) shows the corresponding location on the following date. The historical observations are concentrated below the prevailing forward rate in both models, but the locations of volatilities differ to a greater extent. While the scenarios in Component_HS model are take better into account the prevailing shape of the volatility smile on the estimation day (*t*), the scenarios do not contain sufficiently large volatility jumps to match the actual change to *t+1*, which leads to a VaR exception.



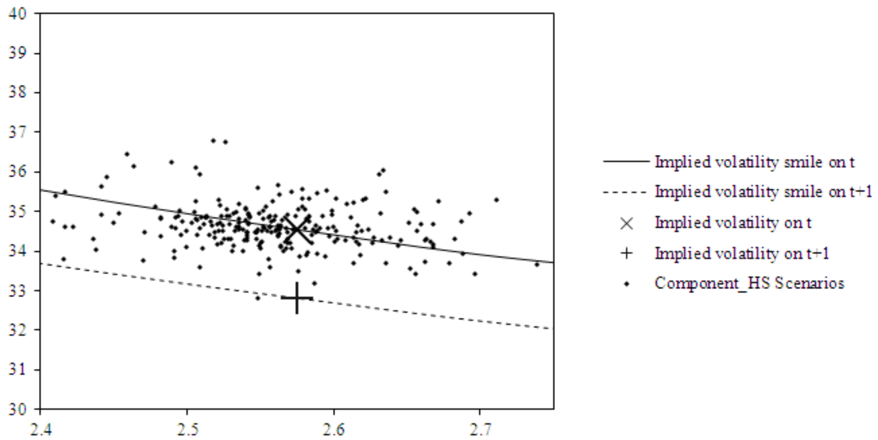




Table 12 shows the p-values from Kupiec (1995) test. The p-values represent the probability of a obtaining an observation that is even less likely than the number of hits found in Table 10 with 228 observations and 95% VaR confidence level if the model in question was accurate. In short, the results from proportion of failures test are similar to those of regulatory test. Consequently, Direct_TW with both smile methods and Proxy_TW with moneyness-dependent smile method, as well as Proxy_HS and Direct_HS with moneyness-dependent implied volatility changes pass the test with each of the swaption contracts. Additionally, Component_HS with moneyness-dependent smile methods perform rather well, but still yields unacceptable test results for 5x5 and 10x5 swaptions with specific strike levels.

In contrast to results from tests of unconditional coverage, not a single one of the studied models can be deemed accurate based on Christoffersen's (1998) test of conditional coverage. This, however, is not that surprising as models based on historical simulation usually suffer from clustering of exceptions, as for instance Pritsker (2001) points out. The explanation arises again from the slow responsiveness to changing market conditions. On the contrary to HS models, unreported results for FHS models reveal that nearly all of them would pass the test of conditional coverage resulting from their faster responsiveness. Also TW models appear to perform slightly better than HS models due to their relatively more dynamic nature. Nevertheless, neither of the TW models pass the test with all swaption contracts. The p-values of Christoffersen's (1998) test are presented in Table 13. The logic is equal to that of Kupiec's (1995) test p-values: the lower the figure, the more apparent it becomes that the null hypothesis of model accuracy should be rejected.

## Table 10:
## VaR exceptions

The table shows the number of VaR exceptions for each of the series when estimated for long receiver swaption positions. VaR confidence level is 95% and the number of daily observations is 228. Consequently, the expected proportion of exceptions is 5%, i.e. number of hits should be approximately 11.

| | Proxy_HS | | Proxy_TW | | Direct_HS | | Direct_TW | | Component_HS | | Proxy_FHS | | Direct_FHS | | Component_FHS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | | | | | | | | | | | | |
| ATM | 9 | | 10 | | 12 | | 14 | | 13 | | 16 | | 21 | | 28 | |
| ATM +25 bps | 14 | 13 | 15 | 16 | 14 | 13 | 17 | 13 | 13 | 13 | 25 | 28 | 31 | 24 | 26 | 25 |
| ATM +50 bps | 11 | 10 | 14 | 14 | 11 | 10 | 17 | 13 | 12 | 11 | 23 | 23 | 26 | 22 | 28 | 23 |
| ATM +100 bps | 10 | 10 | 15 | 15 | 10 | 10 | 15 | 13 | 10 | 10 | 21 | 21 | 24 | 21 | 24 | 22 |
| ATM +200 bps | 7 | 7 | 12 | 13 | 7 | 8 | 12 | 13 | 8 | 8 | 20 | 21 | 21 | 21 | 22 | 21 |
| Panel B: 1x5 | | | | | | | | | | | | | | | | |
| ATM | 9 | | 9 | | 13 | | 13 | | 17 | | 16 | | 21 | | 27 | |
| ATM +25 bps | 12 | 12 | 12 | 12 | 14 | 12 | 13 | 12 | 15 | 15 | 22 | 22 | 25 | 22 | 28 | 26 |
| ATM +50 bps | 14 | 14 | 12 | 13 | 15 | 12 | 14 | 12 | 15 | 14 | 23 | 24 | 28 | 23 | 28 | 25 |
| ATM +100 bps | 12 | 12 | 12 | 13 | 14 | 10 | 14 | 12 | 15 | 10 | 22 | 21 | 25 | 21 | 26 | 21 |
| ATM +200 bps | 9 | 9 | 11 | 12 | 11 | 11 | 13 | 13 | 12 | 9 | 21 | 22 | 22 | 22 | 23 | 22 |
| Panel C: 5x2 | | | | | | | | | | | | | | | | |
| ATM | 8 | | 8 | | 14 | | 11 | | 17 | | 12 | | 19 | | 23 | |
| ATM +25 bps | 14 | 13 | 10 | 10 | 21 | 15 | 14 | 10 | 22 | 14 | 19 | 19 | 28 | 22 | 23 | 21 |
| ATM +50 bps | 13 | 12 | 10 | 10 | 22 | 12 | 16 | 10 | 21 | 13 | 21 | 21 | 27 | 18 | 23 | 20 |
| ATM +100 bps | 14 | 13 | 11 | 12 | 20 | 10 | 14 | 10 | 21 | 11 | 21 | 22 | 26 | 17 | 24 | 18 |
| ATM +200 bps | 9 | 9 | 7 | 8 | 10 | 10 | 9 | 10 | 18 | 11 | 10 | 11 | 13 | 14 | 24 | 16 |
| Panel D: 5x5 | | | | | | | | | | | | | | | | |
| ATM | 10 | | 8 | | 12 | | 8 | | 19 | | 14 | | 19 | | 22 | |
| ATM +25 bps | 10 | 10 | 8 | 10 | 17 | 14 | 10 | 11 | 19 | 19 | 18 | 18 | 23 | 19 | 25 | 20 |
| ATM +50 bps | 11 | 13 | 10 | 10 | 19 | 14 | 12 | 9 | 19 | 19 | 19 | 19 | 23 | 20 | 25 | 20 |
| ATM +100 bps | 14 | 14 | 10 | 10 | 19 | 11 | 13 | 10 | 19 | 13 | 19 | 19 | 24 | 19 | 24 | 18 |
| ATM +200 bps | 11 | 11 | 8 | 9 | 12 | 12 | 9 | 12 | 17 | 11 | 12 | 14 | 14 | 14 | 24 | 15 |
| Panel E: 10x2 | | | | | | | | | | | | | | | | |
| ATM | 6 | | 5 | | 9 | | 8 | | 17 | | 14 | | 15 | | 23 | |
| ATM +25 bps | 6 | 6 | 7 | 7 | 11 | 10 | 8 | 8 | 17 | 14 | 15 | 15 | 18 | 17 | 26 | 21 |
| ATM +50 bps | 10 | 10 | 7 | 8 | 14 | 9 | 14 | 7 | 18 | 13 | 17 | 19 | 24 | 16 | 26 | 19 |
| ATM +100 bps | 10 | 10 | 8 | 8 | 14 | 9 | 14 | 8 | 18 | 9 | 19 | 19 | 25 | 18 | 26 | 14 |
| ATM +200 bps | 9 | 9 | 8 | 7 | 9 | 9 | 9 | 9 | 16 | 9 | 12 | 13 | 15 | 15 | 26 | 12 |
| Panel F: 10x5 | | | | | | | | | | | | | | | | |
| ATM | 6 | | 6 | | 13 | | 9 | | 19 | | 14 | | 19 | | 23 | |
| ATM +25 bps | 10 | 10 | 7 | 7 | 14 | 14 | 10 | 9 | 20 | 17 | 15 | 15 | 19 | 19 | 24 | 21 |
| ATM +50 bps | 13 | 13 | 8 | 10 | 17 | 13 | 12 | 8 | 21 | 17 | 18 | 17 | 22 | 17 | 26 | 20 |
| ATM +100 bps | 14 | 13 | 9 | 9 | 18 | 13 | 13 | 8 | 21 | 14 | 18 | 19 | 24 | 16 | 29 | 16 |
| ATM +200 bps | 10 | 10 | 8 | 8 | 12 | 12 | 9 | 9 | 20 | 11 | 14 | 15 | 17 | 16 | 27 | 15 |

## Table 11:
## Basel test results

The table displays results from the regulatory Basel test. The test result is based on the number of VaR exceptions, and to obtain a "Green" flag, 16 hits are allowed when the number of observations is 229 and VaR confidence level is 95%.

| | Proxy_HS | | Proxy_TW | | Direct_HS | | Direct_TW | | Component_HS | | Proxy_FHS | | Direct_FHS | | Component_FHS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | | | | | | | | | | | | |
| ATM | Green | | Green | | Green | | Green | | Green | | Green | | Yellow | | Red | |
| ATM +25 bps | Green | Green | Green | Green | Green | Green | Yellow | Green | Green | Green | Red | Red | Red | Yellow | Red | Red |
| ATM +50 bps | Green | Green | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Yellow | Red | Yellow | Red | Yellow |
| ATM +100 bps | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow |
| ATM +200 bps | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow |
| Panel B: 1x5 | | | | | | | | | | | | | | | | |
| ATM | Green | | Green | | Green | | Green | | Yellow | | Green | | Yellow | | Red | |
| ATM +25 bps | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Yellow | Red | Yellow | Red | Red |
| ATM +50 bps | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Yellow | Red | Yellow | Red | Red |
| ATM +100 bps | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Yellow | Red | Yellow | Red | Yellow |
| ATM +200 bps | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow |
| Panel C: 5x2 | | | | | | | | | | | | | | | | |
| ATM | Green | | Green | | Green | | Green | | Yellow | | Green | | Yellow | | Yellow | |
| ATM +25 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Green | Yellow | Yellow | Red | Yellow | Yellow | Yellow |
| ATM +50 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Green | Yellow | Yellow | Red | Yellow | Yellow | Yellow |
| ATM +100 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Green | Yellow | Yellow | Red | Yellow | Yellow | Yellow |
| ATM +200 bps | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Green | Green | Green | Green | Green | Yellow | Green |
| Panel D: 5x5 | | | | | | | | | | | | | | | | |
| ATM | Green | | Green | | Green | | Green | | Yellow | | Green | | Yellow | | Yellow | |
| ATM +25 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow | Red | Yellow |
| ATM +50 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow | Red | Yellow |
| ATM +100 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Green | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow |
| ATM +200 bps | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Green | Green | Green | Green | Green | Yellow | Green |
| Panel E: 10x2 | | | | | | | | | | | | | | | | |
| ATM | Green | | Green | | Green | | Green | | Yellow | | Green | | Green | | Yellow | |
| ATM +25 bps | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Yellow | Red | Yellow |
| ATM +50 bps | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Green | Yellow | Yellow | Yellow | Green | Red | Yellow |
| ATM +100 bps | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Green | Yellow | Yellow | Red | Yellow | Red | Green |
| ATM +200 bps | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Green | Red | Green |
| Panel F: 10x5 | | | | | | | | | | | | | | | | |
| ATM | Green | | Green | | Green | | Green | | Yellow | | Green | | Yellow | | Yellow | |
| ATM +25 bps | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Yellow | Green | Green | Yellow | Yellow | Yellow | Yellow |
| ATM +50 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow | Red | Yellow |
| ATM +100 bps | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Green | Yellow | Yellow | Yellow | Green | Red | Green |
| ATM +200 bps | Green | Green | Green | Green | Green | Green | Green | Green | Yellow | Green | Green | Green | Yellow | Green | Red | Green |

## Table 12:
## Kupiec test results

The table shows p-values of Kupiec's proportion of failures test. The p-values refer to probability of getting the number of hits presented in Table 10 with given number of observations and VaR confidence level. Consequently, the lower the p-value, the more likely it is that the model in question fails to provide accurate estimates. Series that should be rejected at 95% test confidence level are tagged with a single asterisk (*) and results that should be rejected at 99% confidence level are marked with a double asterisk (**).

| | Proxy_HS | | Proxy_TW | | Direct_HS | | Direct_TW | | Component_HS | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | | | | | | |
| ATM | 0.450 | | 0.664 | | 0.856 | | 0.445 | | 0.634 | |
| ATM +25 bps | 0.445 | 0.634 | 0.296 | 0.186 | 0.445 | 0.634 | 0.112 | 0.634 | 0.634 | 0.634 |
| ATM +50 bps | 0.903 | 0.664 | 0.445 | 0.445 | 0.903 | 0.664 | 0.112 | 0.634 | 0.856 | 0.903 |
| ATM +100 bps | 0.664 | 0.664 | 0.296 | 0.296 | 0.664 | 0.664 | 0.296 | 0.634 | 0.664 | 0.664 |
| ATM +200 bps | 0.151 | 0.151 | 0.856 | 0.634 | 0.151 | 0.276 | 0.856 | 0.634 | 0.276 | 0.276 |
| Panel B: 1x5 | | | | | | | | | | |
| ATM | 0.450 | | 0.450 | | 0.634 | | 0.634 | | 0.112 | |
| ATM +25 bps | 0.856 | 0.856 | 0.856 | 0.856 | 0.445 | 0.856 | 0.634 | 0.856 | 0.296 | 0.296 |
| ATM +50 bps | 0.445 | 0.445 | 0.856 | 0.634 | 0.296 | 0.856 | 0.445 | 0.856 | 0.296 | 0.445 |
| ATM +100 bps | 0.856 | 0.856 | 0.856 | 0.634 | 0.445 | 0.664 | 0.445 | 0.856 | 0.296 | 0.664 |
| ATM +200 bps | 0.450 | 0.450 | 0.903 | 0.856 | 0.903 | 0.903 | 0.634 | 0.634 | 0.856 | 0.450 |
| Panel C: 5x2 | | | | | | | | | | |
| ATM | 0.276 | | 0.276 | | 0.445 | | 0.903 | | 0.112 | |
| ATM +25 bps | 0.445 | 0.634 | 0.664 | 0.664 | 0.009** | 0.296 | 0.445 | 0.664 | 0.004** | 0.445 |
| ATM +50 bps | 0.634 | 0.856 | 0.664 | 0.664 | 0.004** | 0.856 | 0.186 | 0.664 | 0.009** | 0.634 |
| ATM +100 bps | 0.445 | 0.634 | 0.903 | 0.856 | 0.018* | 0.664 | 0.445 | 0.664 | 0.009** | 0.903 |
| ATM +200 bps | 0.450 | 0.450 | 0.151 | 0.276 | 0.664 | 0.664 | 0.450 | 0.664 | 0.063 | 0.903 |
| Panel D: 5x5 | | | | | | | | | | |
| ATM | 0.664 | | 0.276 | | 0.856 | | 0.276 | | 0.034* | |
| ATM +25 bps | 0.664 | 0.664 | 0.276 | 0.664 | 0.112 | 0.445 | 0.664 | 0.903 | 0.034* | 0.034* |
| ATM +50 bps | 0.903 | 0.634 | 0.664 | 0.664 | 0.034* | 0.445 | 0.856 | 0.450 | 0.034* | 0.034* |
| ATM +100 bps | 0.445 | 0.445 | 0.664 | 0.664 | 0.034* | 0.903 | 0.634 | 0.664 | 0.034* | 0.634 |
| ATM +200 bps | 0.903 | 0.903 | 0.276 | 0.450 | 0.856 | 0.856 | 0.450 | 0.856 | 0.112 | 0.903 |
| Panel E: 10x2 | | | | | | | | | | |
| ATM | 0.072 | | 0.029* | | 0.450 | | 0.276 | | 0.112 | |
| ATM +25 bps | 0.072 | 0.072 | 0.151 | 0.151 | 0.903 | 0.664 | 0.276 | 0.276 | 0.112 | 0.445 |
| ATM +50 bps | 0.664 | 0.664 | 0.151 | 0.276 | 0.445 | 0.450 | 0.445 | 0.151 | 0.063 | 0.634 |
| ATM +100 bps | 0.664 | 0.664 | 0.276 | 0.276 | 0.445 | 0.450 | 0.445 | 0.276 | 0.063 | 0.450 |
| ATM +200 bps | 0.450 | 0.450 | 0.276 | 0.151 | 0.450 | 0.450 | 0.450 | 0.450 | 0.186 | 0.450 |
| Panel F: 10x5 | | | | | | | | | | |
| ATM | 0.072 | | 0.072 | | 0.634 | | 0.450 | | 0.034* | |
| ATM +25 bps | 0.664 | 0.664 | 0.151 | 0.151 | 0.445 | 0.445 | 0.664 | 0.450 | 0.018* | 0.112 |
| ATM +50 bps | 0.634 | 0.634 | 0.276 | 0.664 | 0.112 | 0.634 | 0.856 | 0.276 | 0.009** | 0.112 |
| ATM +100 bps | 0.445 | 0.634 | 0.450 | 0.450 | 0.063 | 0.634 | 0.634 | 0.276 | 0.009** | 0.445 |
| ATM +200 bps | 0.664 | 0.664 | 0.276 | 0.276 | 0.856 | 0.856 | 0.450 | 0.450 | 0.018* | 0.903 |

## Table 13:
## Christoffersen test results

The table presents p-values of Christoffersen's test of conditional coverage. The p-values refer to probability of obtaining a sample that is less likely than that of observed given that the null hypothesis of accurate model is true. Series that should be rejected at 95% test confidence level are tagged with a single asterisk (*) and results that should be rejected at 99% confidence level are marked with a double asterisk (**).

| | Proxy_HS | | Proxy_TW | | Direct_HS | | Direct_TW | | Component_HS | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | | | | | | |
| ATM | 0.350 | | 0.337 | | 0.137 | | 0.878 | | 0.190 | |
| ATM +25 bps | 0.255 | 0.190 | 0.992 | 0.895 | 0.255 | 0.190 | 0.787 | 0.764 | 0.190 | 0.190 |
| ATM +50 bps | 0.094 | 0.062 | 0.878 | 0.878 | 0.094 | 0.062 | 0.147 | 0.764 | 0.137 | 0.094 |
| ATM +100 bps | 0.062 | 0.062 | 0.331 | 0.331 | 0.062 | 0.062 | 0.331 | 0.190 | 0.062 | 0.062 |
| ATM +200 bps | 0.012* | 0.012* | 0.137 | 0.190 | 0.012* | 0.022* | 0.137 | 0.190 | 0.022* | 0.022* |
| Panel B: 1x5 | | | | | | | | | | |
| ATM | 0.389 | | 0.350 | | 0.764 | | 0.764 | | 0.515 | |
| ATM +25 bps | 0.247 | 0.247 | 0.652 | 0.652 | 0.255 | 0.247 | 0.764 | 0.652 | 0.331 | 0.331 |
| ATM +50 bps | 0.255 | 0.255 | 0.652 | 0.764 | 0.331 | 0.652 | 0.878 | 0.652 | 0.331 | 0.255 |
| ATM +100 bps | 0.652 | 0.652 | 0.652 | 0.764 | 0.047* | 0.337 | 0.878 | 0.652 | 0.071 | 0.337 |
| ATM +200 bps | 0.350 | 0.350 | 0.544 | 0.652 | 0.544 | 0.544 | 0.764 | 0.764 | 0.137 | 0.350 |
| Panel C: 5x2 | | | | | | | | | | |
| ATM | 0.266 | | 0.266 | | 0.255 | | 0.544 | | 0.787 | |
| ATM +25 bps | 0.255 | 0.764 | 0.443 | 0.443 | 0.964 | 0.331 | 0.878 | 0.443 | 0.919 | 0.878 |
| ATM +50 bps | 0.764 | 0.652 | 0.443 | 0.443 | 0.919 | 0.652 | 0.895 | 0.443 | 0.964 | 0.764 |
| ATM +100 bps | 0.255 | 0.764 | 0.544 | 0.652 | 0.847 | 0.443 | 0.878 | 0.443 | 0.964 | 0.544 |
| ATM +200 bps | 0.350 | 0.350 | 0.194 | 0.266 | 0.443 | 0.443 | 0.350 | 0.443 | 0.201 | 0.544 |
| Panel D: 5x5 | | | | | | | | | | |
| ATM | 0.062 | | 0.266 | | 0.017* | | 0.266 | | 0.267 | |
| ATM +25 bps | 0.062 | 0.062 | 0.266 | 0.062 | 0.147 | 0.047* | 0.443 | 0.094 | 0.267 | 0.267 |
| ATM +50 bps | 0.009** | 0.029* | 0.443 | 0.443 | 0.267 | 0.047* | 0.017* | 0.350 | 0.267 | 0.267 |
| ATM +100 bps | 0.047* | 0.047* | 0.443 | 0.443 | 0.267 | 0.009** | 0.029* | 0.443 | 0.267 | 0.029* |
| ATM +200 bps | 0.009** | 0.009** | 0.266 | 0.350 | 0.017* | 0.017* | 0.350 | 0.137 | 0.147 | 0.009** |
| Panel E: 10x2 | | | | | | | | | | |
| ATM | 0.133 | | 0.084 | | 0.038* | | 0.266 | | 0.147 | |
| ATM +25 bps | 0.133 | 0.133 | 0.194 | 0.194 | 0.009** | 0.062 | 0.266 | 0.266 | 0.029* | 0.047* |
| ATM +50 bps | 0.005** | 0.005** | 0.194 | 0.266 | 0.047* | 0.002** | 0.255 | 0.194 | 0.046* | 0.029* |
| ATM +100 bps | 0.005** | 0.062 | 0.266 | 0.266 | 0.047* | 0.038* | 0.255 | 0.266 | 0.201 | 0.038* |
| ATM +200 bps | 0.038* | 0.038* | 0.266 | 0.194 | 0.038* | 0.038* | 0.350 | 0.350 | 0.104 | 0.038* |
| Panel F: 10x5 | | | | | | | | | | |
| ATM | 0.133 | | 0.133 | | 0.029* | | 0.038* | | 0.267 | |
| ATM +25 bps | 0.062 | 0.062 | 0.194 | 0.194 | 0.047* | 0.047* | 0.062 | 0.038* | 0.344 | 0.147 |
| ATM +50 bps | 0.029* | 0.029* | 0.022* | 0.062 | 0.147 | 0.029* | 0.137 | 0.022* | 0.432 | 0.147 |
| ATM +100 bps | 0.047* | 0.029* | 0.350 | 0.038* | 0.201 | 0.029* | 0.190 | 0.266 | 0.432 | 0.047* |
| ATM +200 bps | 0.062 | 0.062 | 0.266 | 0.266 | 0.137 | 0.137 | 0.350 | 0.350 | 0.344 | 0.094 |

To summarize the backtesting results from the first stage, none of the considered models pass all of the tests with each of the swaption series. However, as the regulatory test can be regarded as the ultimate limit, some models still perform well enough to be deemed accurate. Consequently, TW models are left on the table for further investigation. Thus, the backtesting is continued in the second stage more thoroughly to study the differences between the smile methods. Moreover, the relevance of using actual changes versus proxy changes in implied volatilities is also studied in the following part.

## 5.3 Stage two backtesting results

While a VaR estimate only indicates the loss figure that is not expected to be exceeded with a given confidence level, it does not provide any approximation about the absolute level of profit and loss for the position, and hence, determining whether an individual VaR figure is correct or not is not feasible. Consequently, unless there are distinctive differences in the backtesting results with respect to observed VaR exceptions, it is not possible to outright declare one model better than the other. On the other hand, an expected shortfall figure does provide a direct estimate about the loss figure on the condition that the respective VaR level is breached. Hence, it is possible to compare the relative accuracy of given models using the data from VaR exceptions and corresponding expected shortfall estimates and actual profit and loss figures. After all, a VaR and an expected shortfall measures are based on the same distribution.

Table 14 presents the mean absolute deviations and mean squared error figures of the expected shortfall estimates from the days when the respective VaR estimates are exceeded. The table also shows the p-values of Hansen’s (2005) SPA test for each of the test series. A low p-value suggests that the accuracy of the alternative model is likely to outperform that of the benchmark model.

In the first test set up computed for Proxy_TW models, the fixed smile method is the benchmark and the random smile method represents the alternative. The results based on the both MAD and MSE figures suggest that only in two out of 24 considered cases the random smile method yields more accurate results, which suggests that using OTM implied volatility data would not provide a measurable enhancement to risk estimation precision.

Additionally, Table 14 presents the results for comparison between Proxy_TW and Direct_TW, which provide insight whether there is a measurable difference between models utilizing either changes in implied volatilities for a given contract versus using approximated changes derived by computing the difference between different contracts but with equal strike rates with respect to their corresponding ATM rates. The latter is used as the benchmark model in the tests. Since the alternative model, i.e. Direct_TW, yields more accurate expected shortfall estimate for only two of the 24 studied contracts when tested with MAD and for none when the differences are measured using MSE, it is unlikely that the more complicated method would provide any measurable enhancement in VaR estimation over the benchmark model.

To summarize the stage two backtesting results, it appears that including moneyness considerations into VaR estimation does not have a measurable impact on the precision of the estimates from either of the possible perspectives. First, the difference between using implied volatility changes from a moneyness-dependent level versus ATM level is negligible. Second, there is no observable enhancement in the estimates when taking moneyness changes into consideration also from the viewpoint of which figures to use in compiling the distribution for volatility fluctuations.

However, while the number of VaR exceptions can be considered somewhat high from the viewpoint of model accuracy, the available sample for this test is apparently nonetheless insufficient. The p-values end up almost automatically to one or very close to zero on the basis of even slightest difference in the observed average differences in ES values. Hence, one should be careful when drawing conclusions based on these results.

**Table 14:**
**Loss function test results**

The table presents the figures from loss function based backtest. MAD stands for mean absolute deviation and MSE for mean squared error. The error terms are computed from the differences between the expected shortfall measures and the actual profit and loss figures from the days when VaR estimates of both of the models being compared are exceeded. A high p-value suggests that the benchmark model is likely to perform at least as well as the alternative model, while low p-values indicate that the alternative is preferable. P-values less than 5% and 1% are highlighted with single- and double-asterisks respectively.

| | MAD | | | MSE | | | MAD | | | MSE | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Proxy_TW ATM | Proxy_TW Smile | p-value | Proxy_TW ATM | Proxy_TW Smile | p-value | Proxy_TW Smile | Direct_TW Smile | p-value | Proxy_TW Smile | Direct_TW Smile | p-value |
| Panel A: 1x2 | | | | | | | | | | | | |
| ATM +25 bps | 1 791 | 1 784 | 0.465 | 4 691 248 | 5 018 561 | 1.000 | 1 623 | 1 634 | 1.000 | 4 508 056 | 4 101 386 | 0.229 |
| ATM +50 bps | 2 497 | 2 388 | 0.154 | 9 346 271 | 9 947 471 | 1.000 | 2 304 | 2 538 | 1.000 | 9 781 922 | 10 590 140 | 1.000 |
| ATM +100 bps | 3 662 | 3 464 | 0.032* | 19 782 234 | 18 912 111 | 0.187 | 3 257 | 3 391 | 1.000 | 18 265 144 | 18 540 222 | 1.000 |
| ATM +200 bps | 4 100 | 4 168 | 1.000 | 25 507 987 | 27 779 159 | 1.000 | 4 181 | 4 078 | 0.150 | 27 089 282 | 24 875 401 | 0.130 |
| Panel B: 1x5 | | | | | | | | | | | | |
| ATM +25 bps | 3 248 | 3 458 | 1.000 | 18 209 659 | 19 901 631 | 1.000 | 3 458 | 3 208 | 0.178 | 19 901 631 | 20 695 984 | 1.000 |
| ATM +50 bps | 4 244 | 4 364 | 1.000 | 29 460 660 | 32 885 584 | 1.000 | 4 364 | 3 943 | 0.069 | 32 885 584 | 24 441 413 | 0.008** |
| ATM +100 bps | 4 748 | 4 685 | 0.398 | 36 797 964 | 40 464 554 | 1.000 | 4 685 | 4 217 | 0.058 | 40 464 554 | 32 527 378 | 0.011* |
| ATM +200 bps | 4 388 | 4 869 | 1.000 | 44 478 463 | 51 846 733 | 1.000 | 5 033 | 5 033 | 1.000 | 51 421 159 | 52 842 976 | 1.000 |
| Panel C: 5x2 | | | | | | | | | | | | |
| ATM +25 bps | 3 185 | 3 206 | 1.000 | 13 743 628 | 14 561 727 | 1.000 | 3 206 | 3 058 | 0.247 | 14 561 727 | 14 771 408 | 1.000 |
| ATM +50 bps | 3 364 | 3 310 | 0.218 | 18 329 401 | 18 787 129 | 1.000 | 3 310 | 3 373 | 1.000 | 18 787 129 | 17 672 368 | 0.175 |
| ATM +100 bps | 3 467 | 3 572 | 1.000 | 20 139 085 | 21 947 255 | 1.000 | 3 655 | 3 486 | 0.285 | 23 385 287 | 19 008 059 | 0.053 |
| ATM +200 bps | 2 120 | 2 395 | 1.000 | 6 345 496 | 7 848 425 | 1.000 | 2 649 | 3 021 | 1.000 | 9 309 409 | 13 189 941 | 1.000 |
| Panel D: 5x5 | | | | | | | | | | | | |
| ATM +25 bps | 4 636 | 4 365 | 0.254 | 29 697 447 | 30 744 902 | 1.000 | 4 743 | 4 870 | 1.000 | 32 521 562 | 41 985 061 | 1.000 |
| ATM +50 bps | 7 820 | 7 451 | 0.009** | 87 842 026 | 84 897 226 | 0.012* | 7 639 | 7 877 | 1.000 | 90 647 385 | 92 675 950 | 1.000 |
| ATM +100 bps | 8 635 | 8 452 | 0.189 | 112 122 760 | 114 422 442 | 1.000 | 8 452 | 8 354 | 0.393 | 114 422 442 | 101 353 898 | 0.086 |
| ATM +200 bps | 8 583 | 8 510 | 0.413 | 89 097 419 | 90 696 907 | 1.000 | 8 737 | 8 442 | 0.366 | 92 994 807 | 103 179 846 | 1.000 |
| Panel E: 10x2 | | | | | | | | | | | | |
| ATM +25 bps | 3 934 | 3 887 | 0.207 | 22 998 189 | 22 331 190 | 0.071 | 3 887 | 3 697 | 0.365 | 22 331 190 | 25 737 409 | 1.000 |
| ATM +50 bps | 4 224 | 4 097 | 0.091 | 32 297 413 | 30 642 421 | 0.01** | 4 097 | 4 184 | 1.000 | 30 642 421 | 31 052 528 | 1.000 |
| ATM +100 bps | 4 659 | 4 648 | 0.475 | 40 746 865 | 41 217 429 | 1.000 | 4 648 | 4 619 | 0.457 | 41 217 429 | 38 028 414 | 0.115 |
| ATM +200 bps | 5 432 | 4 947 | 0.097 | 43 924 967 | 42 287 823 | 0.203 | 4 947 | 4 931 | 0.484 | 42 287 823 | 45 941 317 | 1.000 |
| Panel F: 10x5 | | | | | | | | | | | | |
| ATM +25 bps | 9 186 | 8 971 | 0.138 | 130 347 177 | 127 994 504 | 0.268 | 8 971 | 8 702 | 0.426 | 127 994 504 | 156 034 427 | 1.000 |
| ATM +50 bps | 11 269 | 11 442 | 1.000 | 235 996 106 | 237 492 998 | 1.000 | 11 442 | 11 341 | 0.424 | 237 492 998 | 249 683 343 | 1.000 |
| ATM +100 bps | 11 810 | 11 818 | 1.000 | 314 687 050 | 323 236 891 | 1.000 | 11 818 | 11 516 | 0.205 | 323 236 891 | 298 499 263 | 0.096 |
| ATM +200 bps | 11 665 | 11 690 | 1.000 | 286 968 367 | 302 756 651 | 1.000 | 11 690 | 11 311 | 0.319 | 302 756 651 | 334 345 751 | 1.000 |

# 6 CONCLUSION

VaR has gained a resilient position as a tool for market risk measurement in financial institutions and also as a basis for capital requirements mandated by financial regulators. For these reasons, an ongoing development of VaR models carries its role in safeguarding the stability of individual financial actors and in a larger scale also the soundness of the economy as a whole. This paper contributes to that development with a rather specific approach through studying the significance of accounting for asynchronous shifts in the implied volatility smile when assessing the market risk of a swaption contract. Using a sample of swaption daily profit and loss figures with corresponding VaR estimates created by using a variety of models allows me to test for the impact of incorporating the moneyness-dependent volatility scenarios into VaR modelling on VaR estimation accuracy. This section summarizes my findings, draws the conclusions and finally presents some suggestions for further research.

## 6.1 Summary of results and conclusions

The main findings of this paper are recapitulated in Table 15. Overall, the results concerning the applicability of historical simulation for swaption VaR measurement are mixed. For the most part, the validity of VaR models based on historical simulation is rejected by the implemented backtests due to both too numerous and too clustered VaR exceptions. While it is difficult to provide an utterly conclusive explanation for this finding, it is obvious that VaR models based on historical simulation require on average a substantially longer observation period than the one used in this study, as for instance Pritsker (2000) points out. Nevertheless, HS and TW models perform adequately well in terms of regulatory backtest and could hence be used without additional penalty in VaR based capital charge. However, one should still bear in mind that even if the backtests fail to reject the validity of a given VaR model it does not suggest that the model actually performs well in any given situation.

The additional key objective of this paper is to provide insight about the relevance of embodying volatility smile considerations into VaR estimation process. The second stage backtesting results suggest that employing moneyness-dependent implied volatility changes does not result into improved VaR estimation accuracy when both of the comparable models can be deemed sufficiently accurate with respect to number of VaR exceptions generated during the backtesting period. This finding is in line with the observation that the fluctuations in OTM swaption volatilities closely resemble those occurring in ATM level. The second set of backtests also shows that utilizing swaption contract specific changes instead of moneyness specific changes in implied volatilities does not significantly improve the estimation precision either.

**Table 15:**
**Summary of results**

This table presents a summary of the hypotheses and corresponding findings of this study.

| Hypothesis | Results |
|---|---|
| **H1:** VaR models based on historical simulation generate acceptable estimates of swaption's market risk | Partial support with mixed evidence. While basic historical simulation and time weighted models with both ATM and moneyness-dependent implied volatility changes pass the regulatory backtest with each considered swaption position, none of the examined models is able to pass all of the implemented backtests. |
| **H2:** A model that employs OTM level implied volatility changes in swaption VaR estimation does not generate more accurate estimates than a model with fixed smile approach. | Strong support. VaR estimates based on a distribution of implied volatility changes from ATM level are at least as accurate as the estimates generated using moneyness-dependent implied volatility changes when the model generates an otherwise acceptable number of exceptions. |

Consequently, it is unlikely that incorporating more advanced volatility smile methods into VaR estimation would result in any observable enhancement in VaR precision compared to using merely ATM level changes that are observed between ATM swaptions on successive days. Moreover, the trivial differences in the volatility components of VaR estimates suggest that using ATM volatility changes would be sufficient even if the interest rate risk was hedged. This suggests that there is no need to retrieve a sample of historical OTM swaption volatilities for VaR computations, which is positive news for risk managers as the availability of historical ATM swaption volatilities vastly exceeds that of OTM volatilities. However, the results from the first set of backtests show limited evidence that the number of exceptions are more often within the acceptable boundaries for moneyness-dependent VaR methods when contract and moneyness specific volatility changes are applied instead of merely moneyness specific volatility changes. Hence, if one wishes to go the extra mile of including volatility smile's shape considerations into VaR calculations by using a history of moneyness-dependent changes in implied volatilities, it is advisable to utilize the information content available from the historical smile data and to also use actual volatility changes instead of proxy-changes.

Nevertheless, the observations do not propose that a risk manager could overlook the smile changes as a source of risk, since the absence of major shape changes during this observation period does not connote that they were nonexistent at all times. Hence, rather than including smile changes in daily VaR estimation, it might be more efficient to assess the smile risk through stress testing with appropriate scenarios in which different level shifts, convexity changes and rotations are covered. The scenarios' effect on the present value of the position provides the manager insight about the position's risk status on the level of detail that VaR estimates are unable to deliver. Moreover, the stress tests could expose some risks that VaR fails to disclose, which also highlights the importance of stress tests as a necessary complement to VaR measurement regardless of how sophisticated the VaR system is.

Linking the findings presented in this paper to previous studies is unfortunately inconceivable for the most part due to lack of published comparable research. A number of studies with regards to different VaR models and their backtesting results are naturally available, but those implemented for swaption VaR measurement are conspicuously absent. From that point of view, this study is able to contribute novel insight into practical implementation of VaR systems with regards to swaption market risk measurement. However, the mixed results highlight the need for further research in VaR based risk estimation methods as there obviously appears to be room for improvement in the measurement precision with the models presented in this study.

A natural extension to this study would be to implement analogous VaR estimations with equal model specifications but with a significantly longer historical observation period if possible. As there does not appear to be fundamental issues with the covered models, using a wider spectrum of historical returns should add to the accuracy of any of the described models.

Additionally, as the interest rates move close to zero, log-volatilities have a tendency to become extremely high. As a result, using normal volatilities becomes a viable alternative in swaption modeling as they remain more stable in time. Consequently, developing swaption VaR estimation models based on utilizing a pricing model that assumes normally distributed interest rate changes as opposed to log-normal moves assumed in Black model might present a fruitful avenue for future research in the area of market risk management. A feasible alternative for such pricing model is Bachelier's (1900) model or some of its more recent modifications.

## REFERENCES


Alexander, C. 2008, *Market Risk Analysis Volume IV,* John Wiley & Sons, England.

Ammann, M. & Reich, C. 2001, *VaR for Nonlinear Financial Instruments - Linear Approximation or Full Monte-Carlo?*, Working Paper No 8/01 edn, University of Basel, WWZ/Department of Finance.

Andersen, T. & Lund, J. 1997, "Estimating Continuous-Time Stochastic Volatility of the Short-Term Interest Rate", *Journal of Econometrics,* , no. 77, pp. 343-377.

Angelidis, T. & Benos, A. 2006, "Value-at-risk for Greek stocks", *Multinational Finance Journal,* vol. 12, pp. 67-105.

Angelidis, T., Benos, A. & Degiannakis, S.A. 2007, "A robust VaR model under different time periods and weighting schemes", *Review of Quantitative Finance and Accounting,* vol. 28, no. 2, pp. 187-201.

Artzner, P., Delbaen, F., Eber, J.-. & Heath, D. 1999, "Coherent Measures of Risk", *Mathematical Finance,* vol. 9, no. 3, pp. 203-228.

Bachelier, L. 1900, "Théorie de la spéculation", *Annales Scientifiques de l'École Normale Supérieure,* vol. 3, no. 17, pp. 21-86.

Ball, C. & Torous, N. 1999, "The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence", *Journal of Finance,* vol. 54, no. 6, pp. 2339–2359.

Barone-Adesi, G., Bourgoin, F. & Giannopoulos, K. 1998, "Don't look back", *Risk,* vol. 11, pp. 100-104.

Barone-Adesi, G. & Giannopoulos, K. 2001, "Non parametric VaR Techniques. Myths and Realities", *Economic Notes,* vol. 30, pp. 167-181.

Barone-Adesi, G., Giannopoulos, K. & Vosper, L. 2002, "Backtesting Derivative Portfolios with Filtered Historical Simulation (FHS)", *European Financial Management,* vol. 8, no. 1, pp. 31-58.

Barone-Adesi, G., Giannopoulos, K. & Vosper, L. 1999, "VaR without correlations for non.linear portfolios", *Journal of Futures Markets,* vol. 19, pp. 583-602.

Bianchetti, M. 2010, "Two Curves, One Price: Pricing & Hedging Interest Rate Derivatives: Decoupling Forwarding and Discounting Yield Curve ", *Risk,* vol. 8, pp. 66-72.

Billio, M. & Pelizzon, L. 2000, "f-Risk: a multivariate switching regime approach", *Journal of Empirical Finance,* vol. 7, no. 5, pp. 531-554.

Black, F. 1976, "The Pricing of Commodity Contracts", *Journal of Financial Economics,* vol. 3, pp. 167-179.

Boudoukh, J., Richardson, M., Stanton, R. & Whitelaw, R. 2004, "Long horizon value at risk in a mark-to-market environment", *Journal of Investment Management,* vol. 2, no. 3, pp. 1-6.

Boudoukh, J., Richardson, M. & Whitelaw, R. 1998, "The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk", *Risk,* vol. 11, no. 5, pp. 64-67.

Britten-Jones, M. & Schaeffer, S.M. 1999, "Non-linear Value-at-Risk", *European Finance Review,* vol. 2, no. 2, pp. 161-187.

Butler, J.S. & Schachter, B. 1998, "Estimating Value at Risk with a Precision Measure by Combining Kernel Estimation with Historical Simulation", *Review of Derivatives Research,* vol. 1, pp. 371-390.

Christofferssen, P. 1998, "Evaluating Interval Forecasts", *International Economic Review,* vol. 39, pp. 841-862.

Corb, H. 2012, *Interest Rate Swaps and Other Derivatives,* Columbia University Press, New York.

Derman, E. 1999, "Regimes of volatility", *Risk,* vol. 12, no. 4, pp. 55-59.

Dowd, K. 2002, *An introduction to market risk measurement,* John Wiley & Sons, England.

Eberlein, E., Keller, U. & Prause, K. 1998, "New insights into smile, mispricing and value at risk: the hyperbolic model", *Journal of Business,* vol. 71, pp. 371-405.

Engle, R. 1982, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica,* vol. 50, pp. 987-1007.

Engle, R. & Manganelli, S. 2001, *Value at risk models in finance*, European Central Bank Working paper no. 75 edn.

Fama, E. 1965, "The Behavior of Stock Market Prices", *Journal of Business,* vol. 38, pp. 34-105.

Giot, P. & Laurent, S. 2004, "Modeling daily Value-at-Risk using realized volatility and ARCH type models", *Journal of Empirical Finance,* vol. 11, no. 3, pp. 379-398.

Hagan, P.S., Kumar, D., Lesniewski, A.S. & Woodward, D.E. 2002, "Managing smile risk", *WILMOTT Magazine,* vol. 9, pp. 84-108.

Hansen, P.R. 2005, "A Test for Superior Predictive Ability", *Journal of Business & Economic Statistics,* vol. 23, no. 4, pp. 365-380.

Hendricks, D. 1996, "Evaluating Value-at-Risk Models Using Historical Data", *Economic Policy Review,* vol. 2, no. 1.

Hull, J.C. 2008, *Options, Futures and Other Derivatives* 7th edn, Prentice Hall.

Hull, J.C. & White, A. 1998, "Incorporating volatility updating into the historical simulation method for value at risk", *Journal of Risk,* vol. 1, no. 1, pp. 5-19.

ISDA, *Interest Rate Trade Repository Report* [Homepage of International Swaps and Derivatives Association, Inc.], [Online]. Available: www.isda.org.

Jamshidian, F. & Zhu, Y. 1996, "Scenario Simulation: Theory and methodology", *Finance and Stochastics,* vol. 1, no. 1, pp. 43-67.

Jorion, P. 2007, *Value at Risk, The New Benchmark for Managing Financial Risk,* 3rd edn, McGraw-Hill, United States.

Kritzman, M. & Rich, D. 2002, "The Mismeasurement of Risk", *Financial Analysts Journal,* vol. 58, no. 3, pp. 91-99.

Linsmeier, T.J. & Pearson, N.D. 1996, *Risk Measurement: An introduction to value-at-risk*, Technical report 96-04 edn, OFOR,University of Illinois, Urbana-Champaign, IL.

Longerstaey, J. 1996, *RiskMetrics - Technical Document*, J.P. Morgan.

Longstaff, F., Santa-Clara, P. & Schwartz, E. 2000, " Throwing Away A Billion Dollars:The Cost of Suboptimal Exercise Strategies in the Swaptions Market", *Journal of Financial Economics,* vol. 62, no. 1, pp. 39-66.

Malz, A.M. 2001, "Vega risk and the smile", *Journal of Risk,* vol. 3, no. 2, pp. 41-63.

Mandelbrot, B. 1963, "The Variations of Certain Speculative Prices", *Journal of Business,* vol. 36, pp. 394-419.

Mercurio, F. 2008, "Cash-settled swaptions and no-arbitrage", *Risk,* vol. 21, no. 2, pp. 96-98.

Modigliani, F. & Miller, M. 1958, "The Cost of Capital, Corporation Finance and the Theory of Investment", *American Economic Review,* vol. 48, pp. 261-297.

Pallavicini, A. & Tarenghi, M. 2010, *Interest-rate modeling with multiple yield curves*, Working Paper edn, SSRN eLibrary.

Politis, D. & Romano, J. 1994, "The Stationary Bootstrap", *Journal of the American Statistical Association,* vol. 89, no. 428, pp. 1303-1313.

Politis, D., White, H. & Patton, A. 2009, "Correction: Automatic Block-Length Selection for the Dependent Bootstrap", *Econometric Reviews,* vol. 28, no. 4, pp. 372-375.

Pritsker, M. 2001, *The Hidden Dangers of Historical Simulation*, Finance and Economics Discussion Series 27. edn, Board of Governors of the Federal Reserve System, Washington, D.C.

Shao, X., Lian, Y. & Yin, L. 2009, "Forecasting Value-at-Risk using high frequency data: The realized range model", *Global Finance Journal,* vol. 20, no. 2, pp. 128-136.

Smith, C.W. & Stulz, R.M. 1985, "The Determinants of Firms' Hedging Policies", *The Journal of Financial and Quantitative Analysis,* vol. 20, no. 4, pp. 391-405.

Taleb, N. 1997, *Dynamic hedging: managing vanilla and exotic options,* John Wiley & Sons, Inc.

Trolle, A. & Schwartz, E. 2012, *The Swaption Cube*, working papers series edn, Available at SSRN: http://ssrn.com/abstract=1698910.

Van den Goorbergh, R. & Vlaar, P. 1999, *Value at risk analysis of stock returns: historical simulation, variance techniques or tail index estimation?*, Manuscript edn, Tilburg University.

van Deventer, D., Imai, K. & Mesler, M. 2005, *Advanced Financial Risk Management,* John Wiley & Sons, Asia.

Vlaar, P. 2000, "Value at risk models for Dutch bond portfolios", *Journal of Banking and Finance,* vol. 24, pp. 1131-1154.

West, G. 2005, "Calibration of the SABR model in illiquid markets", *Applied Mathematical Finance,* vol. 12, no. 4, pp. 371-385.

Wiener, Z. 1999, *Introduction to VaR (Value-at-Risk), Risk Management and Regulation in Banking,* Kluwer Academic Publishers, Boston.

## APPENDIX A: ABBREVIATIONS AND TERMINOLOGY

This appendix provides a list of frequently used abbreviations and terminology used in this study. The list contains also short descriptions of selected model names and prefixes used in the tables and result discussion.

**ATM**: at-the-money. A situation when the strike price of an option equals to the price of the underlying security. For a swaption, the term "price" refers to the fixed rate (also called as "swap rate") of the underlying interest rate swap, and for an ATM swaption, the strike rate is equal to the forward rate that would apply between the maturity and the tenor of the underlying swap so that the present value of the swap would be zero on the swaption maturity.
In the result tables, "ATM" refers to the VaR smile method in which the historical implied volatility changes are observed from the ATM level. See Fixed smile method.
**"Component_"-prefix**: in the result tables, refers to VaR model that utilizes *Skew dependent smile method*.
**"Direct_"-prefix**: refers to VaR models where the historical changes in implied volatilities that have occurred for swaption contracts over a 1-day holding period are defined in a fashion that accounts for the simultaneous change in the interest rates. For instance, a contract whose strike rate is initially ATM + 25 bps on day *t,* the moneyness on *t+1* may not be ATM +25 bps anymore, but more or less if the interest rates have changed. Hence, a new implied volatility for *t+1* is first determined using the SABR model, and the actual change in the implied volatility is the difference between the implied volatility given by the SABR model and the implied volatility observed on *t.* See also "Proxy_" -prefix.
**FHS:** filtered historical simulation. A variation of historical simulation in which the historical risk factor returns are scaled by a factor of prevailing volatility estimate divided by the respective historical volatility estimates.
**Fixed smile method**: a VaR method in which the changes in the implied volatilities for swaptions with a given maturity-tenor -pair are assumed to change in parallel fashion, i.e. the changes for OTM swaptions are assumed to equal changes that occur on ATM level. Hence, the distribution of historical changes is derived from the ATM level only.

**HS:** historical simulation. A VaR model where the distribution of profit and loss scenarios is created through applying historical observed changes in risk factors to the prevailing values of those risk factors.

**Implied volatility**: the value of the volatility of the underlying instrument to be used as an input in an option pricing model in order to achieve a theoretical price that equals to the current market price of the given option.

**IR**: interest rate, (used with VaR components).

**IRS**: interest rate swap. An interest rate derivative instrument in which two parties agree to exchange interest rate cash flows. The cash flows are based on a given notional amount from a fixed rate to a floating or vice versa (or one floating to another floating). The floating rate is often indexed to a reference rate such as Euribor.

**MAD**: mean absolute deviation.

**Moneyness**: the relative position of the current price of the underlying security of an option with respect to the strike price of the option.

**MSE**: mean squared error.

**OTM**: out-of-the-money. A situation when the strike price of an option is different from the price of the underlying security. See also ATM.

**Payer swaption**: see swaption.

**"Proxy_" -prefix:** refers to VaR models where the historical changes in implied volatilities that have occurred for swaption contracts over a 1-day holding period are defined as changes between implied volatilities for swpation contracts with equal moneyness. For instance, the change in implied volatility for a swaption whose strike rate is ATM +25 bps made on day *t* is defined as the difference between implied volatility for an ATM +25 bps on *t+1* less the observed implied volatility for an ATM +25 bps on *t*. Consequently, the computed change in implied volatility is does not actually refer to the change that occurred for an ATM +25 bps made on *t* if the underlying swap rate changed also between *t* and *t+1*. See also "Direct_" -prefix.

**Random smile method**: a VaR method in which the changes in the implied volatilities for swaptions with a given maturity-tenor -pair are not assumed to move in a parallel fashion. Rather, the historical observations used for compiling the distribution of implied volatility changes depends on the moneyness of the given swaption contract. For instance, in case of an ATM +50 bps swaption, the implied volatility changes are taken from changes that have occurred at ATM +50 bps level (instead of using changes that have occurred on ATM level as in Fixed smile method).

**Receiver swaption**: see swaption.

**SABR model**: a stochastic volatility model used for modeling the volatility smile in interest rate derivative markets. The model was developed by Hagan et al. (2002). The name stands for "stochastic alpha, beta, rho" where the Greek letters refer to the parameters of the model.

**Skew dependent smile method**: a VaR model where the historical interest rate changes are separated from respective historical changes in implied volatilities. Under this smile method, the interest rate - implied volatility -pair scenarios are created in two stages. First, a given interest rate change ("shock") is applied to the prevailing rate, and for the new point, an estimate of implied volatility is computed using the SABR model. Then the respective implied volatility scenario that occurred simultaneously with the interest rate change is applied to the scenario point created in the previous step, which results into the final scenario to be used.

**Smile**: in the result tables, "Smile" refers to the smile method in which historical implied volatility changes are observed from a moneyness level that depends on the moneyness of the swaption in question. See Random smile method.

Generally, smile (or skew or smirk) refers to the effect observed on implied volatilities when plotted against different strike rates and/or maturities.

**Swaption**: an option to enter into an interest swap agreement. Swaption is called *a receiver* when the holder has an option to enter into an IRS where she receives the fixed rate and pays the floating rate. Accordingly, *a payer* swaption grants the holder the right to enter into an IRS where she pays the fixed rate and receives the floating rate.

**TW**: time weighted historical simulation. A variation of historical simulation in which more recent risk factor changes receive more weight than those occurred earlier.

**VaR**: Value at Risk, the worst expected loss that a portfolio may suffer during a specified period under normal market conditions with a specified level of confidence.

**VaR component**: VaR figure with respect to a single risk factor, e.g. IR or implied volatility.

# APPENDIX B: RISK FACTOR STATISTICS AND FIGURES

This appendix presents additional statistics and graphs about risk factors covered in 4.1.1.

**Table 16:**
**Risk factor summary statistics**

The table presents the summary statistics of daily changes in swaption implied volatilities and underlying forward swap rates (F). The changes in F are measured in absolute terms and presented in basis points, whereas the statistics for implied volatilities are computed using log-changes and values are shown as percentages. The observation period is from March 8, 2011 to February 1, 2013 and the number of observations is 479.

| | F | ATM | +25bps | +50bps | +100bps | +200bps |
|---|---|---|---|---|---|---|
| Panel A: 1x2 | | | | | | |
| Average | -0.42 | 0.14 | 0.13 | 0.13 | 0.13 | 0.12 |
| Median | -0.32 | 0.19 | 0.12 | 0.25 | 0.24 | 0.21 |
| Std dev | 4.99 | 3.98 | 3.87 | 4.16 | 4.57 | 4.93 |
| Skewness | -0.31 | 0.01 | 0.02 | -0.10 | -0.23 | -0.63 |
| Kurtosis | 4.93 | 4.36 | 4.52 | 4.25 | 5.99 | 8.84 |
| Panel B: 1x5 | | | | | | |
| Average | -0.41 | 0.12 | 0.12 | 0.12 | 0.13 | 0.12 |
| Median | -0.18 | 0.24 | 0.18 | 0.10 | 0.14 | 0.19 |
| Std dev | 4.97 | 2.85 | 2.82 | 2.98 | 3.42 | 4.46 |
| Skewness | -0.12 | -0.24 | -0.21 | -0.18 | -0.21 | -0.09 |
| Kurtosis | 3.66 | 4.39 | 4.34 | 4.73 | 7.78 | 10.68 |
| Panel C: 5x2 | | | | | | |
| Average | -0.34 | 0.09 | 0.09 | 0.08 | 0.08 | 0.08 |
| Median | -0.51 | 0.13 | 0.14 | 0.10 | 0.09 | 0.05 |
| Std dev | 5.36 | 2.26 | 2.25 | 2.25 | 2.26 | 2.38 |
| Skewness | -0.09 | -0.12 | -0.10 | -0.13 | -0.15 | -0.10 |
| Kurtosis | 3.86 | 4.39 | 4.27 | 4.20 | 4.07 | 4.12 |
| Panel D: 5x5 | | | | | | |
| Average | -0.31 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 |
| Median | -0.36 | 0.10 | 0.04 | 0.07 | 0.05 | 0.06 |
| Std dev | 5.51 | 2.18 | 2.14 | 2.17 | 2.18 | 2.23 |
| Skewness | -0.13 | -0.01 | 0.07 | 0.01 | 0.02 | 0.13 |
| Kurtosis | 3.94 | 4.55 | 4.40 | 4.35 | 4.22 | 5.12 |
| Panel E: 10x2 | | | | | | |
| Average | -0.29 | 0.07 | 0.06 | 0.06 | 0.06 | 0.06 |
| Median | -0.30 | 0.15 | 0.12 | 0.11 | 0.06 | 0.01 |
| Std dev | 5.83 | 2.05 | 2.05 | 2.08 | 2.10 | 2.12 |
| Skewness | -0.14 | -0.07 | -0.07 | -0.01 | 0.01 | 0.07 |
| Kurtosis | 4.76 | 4.27 | 4.28 | 4.32 | 4.43 | 4.63 |
| Panel F: 10x5 | | | | | | |
| Average | -0.30 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| Median | -0.23 | 0.11 | 0.09 | 0.08 | 0.08 | 0.12 |
| Std dev | 5.97 | 2.14 | 2.10 | 2.10 | 2.09 | 2.05 |
| Skewness | -0.21 | 0.12 | 0.21 | 0.19 | 0.20 | 0.35 |
| Kurtosis | 5.09 | 5.14 | 4.96 | 5.13 | 4.98 | 5.25 |

**Table 17:**
**Risk factor correlations during the backtesting period**

The table presents correlations between the risk factors under the backtesting period from March 5, 2012 to February 1, 2013. Changes in forward swap rate (F) are measured in absolute terms whilst changes in implied volatilities are measured using log-returns.

| | F | ATM | +25bps | +50bps | +100bps | +200bps |
|---|---|---|---|---|---|---|
| Panel A: 1x2 | | | | | | |
| F | 1.000 | -0.709 | -0.684 | -0.574 | -0.468 | -0.385 |
| ATM | -0.709 | 1.000 | 0.988 | 0.823 | 0.690 | 0.562 |
| +25bps | -0.684 | 0.988 | 1.000 | 0.819 | 0.692 | 0.565 |
| +50bps | -0.574 | 0.823 | 0.819 | 1.000 | 0.887 | 0.815 |
| +100bps | -0.468 | 0.690 | 0.692 | 0.887 | 1.000 | 0.880 |
| +200bps | -0.385 | 0.562 | 0.565 | 0.815 | 0.880 | 1.000 |
| Panel B: 1x5 | | | | | | |
| F | 1.000 | -0.735 | -0.721 | -0.621 | -0.433 | -0.264 |
| ATM | -0.735 | 1.000 | 0.995 | 0.820 | 0.577 | 0.337 |
| +25bps | -0.721 | 0.995 | 1.000 | 0.822 | 0.586 | 0.349 |
| +50bps | -0.621 | 0.820 | 0.822 | 1.000 | 0.877 | 0.706 |
| +100bps | -0.433 | 0.577 | 0.586 | 0.877 | 1.000 | 0.867 |
| +200bps | -0.264 | 0.337 | 0.349 | 0.706 | 0.867 | 1.000 |
| Panel C: 5x2 | | | | | | |
| F | 1.000 | -0.775 | -0.768 | -0.749 | -0.712 | -0.595 |
| ATM | -0.775 | 1.000 | 0.988 | 0.987 | 0.946 | 0.815 |
| +25bps | -0.768 | 0.988 | 1.000 | 0.978 | 0.948 | 0.818 |
| +50bps | -0.749 | 0.987 | 0.978 | 1.000 | 0.982 | 0.890 |
| +100bps | -0.712 | 0.946 | 0.948 | 0.982 | 1.000 | 0.951 |
| +200bps | -0.595 | 0.815 | 0.818 | 0.890 | 0.951 | 1.000 |
| Panel D: 5x5 | | | | | | |
| F | 1.000 | -0.843 | -0.835 | -0.836 | -0.816 | -0.763 |
| ATM | -0.843 | 1.000 | 0.977 | 0.987 | 0.959 | 0.894 |
| +25bps | -0.835 | 0.977 | 1.000 | 0.967 | 0.940 | 0.876 |
| +50bps | -0.836 | 0.987 | 0.967 | 1.000 | 0.990 | 0.941 |
| +100bps | -0.816 | 0.959 | 0.940 | 0.990 | 1.000 | 0.969 |
| +200bps | -0.763 | 0.894 | 0.876 | 0.941 | 0.969 | 1.000 |
| Panel E: 10x2 | | | | | | |
| F | 1.000 | -0.825 | -0.813 | -0.812 | -0.795 | -0.772 |
| ATM | -0.825 | 1.000 | 0.998 | 0.990 | 0.972 | 0.941 |
| +25bps | -0.813 | 0.998 | 1.000 | 0.989 | 0.970 | 0.934 |
| +50bps | -0.812 | 0.990 | 0.989 | 1.000 | 0.986 | 0.959 |
| +100bps | -0.795 | 0.972 | 0.970 | 0.986 | 1.000 | 0.976 |
| +200bps | -0.772 | 0.941 | 0.934 | 0.959 | 0.976 | 1.000 |
| Panel F: 10x5 | | | | | | |
| F | 1.000 | -0.836 | -0.832 | -0.826 | -0.807 | -0.772 |
| ATM | -0.836 | 1.000 | 0.995 | 0.986 | 0.957 | 0.910 |
| +25bps | -0.832 | 0.995 | 1.000 | 0.982 | 0.953 | 0.907 |
| +50bps | -0.826 | 0.986 | 0.982 | 1.000 | 0.987 | 0.955 |
| +100bps | -0.807 | 0.957 | 0.953 | 0.987 | 1.000 | 0.980 |
| +200bps | -0.772 | 0.910 | 0.907 | 0.955 | 0.980 | 1.000 |

**Figure 12:**
**Risk factor sensitivities**

The figures on the left illustrate swaption sensitivities to interest rate changes (modified delta) during the backtesting period and the figures on the right depict sensitivities with respect to volatility changes (numerical vega). The modified delta represents a change in the present value of a contract when interest rates are shifted by 1 bps and the numerical vega denotes the value change resulting from 100 bps shift in implied volatilities. Both measures are computed for long receiver swaption positions. The values do not reflect actual price sensitivities accurately as the computations do not entail the interaction between interest rates and implied volatilities. Instead, the figures provide an illustration about the impact of changes in moneyness and time to maturity.

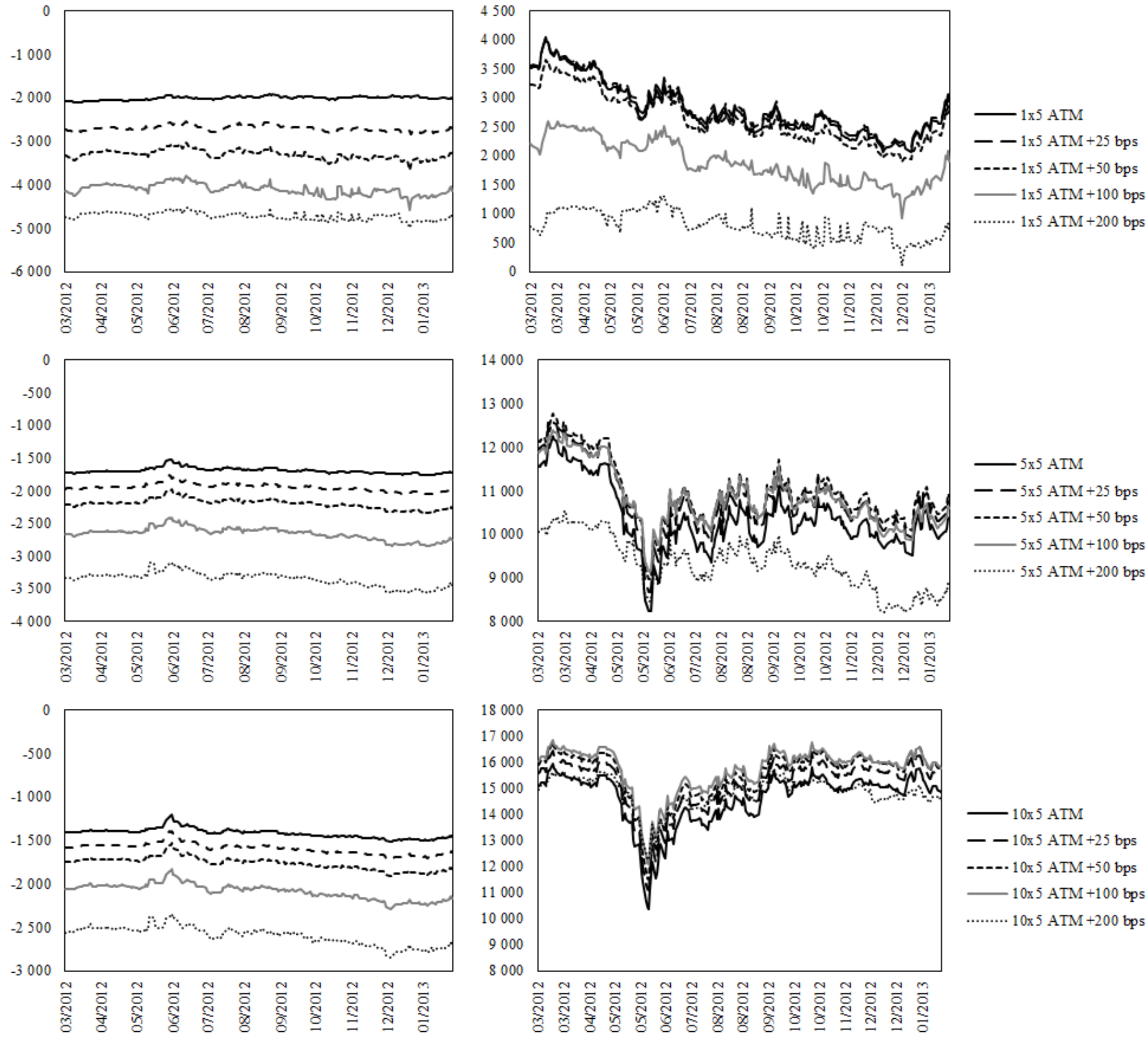
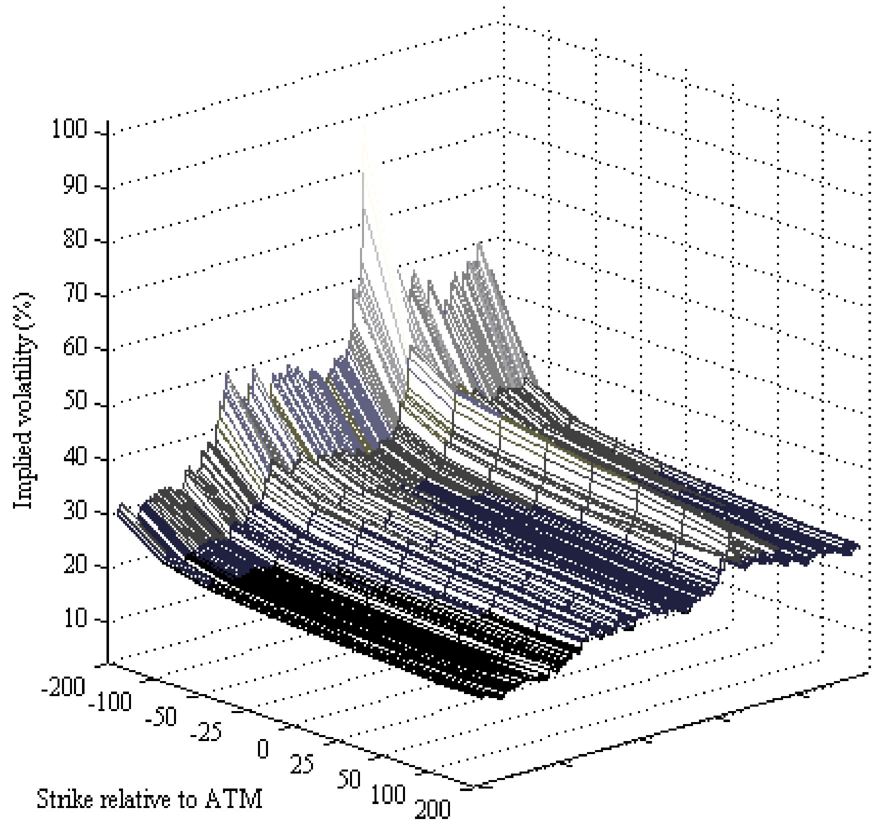


**Figure 13:**
**Implied volatility skew evolution**

The figure shows how the skew for 5x5 swaption evolves during the observation period. While the level of the implied volatilities varies to some extent, the shape of the skew stays quite unchanged especially for strikes above ATM level.



# APPENDIX C: ADDITIONAL VAR STATISTICS

This appendix presents the VaR exception results for short receiver swaption positions and average VaR components for Proxy_HS models with fixed and random smile methods.

## Table 18: VaR exceptions for short positions

The table shows the number of VaR exceptions for each of the series when estimated for short receiver swaption positions. VaR confidence level is 95% and the number of daily observations is 228. Consequently, the expected proportion of exceptions is 5%, i.e. number of hits should be approximately 11.

| | Proxy_HS | | Proxy_TW | | Direct_HS | | Direct_TW | | Component_HS | | Proxy_FHS | | Direct_FHS | | Component_FHS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | | | | | | | | | | | | |
| ATM | 4 | | 14 | | 6 | | 13 | | 7 | | 10 | | 18 | | 20 | |
| ATM +25 bps | 3 | 3 | 12 | 13 | 3 | 3 | 12 | 12 | 7 | 5 | 9 | 9 | 12 | 11 | 19 | 17 |
| ATM +50 bps | 3 | 3 | 10 | 12 | 3 | 3 | 10 | 13 | 5 | 3 | 9 | 10 | 13 | 12 | 16 | 13 |
| ATM +100 bps | 3 | 2 | 11 | 10 | 3 | 3 | 10 | 11 | 4 | 2 | 10 | 10 | 11 | 10 | 12 | 10 |
| ATM +200 bps | 2 | 2 | 12 | 12 | 2 | 2 | 12 | 12 | 2 | 2 | 10 | 11 | 11 | 10 | 11 | 9 |
| Panel B: 1x5 | | | | | | | | | | | | | | | | |
| ATM | 1 | | 11 | | 5 | | 8 | | 8 | | 7 | | 13 | | 18 | |
| ATM +25 bps | 1 | 1 | 10 | 10 | 3 | 3 | 9 | 8 | 7 | 5 | 6 | 6 | 9 | 9 | 16 | 13 |
| ATM +50 bps | 0 | 0 | 8 | 8 | 1 | 2 | 5 | 6 | 4 | 3 | 5 | 5 | 8 | 9 | 18 | 12 |
| ATM +100 bps | 1 | 0 | 9 | 9 | 1 | 0 | 7 | 8 | 4 | 0 | 8 | 7 | 8 | 7 | 14 | 8 |
| ATM +200 bps | 1 | 1 | 9 | 9 | 1 | 1 | 9 | 9 | 1 | 0 | 10 | 10 | 11 | 10 | 11 | 11 |
| Panel C: 5x2 | | | | | | | | | | | | | | | | |
| ATM | 2 | | 9 | | 3 | | 9 | | 11 | | 5 | | 8 | | 15 | |
| ATM +25 bps | 2 | 2 | 4 | 6 | 2 | 3 | 4 | 8 | 10 | 8 | 2 | 3 | 5 | 6 | 17 | 15 |
| ATM +50 bps | 2 | 2 | 5 | 8 | 3 | 3 | 6 | 8 | 12 | 6 | 6 | 4 | 9 | 10 | 19 | 14 |
| ATM +100 bps | 1 | 2 | 10 | 11 | 3 | 3 | 10 | 11 | 11 | 5 | 8 | 6 | 12 | 10 | 21 | 12 |
| ATM +200 bps | 12 | 12 | 28 | 15 | 15 | 7 | 28 | 13 | 9 | 2 | 28 | 22 | 33 | 15 | 23 | 11 |
| Panel D: 5x5 | | | | | | | | | | | | | | | | |
| ATM | 2 | | 9 | | 4 | | 9 | | 10 | | 7 | | 9 | | 16 | |
| ATM +25 bps | 2 | 2 | 8 | 9 | 4 | 3 | 8 | 8 | 10 | 7 | 5 | 5 | 8 | 10 | 17 | 15 |
| ATM +50 bps | 2 | 2 | 7 | 8 | 3 | 4 | 7 | 7 | 12 | 5 | 6 | 6 | 10 | 10 | 18 | 14 |
| ATM +100 bps | 2 | 2 | 12 | 9 | 3 | 3 | 12 | 9 | 12 | 4 | 9 | 8 | 11 | 11 | 19 | 13 |
| ATM +200 bps | 4 | 4 | 23 | 12 | 10 | 4 | 23 | 11 | 8 | 3 | 21 | 18 | 22 | 13 | 18 | 10 |
| Panel E: 10x2 | | | | | | | | | | | | | | | | |
| ATM | 2 | | 7 | | 6 | | 8 | | 10 | | 5 | | 9 | | 18 | |
| ATM +25 bps | 2 | 2 | 7 | 6 | 6 | 5 | 8 | 6 | 10 | 8 | 5 | 5 | 11 | 11 | 19 | 18 |
| ATM +50 bps | 1 | 2 | 6 | 8 | 6 | 6 | 6 | 8 | 11 | 7 | 5 | 5 | 8 | 8 | 19 | 15 |
| ATM +100 bps | 4 | 3 | 8 | 7 | 7 | 6 | 8 | 7 | 9 | 6 | 5 | 5 | 12 | 8 | 20 | 10 |
| ATM +200 bps | 7 | 7 | 20 | 9 | 8 | 7 | 20 | 9 | 8 | 5 | 18 | 20 | 26 | 10 | 20 | 6 |
| Panel F: 10x5 | | | | | | | | | | | | | | | | |
| ATM | 1 | | 7 | | 5 | | 7 | | 7 | | 6 | | 7 | | 15 | |
| ATM +25 bps | 1 | 2 | 7 | 7 | 4 | 5 | 7 | 7 | 7 | 5 | 6 | 6 | 7 | 9 | 18 | 13 |
| ATM +50 bps | 1 | 2 | 6 | 9 | 5 | 5 | 6 | 9 | 7 | 5 | 6 | 6 | 8 | 8 | 19 | 12 |
| ATM +100 bps | 2 | 2 | 8 | 8 | 5 | 5 | 9 | 8 | 8 | 5 | 8 | 8 | 10 | 9 | 18 | 9 |
| ATM +200 bps | 6 | 6 | 16 | 8 | 7 | 5 | 16 | 8 | 8 | 4 | 15 | 14 | 18 | 8 | 16 | 7 |

**Table 19:**
**Average VaR components**

The table presents average VaR components for Proxy_HS models with fixed ("ATM") and random ("Smile") smile methods. IR stands for interest rate and Vol for implied volatility as defined in 4.2.5. Theoretical Vol refers to implied volatility component obtained through applying the theoretical volatility shock in a similar fashion as in skew dependent VaR model (2.8.5) but without the according interest rate change.

| | IR | Vol | | Theretical Vol | |
|---|---|---|---|---|---|
| | ATM | ATM | Smile | ATM | Smile |
| Panel A: 1x2 | | | | | |
| ATM | 5 131 | 2 088 | | 1 716 | |
| ATM +25 bps | 8 674 | 2 099 | 2 091 | 1 727 | 1 883 |
| ATM +50 bps | 10 803 | 1 771 | 1 861 | 1 460 | 1 762 |
| ATM +100 bps | 13 305 | 1 095 | 1 185 | 908 | 1 207 |
| ATM +200 bps | 14 691 | 394 | 411 | 330 | 439 |
| Panel B: 1x5 | | | | | |
| ATM | 16 214 | 5 846 | | 4 051 | |
| ATM +25 bps | 23 822 | 5 814 | 5 959 | 4 031 | 4 362 |
| ATM +50 bps | 29 371 | 5 209 | 5 234 | 3 626 | 4 292 |
| ATM +100 bps | 36 326 | 3 598 | 3 672 | 2 529 | 3 549 |
| ATM +200 bps | 41 333 | 1 471 | 1 749 | 1 054 | 1 935 |
| Panel C: 5x2 | | | | | |
| ATM | 7 534 | 4 787 | | 2 841 | |
| ATM +25 bps | 9 836 | 4 871 | 4 801 | 2 891 | 2 984 |
| ATM +50 bps | 10 863 | 4 846 | 4 815 | 2 878 | 3 303 |
| ATM +100 bps | 12 439 | 4 648 | 4 824 | 2 765 | 4 134 |
| ATM +200 bps | 11 886 | 3 908 | 4 111 | 2 335 | 4 502 |
| Panel D: 5x5 | | | | | |
| ATM | 19 198 | 11 797 | | 6 203 | |
| ATM +25 bps | 22 961 | 11 909 | 11 671 | 6 263 | 7 102 |
| ATM +50 bps | 26 977 | 11 859 | 11 762 | 6 238 | 7 892 |
| ATM +100 bps | 31 945 | 11 293 | 11 412 | 5 950 | 9 641 |
| ATM +200 bps | 34 361 | 9 331 | 9 294 | 4 944 | 10 446 |
| Panel E: 10x2 | | | | | |
| ATM | 7 826 | 5 129 | | 3 213 | |
| ATM +25 bps | 8 897 | 5 169 | 5 090 | 3 239 | 3 497 |
| ATM +50 bps | 10 592 | 5 190 | 5 260 | 3 252 | 4 014 |
| ATM +100 bps | 12 527 | 5 104 | 5 148 | 3 202 | 4 585 |
| ATM +200 bps | 13 617 | 4 710 | 4 597 | 2 962 | 5 391 |
| Panel F: 10x5 | | | | | |
| ATM | 18 930 | 12 461 | | 6 544 | |
| ATM +25 bps | 21 610 | 12 556 | 12 433 | 6 596 | 6 893 |
| ATM +50 bps | 25 645 | 12 594 | 12 177 | 6 618 | 7 885 |
| ATM +100 bps | 30 419 | 12 363 | 12 024 | 6 505 | 10 218 |
| ATM +200 bps | 33 399 | 11 335 | 10 584 | 5 987 | 12 794 |