

Derivative market: efficient option pricing models and predictive informational content

Martins Feldmanis^α

Juris Rumba^β

Abstract

In this study we have examined the informational content of OMXS30 index European style call and put Options which are traded on the OMX Swedish stock exchange by applying extensions of the BS model. Firstly, we use two models (Gram-Charlier expansion and Model-Free) to obtain robust implied higher order moment estimates. Furthermore, we develop methodology which allows us to control for higher order moments and market based factors simultaneously. We find that the incorporation of skewness and kurtosis decreases structural pattern (smile) for Gram-Charlier implied volatility surface and that the market based factors, in particular liquidity, affect option pricing. Secondly, we observe that in the money put/call volume ratio has significant and robust predictive power in comparison to other types of volume ratios (including popular market sentiment put/call ratio). Finally, we test whether the implied volatility (controlled for skewness and kurtosis) is informative of the future realized volatility of the underlying. Our results confirm the existence of volatility premium which is increasing with time horizon. Furthermore by examining non-linear predictive relationship using quantile regression technique, we find that in times of market turmoil (high realized volatility) the implied volatility premium is negative.

JEL Classification: G12, G13, G14

Keywords: Volatility surface; Implied volatility and its informational content; Put/Call volume ratio, Model-Free implied volatility; Gram-Charlier Expansion

Tutor: Roméo Tédongap

Date: December 10, 2012, Stockholm

^α40280@student.hhs.se

^β40269@student.hhs.se

Table of contents

1	Introduction	4
2	Literature review.....	6
2.1	Model based explanation	6
2.2	Market based explanation	7
2.3	Role of derivative markets	9
3	Data description	13
3.1	Nasdaq OMX index.....	13
3.2	Option database and Swedish market.....	15
3.3	Risk free rate and Dividend yield	15
3.4	Data cleansing	16
4	Methodology.....	18
4.1	Implied skewness and kurtosis estimation	18
4.1.1	Gram-Charlier series expansion	18
4.1.2	Gram-Charlier adjusted volatility (GC volatility)	22
4.1.3	Model Free Implied Measures	23
4.2	Regression analysis of the Gram-Charlier IV.....	25
4.3	Predictability	27
4.3.1	Index Return Forecastability	27
4.3.2	Realized volatility predictability using IV	29
5	Empirical results.....	30
5.1	Robustness of GC measures.....	30
5.2	Regression analysis of the IV smile	31
5.3	Index return predictability	35
5.4	Realized volatility predictability.....	37
6	Conclusion.....	41
7	Limitations and suggestions for further research.....	42
8	References	44
	Appendices.....	53

The list of figures

Figure 1: Distribution of OMXS30 index returns.....	14
Figure 2: OMXS30 Index and GC IV.....	14
Figure 3: Cubic interpolation for call BS IV	25
Figure 4: Scatter plot of put and call volume ratio (ITM and All) and realized returns	36
Figure 5: Non-linearity of short term horizon (call and put options)	39
Figure 6: Non-linearity of medium term horizon (call and put options)	40
Figure 7: Non-linearity of Long term horizon (call and put options)	40

The list of tables

Table 1: F-Test values for joint test.....	32
Table 2: Unbiasedness test	38
Table 3: Average volatility premium.....	41

Abbreviations

IV – implied volatility

GC – Gram Charlier expansion

MF – Model Free

MF IV – Model Free implied volatility

GC IV – Gram Charlier implied volatility

BS – Black and Scholes

OTM – out of the money (option)

ITM – in the money (option)

DTM – days to maturity

OMXS30 – OMX Stockholm 30 Index

1 Introduction

Since the introduction of Black and Scholes (B-S) European style Call and Put option pricing framework in 1973, it is possible to observe that deep in the money options and deep out of the money options exhibit on average higher implied volatility than at the money options which have resulted in option implied volatility smile (Hentschel, 2003; Parasuraman & Ramudu, 2011). However, this deviation has been especially pronounced after 1987 S&P crash as shown by Bates (2000). In addition, BS implied volatility smile (and other patterns such as smirk or grin) are not unique phenomenon for one market or asset class due to the fact it has been documented across several markets within different asset classes as shown by Beneder & Elkenbracht-huizing (2003), Deuskar, Gupta, & Subrahmanyam (2007) and Jackwerth (2004).

Consequently, the academic literature has tried explaining this shape phenomenon via model based explanation or market based explanation. The research focused on the pricing models implies that the Black and Scholes model is not the most efficient predictor of the European Call and Put option prices, because one of the assumptions for BS pricing model is that the underlying return distribution is normally distributed. However, it is common to observe that the distribution has excess skewness and kurtosis (Derman, Kani, & Zou, 1996) which introduces bias in the pricing procedure. As a result, there have been different propositions to introduce new option pricing models which would take into account the non-normality of the distribution such as Gram-Charlier expansion (Corrado & Su, 1997) as well as various versions of stochastic volatility models have been considered by Bakshi, Cao, & Chen (1997), Heston (1993) and others. The market based perspective on the other hand introduces additional factors that determine the price of European style options. For instance, market maker inventory perspective indicates that the associated costs for the market maker to hold assets and provide corresponding liquidity to the market leads to different option prices and different volatilities (Demsetz, 1968). The adverse selection view within different market participants leads to situation where the informed market participants can exploit uninformed market participants (in particular market makers) and thus the market maker demands compensation for these risks (Bartram, Fehle, & Shrider, 2008). As a result, from this perspective the main determinant of the correct price of the call and put options are investor heterogeneity, their beliefs and the supply and demand in the market, which is also intrinsically linked to liquidity considerations

Another research area has concentrated on the informational efficiency of the derivative markets. This research area is driven by the notion that the information extracted from the derivative

market, where smaller and more sophisticated group of investors is concentrated, is more efficient than the information obtained from the stock market. Grosman (1987) has argued that derivative markets can help to forecast returns of the underlying. For instance, the volume information within derivative market is considered among the practitioners as an indicator of the broader market sentiment and it has been identified that it can provide information about the future returns of the underlying (Kuo & Lin, 2011). Furthermore, it has been found that the implied volatility from the derivative market is more efficient predictor of the realized volatility than the historical volatility (Jiang & Tian, 2005; Pan & Poteshman, 2006) which implies that the derivative market can signal about the future disturbances in the underlying market.

As a result, this paper will cover following areas. Firstly, we examine how the introduction of skewness and kurtosis in the option pricing model affects the volatility surface of the OMX Stockholm 30 (OMXS30) index European style call and put options and whether the market based factors (specifically liquidity considerations), after controlling for skewness and kurtosis, play an important role in option pricing. Hence, our research provides important bridge between market-based and model-based explanations, because we propose methodology to disentangle implied higher order moments for the volatility surface. Secondly, we examine whether it is possible to predict underlying returns using information (put/call volume ratio) from the derivative market. Thirdly, we examine the informational content of the implied volatility by using predictive linear and non-linear OLS regressions to forecast OMXS30 realized volatility. Extension of the non-linearity provides important insights for the market makers and practitioners, and to our knowledge this is the first analytical depiction of such phenomenon. As a result, this paper will contribute to current literature by exploring the factors affecting the volatility smile and uncovering the role of the option market in predicting the volatility and return in the market.

The structure of this paper is as follows: Section 2 – provides academic literature overview, Section 3 – describes dataset, Section 4 – describes methodology, Section 5 – describes empirical findings, Section 6 – concludes and Section 7 – provides limitations and suggestions for further research.

2 Literature review

There have been numerous empirical evidences of observed volatility smile within BS option implied volatility surface. In our opinion, two major research strands have focused on either development of more sophisticated option pricing models or have looked on it from market based perspective, which focuses on inventory view, adverse selection and investor preference (including their belief) heterogeneity. It is important to note that in model-based perspective, existence of option implied smile can simply be an artefact of incorrect model specification.

2.1 Model based explanation

History and development of option pricing

Since the early introduction of Black & Scholes (1973) option pricing model (BS), there have been numerous advancements in contingent claim valuation. The BS model explicitly assumed: 1) there is no arbitrage opportunity 2) perfect and efficient market (costless short-selling, unlimited ability to lend and borrow and no other transaction costs) 3) no dividend for underlying 4) underlying process follows Geometric Brownian motion with constant drift and volatility. Subsequent research has shown that these assumptions are not realistic; therefore another popular extension of BS framework is to introduce market imperfections. Leland (1985) introduces transaction costs. Broadie, Cvitanic, & Soner (1998) explicitly models replicating portfolio constraints, such as gamma (second order derivative with respect to underlying), portfolio size and number of shares constraints, and its implication on the price of related claim. Similarly, Karatzas & Kou (1996) consider various market imperfection issues. In order to evaluate BS model assumption of normality feasibility, Jackwerth & Rubinstein (1996) for the S&P 500 index and Ring (2011) for Swedish stock-market have shown that recovered risk-neutral PDF is not log-normal. In line with non-normality argument Derman et al. (1996) uses volatility surface as a whole and derives local volatilities in the binomial pricing model framework. More recent approach to extend BS model includes inclusion of two log-normal distributions such as Lim, Martin, & Martin (2005) and Melick & Thomas (1997).

Cox & Ross (1976) develop BS analysis further and acknowledge that it is possible to derive analytical expression that can be calculated as its expected value using the risk-free rate. Furthermore the pricing kernel used for deriving option prices must satisfy martingale restriction (the implied underlying value has to be equal to market price), otherwise there will be an arbitrage opportunity as shown by Harrison & Kreps (1979) and Longstaff (1995); yet, they find that option models who relax this

assumption perform better. Gou, Han, & Ryu (2011) and Neumann & Schlag (1996) provide both parametric and non-parametric ways how to estimate martingale restriction.

Another extension of the BS model is to relax the assumption about fixed volatility, underlying process and include possibility of model jumps. Most popular is Heston (1993) stochastic volatility model. Boes, Drost, & Werker (2007) show that overnight price jump account for $\frac{1}{4}$ of total jump risk and thus indicating another possible extension of the BS and stochastic models. Similarly Bakshi et al. (1997) derives model with stochastic jumps, volatility and interest rates and show that they are important for pricing and internal consistency as well as they find that incremental inclusion of stochastic volatility yields the best marginal result with all subsequent stochastic additions gradually improving pricing performance.

Due to availability of computation power another recent focus of academic discussion has been to integrate time series analytic methods in option pricing. Heston & Nandi (2000) develop a closed form option pricing model, where underlying spot price follows GARCH type dynamics. Christoffersen & Jacobs (2004) out of sample analysis indicates that the model which incorporates volatility clustering and leverage effects would deliver the most efficient results as well as they provide detailed discussion on various GARCH-like option pricing models.

Due to vast amount of literature available on this topic there have been a number of meta-studies to evaluate option pricing model performance. A survey of various option pricing models is provided by Broadie & Detemple (2004) who show that models with more detailed structure (liquidity, information arrival) and more tractable models that incorporate stochastic volatility are useful avenues for future research. However, they argue that much of the future advances cannot be solved by increases in computational power alone, because new algorithmic (and conceptual) advances will be required.

2.2 Market based explanation

However, realization of the IV smile can be rooted in market based explanations, because markets are not fully efficient; market makers require compensation for various additional risks, when providing (shorting) out of the money options (which are more difficult to hedge). While institutional and behavioural considerations might indicate inefficiency of the market, they simultaneously provide additional perspective on the IV smile origins.

Inventory model

The inventory model put forward by Demsetz (1968) and Ho & Stoll (1981) and many others suggest that liquidity depends on market makers associated costs to hold assets and provide corresponding liquidity to the market. Comerton-Forde, Hendershott, Jones, Moulton, & Seasholes (2010) empirically examine link between specialist trading firm balance sheets and overall market liquidity measures. They conclude that the overall market liquidity is affected by the specialist trading firms if these firms have large positions in the underlying or have experienced significant losses. Atanasov & Gyoshev (2001) and Bollen & Whaley (2004) provide evidence that institutional end investors have strong preference for long put positions that provide their portfolios with protection. It simultaneously requires that market makers have to short large amounts of put options in absence of perfect hedging. Another approach is to view how institutional investors can benefit from holding call options, for instance, net buying pressure would not matter if market makers would be able to perfectly replicate the option strategies; yet, continuous hedging is costly (Martellini & Simsek, 2005). Recent attempts to alleviate hedging costs by reviewing regression based technique of delta hedge ratio for Swedish stock-exchange was done by Kiploks & Lazdins (2011). Longstaff (1995) shows that implied underlying value by pricing kernel is negatively correlated with market based liquidity measures giving support to liquidity based explanation for option smile. Another explanation supporting liquidity-based explanation for smile is the fact that the traded volume for ITM and OTM options usually is smaller than for ATM options. Brenner, Eldor, & Hauser (2001), show that not-traded options on average are priced 21% lower than their exchange counterparts.

Adverse selection view (asymmetric information)

Vijh (1990) show that option spread explained by adverse selection component is relatively small. On the contrary Bartram et al. (2008) and Cao, Chen, & Griffin (2005) show that market makers face significant adverse selection from informed participants. If perfect hedging would be attainable market makers would not require additional compensation, nevertheless in practice even dynamic continuous hedging is costly. Thus, informed investors might trade in options and signal information to market makers. However, the impact from trade to option IV is not clear, because it depends on interpretation by the market makers. If trading volume is driven by informed investors, market maker would require higher premium, while if volume is driven by noise traders they would lower premium to provide larger liquidity.

Other views

Shefrin (1999) notes that underlying market trend (and expectations for reversal) can influence the underlying IV shape, which are especially pronounced in case of large heterogeneity in investor beliefs. Ziegler (2007) shows that some degree of explanation on implied volatility smile can come from heterogeneous preferences across different agents (investors); however, he cautiously notes that existing heterogeneity is unlikely to explain the volatility smile fully. Shefrin (1999) notes that bullish investors on market overbid call options, while bearish investors on market overbid put options resulting in option smile. This notion is extended by Barr (2009), who develop demand-based pricing equilibrium models to capture IV smile. Meyers (1994) notes that daily call put volume ratio (CPR) is signal for optimistic sentiment, while Shefrin (1999) notes that comparison of put and call implied volatilities can signal change in corresponding market sentiment. Figlewski (1989) provides critical assessment of option model valuation and market forces and argues that no model can capture the real underlying forces for demand and supply.

2.3 Role of derivative markets

Further we provide discussion on role of IV for predictability purposes and implemented return strategies based on the information extracted from the derivative market as well as predictability of index returns and realized volatility.

Index return predictability

Index return predictability stems from the fact that derivatives markets play important informational role and due to limited investor heterogeneity (by limited access of noise traders) they provide more superior information. Chakravarty, Gulen, & Mayhew (2004) have estimated that the marginal contribution of the option market to the price discovery process could be around 17%, thus confirming the notion that derivatives play important informational role. The market efficiency is revisited by Battalio & Schultz (2006) who find that rational investors in option market even during Internet stock bubble could synthetically avoid short sale restriction using derivative markets; nevertheless, they find that the synthetic stock prices were very close to the actual stock prices. Finally, L. J. Ho (1993) shows that stock prices with options traded on them 'anticipate' earnings' change earlier than stocks with no options.

However, there are numerous attempts to predict and explain index returns using information from option markets. Grossman (1987) argues that lack of trading in options today can signal information for stock returns. In addition, lack of trading indicates larger difficulties for market makers

to forecast required capital. Vilkov & Xiao (2012) develop tail loss measure derived through extreme value theory and show that using OTM put prices it is possible to predict large negative returns (i.e. market crashes). Balyeat & Erturk (2005) shows that natural log of the ratio for OTM call to equally OTM put price is positively related to the realized return over the life of the options.

Banerjee, Doran, & Peterson (2007) and Dumas, Fleming, & Whaley (1998) use another approach and show that volatility index predicts stock market indices even when controlling for common risk factors. Nevertheless this relationship is not linear and there are two explanations for return-volatility asymmetry: 1) leverage effect and 2) volatility feedback effect. Leverage effect hypothesis argues that with increase in leverage of firm, volatility of the equity should increase, because equity holders bear the maximum risk. On other side volatility-feedback argues that increases in volatility trigger negative effects due to increase in expected returns which is attributable to the drop in current stock prices. Badshah (2010) argues that both explanations fail to explain the asymmetric relationship fully. These theories have been extended and Bekaert & Wu (2000) who argue that firm returns are more correlated in down markets when overall market volatility rises. In addition, Avramov, Chordia, & Goyal, (2006) find that following positive returns informed investors tend to liquidate their positions and thus reducing volatility but after negative return uninformed investors tend liquidate their positions and thus leading to an increase in volatility.

Investment strategies based on option information

One branch of investment strategies are based on extracting stock-specific option implied information and then creating corresponding ranks for options. For example, Bali & Murray (2012), Rehman & Vilkov (2012) and Xing, Zhang, & Zhao (2008) show that stock-specific smirk is able to forecast cross-sectional relative equity returns. Another approach by Pan & Poteshman (2006) provide results that stocks with low put-call traded volume ratio overperform their counterparts. Driessen, Maenhout, & Vilkov (2009) show that within stock correlation is priced and thus it is possible to form profitable investment strategy to exploit it. In addition, by considering the return generation in 'alpha' context, Demiguel, Plyakha, Uppal, & Vilkov (2012) show that using option implied moments (including skewness) it is possible to improve selection of mean variance portfolios.

Furthermore, there have been more direct attempts to utilize options in order to derive profitable return strategies. For example, Bondarenko (2003) show that strategies selling deep-out of the money (OTM) and at the money (ATM) puts would have resulted in substantial profits. From buyers' perspective ATM puts are traded at average excess return of -39%. The rationale for this phenomenon

stem from a) risk-premium b) peso-problem (where event that could have happened did not happen) c) biased beliefs. However, Branger & Schlag (2005) argues that peso-problem is driven by unknown path, rather than unknown terminal stock price thus indicating that high OTM put prices are not overpriced. There have been also attempts to relate option strategy returns to risk free returns. Coval & Shumway (2001) show that expected call returns exceed risk free rate and put returns, while ATM straddle positions consistently lose money indicating that there is some additional factor (such as stochastic volatility priced in). Jones (2006) extends the option return argumentation and states that volatility risk premium rather than volatility per se and jump risk premium contribute to option strategy expected returns; however, they are insufficient in explaining the magnitude of returns. Finally Bookstaber (1981) provides criticism on Black and Scholes (1973) paper and argues that all option returns might be artefacts of non-simultaneous quotations.

Realized volatility predictability

Due to previously outlined argument that within option markets there are more informed investors, there has been empirical discussion whether the option implied volatility is best predictor of the future realized volatility. Generally prior tests concentrate on the following two regressions (equation 1 and 2) to test three distinct hypothesis:

$$\sigma_{t,t+\tau}^{RE} = \beta_0 + \beta_1 \sigma_{t,t+\tau}^{IV} + \beta_2 \sigma_{t,t}^{RE} + \varepsilon_t \quad (1)$$

where: $\sigma_{t-\tau,t}^{RE}$ – historical realized volatility; $\sigma_{t,t+\tau}^{IV}$ – option IV and $\sigma_{t,t+\tau}^{RE}$ – ex-post realized volatility from time period t. The historical forecast can be formed in different ways and it can include different historical time periods (e.g. 500 day forecast) to formulate it.

Informativeness: β_1 is significantly different from zero

Unbiasedness: β_1 is equal to 1 and β_0 is equal to zero

Informational efficiency: No other variables are significant (i.e. β_2 is zero), which incorporates also unbiasedness hypothesis

Similarly conclusions about informativeness and unbiasedness can be done with respect to following regression:

$$\sigma_{t,t+\tau}^{RE} = \beta_0 + \beta_1 \sigma_{t,t+\tau}^{IV} + \varepsilon_t \quad (2)$$

Bates (1995), Neely (2004) and Poteshman (2000) provide overview of different studies and sample period findings. In our opinion there are two main methods that are used to obtain daily estimate from IV curve in BS model: 1) instantaneous variance (using integration) as in Poteshman (2000) and 2) different combinations of IV weighting (vega, volatility elasticities, simple average) as reviewed by Mayhew (1995). As a result, it seems that the majority of studies favour informativeness and reject unbiasedness. There are four interpretations that support the rejection of unbiasedness: 1) it is difficult to implement strategies that will drive prices to proper levels 2) existence of other more profitable strategies for market makers and lack of incentive to profit from this mispricing 3) ability to profit from larger bid-ask spreads and 4) volatility risk premium. Finally, Chernov (2001) decomposes the IV into to risk factors: 1) market risk premium on underlying asset 2) risk of the volatility. He argues that these are the reasons why some studies have found IV as a biased predictor. Another important consideration outlined by Bandi & Perron (2002) is the effect of longer time horizons, in particular, they find that the unbiasedness hypothesis is confirmed for longer horizons. However, these results have been updated with development of improved option pricing models. Jiang & Tian (2005), Shu & Zhang (2004) provide assessment of MF IV and its superior informational content for future realized volatility. Ahoniemi & Lanne (2007) develop improved GARCH-like model that forecasts future realized volatility, which includes cross-effects from both put IV and call IV. Another alternative extension of the models is to adjust historical forecast. Szakmary, Ors, Kim, & Davidson (2003) using broad array of 35 future option markets show that IV performs well regardless whether historical forecast is modelled in historical sense or in GARCH framework. Ederington & Guan (2010) provide evidence for volatility asymmetry asymmetric time series model (EGARCH). They find that after large negative returns both implied and realized volatility rises while after substantial positive returns realized and implied volatility fall sharply beyond the forecasted level by EGARCH. In addition, after small positive returns, volatility remained at the same level while EGARCH predicted decline.

In line with Poteshman (2000), who provides overview of sources of bias for realized volatility estimation, we identify that there are four major types of possible errors: 1) market under study is inefficient 2) improbable events or peso-type problems are observed 3) misspecification of option pricing model 4) econometric problems. Market efficiency has generated so large empirical interest that we do not continue with its discussion and we have addressed already peso-type problem and misspecification of the option pricing model previously in the paper. However, econometric errors merits further discussion, which mainly fall in two broad categories: a) overlapping return problem b) error-in variable problem. Christensen & Prabhala (1998) after considering these two problems conclude

that option implied volatility outperforms historical volatility for predicting realized volatility. We also identify fifth problem which is estimation of realized volatility. While common approach is to use squared returns it is not the only approach used in literature. Other notable approaches include extreme value method that incorporates daily high and low data proposed by (Parkinson, 1980) and developed further by Yang & Zhang (2000) as well as usage of squared residuals from a regression model of returns by Maheu & McCurdy (2002), who consider this as an alternative.

3 Data description

3.1 *Nasdaq OMX index*

Nasdaq OMXS 30 index consists of the 30 most traded stocks in the Stockholm market Exchange, which ensures high liquidity for the underlying index constituents. Index is constructed as market weighted price index. The detailed index calculation formula is given as:

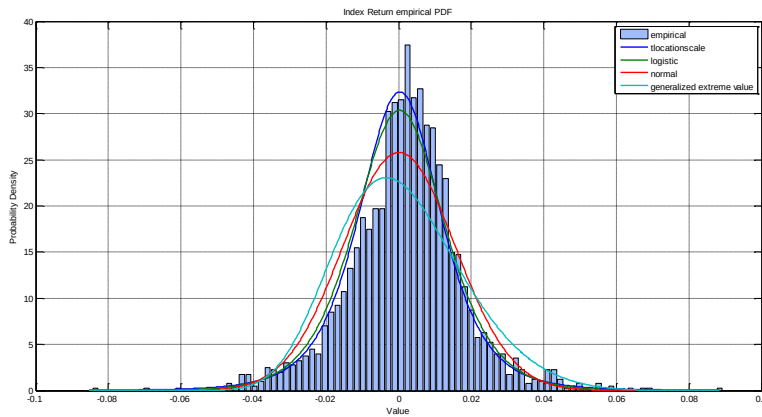
$$I_t = \frac{\sum_{i=1}^n q_{i,t} \cdot p_{i,t} \cdot r_{i,t}}{\sum_{i=1}^n q_{i,t} \cdot (p_{i,t} - d_{i,t}) \cdot r_{i,t} \cdot j_{i,t}} \cdot I_{t-1} \quad (3)$$

where: I_t – the index level at time t ; $q_{i,t}$ – number of shares of company i applied in the index at time t ; $p_{i,t}$ – the price in quote currency of a share in company i at time t ; $d_{i,t}$ – dividend only used for total return indexes and special dividend in price indexes; $r_{i,t}$ – foreign exchange of index quote currency of company i and $j_{i,t}$ – adjustment factor to adjust share price of company i due to corporate actions. The index constituents are reassessed twice a year.

Index return normality

Similarly to the well documented empirical phenomenon that stock returns do not exhibit normal distribution we also observe it in our sample period. The daily index returns as shown in figure below are clearly not normal with slight positive skewness (0.0015) and kurtosis (5.27). The normality of these returns is rejected by Jarque-Bera (JB) test with test value 489. Figure 1 also shows the 4 best fitting distributions (out of 17) sorted by Bayesian Information Criteria (BIC). The best fit is shown by tlocation-scale distribution, which is characterized by three parameters (mean, scale and shape).

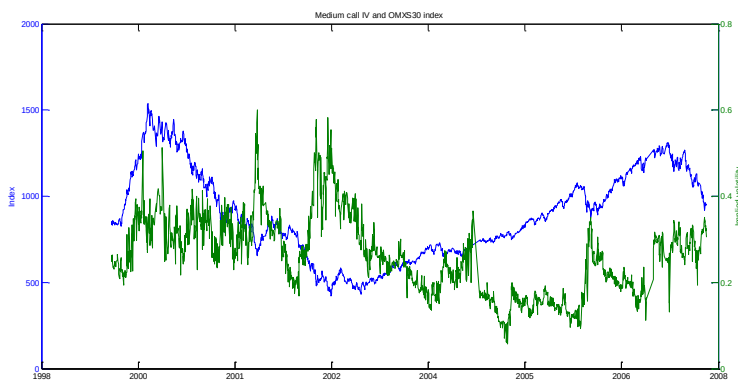
Figure 1: Distribution of OMXS30 index returns



Index level and daily medium-term IV

We also provide index level and GC IV development throughout the sample period, because it provides assessment of the underlying market dynamics. The sample period covers full business cycle, because it starts before 2001 recession and ends before 2009 Global financial crisis. As shown on a smaller sample period by Kiploks & Lazdins (2011) the two are negatively correlated (-0.074) indicating intrinsic relationship between IV and OMXS30 index level. The medium term options are chosen because they better reflect the expectations of underlying asset value development by incorporating both short term and long term investor views.

Figure 2: OMXS30 Index and GC IV



3.2 Option database and Swedish market

In order to classify options and provide distinction between different maturity effects we use following classification: short term (1-30 days); medium term (30-90days) and long term (above 90 days). In Appendix 2 we report the structure of our data-set with respect to option days to maturity. We observe that there is clear distinction at the boundary of 90 days, where the prevalence of options with longer maturities significantly decreases, thus it is chosen as a cut-off point.

We obtain option price data from FINBAS database, which was donated to the SSE (Stockholm School of Economics) by Nasdaq OMX in 2010. We also obtain OMXS30 index values from the same database. Briefly prior our initial sample period on April 27, 1998 the OMXS30 was split in factor 4:1, thus our sample period is likely to be intact from initial changes in market participants trading patterns, expectations and internal risk routines. Our database covers sample period from 1999.08.04 to 2008.02.01 (for both puts and calls). During the sample period there were on average 32 daily call contracts and 36 put contracts. Option data consists of detailed information on option closing bid, ask, high, low, transaction volume, open balance of particular option type. Nordén (2003) provides detailed discussion of Swedish derivative security market design. For the purposes of this paper we outline that options are traded in fully computerised system and trader can submit new orders electronically or via phone (in this case the order would be executed manually). The strike levels above 1000 are set 20 points apart, while strike levels below this threshold are set only 10 points apart. The fourth Friday of the expiration month (if it is a trading day) is the expiration date for derivative contracts. In addition, there exists upstairs market, where trades have to be reported no later than 15 minutes prior the next opening day. This market mainly serves for large-volume transactions. Fong, Madhavan, & Swan (2006) on Australia Stock Exchange find no evidence that competing upstairs market has adverse effects on the downstairs market. In addition, Nordén (2003) provides empirical evidence that this regime change induced reduction in concentration of option holding.

3.3 Risk free rate and Dividend yield

We obtain risk-free rate from Riksbanken and we base our risk-free rate calculations on Swedish National Debt Office (SNDO) issued bills (below 1Y maturity) and bonds (above 1Y in maturity). We match closest risk-free rate to the option time to maturity. Despite the fact that some earlier papers used STIBOR as a proxy for market risk-free rate, we chose to use government debt as a proxy for risk free rate. This approach was motivated by Hull (2009), who explicitly uses government risk-free rate in his textbook and the fact that longer maturity STIBOR rates have limited transaction volumes. Even

more, Hull & White (2012) argues that even in high market uncertainty there should not be done any funding valuation adjustment (the costs between funding of the hedge and risk free rate) for the option valuation. However, this claim is hotly contested by other market participants such as Castagna (2012). Despite the fact that there is limited amount of summary volume statistics for STIBOR, we consider Bernhardsen, Kloster, & Syrstad (2012) market assessment and use SNDO rate. By providing matched interest rate we partially capture risk-free rate dynamics and thus deem that detailed risk free rate modelling is unnecessary and beyond the scope of the paper.

NASDAQ-OMX (2011) does not provide adjustment for 'regular' dividends/cash distributions or capital reduction for the index; however based on the NASDAQ-OMX board judgment ad-hoc index adjustments for extra-ordinary dividends can be incorporated. As OMXS30 is price weighted index and no cash dividends are reinvested. In order to account for dividend yield, we use AFFWALZ (DY) dividend yield obtained from Datastream. Geske (1978) develops pricing model under stochastic dividends and argues that Black-Scholes model is biased and dividend yield is important. In addition, dividend yield is also an important concept in empirical research, for instance, Kalay & Subrahmanyam (1984) find market inefficiency evidence for ITM American calls around pre-dividend dates and Zivney & Alderson (1986) argues for hedged dividend capture by purchasing stock and writing a call on the underlying. Castanias, Chung, & Johnson (1988) using option open interest market data (to review early-exercise) argues that previous results of Kalay & Subrahmanyam (1984) are wrongly classified and that there is no evidence for market inefficiency.

In order to estimate option implied parameters we divide time until maturity (which includes holidays) by 360 in order to reflect the fact that risk-free rate is also earned during non-trading days.

3.4 Data cleansing

Despite the fact that Bakshi et al. (1997) have suggested removing options with very small values (typically deep out of the money) to mitigate the effect of discrete prices on option valuation we chose to retain such options in order to better capture the effect of the volatility smile. We also chose to retain options that do not have traded volume, because such restriction would reduce our sample by half. Jointly with negative bid exclusion our sample does not have options where relative bid-ask spread exceeds 2. The BS IV is 4.5% higher for call options with relative bid-ask spread above 1. Decision to leave such options stems from two arguments: 1) we want to explore how much market liquidity influences option smile 2) due to relative size of the market there might be inefficiencies in market price-discovery and buyer-selling matching, however existence of bid-ask spreads already indicates that

there is interest from the market participants. We want to extend research to incorporate information in options that are further away from the at-the-money point (moneyness of 1).

Another issue is what to do with negative implied volatilities. We chose to exclude them from our sample for two reasons – a) they are unlikely to provide any meaningful insight into analysis of the option smile b) we believe that the approach chosen by Canina & Figlewski (1993) to arbitrary set negative implied values to zero introduces exogenous undue influence on parameter estimates despite the fact that it might reduce the extent of observed volatility smile. They also find no information value in option implied volatility. As contrasted by Christensen & Prabhala (1998) these results are likely to be due to short period and sample selection (prior 1987 crash) biases. In our sample the bid-ask spread for such options was nearly two times larger (7.30 SEK) than for other options thus indicating that these options have obtained negative IV partially due to liquidity concerns.

There are also further issues to consider, when constructing our data sample. Evnine & Rudd (1985) finds that option prices often violate arbitrage constraint, put-call parity and are often substantially mispriced; however, by excluding options with negative and very large implied volatility we reduce obvious arbitrage constraint violators. In addition, leaving such options would not contaminate our sample with sample selection bias, where we impose structural form on arbitrage constraint, implementation costs and realistic arbitrage band. Similarly Hentschel (2003) argues that IV estimation for options far away from the money are contaminated with estimation errors (such as small error in option price will produce large error in IV). He shows that large part of far from the money options 95% confidence interval is in the range of $\pm 6\%$. In addition, Vijn (1990) argues that only small part of option trading has occurred in daily closing option bid-ask price range. Bearing all these issues in mind we proceed with our selected approach further.

As a result we have applied several filters to the data. In our analysis we use only options that have both bid and ask and due to large amount of data we are unable to report number of observations options that have only bid or ask quotes. We also remove options that have: i) negative BS implied volatilities b) have BS volatilities above 100% c) are deep in (above 1.3) or out (below 0.5) of the money. Moneyness is defined for calls as $\text{Index}/\text{Strike}$ and for puts as $\text{Strike}/\text{Index}$. As a result by using this filter we have removed 1741 call and 1519 put options.

We finally note that despite existing evidence of transmission of liquidity from derivative market to stock market, we excluded Nasdaq Nordic OMX bid and ask price quotes for the index, due to fact that the data was available only starting from 2005.

4 Methodology

The methodology is structured in three parts. The first part of the methodology covers the estimation procedures of the implied skewness and kurtosis for the option dataset as well as specifies how individual GC adjusted volatilities are obtained. Second part explains how structural form of the GC IV is assessed and examines whether there are patterns in the implied volatility. Third part focuses on predictability of index returns and realized volatility.

4.1 Implied skewness and kurtosis estimation

The dataset is split into three parts. The short term options have the maturity of 1 up to 30 days, medium term options with maturity from 31 to 90 days and long term options with maturity more than 90 days. The implied skewness and kurtosis are estimated for each day by compiling all the options which are available in that particular day and assigning them to the particular subsample: short, medium and long term. Hence, per one trading day there are three minimization procedures applied, assuming that at least three options per each subsample is traded or bid and ask is available. In order to ensure robustness of our higher order moment estimations, we also perform incremental optimizations (volatility only; volatility and skewness; volatility, skewness and kurtosis). The detailed estimation procedure is explained in the following sections.

4.1.1 Gram-Charlier series expansion

Black-Sholes model has been criticized for its rather strong assumptions and lack of ability to capture well documented empirical phenomenon such as negative skewness and positive excess kurtosis. Despite the fact that empirical literature for asset return modelling with higher order moments proposes other expansion procedures such as Cornish-Fisher expansion, a different expansion was chosen for this paper, because of available explicit put and call option formulas. In this paper we also disregard Chernozhukov, Fernandez-val, & Gallichon (2007) proposed monotonicity rearrangement, because we find little academic support of using this adjustment in smoothing approximating polynomial.

For the purposes of this paper we will employ Corrado & Su (1997) method that incorporates higher order moments (namely skewness and kurtosis). The method is based on Gram-Charlier probability density series expansion that is adjusted by non-normal skewness and kurtosis. The essence of this method is to approximate unknown distribution F in terms of a normal distribution and to recover the parameters for distribution F via inverse Fourier transform. This method was initially

proposed by Jarrow & Rudd (1982) and further used by Abken & Madan (1996), Backus, Foresi, & Wu (2004), Longstaff (1995) and Madan & Milne (1994).

The initial set-up is to re-write the unknown distribution F in terms of standard normal density function scaled by its moments. Variable z is standard parameter in Black and Scholes option pricing formula:

$$z = \frac{\ln\left(\frac{S_t}{S_0}\right) - \left(r - q - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}} \quad (4)$$

where: S_t – underlying spot price; S_0 – exercise price; r – continuously compounding interest rate; q – continuously compounding dividend yield; σ – volatility of underlying asset returns; t – time until expiration of option.

When the true underlying distribution is not known, but is likely to resemble normal distribution it is typical to express the unknown distribution $F(z)$ in terms of its normal density distribution $n(z)$ and $p_n(z)$, which is chosen in such a manner so $F(z)$ has the same first moments as the probability density function of z .

$$F(z) = n(z) \cdot p_n(z) \quad (5)$$

A common way to express $p_n(z)$ is using Hermite polynomials:

$$F(z) = n(z) \left[1 + \frac{\gamma_1}{3!} He_3(z) + \frac{\gamma_2}{4!} He_4(z) \right] \quad (6)$$

Jondeau & Rockinger (1999) using integration show that $\gamma_1 = skew$ and $\gamma_2 = excess\ kurtosis$. In addition, using Hermite polynomial structure it is possible to express them further, in order to obtain final resulting approximation.

$$F(z) = n(z) \left[1 + \frac{skew}{3!} (z^3 - 3z) + \frac{kurt - 3}{4!} (z^4 - 6z^2 + 3) \right] \quad (7)$$

The formula above is referred as Gram-Charlier type A expansion. Jurczenko, Maillet, & Negrea (2002) provides overview on various versions of the Corrado & Su (1997) formulas and points out that original Corrado and Su (1996) paper included minor error. In our paper we use the correct version of the formula. However, another type of expansion that uses similar argumentation is Edgeworth's expansion that retains the number of parameters to estimate, however adds one more Hermite polynomial.

$$F(z) = n(z) \left[1 + \frac{skew}{3!} (z^3 - 3z) + \frac{kurt-3}{4!} (z^4 - 6z^2 + 3) + \frac{skew^2}{72} (z^6 - 15z^4 + 45z^2 - 15) \right] \quad (8)$$

Nevertheless it is shown by Barton & Dennis (1952) and Rubinstein (1998) that such expansion (eq.8) is more limited in its positivity restriction, therefore we focus only on this Gram-Charlier type A (eq.7) expansion. The distinction between Gram-Charlier type A and Edgeworth's expansion is important, because they both are derived in a similar fashion and sometimes they are used interchangeably.

However, by construction normal distribution probability density function is always positive and we need to ensure that our expanding polynomial also satisfies the same properties, because otherwise our resulting distribution $F(z)$ will be with negative density values. Hence there are three parameters that determine positivity of probability density function: z , $skew$ and $kurt$. Therefore there is skewness-kurtosis relationship conditional on z that provides positive density function. Rubinstein (1998) uses binomial tree approach (its continuous time limit is the BS model) with 101 nodes to show this relationship, while Jondeau & Rockinger (1999) use analytical geometry to show this envelope-type relationship. They show that there are two envelopes one larger corresponding to $|z| \leq \sqrt{3}$ and one smaller for $|z| \geq \sqrt{3}$. The envelope plot is given in Appendix 7. For our estimations we chose to bind our solutions within the smaller envelope, because more than 50% of days have at least one option per day that falls outside $|z| \geq \sqrt{3}$. Due to the fact that we are estimating only one combination of $skew$ and $kurt$ per day and subsample that would yield that for some of the options our implied PDF is negative, which would complicate the estimation of individual option GC IV (discussed in next sections of this paper). In addition, usage of proposed smaller boundary is supported by empirically observed return parameters (both in our sample as well as in general). Detailed overview of realized index moments is provided in Appendix 1, which shows that majority of estimated results are in line with the proposed envelope-type boundaries.

Another approach developed by Gallant & Tauchen (1989) is to take expansion term squared and in order to ensure that the density integrates to 1 divide by the expansion integral. However, we do not select this approach, because such solution would yield only numerical estimation procedures for call formulas. Straja (2003) also provides discussion on problems associated with introducing other terms to overcome positivity boundary such as: 1) more difficult search 2) can reduce positivity boundary 3) additional terms can induce numerical instability or multicollinearity between parameters.

We also outline that such methods are likely to be computationally more cumbersome while deriving put and call pricing kernel.

Another aspect which is important for the resulting density expansion is multimodality. While it can be argued that in any given time there are several most likely paths (such as several most likely scenarios of future that are priced in stock price/returns) that the stock returns can take thus giving theoretical support for multimodal densities. As outlined by Jondeau & Rockinger (1999) GC density expansion for some values of *skew* and *kurt* can yield multimodal density. However, more explicit approach to model multimodality is taken by mixture of log-normals approach (Lim et al., 2005; Melick & Thomas, 1997).

Applying the density $F(z)$ to derive a theoretical call price as the present value of an expected payoff at option expiration yields the following expression:

$$C_{GC} = e^{-rt} \int_K^{+\infty} (S_t - K) F(z) dz (S_t) \quad (9)$$

We also note that integral evaluation of $F(z)$ is stock-price varying and resulting formula of the Gram-Charlier expansion is:

$$C_{GC} = C_{BS} + skew \cdot Q_3 + (kurt - 3) \cdot Q_4 \quad (10)$$

$$Q_3 = \frac{1}{3!} S_0 \sigma \sqrt{t} ((2\sigma\sqrt{t} - d)n(d) + \sigma^2 t N(d)) \quad (11)$$

$$Q_4 = \frac{1}{4!} S_0 \sigma \sqrt{t} ((d^2 - 1 - 3\sigma)\sqrt{t}(d - \sigma\sqrt{t})n(d) + \sigma^2 t^{3/2} N(d)) \quad (12)$$

$N(d)$ – standard cumulative normal distribution, $n(d)$ – standard normal density and other terms have been defined previously in eq.4.

The Gram-Charlier option formula nests the Black and Scholes model, because terms Q_3 and Q_4 will disappear if the higher order moments are set to 0 and 3 respectively (as it is the case for normal distribution). In order to estimate the unknown standard deviation (ISD), skewness (ISK) and kurtosis (IKT) parameters we follow the approach of Corrado & Su (1997) and minimize the sum of squares.

$$\min_{ISD, ISK, IKT} \sum_{j=1}^N \left[C_{obs,j} - [C_{BS,j}(ISD) + ISK \cdot Q_3 + (IKT - 3) \cdot Q_4] \right]^2 \quad (13)$$

For our estimations we employ Matlab integrated routine that uses medium-scale estimation method (i.e. it is computationally more intensive, because it stores full Hessian matrices that are approximated with finite difference method in memory to build up information on the objective function curvature) with Interior-point algorithm. We avoid Nelder-Mead simplex algorithm (fminsearch), because it does not use derivatives, but our optimization problem is nonlinear. In an unreported test for options with 28 days until maturity we see that usage of derivatives improves RMSE of option price values by nearly 30% and is much less sensitive to the starting values. For estimations we provide the following set of starting values ($\sigma=0.2$; $s=-0.5$; $e^{xcess}=0$). For our objective function we have chosen equal weight assignment to all options in that particular day, because despite possible errors in deep ITM options we want to capture all information from them. Mayhew (1995) proposed vega-weighted approach is likely to be biased, because it will give most weight to the ATM options, which goes in contradiction with our stated objective to obtain overall market outlook and previously outlined rationale to retain all options. Christoffersen & Jacobs (2001) do not promote any particular loss function (RMSE, %RMSE or IV MSE); however, they argue that usage of loss-function must be consistent in estimation and model evaluation. In our paper we consistently use RMSE as the loss function for option specific GC volatilities and daily GC volatility estimations.

Put values are obtained using put-call parity, which rests on assumption that there are no short-sale constraints and non-arbitrage considerations. However, within the scope of this paper, we believe that usage of explicit put formula would not likely to increase accuracy of the estimates.

4.1.2 Gram-Charlier adjusted volatility (GC volatility)

The implied volatility for each option is estimated using Gram-Charlier expansion. However, Gram – Charlier expansion is modified in order to accommodate the calculated daily skewness and kurtosis using our optimization from previous steps. On each particular day we match short/medium/long maturity options with our daily short/medium/long estimated skewness and kurtosis. We adjust each type of option with its corresponding skewness and kurtosis estimates in order to reflect underlying differences in option IV structure. In addition, the underlying higher-order moment estimates are likely to reflect market expectations about different future horizons.

$$\min_{\widehat{\sigma}_{GC}} [C_{obs} - (C_{BS}(\widehat{\sigma}_{GC}) + skew_{d,h} \cdot Q_3 + (kurt_{d,h} - 3) \cdot Q_4)]^2 \quad (14)$$

Hence, the only unknown in this equation is the volatility for each option $\widehat{\sigma}_{GC}$. The implied volatility, by using this kind of estimation technique, would be better estimate with respect to higher

order moments in the distribution, which the simple Black-Scholes model does not take into account, because it assumes log-normal distribution.

Our rationale to reverse-engineer option implied volatility using daily skewness and kurtosis is done in order to test volatility patterns after adjusting for higher order moments. In addition, this allows disentangling the effects to volatility smile from model-based explanation and market-based explanation. To our knowledge this reverse-engineering is not documented in the current literature.

4.1.3 Model Free Implied Measures

Thus in order to avoid undue influences from the pricing kernel we use model free implied methodology, which is not based on specific option pricing model with mild assumptions that the underlying stochastic process does not have price jumps. Barr (2009) argues that MF IV provides more robust test of market efficiency and informational content as opposed to BS case where IV jointly provides test of market informational content and corresponding market model (BS). An important distinction can be drawn about the method is that the structure of option prices can be recovered through the usage of implied volatility smile (curve). The basic insight while decomposing model-free parameters stems from the Bakshi & Madan (2000) finding that option (even exotic option claims) can be reproduced through unbundling of underlying options to make all option prices internally consistent. In order to construct model free implied measures we follow the approach of Bakshi, Kapadia, & Madan (2003) and provide brief summary of the key formulas used in our estimations.

First we set that $R(t, \tau) = \log(S_{t+1}) - \log(S_t)$ and $\mu(t, \tau) = E_t^Q[R(t, \tau)]$.

1. The risk-neutral return skewness is given by:

$$SKEW(t, \tau) = \frac{E_t^Q \left[(R(t, \tau) - \mu(t, \tau))^3 \right]}{\left(E_t^Q \left[(R(t, \tau) - \mu(t, \tau))^2 \right] \right)^{3/2}} = \frac{e^{rt} W(t, \tau) - 3\mu(t, \tau) \cdot e^{rt} \cdot v(t, \tau) + 2\mu(t, \tau)^2}{(e^{rt} \cdot v(t, \tau) - \mu(t, \tau)^2)^{3/2}} \quad (15)$$

2. The price of volatility contract

$$v(t, \tau) = \int_{S_t}^{\infty} \frac{2 \left(1 - \log \left[\frac{K}{S_t} \right] \right)}{K^2} C_t(\tau; K) dK + \int_0^{S_t} \frac{2 \left(1 - \log \left[\frac{S_t}{K} \right] \right)}{K^2} P_t(\tau; K) dK \quad (16)$$

In order to estimate volatility contract OTM calls and OTM puts are used, which are weighted by strike dependent amount. Notably, the more pronounced fourth moment will result as increased prices

for deep OTM options, therefore weighting scheme that assigns more weight to far out of the money options does have intuitive explanation. Bakshi et al. (2003) also argues that usage of OTM calls is warranted due to small early-exercise risk premium, this argument does not apply to Swedish option market due to option type (European). Similarly as for GC IV estimation, we account for future continuous dividends that are paid until option maturity. However, it is important to note that this in essence is variance contract and square root is taken in order to obtain volatility.

In this paper MF IV is estimated separately using prevailing market call and put options and their implied counterparts (derived from put call parity). This distinction between call MF IV and put MF IV is done in order to differentiate informational content from call and put options. Similarly Britten-Jones & Neuberger (2000) have used only call side of the option volatility smile therefore we consider that our slight modification of Bakshi et al. (2003) is applicable.

Our estimations are subject to truncation and discretization errors and as shown by Jiang & Tian (2005) they are likely to be negligible if moneyness range is at least ± 2 sample deviations of moneyness and if step size smaller than 0.3 of sample standard deviation of moneyness. For MF IV estimation procedure we have selected moneyness range $[0.3; 3]$ and 1001 discretization intervals thus making truncation and discretization error impact negligible. Selected boundaries are larger for our estimation thus mitigating the effect even further.

3. The price of cubic contract

$$W(t, \tau) = \int_{S_t}^{\infty} \frac{6 \log \left[\frac{K}{S_t} \right] - 3 \left(\log \left[\frac{K}{S_t} \right] \right)^2}{K^2} C_t(\tau; K) dK - \int_0^{S_t} \frac{6 \log \left[\frac{S_t}{K} \right] + 3 \left(\log \left[\frac{S_t}{K} \right] \right)^2}{K^2} P_t(\tau; K) dK \quad (17)$$

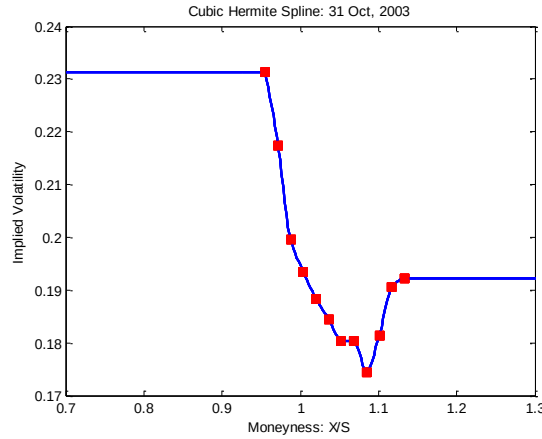
Where r – risk free interest rate, t – is the time today, τ – is the time at expiration, K – strike price, S – current stock price, C – call price and P – put price.

$$\mu(t, \tau) = e^{rt} - 1 - \frac{e^{rt}}{2} v(t, \tau) - \frac{e^{rt}}{6} W(t, \tau) - \frac{e^{rt}}{24} X(t, \tau) \quad (18)$$

In order to estimate MF parameters, implied volatility curve is used to exemplify the differences of options at different strike prices. Due to the fact that we cannot observe option prices at continuous spectrum, but rather discrete increments, we need to provide interpolation function. We use 1001 points for integral approximation and approximate the IV curve using Cubic Hermite Interpolating Polynomial. Its advantage stems from the fact that this interpolation has fewer oscillations and

overshoots if data is not smooth. For moneyness larger than any existing observation, the implied volatility takes on the corresponding implied volatility level. This preview of IV curve interpolation is critical, because the accuracy of our interpolation directly influences MF IV and MF skewness and kurtosis estimates.

Figure 3: Cubic interpolation for call BS IV



4.2 Regression analysis of the Gram-Charlier IV

Implied volatility smile explanatory analysis

In order to make our analysis easier and more intuitive we define moneyness for calls as Index/Strike and for puts as Strike/Index. In such manner our resulting calls and puts are both in the money if their respective moneyness is above 1.

We continue further with regression analysis in order to determine the factors besides the implied skewness and kurtosis which would explain the implied volatility inconsistency (Smile or Smirk) with respect to option moneyness. As the result, the following model is proposed:

$$\begin{aligned} Implvol_{it}^{adj} = & \beta_0 + \beta_{1:5,it}(Moneyness^1 * Call) + \beta_{6:10,it}(Moneyness^1 * Put) \\ & + \beta_{11,it}(Volume * Call) + \beta_{12,it}(Volume * Put) + \beta_{13,it}(Spread: Bid Ask * Call) \\ & + \beta_{14,it}(Spread: Bid Ask * Put) + \beta_{15,it}(Time to maturity) + \varepsilon_{it} \end{aligned} \quad (19)$$

¹: Moneyness: xOTM, OTM, ATM, ITM and xITM (separate coefficient reported for each category)

²: In regression we also control for monthly time fixed effects

We control for the moneyness as it has an impact on the volatility smile as noted by many academic researches Dumas et al. (1998), Verma, Ertosun, Wang, Ambruster, & Giesecke (2008) and

others. We introduce interaction variable for call and put options with moneyness. This allows us to control for different level of volatility across moneyness and option type. We introduce dummy variable, when there is corresponding turnover for particular option in the market, which is warranted by the fact that marginal effect from larger turnover (e.g. 1 SEK or even 10% in log-specification) was economically insignificant and unnecessary for the purposes of this paper. As a result, dummy variable has a simple and intuitive interpretation as the effect on GC IV if the option is traded. We have interaction variable for trade and option type to disentangle the different effect of turnover for different types of options. Nordén & Xu (2012) show that for Swedish market steepness option market smirk significantly depends on relative liquidity for different options with different moneyness levels. However, they use only BS model, which fails to capture the non-normality of returns therefore our approach is superior, because we are able to disentangle both higher-order moment driven information as well as more market based explanation.

The bid ask spread shows the market liquidity and illustrates the costs for the investor to buy or sell the option in the market. Similarly we introduce the interaction variable for spread and option type to account for different effects between different types of options. As argued previously higher spread can result in higher volatility due to the fact the investor might demand compensation (higher the volatility) because the costs associated with the liquidation of his/her position is higher. We also control for time to maturity because as shown by Goncalves & Guidolin (2004) and Manuel & Chang (1995), the IV is a function of time to maturity and is dynamically changing over time.

In this analysis it is not applicable to use well-documented curvature and regression measures, because ex-ante we do not know the resulting volatility pattern. As noted by Verma et al. (2008) first three factors capture large variation of the volatility of the BS IV (level, slope and curvature), we use only two factors (level and slope) yet curvature is captured by using interaction variables because estimation of curvature would extend the analysis to properly identify the right curvature parameter. In addition, classifying options as xITM, ITM, ATM, xOTM and OTM and using interaction variables will capture large part of the curvature effect with more intuitive and linear interpretation

While logarithms of IV are noted by Verma et al. (2008) to be more normally distributed than the volatilities themselves we continue to use implied volatilities, because a) we have detailed sub-sample selection b) in our opinion it allows more convenient estimated coefficient interpretation and c) Jarque-Bera test indicated that the logarithm of the implied volatilities is less normally distributed ($4.2e^5$ instead of $2.6e^5$). While usage of OLS under non-normality is sub-optimal we still use OLS estimators in

order to provide consistent comparison with other empirical studies such as Dumas et al. (1998) and Shie (2009). In addition, exogeneity assumption of OLS estimates $E(X\varepsilon) = 0$ is corroborated by a simple correlation analysis.

We report clustered standard errors and their corresponding t-statistic in order to account for correlation within groups (individual options). Given our sample size and number of options our clustered group errors are unlikely to suffer from small cluster problem (Angrist & Pischke, 2008). For additional robustness and theoretical support of time-varying volatility we implement monthly time fixed effects with option level clustered standard errors. In addition, for robustness and comparison purposes we also apply the same regression specification to BS implied volatility.

4.3 Predictability

4.3.1 Index Return Forecastability

We provide extension of Shefrin (1999) and Meyers (1994) claims that bullish investors overbid call options, while bearish investors overbid put options and thus turnover of these instruments can forecast future market movements.

There have been informational studies such as Pan & Poteshman (2006) as well as Kuo & Lin (2011), who use unique databases and find that non-public open-buy (buyer initiated transactions to open new positions) put-call volume ratio has predictive power for the underlying. Similar predictability results have been found for Finnish stock Exchange by Martikainen & Puttonen (1996), who use publicly available put-call volume turnover data.

We provide an important and to our knowledge undocumented extension of these findings by focusing explicitly for the ITM options. Other well-documented examples of market predictability (January effect, Monday effect) have gradually decreased over time thus our results are robust extension of the existing studies, because investors might have set-up strategies that utilizes all put/call ratios. This argument is strengthened by the fact that if an individual investor has superior information and is willing to make directional bet it costs more to use ITM option thus implicitly signalling about strength of the belief. The usage of larger capital can be interpreted as a signal by informed/confident investor. In a simple example¹ a put option that is 5% in-the-money costs 6.4 SEK while similar out-of-the-money options costs mere 1.4 SEK. As a result risk-averse investors are required to commit larger capital ex-ante and they are also exposed to negative price development in the underlying (because ITM

¹ 30 days until maturity; 30% annual volatility and 3% CCIR

option has priced in the intrinsic value of the underlying). This argument is further enhanced by De Long, Shleifer, Summers, & Waldmann (1990) who show that rational/informed investors might reduce their positions when they expect that prices will increase due to noise trader risk. As the result, the following model is tested:

$$r_{t+\tau}^{index} = \beta_0 + \beta_{1,t}(Ratio) + \varepsilon_t \quad (20)$$

where $Ratio = \frac{SEK \text{ volume of traded ITM puts}}{SEK \text{ volume of traded ITM calls}}$

While constructing market signal ratio we use cut-off point as 1 to determine whether instrument is in the money or not. While this is likely to introduce some selection bias from trading in ATM options it is driven by our data availability, where ITM options (as classified previously above 1.03) constitute only 14% and 21% of total number for puts and calls respectively. This also supports the notion that market depth might be insufficient for particular instrument and informed/confident traders might have to use other imperfect substitutes.

For our regression purposes we will use both overlapping and non-overlapping returns. For overlapping returns we use Newey-West heteroskedasticity and Autocorrelation consistent standard errors in order to mitigate the overlapping data problem. Han (2011) provided rule of thumb for number of lags Newey-West SE calculation $L = 0.75 * (N)^{1/3}$ we use default Matlab implemented approach $L = 4 * (N/100)^{2/9}$. In addition, the usage of Newey-West autocorrelation errors in the current form is somewhat supported by the significant lag-length (as shown by unreported BIC test) rather than usage of full number of overlapping observations. However, this approach deals with biased standard errors, so in order to mitigate the problem with potentially biased OLS coefficient estimates we also consider necessity to include lagged dependent return variable. Achen (2001) and Keele & Kelly (2006) discuss inclusion of lagged dependent variable versus risks of OLS estimate biases and conclude that inclusion of lagged dependent variable is warranted if the process is strongly dynamic (the returns have path-dependency i.e. tomorrow returns have strong relationship to today's returns). We do not include lagged return variable, which is in line with Bali & Peng (2006) who find that bias is not significant, when independent variable follows AR (1) process with 16 years of data. We consider that our 10 year sample likely will also not have such bias.

As a result, we expect that larger Ratio (indicating larger volume for put options) to be associated with more negative returns. We test this hypothesis using different time-horizons and we also indirectly provide support for out-of-sample performance using strictly non-overlapping data.

4.3.2 Realized volatility predictability using IV

We also examine the realized volatility predictability using IV. Due to the fact that we do not have high-frequency data available we compute realized volatility in the following way:

$$\sigma^{RE} = \sqrt{\frac{360}{n} \sum_{t=1}^n (r_t)^2} \quad (21)$$

Christensen & Prabhala (1998) note that usage of overlapping time series data is likely to induce biased estimation results thus we provide both overlapping sample analysis as well as non-overlapping sample analysis. Consistent with previous structure we provide results for realized volatility prediction in short, medium and long time horizons. For estimation purposes we use both MF IV and GC IV. Another facet to consider is variance ratio of 1 which is implicit in such calculation. So for example variance ratio that is larger than one indicates that there is positive serial correlation and ratio smaller than one indicates there is mean reversion during longer interval. However, if market participants are rational they would also correspondingly scale their annual volatility to reflect their opinion on short term horizons. In our empirical data we see that for overlapping sample annual variance ratio of 1 is rejected for short (15 DTM) and medium (41 DTM) subsample horizons, while it cannot be rejected for longer (235 DTM) subsample.

As already discussed previously we test two different empirical specifications to infer conclusions about informativeness, unbiasedness and informational efficiency. We start by estimating simple regression about informativeness and unbiasedness. In addition, this regression specification will provide insight on variance premium.

$$\sigma_{t,t+\tau}^{RE} = \beta_0 + \beta_1 \sigma_{t,t+\tau}^{IV} + \varepsilon_t \quad (22)$$

where: $\sigma_{t-\tau,t}^{RE}$ – historical realized volatility and $\sigma_{t,t+\tau}^{IV}$ – option IV

Second regression specification is to simultaneously test for informativeness, unbiasedness and informational efficiency using historical in-sample volatility forecast. We use the same sample period length for estimating this alternative forecast to avoid unduly impact from time-horizon selection problems. Thus our second regression specification is:

$$\sigma_{t,t+\tau}^{RE} = \beta_0 + \beta_1 \sigma_{t,t+\tau}^{IV} + \beta_2 \sigma_{t-\tau,t}^{RE} + \varepsilon_t \quad (23)$$

where: $\sigma_{t-\tau,t}^{RE}$ – historical realized volatility; $\sigma_{t,t+\tau}^{IV}$ – option IV and $\sigma_{t,t+\tau}^{RE}$ – ex-post realized volatility from time period t (which is formed in the previous τ days). Detailed hypothesis descriptions associated with regression specifications (eq.22 and eq.23) are provided in Section 2.3.

Maheu & McCurdy (2002) provide evidence for non-linear realized volatility relationship (between volatility forecast using ARMAX framework and realized volatility in FX markets) and we are inspired by this notion and provide to our knowledge undocumented extension to option market implied forecasts (IV). We test these hypotheses using different time-horizons and we also indirectly provide support for out-of-sample performance using strictly non-overlapping data. We use quantile regression developed by Koenker & Hallock (2001), in order to estimate non-linearity for the first specification. As a result, we would obtain evidence on how unbiased is the forecast in different volatility-regimes as well as how the implied volatility premium is distributed across different realized volatility levels:

$$\min_{\beta} [p_t(y_i - \beta x_i)] \quad (24)$$

where: p_t is the absolute value function that yields the sample quantile as its solution.

5 Empirical results

5.1 *Robustness of GC measures*

Jondeau & Rockinger (1999) provided discussion of positive boundary for GC volatility measures, therefore we proceeded with these values. RMSE of our incremental estimations are reported in Appendix 6, The results show that adding higher order moments allows to better explain daily option observations by about 30% (when compared to single volatility estimation) with the largest part coming from inclusion of kurtosis. We also report the RMSE development through time, which indicates that in periods of larger market uncertainty the estimation error increases. We also performed unreported optimization procedure (with following boundaries: $skew \in (-1.05; 1.05)$ and $\in (3; 7)$), which was able to yield significantly lower RMSE. These results are somewhat expected, because positive density requirement per se do not exclude better fit of other even implausible PDFs. However, we believe that positivity constraint should be considered, in order to obtain empirically sound results of higher order moments.

In our estimations we require that there are at least 4 options. While estimating MF IV there were less estimated daily parameters if the moneyness level is the same (i.e. maturity differs). For MF IV

average of maturity and IV was taken and then it was used in the resulting MF moment calculation. However, we observe that obtained volatilities using both methodologies are surprisingly similar, and they are reported in Appendix 5. Even though the possible error was unlikely to be significant in relative terms, we were surprised to see such robustness (and persistence) of both measures. We also observe that when comparing higher order moments the situation is slightly different, because the similarity of measures becomes less persistent. Due to our unreported optimizations we can contrast that these results are direct result of the positive boundary (envelope type) requirement, which warrants further research into this topic.

5.2 Regression analysis of the IV smile

In order to provide analysis of the IV smile we separate the analysis in two distinct parts: 1) analysis of the IV shape structural form and 2) analysis of the model-based and market-based explanation. For detailed results please refer to Appendix 4.

Implied volatility smile steepness

This section shows the results of the implied volatility for each option and sorts them based on the days until the maturity-short term, medium term and long term. We provide simple average analysis (to recover term structure implications and provide intuitive overview of results in Appendix 3: GC and BS volatility plot comparison and Appendix 2: Summary statistics of call and put Implied Volatilities) as well as provide regression analysis (to recover possible shape effects) with time fixed effects.

The implied volatility for Call options obtained from GC expansion on average (with several exceptions) is higher across all moneyness levels in comparison to the implied volatility obtained from BS. The average difference ranges from 1% to 4%, except for short term deep ITM call options where the BS implied volatility is by 13% higher than the GC implied volatility, which is explained by the fact that GC IV implied curve is relatively flat. The results for put options is slightly different, we can observe that the GC implied volatility for OTM options is substantially lower in comparison to the BS implied volatility but for ITM options GC volatility is higher. In appendix 3 we provide simple graphical analysis of the average implied volatility levels. However, the graphical results are puzzling, because it seems that the average volatility becomes much flatter for ITM calls and OTM puts. Such results are also puzzling in the context of MF IV construction and in order to quantify them in a more robust way we proceed with regression analysis outlined in section 4.2.

We observe that the coefficients which evaluate the steepness of GC implied volatility for call options are substantially less significant and smaller in magnitude in comparison to the coefficients which evaluate the BS implied volatility. This indicates that the inclusion of higher order moments reduces the smile pattern. The same findings also apply to put options. Consequently, from regression analysis perspective, it is possible to infer that the smile is less pronounced when the implied volatility is adjusted by higher order moments (skewness and kurtosis) and regression specification includes several control variables (which is consistent with documented examples in literature). In addition, we perform F-test results with hypothesis that in the regression specification all coefficients on interacted moneyness variables are zero. These results show that there is significant improvement in coefficients; however, there is still some dependency on the moneyness for the IV level.

Table 1: F-Test values for joint test

	All	Short	Medium	Long
BS	189.0	219.9	212.2	138.7
GC	11.7	34.8	19.4	22.3

5% critical value for F-test with correct df is 1.83

Implied volatility smile explanatory analysis

When examining the merit for market based explanation of the smile, it should be noted that the bid-ask spreads tend to increase with moneyness, which is explained by the fact that such options have larger price quotations which allows charging higher bid-ask spreads in absolute terms. Consistent with previous expectations, options closer to ATM have larger volume.

Using fully-specified regression models with monthly fixed effects and clustered standard errors an interesting pattern emerges. Firstly, call options when they are traded on average are less expensive in comparison to put options and the finding is robust across BS IV and GC IV measures. The magnitude of the volume effect is statistically significant, yet the economic significance is questionable because coefficients for Calls and Puts are 0.22% and 0.58%. In unreported regression the volume effect using log value of trade volume yielded even smaller results showing robustness of previous findings. Nevertheless, it indicates that the turnover has different implications for different types of options and thus should be analysed in more detail. First implication is that options exhibit specific demand driven phenomenon, because, if there is trade, the options *ceteris paribus* are more expensive. The negative coefficient for short-horizon subsample is explained by the fact that in the short-term there might be larger sell-pressure, because option holders want to close their positions and not deal with settlement procedures. Secondly, the difference between puts and calls is explained by the well documented

evidence by Atanasov & Gyoshev (2001) and Bollen & Whaley (2004) where institutional end-investors prefer to have long positions in puts to hedge their position in the underlying assets. This lowers the price elasticity for put option buyers and thus they are willing to pay higher premiums to acquire these options, which results in larger coefficient for ex-post impact of trade. Our results do not contradict Brenner, Eldor, & Hauser (2001), because we identify existence of demand-pull rather than examination of the counterfactual options (not-traded vs. traded).

Another interesting finding is the impact of bid-ask spread for call and put option on implied volatility. We can observe that for both types of options the Bid-Ask spread has positive effect on the IV. Interacted bid-ask spread variable is expressed in log terms, thus allowing a slightly different interpretation. Usage of log transformation is done in order to mitigate large non-normality in the data. While the transformed data still rejects test for normal distribution transformed data is much closer to normality. As a result, doubling (i.e. 100% increase) of the bid-ask spread will have *ceteris paribus* increase for IV of 0.6% and 1.3% for calls and puts respectively. These results vary within sub samples; however, the underlying economic magnitude stays similar and it possible to see that the IV for call options is less affected by the increase in the Bid-Ask spread in comparison to put options, which is especially pronounced for the long-term options. The increase in Bid-Ask spread increase is associated with an increase in demand or decline in supply which consequently realizes in higher implied volatility/price (Longstaff, 1995). To our mind this has been caused by two market based factors. First, it shows that there are different buying patterns for put and call options. We argue that the demand for call options is mainly driven by informed traders which are not willing to pay high price to acquire overvalued options which results in lower demand and liquidity. In addition, they would favour derivatives instead of plain long positions in the underlying due to the fact it offers the leverage and demands less capital. As the result, underwriter has less bargaining power which results in lower implied volatility/price when spread increases. The demand for put options is mainly driven by investors who want to hedge their exposure. This implies that the put option price/volatility is more inelastic due to the fact the investors are willing to acquire also overvalued put options. As the result, the underwriter has higher bargaining power with respect to investors who want to acquire put options. This implies that in case of increase in demand for put options the spread would widen because the underwriter would not supply enough put options to cover the demand which in the end would result in higher implied volatility and price of the put options. The reasoning for lower liquidity for call options is also supported in the data: the average bid-ask spread for call options is 3.15 SEK, while for put options it is 2.63 SEK. This finding is in line with standard liquidity argumentation that liquidity risk must be compensated.

Secondly, it shows important insight for market makers, who require larger premium if the option is in the money and is expected to end in the money as well.

Finally, it should be noted that time until maturity has negative effect on implied volatility, which shows that time risk is decreasing as options move closer to expiration. While the effect is statistically significant it is weakly economically significant, because *ttm* variable is expressed as fraction of a year and hence in GC IV setting 30 day increase for maturity *ceteris paribus* will increase the volatility by only 0.1%. Nevertheless, consistent with previous empirical findings and IV regression methodology it is important to provide control for maturity effects. While this result preliminary contradicts the existing findings omission of it would yield to omitted variable bias and secondly we believe that the documented small effect is likely to be due to model specification, because model-based explanation coefficients (i.e. introduction of higher order moments) subsume more information and leave time until maturity effect as economically insignificant. On the same token the large effect for short-horizon effect is driven by the fact that 1 day increase *ceteris paribus* will result in IV increase of 0.3%. This confirms the notion that in the short-term markets are unpredictable and this effect might be manifestation of possible jump (large movement) risk. In Table 3: Average volatility premium, we show the average volatility premium for different options where we could observe that the implications from market-based factors would be larger in the context of volatility premium. As a result, it is possible to conclude that the market based factors (liquidity risks) are priced in by the market maker.

While these findings showed important insight into derivative markets, they still in our opinion are relatively small and the largest extent of volatility is driven by the underlying market dynamics, which give rise to model-based explanation. However, distinction between model-based and market-based explanation of the IV surface is somewhat limited, because we do not have access to the individual trade-level and investor-level data. Model based explanation in our opinion is more robust, because in stock markets there is limited empirical support for strong economic effect from bid-ask spreads and turnover (liquidity measures) to the stock return dynamics (returns and volatility). In addition, direction of causality between trading volume and returns as found by Sum (2012) is simultaneous, while bid-ask spread is influenced by other factors such as adverse selection (information-based trading), inventory and order processing. Nevertheless, we document that bid-ask spread is larger for ITM options and we argue that it is direct result of the economic incentive from market makers to make bid-ask spread comparable in relative terms and thus lending support to causality running from moneyness to bid-ask spread. As a result previously outlined explanations should be interpreted with

care and it should be noted that the causality is also subject to seasonal variations as outlined by Naes, Skjeltorp, & Ødegaard (2011).

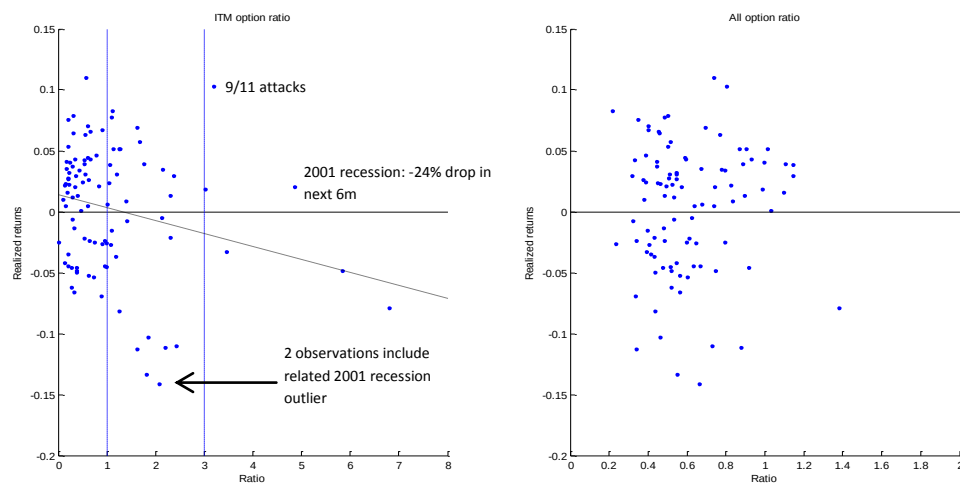
5.3 Index return predictability

Due to the market force argumentation provided earlier we examine whether ITM puts and ITM calls provide predictability of index movements. Our results show that the explanatory power of these regressions is quite high with R-square ranging around 0.23. In order to provide better robustness of the results, we provide three different regression specifications: 1) pooled overlapping data 2) pooled overlapping data (independent variable) and daily index returns (dependent variable) with no overlapping 3) exclusively non-overlapping data for volume and returns. For detailed results, please refer to Appendix 8.

All estimated results show that there is robust statistical significance of the predictability for the index returns. Using non-overlapping data for volume in specification 3, it is possible to observe that the results are robust and hold for different indicator selections; however, the predictability power decreases, when predictable returns have larger horizon than SEK volume indicator (Specifications 5 and 6). As a result, the average daily coefficient (excluding specifications 5 and 6) is -0.00049. As outlined by Huang (2012) there are theoretical bounds on predictability coefficients and that in literature the predictability coefficient is generally below 1%. Our results with larger time-frames show that there is even larger predictability (smallest statistically significant R-square is 1.1% while the largest is 4%). Robertsson (2000) shows that for Swedish stock market maximum predictable tactical portfolio allocation is as high as 25%; yet, these results are based on different industry sectors rather than pure index predictability. Hence, while these results might seem very small, in perspective such predictability is considerable. Nevertheless, as examined further using graphical analysis (scatter plot), these results partially are likely to be driven by three different market conditions. Scatter plot consists of the predictor variable and realized returns for the non-overlapping data with the highest R^2 . We compare these results with other previously discussed market sentiment indicator-simple put/call ratio. As a result, three patterns emerge, which are 'normal' market conditions where returns are likely to be small and ITM put-call ratio is below 1, 'elevated market' where there is larger uncertainty and larger probability of negative returns and ITM put-call ratio is between 1 and 3 and 'negative return' where negative returns are nearly certain and ITM put-call ratio is above 3. In the following chart there are also two outliers, one of them is driven by 2001 September terrorist attacks, where there was large demand for puts but ex-post there was no another attack and another is driven by subsequent 24% drop within

6month period. When the ITM put/call ratio is compared with total put/call ratio, we can observe that ITM put/call ratio has larger range (0.1-7 and 0.2-1.4 respectively) and it has virtually no predictability. The conditional averages for these regimes are 0.9%, -0.8% and -5.3%; however, first two regimes means are not statistically different from zero, yet the difference in our opinion is economically significant. Our regression results include these outliers, thus leading to underestimation of the true predictability power.

Figure 4: Scatter plot of put and call volume ratio (ITM and All) and realized returns



As a result both extensive regression results in combination with provided graphical analysis and argumentation why ITM options are likely to be better proxy provide in our opinion sound extension of the market sentiment index.

Another facet to consider while using investor sentiment as predictor variable is causality. Despite the fact that Wang, Keswani, & Taylor (2006) using Granger causality tests argue that returns and volatility per se cause sentiment rather than vice-versa, we believe that our extension of informed and committed (capital-wise) investors circumvents this problem by narrowing down scope of signal to ITM options only. In our dataset out of all options 53% have volume data (i.e. were traded) and 26% of traded options have been in the money and 36% of total volume traded has been within ITM options. This shows that investors do trade more in ITM options.

In order to provide additional insight of our findings, we briefly discuss several unreported findings of different hypothesis. Firstly, we test whether usage of OTM puts and OTM calls provides better predictability, because it would imply the opinion/forecast of very optimistic investors (who expect large movement in the underlying). However, the results were mixed (different significance

across different specifications) indicating potential risk of data mining. These less significant results for very optimistic investor indicate that they are less informed or even arguably rational, leading to conclusion that they are also noise traders. Additional explanation stems from price perspective, because usage of low-price options (OTM) are less likely to pickup investor sentiment, because option prices are much lower and corresponding ratio signal is likely to be contaminated more with noise traders. On the contrary usage of high-priced option SEK volume allows to pick-up effect from increase in options traded (i.e. they more easily transform into change in predictor variable). Secondly, we perform also total put/call ratio analysis and we find no statistically significant results. Lack of significant results can be partially explained by the data in-balance in our sample, because the average price with moneyness above 1 for calls and puts is 69 SEK and 60 SEK, while the average option price between all calls and puts is 37 and 27. The resulting difference is nearly double in relative terms thus it is likely to bias the results. In addition, the rejection rationale of all put-call volume ratio was already outlined previously in Section 4.3.1. Thirdly, we test claim whether the ratio of daily IV from puts and calls helps to predict future returns and we find no statistically significant results. Such hypothesis is supported by the fact that in a demand-driven environment puts and calls are likely to be more expensive if there is large demand for them. However, as discussed previously in Section 5.2 there is no economical significance from increasing traded SEK volume, which translated into no predictability power for the underlying of the IV ratio.

5.4 Realized volatility predictability

In order to provide realized volatility predictability analysis we provide analysis in two parts: 1) focusing on well-documented hypothesis (unbiasedness, informativeness and informational efficiency) in literature; 2) provide extension of predictability by incorporating non-linear effects and its implications for volatility premium.

Realized volatility predictability

The results for realized volatility predictability are mixed; yet, they fail to reject the informativeness hypothesis ($H_0: \beta_{IV} \neq 0$). We do not report individual t-statistic for this hypothesis, but they were statistically significant in all cases. In addition, using simple significance tests (t-statistics) we do not reject the hypothesis that ($H_0: \beta_{IV} = 1$), which suggests some degree of unbiasedness of IV forecast (for details refer to Appendix 9). However, it is important to distinguish that hypothesis ($H_0: \beta_{IV} = 1$) is strictly speaking not unbiasedness test and thus we need to provide joint hypothesis test (discussed further in this section). The results across time-horizons are mixed, because for short-term options

(overlapping and non-overlapping data samples) there is support for hypothesis ($\beta_{IV} = 1$) for GC IV but there is no support for MF IV. Despite evidence by Jiang & Tian (2005) that MF IV has superior informational content (for 60, 120 and 180 days) our estimates show slightly different picture specifically in the short-run. For medium and long term horizons we report only the overlapping observations, because otherwise the resulting observations will be too small for regression analysis. For these horizons we also fail to reject the hypothesis that coefficients on implied volatility are equal to 1 for GC IV forecast, but reject for MF IV. We argue that this phenomenon is due to the fact that in longer term the GC forecast has strong model assumptions in place. However, due to conventional practice of quoting option price in volatility terms this result simultaneously implies that market makers require larger premium for the long-term options thus rendering forecasts biased. In addition, there is no strong emerging pattern between put and call IV results showing that there is no evidence for specific preference for put or call options by informed, rational investors. Our results contradict results provided by Bandi & Perron (2002) and we argue that our results support stronger economic intuition for existence of volatility risk premium as opposed to long-term market predictability.

As a result, we also report joint significance test explicitly testing unbiasedness hypothesis $H_0: \beta_1^{IV} = 1$ and $\beta_0 = 0$. Results in Table 2 show that only for short-term call GC IV we fail to reject H_0 ; however, for the medium and longer term horizon options we reject the unbiasedness hypothesis.

Table 2: Unbiasedness test

	Short-ov		Medium-ov		Long-ov		Short-non	
	GC	MFIV	GC	MFIV	GC	MFIV	GC	MFIV
Calls	1.78	7.75	8.35	21.94	96.61	37.52	2.18	36.54
Puts	3.81	10.82	5.34	15.53	105.73	58.52	4.57	5.94

Note: F-test critical value for ov (overlapping) is 3.0 and for non (non-overlapping) is 3.1

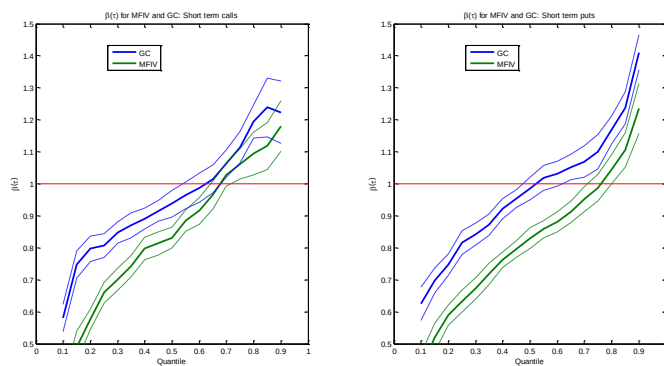
The results, while incorporating alternative forecast (historical volatility), leads to similar conclusions (for details refer to Appendix 10). The results of alternative forecast (historic volatility) in GC IV regression specifications show that estimated coefficients are small and statistically insignificant, while the estimated coefficients for MF IV reject the hypothesis ($H_0: \beta_{hist} = 0$), thus showing that historical forecast has some degree of informative content when compared with MF IV. The effect for longer time horizons for MF IV shows that both historical forecast and MF IV forecast have increasingly smaller informative power (as judged by point estimates, reported t-statistics and adjusted r-square). In addition, there is decreasing informative power of GC IV forecasts for longer time horizons (as judged by coefficient magnitude). While different analysis of subsamples could lead to different conclusions, in our

opinion they partially support the previous findings that with longer time horizons there is smaller informative power for MF and GC IV forecasts. It is important to note that point estimate for IV in GC or MF setting within non-overlapping data is larger for put options than for call options indicating that put options are used by more informed, rational investors and are more important while gauging the future realized volatility. However, the implications for put implied volatility risk premium are more puzzling, because such finding would simultaneously imply that risk premium is smaller for put options. As the joint unbiasedness test is insignificant and judging by the coefficient magnitude there is no empirical support for informational efficiency of the IV forecast.

Non-linear effects and volatility premium

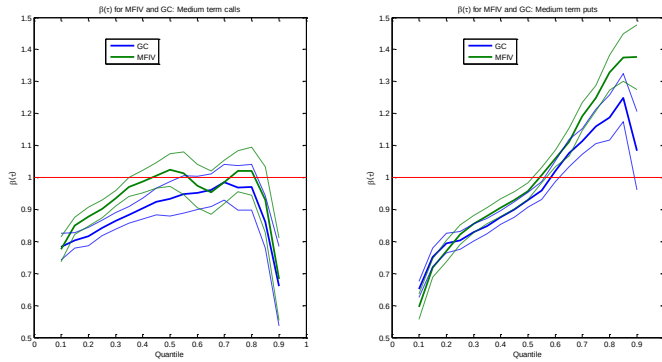
We further extend our analysis by considering non-linear effects using quantile regression and its implication to volatility premium.

Figure 5: Non-linearity of short term horizon (call and put options)



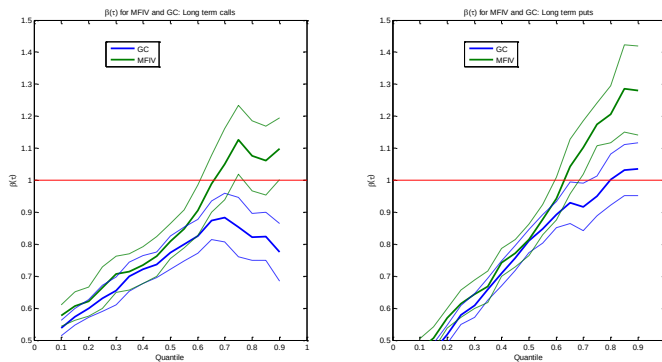
Our results for non-linear effects are somewhat more striking. The estimated coefficient for GC and MF IV is increasing with quantile and it exceeds 1 (being unbiased estimate) for large realized volatilities. This shows that in case of distress the IV is less reliable for volatility prediction. For the short-term options there is larger gap in estimated coefficients of GC and MF; however, there is no evident discrepancy between both puts and call options.

Figure 6: Non-linearity of medium term horizon (call and put options)



For medium-term option horizon the estimated horizon is slowly increasing for call options, and the large drop for medium term options in top decile are associated with the fact that during 2001 and 2002 September, while there was widespread fear of the possible terrorist attack (and its repetition in 2002) there was also belief by rational investors that this slump will be temporary. Hence, to exploit this market situation, rational investors had to acquire medium term call options (short term call options would inherit the risk of slow market reversal) which increased the demand and consequently the price/volatility. The results for put-options indicate that the risk of associated slump was already priced and thus there was no associated drop in coefficient for short-term put options. In addition, exercising profitable strategy from call options is easier, because investor could hold option until maturity. For put option profitable strategy would require selling option when the market has reached the lowest point, which could be difficult if the traders in the option market believe it is the lowest point and there is thin trading volume by other investors or market makers.

Figure 7: Non-linearity of Long term horizon (call and put options)



The long-term options also seem to confirm the previously outlined trend of increasing coefficient estimates; however, with them there is also larger uncertainty at the high volatility, which is driven by the fact that there is larger uncertainty about the unbiasedness of the forecast. Thus IV is less informative forecast for longer time horizons.

Our results are not driven by time variation as documented by Levy (2010) , who document strong monthly negative premium for month of July. In unreported analysis of short term GC calls the results had statistically significant support for monthly effects, yet the size of these coefficients were below 0.05, which reject the notion of strong economically significant effect.

Our findings slightly contradict the existing documented stable premium for underwriting put options and we argue that our results more accurately reflect the time until maturity effect by using directly options available in the market. In comparison Maidel & Sahlin (2010) find much lower stable volatility premium through time that would imply more stable forecast of the IV. Table 3 shows simple volatility premium calculated as difference between IV and realized volatility. Further analysis of volatility premium shows that close to expiry options have negative volatility premium as the time risk has decreased substantially. The results support the notion that volatility premium is larger for larger time horizons. Previously outlined quantile findings in our opinion are important extension of the previously documented debate about efficiency of the IV forecast. In addition, these results use directly matched volatility to each period as opposed to other researchers who estimate synthetic implied 30day forecast from available options in the market.

Table 3: Average volatility premium

	Short		Medium		Long	
	GC	MFIV	GC	MFIV	GC	MFIV
Calls	0.41%	0.99%	1.55%	3.41%	2.60%	4.64%
Puts	-0.05%	0.53%	0.99%	2.24%	1.84%	3.41%

Note: Short term option exclude options below 10 trading days

6 Conclusion

In this study we have examined the informational content of OMXS30 index European style Call and Put Options which are traded on the OMX Nordic stock exchange by applying extensions of the BS models which incorporates higher order moments.

Firstly, we have found that by incorporating higher order moments in the option pricing framework using GC estimation procedure it is possible to obtain skewness-kurtosis adjusted IV which has smaller structural pattern. While the results showed no statistical significance for claim that there is no structural dependency on moneyness, the resulting tests showed even 10 fold improvement in F-statistics for model structural dependency in comparison to BS IV. In addition, we have examined market-based explanation and found that the liquidity based explanations are important for IV; yet, the economical effect is less pronounced. Nevertheless, if we consider volatility risk premium, the magnitude of this effect is economically considerable.

Secondly, we showed that the usage of ITM put/call ratio has significant return predictability power, and these findings are robust across different model specifications. Furthermore, these results is more reliable (robust across different model specification) as well as statistically and economically more significant in comparison to OTM put/call ratio and put/call ratio.

Thirdly, we have showed that GC IV is informative, and for shorter horizons unbiased forecast for the future realized volatility. As there is no strong support for unbiasedness the informationally efficient hypothesis is therefore rejected. In addition, we have extended the existing forecasting framework by incorporating non-linear effects, which has showed that GC IV is biased in large volatility market regimes. This has large implications for the market makers or investors seeking profit and diversification possibilities by applying this alternative strategy. Contrary to widely held belief among practitioners, our results suggest that the volatility premium is subject to time-variation and might entail significant losses in market turmoil.

To conclude, we can state that the above findings support the argument that higher order moments of the distribution affect option pricing along with market based factors. Moreover, one could use the information from the derivative market, where on average more informed and sophisticated traders are concentrated, to obtain more efficient forecasts of the underlying asset returns and volatility.

7 Limitations and suggestions for further research

Despite the fact that there is narrow estimation range it is possible that there are several resulting GC volatility, skewness and kurtosis solutions with comparable RMSE. However, we have addressed this

issue by using minimization procedure in Matlab with partial derivatives and thus we believe that the amount of error in the estimation procedure is limited.

Corrado (2007) shows that it is possible to extend Gram-Charlier density expansion and incorporate Martingale Restriction in option pricing model and thus it would provide additional extension of the option pricing model.

One important extension not discussed in our paper is predictability of higher-order moments using GC and MF measures. Nevertheless, this area of research is subject to sample selection bias and difficulties in realized moment measurements techniques. Despite the fact that Diavatopoulos, Doran, Fodor, & Peterson (2008) provide predictability of higher order moments using MF IV measures we still believe that there is possibility to extend these measures into GC area.

Another possible extension of the research could be the application of the high frequency data. This would allow obtaining more accurate estimates of the realized volatility of underlying which would result in more accurate estimates of the volatility premium and predictability.

In addition, it is possible to provide joint MF forecast for realized volatility, which would incorporate information from both derivatives: calls and puts. However, our subsample selection shows that there is little difference in predictive power between puts and calls, if anything the former being more efficient predictor. Similarly in line with Vilkov & Xiao (2012) it is possible to provide extreme-value theory modelling of put IV and complement our ITM put/call ratio analysis.

8 References

- Abken, P. A., & Madan, D. B. (1996). Estimation of Risk-Neutral and Statistical Densities by Hermite Polynomial Approximation: With an Application to Eurodollar Futures Options. *Working Paper 96-5, Federal Reserve Bank of Atlanta*.
- Achen, C. H. (2001). Why Lagged Dependant Variable Can Surpress the Explanatory Power of Other Independant Variables. *Unpublished Manuscript, University of Michigan*.
- Ahoniemi, K., & Lanne, M. (2007). Joint Modeling of Call and Put Implied Volatility. *MPRA Paper No. 6318, Helsinki School of Economics*.
- Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion* (1st ed.). (Princeton University Press, Princeton).
- Atanasov, V., & Gyoshev, S. (2001). Why Large Financial Institutions Buy Long-Term Put Options from Companies. *Working Paper, Penn State University*.
- Avramov, D., Chordia, T., & Goyal, A. (2006). The Impact of Trades on Daily Volatility. *Review of Financial Studies*, 19(4), 1241–1277.
- Backus, D., Foresi, S., & Wu, L. (2004). Accounting for Biases in Black-Scholes. *Unpublised Manuscript, Stern School of Business*.
- Badshah, I. U. (2010). Quantile Regression Analysis of the Asymmetric Return-Volatility Relation. *Working Paper, Hanken School of Economics*.
- Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical Performance of Alternative Pricing Models Option. *The Journal of Finance*, 52(5), 2003–2049.
- Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies*, 16(1), 101–143.
- Bakshi, G., & Madan, D. (2000). Spanning and Derivative-Security Valuation. *Journal of Financial Economics*, 55(2), 205–238.
- Bali, T. G., & Murray, S. (2012). Does Risk-Neutral Skewness Predict the Cross-Section of Equity Option Portfolio Returns? *Working Paper, Georgetown University*.
- Bali, T. G., & Peng, L. (2006). IS THERE A RISK – RETURN TRADE-OFF ? EVIDENCE FROM HIGH-FREQUENCY DATA. *Journal of Applied Econometrics*, 21(8), 1169–1198.
- Balyeat, B. R., & Erturk, B. (2005). Option Prices as Predictors of Aggregate Stock Returns. *Working Paper, Xavier University*.
- Bandi, F. M., & Perron, B. (2002). A Fundamentally Different Interpretation of the Relation Between Implied and Realized Volatility. *Working Paper, University of Chicago*.

- Banerjee, P. S., Doran, J. S., & Peterson, D. R. (2007). Implied volatility and future portfolio returns. *Journal of Banking & Finance*, 31(10), 3183–3199.
- Barr, K. J. (2009). *The implied volatility bias and option smile: is there a simple explanation?* Thesis, Iowa State University.
- Barton, D. E., & Dennis, K. E. (1952). The Conditions Under Which Gram-Charlier and Edgeworth Curves are Positive Definite and Unimodal. *Biometrika*, 39(3/4), 425–427.
- Bartram, S. M., Fehle, F., & Shrider, D. G. (2008). Does adverse selection affect bid–ask spreads for options? *The Journal of Futures Markets*, 28(5), 417–437.
- Bates, D. S. (1995). Testing Option Pricing Models. *NBER Working Paper No. 5129*.
- Bates, D. S. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94(1-2), 181–238.
- Battalio, R., & Schultz, P. (2006). Options and the Bubble. *The Journal of Finance*, 61(5), 2071–2102.
- Bekaert, G., & Wu, G. (2000). Asymmetric Volatility and Risk in Equity Markets. *Review of Financial Studies*, 13(1), 1–42.
- Beneder, R., & Elkenbracht-huizing, M. (2003). Foreign Exchange options and the Volatility Smile. *Medium Econometrische Toepassingen*, 11(2), 30–36.
- Bernhardsen, T., Kloster, A., & Syrstad, O. (2012). Risk premiums in NIBOR and Other Countries' Interbank Lending Rates. *Staff Memo No. 21, Norges bank*.
- Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3), 637–654.
- Boes, M., Drost, F. C., & Werker, B. J. M. (2007). The Impact of Overnight Periods on Option Pricing The Impact of Overnight Periods on Option Pricing. *The Journal of Financial and Quantitative analysis*, 42(2), 517–533.
- Bollen, N. P. B., & Whaley, R. E. (2004). Does Net Buying Pressure Affect the Shape of Implied Volatility Functions? *The Journal of Finance*, 59(2), 711–754.
- Bondarenko, O. (2003). Why are Put Options So Expensive? *Working Paper, University of Illinois at Chicago*.
- Bookstaber, R. M. (1981). Observed Option Mispricing and Nonsimultaneity of Stock and Option Quatations. *The Journal of Business*, 54(1), 141–155.
- Branger, N., & Schlag, C. (2005). Put Options Are Not Too Expensive: An Analysis of Path Peso Problems. *Working Paper, Goethe University Frankfurt*.

- Brenner, M., Eldor, R., & Hauser, S. (2001). The Price of Options Illiquidity. *The Journal of Finance*, 56(2), 789–805.
- Britten-Jones, M., & Neuberger, A. (2000). Option Prices , Implied Price Processes , and Stochastic Volatility. *The Journal of Finance*, 55(2), 839–866.
- Broadie, M., Cvitanic, J., & Soner, H. M. (1998). Optimal Replication of Contingent Claims under Portfolio Constraints. *The Review of Financial Studies*, 11(1), 59–79.
- Broadie, M., & Detemple, J. B. (2004). Option Pricing: Valuation Models and Applications. *Management Science*, 50(9), 1145–1177.
- Canina, L., & Figlewski, S. (1993). The Informational Implied Volatility Content of Implied Volatility. *The Review of Financial Studies*, 6(3), 659–681.
- Cao, C., Chen, Z., & Griffin, J. M. (2005). Informational Content of Option Volume Prior to Takeovers. *The Journal of Business*, 78(3), 1073–1109.
- Castagna, A. (2012). Yes , FVA is a Cost for Derivatives Desks. *Working Paper, Iason Ltd.*
- Castanias, R., Chung, K.-Y., & Johnson, H. (1988). Dividend Spreads. *The Journal of Business*, 61(3), 299–319.
- Chakravarty, S., Gulen, H., & Mayhew, S. (2004). Informed Trading in Stock and Option Markets. *The Journal of Finance*, 59(3), 1235–1257.
- Chernov, M. (2001). Implied Volatilities as Forecasts of Future Volatility, Time-Varying Risk Premia, and Returns Variability. *Working Paper, Columbia Business School.*
- Chernozhukov, V., Fernandez-val, I., & Gallichon, A. (2010). Rearranging Edgeworth-Cornish-fisher expansions. *Economic Theory*, 42(2), 419–435.
- Christensen, B. J., & Prabhala, N. R. (1998). The Relation Between Implied and Realized Volatility. *Journal of Financial Economics*, 50, 125–150.
- Christoffersen, P., & Jacobs, K. (2001). The Importance of the Loss Function in Option Pricing. *Unpublished Manuscript, McGill University and Cirano.*
- Christoffersen, P., & Jacobs, K. (2004). Which GARCH Model for Option Valuation? *Management Science*, 50(9), 1204–1221.
- Comerton-Forde, C., Hendershott, T., Jones, C. M., Moulton, P. C., & Seasholes, M. S. (2010). Time Variation in Liquidity : The Role of Market-Maker Inventories and Revenues. *The Journal of Finance*, 65(1), 295–331.
- Corrado, C. (2007). The Hidden Martingale Restriction in Gram-Charlier Option Prices. *Journal of Futures Markets*, 27(6), 517–534.

- Corrado, C. J., & Su, T. (1997). Implied Volatility Skews and Stock Return Skewness and Kurtosis Implied by Stock Option Prices. *The European Journal of Finance*, 3(1), 73–85.
- Coval, J. D., & Shumway, T. (2001). Expected Option Returns. *The Journal of Finance*, 56(3), 983–1009.
- Cox, J. C., & Ross, S. A. (1976). The Valuation Of Options For Alternative Stochastic Processes. *Journal of Financial Economics*, 3, 145–166.
- De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *The Journal of Political Economy*, 98(4), 703–738.
- Demiguel, V., Plyakha, Y., Uppal, R., & Vilkov, G. (2012). Improving Portfolio Selection Using Option-Implied Volatility and Skewness. *Working Paper, London Business School*.
- Demsetz, H. (1968). The cost of transacting. *The Quarterly Journal of Economics*, 82(1), 33–53.
- Derman, E., Kani, I., & Zou, J. Z. (1996). The Local Volatility Surface: Unlocking the Information in Index Option Prices. *Financial Analyst Journal*, 52(4), 25–36.
- Deuskar, P., Gupta, A., & Subrahmanyam, M. G. (2007). The Economic Determinants of Interest Rate Option Smiles. *Working paper, University of Illinois at Urbana-Champaign*.
- Diavatopoulos, D., Doran, J. S., Fodor, A., & Peterson, D. R. (2008). The Information Content of Implied Skewness and Kurtosis Changes Prior to Earnings Announcements for Stock and Option Returns. *Working Paper, Florida State University*.
- Driessen, J., Maenhout, P. J., & Vilkov, G. (2009). The Price of Correlation Risk: Evidence from Equity Options. *The Journal of Finance*, 64(3), 1377–1406.
- Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied Volatility Functions: Empirical Tests. *The Journal of Finance*, 53(6), 2059–2106.
- Ederington, L. H., & Guan, W. (2010). How asymmetric is U.S. stock market volatility? *Journal of Financial Markets*, 12(2), 225–248.
- Evnine, J., & Rudd, A. (1985). Index Options : The Early Evidence. *The Journal of Finance*, 40(3), 743–756.
- Figlewski, S. (1989). What Does an Option Pricing Model Tell Us about Option Prices? *Financial Analyst Journal*, 45(5), 12–15.
- Fong, K., Madhavan, A., & Swan, P. L. (2006). Upstairs, Downstairs: Does the Upstairs Market Hurt the Downstairs. *Working Paper, School of Banking and Finance, University of New South Wales*.
- Gallant, R. A., & Tauchen, G. (1989). Semiparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications. *Econometrica*, 57(5), 1091–1120.

- Geske, R. (1978). The Pricing of Options with Stochastic Dividend Yield. *The Journal of Finance*, 33(2), 617–625.
- Goncalves, S., & Guidolin, M. (2004). Predictable Dynamics in the S&P 500 Index Options Implied Volatility Surface. *Working Paper, Universite de Montreal*.
- Gou, B., Han, Q., & Ryu, D. (2011). Non-parametric Tests for the Martingale Restriction: A New Approach. *Working Paper, Xiamen University*.
- Grossman, S. J. (1987). An Analysis Of The Implications For Stock And Futures Price Volatility Of Program Trading And Dynamic Hedging Strategies. *NBER Working Paper: 2357*.
- Han, E. (2011). Econometrics Honor ' s Exam Review Session. *Lecture presented in Economics, Harward Univeristy, Cambridge*.
- Harrison, J. M., & Kreps, D. M. (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory*, 408, 381–408.
- Hentschel, L. (2003). Errors in Implied Volatility. *The Journal of Financial and Quantitative analysis*, 38(4), 779–810.
- Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2), 327–343.
- Heston, S. L., & Nandi, S. (2000). A Closed-Form GARCH Option Valuation Model. *The Review of Financial Studies*, 13(3), 585–625.
- Ho, L.-C. J. (1993). Option Trading and the Relation Between and Earnings : Cross-Sectional Analysis. *The Accounting Review*, 68(2), 368–384.
- Ho, T., & Stoll, H. (1981). Optimal dealer Pricing Under Transactions and Return Uncertainty. *The Journal of Financial Economics*, 9(1), 47–73.
- Huang, D. (2012). What Is the Maximum Return Predictability Permitted by Asset Pricing Models? *Working Paper, Washington University in St. Louis*.
- Hull, J. (2009). *Options, Futures, & Other Derivatives* (7th ed.). Prentice Hall.
- Hull, J., & White, A. (2012). Is FVA a Cost for Derivatives Desks ? *Risk magazine*, 83–85.
- Jackwerth, J. C. (2004). Option-implied risk-neutral distributions and risk aversion. *Research Foundation of AIMR*.
- Jackwerth, J. C., & Rubinstein, M. (1996). Recovering Probability Distributions Option Prices from Option Prices. *The Journal of Finance*, 51(5), 1611–1631.

- Jarrow, R., & Rudd, A. (1982). Approximate Option Valuation For Arbitrary Stochastic Processes. *Journal of Financial Economics*, 10, 347–369.
- Jiang, G. J., & Tian, Y. S. (2005). The Model-Free Implied Volatility and Its Information Content. *Review of Financial Studies*, 18(4), 1305–1342.
- Jondeau, E., & Rockinger, M. (1999). Estimating Gram-Charlier Expansions with Positivity Constraints. *Working paper No. 56, Banque de France*.
- Jones, C. S. (2006). A Nonlinear Factor Analysis of S&P 500 Index Option Returns. *The Journal of Finance*, 61(5), 2325–2363.
- Jurczenko, E., Maillet, B., & Negrea, B. (2002). Skewness and Kurtosis Implied by Option Prices : A Second Comment. *Unpublished Manuscript, University Paris*.
- Kalay, A., & Subrahmanyam, M. G. (1984). The Ex-Dividend Day Behavior of Option Prices. *The Journal of Business*, 57(1), 113–128.
- Karatzas, I., & Kou, S. G. (1996). On the Pricing of Contingent Claims under Constraints. *The Annals of Applied Probability*, 6(2), 321–369.
- Keele, L., & Kelly, N. J. (2006). Dynamic Models for Dynamic Theories : The Ins and Outs of Lagged Dependent Variables. *Political Analysis*, 14(2), 186–205.
- Kiploks, J., & Lazdins, J. (2011). *Hedging Effectiveness of Index Options in Sweden*. Thesis, Stockholm School of Economics.
- Koenker, R., & Hallock, K. F. (2001). Quantile Regression. *Journal of Economic Perspectives*, 15(4), 143–156.
- Kuo, W., & Lin, T.-Y. (2011). Does the Put-Call Ratio Forecast Market Returns ? Evidence from an Emerging Market. *International Research Journal of Finance and Economics*, 1(69), 95–107.
- Leland, H. E. (1985). Option Pricing and Replication with Transaction costs. *The Journal of Finance*, 40(5), 1283–1301.
- Levy, H. (2010). Volatility Risk Premium , Market Sentiment and Market Anomalies. *Working paper, The University of Melbourne*.
- Lim, G. C., Martin, G. M., & Martin, V. L. (2005). Parametric pricing of higher order moments in S&P500 options. *Journal of Applied Econometrics*, 20(3), 377–404.
- Longstaff, F. A. (1995). Option Pricing and the Martingale Restriction. *The Review of Financial Studies*, 8(4), 1091–1124.
- Madan, D., & Milne, F. (1994). Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance*, 4(3), 223–245.

- Maheu, J. M., & McCurdy, T. H. (2002). Nonlinear Features of Realized FX Volatility. *The Review of Financial Studies*, 15(4), 668–681.
- Maidel, S., & Sahlin, K. (2010). Capturing the volatility premium through call overwriting. (*Russel Research, Russel Investments*).
- Manuel, J., & Chang, P. H. K. (1995). Testing the Expectations Hypothesis on the Term Structure of Volatilities in Foreign Exchange Options. *The Journal of Finance*, 50(2), 529–547.
- Martellini, L., & Simsek, K. (2005). Structured Forms of Investment Strategies in Institutional Investors' Portfolios. *Working Paper, EDHEC*.
- Martikainen, T., & Puttonen, V. (1996). Finnish day-of-the week effects. *Journal of Business Finance & Accounting*, 23(7), 1019–1032.
- Mayhew, S. (1995). Implied Volatility. *Financial Analyst Journal*, 51(4), 8–20.
- Melick, W. R., & Thomas, C. P. (1997). Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis. *The Journal of Financial and Quantitative analysis*, 32(1), 91–115.
- Meyers, T. A. (1994). *The Technical Analysis Course by Thomas A. Meyers* (2nd ed., pp. 89–105). New York: McGraw-Hill Trade.
- Naes, R., Skjeltorp, J. A., & Ødegaard, B. A. (2011). Stock Market Liquidity and the Business Cycle. *The Journal of Finance*, 66(1), 139–176.
- NASDAQ-OMX. (2011). OMX Stockholm 30 Index. *Rules for the Construction and Maintenance of the OMX Stockholm 30 Index*. Retrieved from https://indexes.nasdaqomx.com/docs/Methodology_OMXS30.pdf
- Neely, C. J. (2004). Forecasting Foreign Exchange Volatility: Why Is Implied Volatility Biased and Inefficient? And Does It Matter? *Working Paper 2002-017D, Federal Reserve Bank of St. Louis*.
- Neumann, M., & Schlag, C. (1996). Martingale Restrictions and Implied Distributions for German Stock Index Option Prices. *Working paper No. 196, University at Karlsruhe*.
- Nordén, L. (2003). Option Exchange Design: Concentration of Trading and Open Interest at the Swedish Index Options Market. *Working Paper, Stockholm University*.
- Nordén, L., & Xu, C. (2012). Option happiness and liquidity: Is the dynamics of the volatility smirk affected by relative option liquidity? *Journal of Futures Markets*, 32(1), 47–74.
- Pan, J., & Poteshman, A. M. (2006). The Information in Option Volume for Future Stock Prices. *Review of Financial Studies*, 19(3), 871–908.

- Parasuraman, N. R., & Ramudu, P. J. (2011). Historical and implied volatility: an investigation into nifty futures and options. *Australian Journal of Business and Management Research*, 1(7), 112–120.
- Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, 53(1), 61–65.
- Potesman, A. M. (2000). Forecasting Future Variance From Option Prices. *Working Paper, University of Illinois at Urbana-Champaign*.
- Rehman, Z., & Vilkov, G. (2012). Risk-Neutral Skewness : Return Predictability and Its Risk-Neutral Skewness : Return Predictability and Its Sources. *Working Paper, Goethe University Frankfurt*.
- Ring, J. (2011). *The Implied Probability Distribution: A Study of Swedish Equity Index Options*. Stockholm School of Economics.
- Robertsson, G. (2000). Conditioning Information in Tactical Asset Allocation. *Unpublished Manuscript, Stockholm School of Economics*.
- Rubinstein, M. (1998). Edgeworth Binomial Trees. *Journal of Derivatives*, 5(3), 20–27.
- Shefrin, H. (1999). Irrational Exuberance and Option Smiles. *Financial Analyst Journal*, 55(6), 91–103.
- Shie, F.-S. (2009). Constructing the Implied Volatility under the Bull / Bear Markets : The Case of S & P 500 Index Options. *International Research Journal of Finance and Economics*, 1(34), 96–106.
- Shu, J., & Zhang, J. E. (2004). The Relationship Between Implied and Realized Volatility of S & P 500 Index. *Wilmott Magazine*, 83–91.
- Straja, S. R. (2003). Derivative Valuation – Gram-Charlier. *Unpublished Manuscript, Montgomery Investment Technology*.
- Sum, V. (2012). Stock Market Returns and Liquidity : Dynamic Relationships and Causality. *Working Paper, University of Maryland-Eastern Shore*.
- Szakmary, A., Ors, E., Kim, K. J., & Davidson, W. N. (2003). The predictive power of implied volatility: Evidence from 35 futures markets. *Journal of Banking & Finance*, 27(11), 2151–2175.
- Verma, S., Ertosun, G. M., Wang, W., Ambruster, B., & Giesecke, K. (2008). Implied Volatility Modeling. *Working Paper, Northwestern University*.
- Vijh, A. M. (1990). Liquidity of the CBOE Equity Options. *The Journal of Finance*, 45(4), 1157–1179.
- Vilkov, G., & Xiao, Y. (2012). Option-Implied Information and Predictability of Extreme Returns. *Working Paper, Goethe University Frankfurt*.
- Wang, Y.-H., Keswani, A., & Taylor, S. J. (2006). The Relationships Between Sentiment, Returns and Volatility. *International Journal of Forecasting*, 22(1), 109–123.

- Xing, Y., Zhang, X., & Zhao, R. (2008). What Does Individual Option Volatility Smirk Tell Us About Future Equity Returns? *Working Paper, Rice University*.
- Yang, D., & Zhang, Q. (2000). Drift-Independent Volatility Estimation Based on High , Low , Open , and Close Prices. *Journal of Business*, 73(3), 477–491.
- Ziegler, A. (2007). Why Does Implied Risk Aversion Smile? *The Review of Financial Studies*, 20(3), 859–904.
- Zivney, T. L., & Alderson, M. J. (1986). Hedged Dividend Capture with Stock Index Options. *Financial Managment*, 15(2), 5–12.

Appendix 1: Index higher moment realizations

The figures below show sample estimated and realized volatility, skewness and kurtosis for short, medium and long time horizons. Values are estimated on a rolling basis with 30, 60 and 90 day horizon in order to show dynamics through time of the underlying distribution statistical moments.

Figure 1: Realised volatility

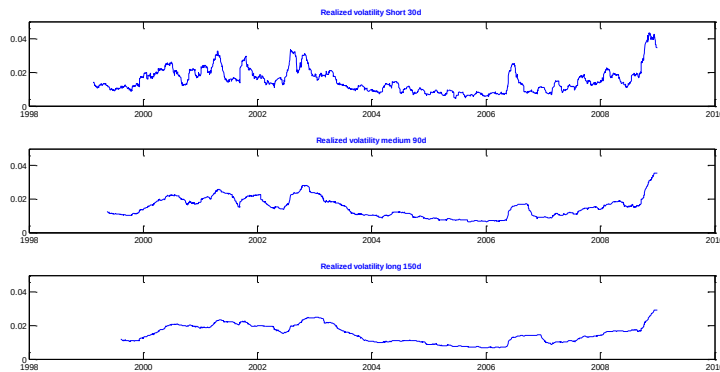


Figure 2: Realised skewness

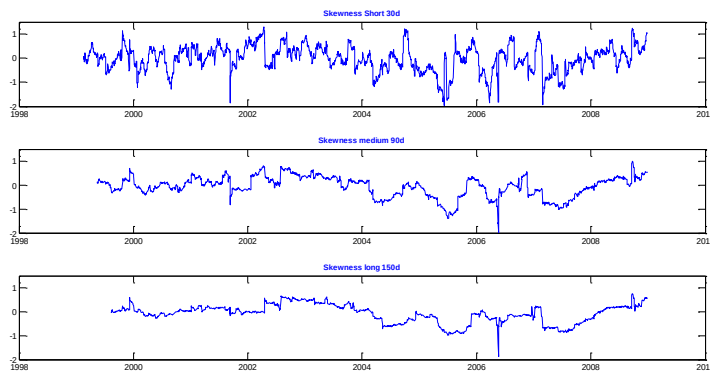
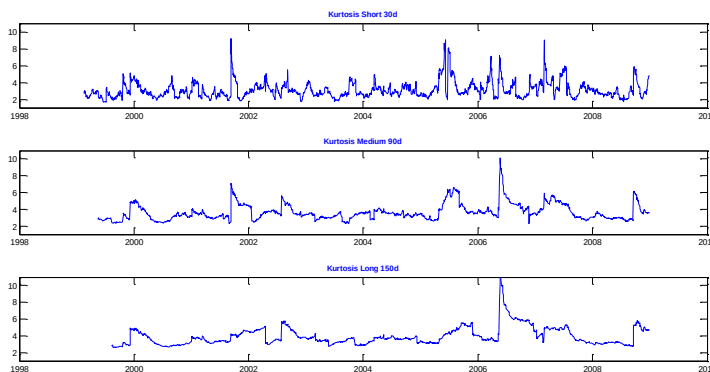


Figure 3: Realised kurtosis



Appendix 2: Summary statistics of call and put Implied Volatilities

The tables below show average estimated GC volatility per option basis (which is adjusted by matched daily skewness and kurtosis). Note that this is not the average of estimated daily volatility. Black-Scholes IV is the average of implied volatility on option basis.

Table 1: Summary of implied volatility for call contracts

		Time span of options			Total	
		Short	Medium	Long	# options	%
xOTM (<0.90)	Black-Scholes IV	35.4%	27.7%	23.0%		
	Gram Charlier IV	36.0%	31.9%	27.7%		
	# options	1514	1877	2322	5713	12%
OTM (0.90-0.97)	Black-Scholes IV	27.2%	22.6%	19.9%		
	Gram Charlier IV	30.1%	25.9%	23.1%		
	# options	6774	10109	5306	22189	45%
ATM (0.97-1.03)	Black-Scholes IV	25.1%	22.7%	20.4%		
	Gram Charlier IV	27.2%	25.0%	23.0%		
	# options	7733	12112	5600	25445	52%
ITM (1.03-1.10)	Black-Scholes IV	32.8%	28.5%	23.3%		
	Gram Charlier IV	30.2%	29.7%	25.5%		
	# options	2946	5031	3130	11107	23%
xITM (>1.10)	Black-Scholes IV	44.7%	35.7%	26.3%		
	Gram Charlier IV	31.0%	34.1%	27.7%		
	# options	455	741	2078	3274	7%
Total	# options	19422	29870	18436	67728	
	# options, %	39%	61%	37%		

Table 2: Summary of implied volatility for put contracts

		Time span of options			Total	
		Short	Medium	Long	# options	%
xOTM (<0.90)	Black-Scholes IV	41.0%	30.4%	24.3%		
	Gram Charlier IV	23.9%	24.6%	22.7%		
	# options	3628	5710	6968	16306	33%
OTM (0.90-0.97)	Black-Scholes IV	30.7%	24.7%	22.3%		
	Gram Charlier IV	26.3%	24.5%	23.7%		
	# options	7941	12028	6388	26357	53%
ATM (0.97-1.03)	Black-Scholes IV	25.0%	23.0%	22.5%		
	Gram Charlier IV	27.0%	25.4%	25.6%		
	# options	7483	11299	4188	22970	47%
ITM (1.03-1.10)	Black-Scholes IV	30.5%	25.8%	24.4%		
	Gram Charlier IV	35.2%	30.0%	29.0%		
	# options	2437	4584	2014	9035	18%
xITM (>1.10)	Black-Scholes IV	36.2%	26.0%	23.7%		
	Gram Charlier IV	41.9%	31.1%	29.2%		
	# options	172	515	1154	1841	4%
Total	# options	21661	34136	20712	76509	
	# options, %	44%	69%	42%		

Appendix 3: GC and BS volatility plot comparison

The figures below show average estimated GC volatility per option basis (which is adjusted by matched daily skewness and kurtosis). Note that this is not the average of estimated daily volatility. Black-Scholes IV is the average of implied volatility on option basis.

Figure 1: Short term horizon

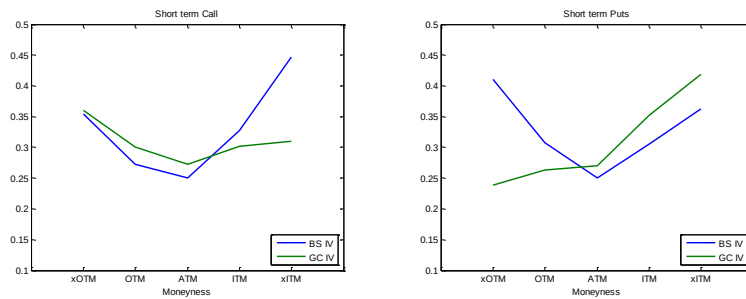


Figure 2: Medium term horizon

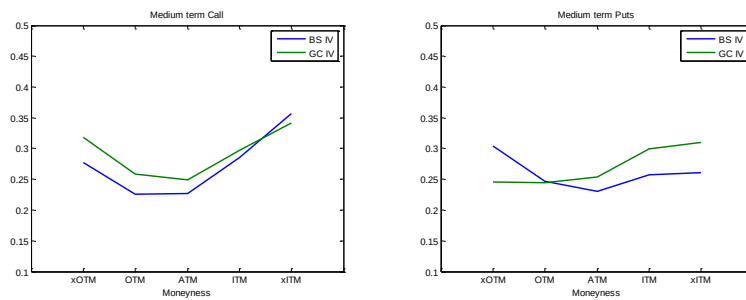
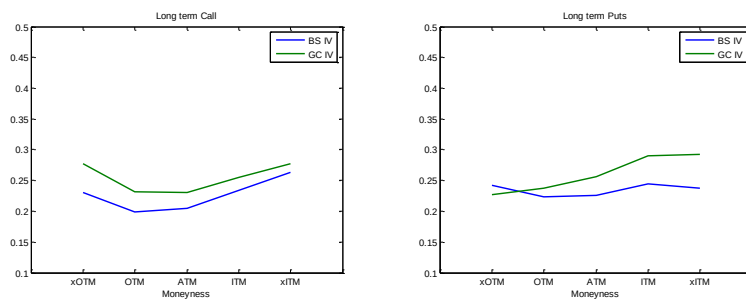


Figure 3: Long term horizon



Appendix 4: GC and BS volatility explanatory regression

Below are provided results for regression specification of estimated slopes for GC skewness and kurtosis adjusted volatility and BS volatility for put and call options with monthly fixed effects. Call_ITM, Call_OTM, Put_ITM and Put_OTM are interaction control variables where moneyness is set in following manner (ITM=in the money, OTM=out of the money and subscript x refers to deep in ITM and OTM). Call_volume, Put_volume, Call_bidask and Put_bidask are option type (Call and Put) interaction variables with *volume_dummy* which is 1 if there is trade for option in that particular day and *Bidasksek* is option bid-ask spread (SEK). In addition, *ttm* refers to time to maturity expressed in year. Option-level clustered errors are reported. 1462 option contracts dropped if their RMSE was above 0.05 SEK.

	Gram-Charlier				Black-Scholes			
	All (1)	Short (2)	Mid (3)	Long (4)	All (5)	Short (6)	Mid (7)	Long (8)
	implvol_adj	implvol_adj	implvol_adj	implvol_adj	implvol	implvol	implvol	implvol
call_xITM	-0.0144 (-0.79)	-0.243*** (-7.44)	-0.111*** (-4.28)	-0.0209 (-0.89)	-0.0851*** (-4.60)	-0.402*** (-13.39)	-0.219*** (-9.10)	-0.105*** (-4.07)
call_ITM	-0.0404* (-2.11)	-0.281*** (-8.45)	-0.138*** (-5.30)	-0.0415 (-1.65)	-0.134*** (-6.76)	-0.504*** (-16.82)	-0.282*** (-11.85)	-0.145*** (-5.21)
call_ATM	-0.0475* (-2.36)	-0.273*** (-7.89)	-0.148*** (-5.39)	-0.0503 (-1.88)	-0.168*** (-8.09)	-0.553*** (-17.76)	-0.316*** (-12.65)	-0.169*** (-5.69)
call_OTM	-0.0478* (-2.26)	-0.258*** (-7.15)	-0.151*** (-5.22)	-0.0563* (-2.00)	-0.185*** (-8.43)	-0.564*** (-17.38)	-0.339*** (-12.95)	-0.186*** (-5.98)
call_xOTM	-0.0498* (-2.19)	-0.262*** (-6.82)	-0.145*** (-4.74)	-0.0696* (-2.28)	-0.197*** (-8.49)	-0.573*** (-16.53)	-0.355*** (-12.70)	-0.218*** (-6.51)
put_xITM	-0.0576** (-3.10)	-0.232*** (-7.07)	-0.165*** (-6.38)	-0.0552* (-2.09)	-0.161*** (-8.56)	-0.496*** (-16.82)	-0.326*** (-13.97)	-0.157*** (-5.75)
put_ITM	-0.0503* (-2.56)	-0.263*** (-7.84)	-0.149*** (-5.48)	-0.0551* (-1.98)	-0.166*** (-8.27)	-0.547*** (-18.14)	-0.318*** (-13.00)	-0.164*** (-5.57)
put_ATM	-0.0539** (-2.62)	-0.288*** (-8.29)	-0.149*** (-5.22)	-0.0528 (-1.79)	-0.169*** (-8.02)	-0.573*** (-18.32)	-0.318*** (-12.36)	-0.163*** (-5.22)
put_OTM	-0.0489* (-2.27)	-0.284*** (-7.81)	-0.149*** (-4.99)	-0.0446 (-1.45)	-0.144*** (-6.50)	-0.538*** (-16.42)	-0.304*** (-11.26)	-0.151*** (-4.62)
put_xOTM	-0.0380 (-1.64)	-0.278*** (-7.26)	-0.145*** (-4.85)	-0.0374 (-1.14)	-0.111*** (-4.67)	-0.503*** (-14.48)	-0.276*** (-10.01)	-0.133*** (-3.77)
call_volume	0.00221* (2.28)	-0.0270*** (-6.00)	-0.00339*** (-4.28)	0.00158 (1.36)	0.000191 (0.20)	-0.0363*** (-9.31)	-0.00694*** (-10.25)	0.000566 (0.54)
put_volume	0.00588*** (6.23)	-0.0154*** (-3.98)	0.00169 (1.75)	0.000781 (0.83)	0.00619*** (6.93)	-0.0161*** (-4.90)	0.00171* (2.16)	0.00212* (2.17)
call_bidask	0.00657*** (6.82)	0.0258*** (18.51)	0.0227*** (12.25)	0.00866*** (4.60)	0.00365*** (3.80)	0.0232*** (18.58)	0.0197*** (11.64)	0.00920*** (5.12)
put_bidask	0.0130*** (10.52)	0.0287*** (18.93)	0.0215*** (8.17)	0.0146*** (4.36)	-0.000480 (-0.45)	0.0193*** (14.79)	0.0168*** (8.35)	0.00778** (2.80)
ttm	-0.0166*** (-7.18)	-1.374*** (-44.05)	-0.128*** (-12.54)	0.00149 (0.62)	-0.0358*** (-13.70)	-1.639*** (-58.39)	-0.169*** (-19.32)	-0.0108*** (-4.92)
_cons	0.307*** (15.40)	0.640*** (18.14)	0.410*** (15.07)	0.276*** (10.59)	0.411*** (19.93)	0.918*** (28.93)	0.565*** (22.82)	0.373*** (12.97)
N	142775	40029	63705	39041	142775	40029	63705	39041
adj. R-sq	0.685	0.657	0.798	0.823	0.675	0.694	0.815	0.814

t statistics in parentheses * p<0.05, ** p<0.01, *** p<0.001

Appendix 5: Comparison of GC and Model Free measures

The figures below provide comparative overview of Gram-Charlier estimated daily measures (including skewness and kurtosis) with Model-Free implied volatility and skewness. Black-line indicates 1:1 match between both measures, however x-axis for skewness scatter plot is rescaled, because this measure does not have a theoretical boundary. While Model-Free estimated skewness is not strictly bounded by the region $[-2; 2]$, it was selected to provide best visual overview for the data and there were only few outliers falling outside this range.

Figure 1: Short term horizon

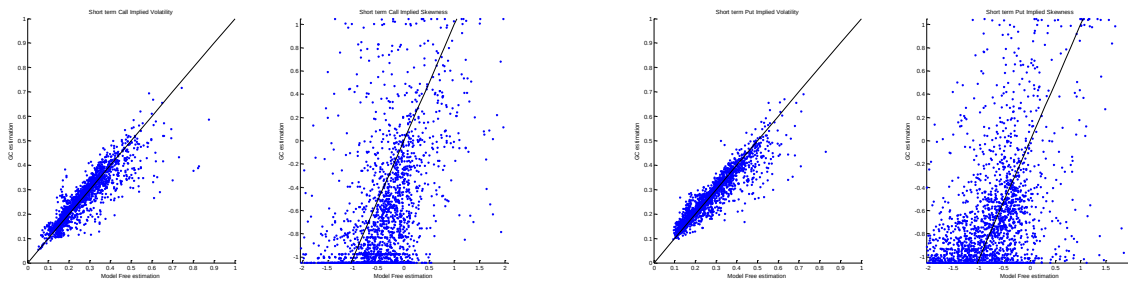


Figure 2: Medium term horizon

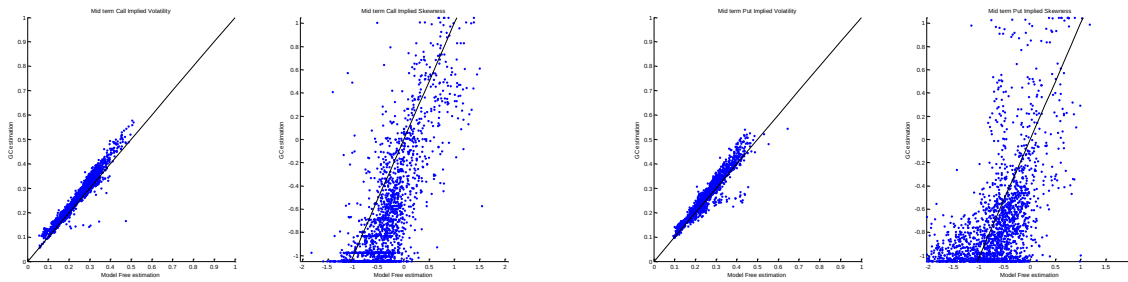
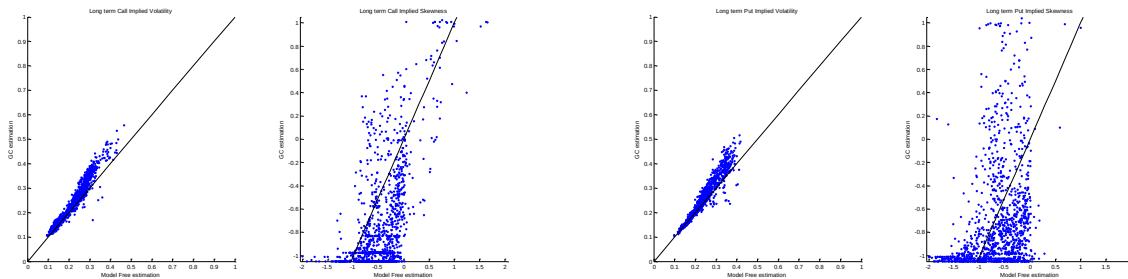


Figure 3: Long term horizon



Appendix 6: Estimation RMSE comparison

Table 1 and Table 2 show estimated RMSE for data used in our sample. Figure 1 shows RMSE development through time for short-term put and call options. Table 3 show estimated RMSE for data without the positivity constraints.

Table 1: RMSE improvements for Call estimation

	Volatility	Skewness	Kurtosis
Short	1.3770	1.3742	1.0798
<i>Improve, %</i>		0%	21%
Medium	1.0079	1.0056	0.7668
<i>Improve, %</i>		0%	24%
Long	1.7728	1.7689	1.3358
<i>Improve, %</i>		0%	24%

RMSE given in SEK

Table 2: RMSE improvements for Put estimation

	Volatility	Skewness	Kurtosis
Short	1.2189	1.1904	0.8267
<i>Improve, %</i>		2%	31%
Medium	0.9354	0.9116	0.5973
<i>Improve, %</i>		3%	34%
Long	1.5751	1.5395	1.0273
<i>Improve, %</i>		2%	33%

RMSE given in SEK

Figure 1: Root Mean Squared Error

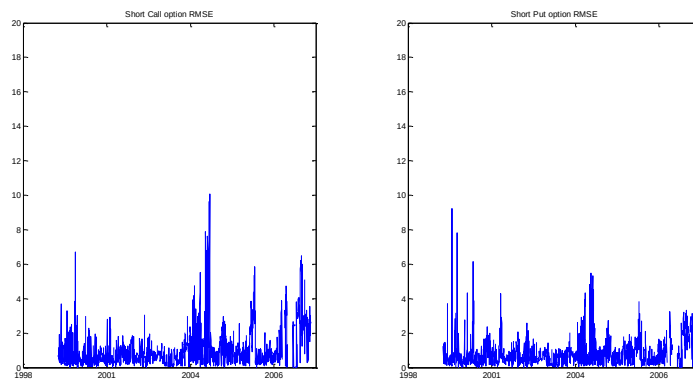


Table 3: RMSE improvements for Call estimation (unconstrained)

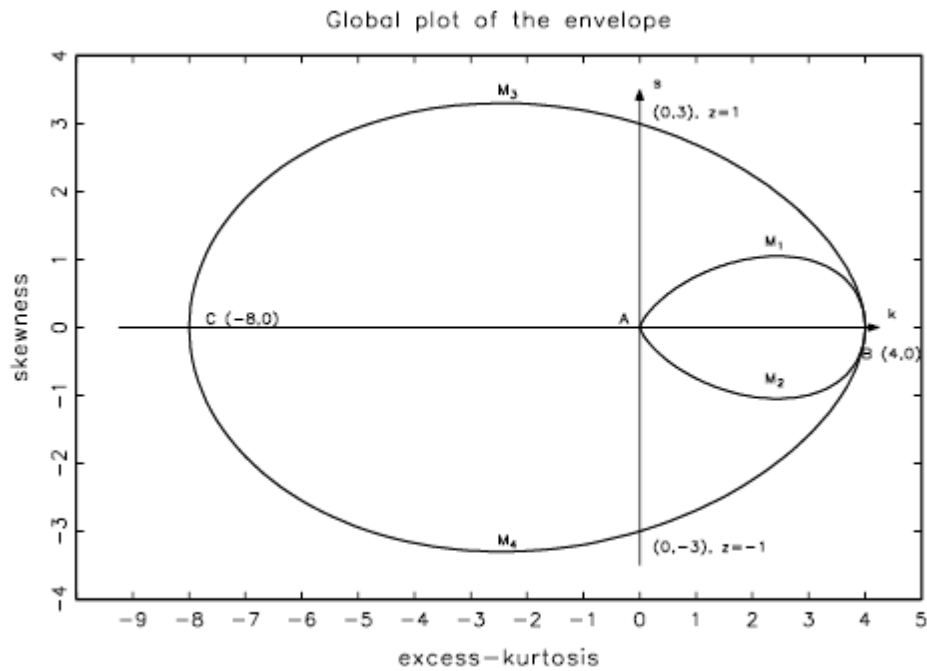
	Volatility	Skewness	Kurtosis
Short	0.4002	0.2987	0.2668
<i>Improve, %</i>		25%	11%
Medium	0.8453	0.7014	0.6804
<i>Improve, %</i>		17%	3%
Long	2.4786	2.2263	2.1864
<i>Improve, %</i>		10%	2%

RMSE given in SEK

Appendix 7: Global envelope plot

The figure below provides theoretical boundaries for skewness and kurtosis combinations that will yield positive density expansion. Large circle provides boundary for skewness/kurtosis combinations if $|z| \leq \sqrt{3}$ and smaller circle provides boundary if $|z| \geq \sqrt{3}$. Source of figure: Jondeau and Rockinger (1999)

Variable z is standard parameter in BS framework with $z = \frac{\ln\left(\frac{S_t}{S_0}\right) - \left(r - q - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}$.



Appendix 8: Put-Call SEK volume ratio return predictability

The tables below provide regression specifications overview of various index return horizons. Table 1 shows predictive and explanatory regressions for different combinations of volume and return aggregation horizons (overlapping). Table 2 shows daily return predictability using different volume aggregation horizons (overlapping data). Table 3 shows non-overlapping results using different volume and return aggregation horizons. For Table 1 and Table 2 Newey-West heteroskedasticity and autocorrelation robust t-stats are reported. For Table 3 standard t-stat is reported. In regression specifications $Ratio = \frac{SEK \text{ volume of traded ITM puts}}{SEK \text{ volume of traded ITM calls}}$

Table 1: Put-call ratio predictability of index return

	(1)	(2)	(3)	(4)	(5)
	5 day	Predictive 20 day	60 day	Explanatory 5 day	20 day
_cons	0.002	0.010	0.109	0.011	0.048
	1.35	2.06	5.40	5.63	10.92
Ratio	-0.001	-0.010	-0.130	-0.007	-0.046
	-2.28	-2.30	-5.32	-6.90	-11.47
adj. R-sq	0.01	0.02	0.09	0.23	0.47

Table 2: Put-call ratio predictability of daily index return

	(1)	(2)	(3)	(4)
Volume horizon	5	10	15	20
_cons	0.002	0.002	0.002	0.002
	2.23	3.41	2.86	2.30
Ratio	-0.0009	-0.0019	-0.0018	-0.0016
	-3.53	-4.51	-3.09	-2.32
adj. R-sq	0.011	0.014	0.009	0.005

Table 3: Put-call ratio predictability of non-overlapping index returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Volume horizon	5	5	10	10	10	20	20	20
Returns horizon	5	10	5	10	15	10	15	20
_cons	0.0042	0.0074	0.0050	0.0074	0.0088	0.0108	0.0142	0.0155
	1.91	1.90	1.61	1.82	1.35	1.76	2.01	1.79
Ratio	-0.0018	-0.0033	-0.0030	-0.0039	-0.0038	-0.0034	-0.0106	-0.0110
	-2.12	-2.08	-1.90	-1.87	-1.07	-0.86	-2.31	-1.98
adj. R-sq	0.011	0.022	0.018	0.018	0.009	0.008	0.053	0.040
N	391	196	195	195	130	98	98	98

Note: This estimation procedure uses strictly non-overlapping data, so for any combination where there is some possible overlap, the largest overlap horizon is chosen in order to avoid this problem.

Appendix 9: Realized volatility prediction

Below are provided results for regression specification of estimated predictability of realized volatility using GC and MF IV measures. Regression specification includes realized volatility as independent variable. *Ratio* is the estimated OLS coefficient on respective type of implied volatility. Table 1 includes Newey-West HAC standard errors and corresponding t-stat value is provided. Estimated t-stats provide hypothesis that estimated *constant* $H_0 = 0$ and *ratio* $H_0 = 1$. Table 2 includes only non-overlapping observations for short-term options, with sampling period as 1st day of period when the options were issued. Estimated t-stats provide hypothesis that estimated *constant* $H_0 = 0$ and *ratio* $H_0 = 1$ with standard OLS errors.

Table 1: Predictability of realized volatility (overlapping observations)

	Calls						Puts					
	Short	Medium		Long			Short	Medium		Long		
	GC (1)	MFIV (2)	GC (3)	MFIV (4)	GC (5)	MFIV (6)	GC (7)	MFIV (8)	GC (9)	MFIV (10)	GC (11)	MFIV (12)
_cons	0.017	0.060	0.052	0.059	0.002	-0.014	0.007	0.052	0.033	0.042	0.014	-0.009
	-0.81	-3.56	-3.48	-3.52	-0.17	0.90	-0.42	-3.58	-2.71	-3.10	-0.85	0.55
Ratio	0.898	0.737	0.859	0.896	0.752	0.874	0.920	0.745	0.911	0.920	0.687	0.810
	1.18	3.81	2.69	1.61	5.23	1.74	1.17	4.33	1.92	1.38	4.78	2.46
adj. R-sq	0.43	0.38	0.48	0.45	0.46	0.42	0.46	0.41	0.54	0.49	0.41	0.38
N	1907	1867	1950	1836	1505	1219	1904	1883	1902	1856	1427	1204

Table 2: Predictability of realized volatility (non-overlapping observations)

	Calls		Puts	
	GC (1)	MFIV (2)	GC (3)	MFIV (4)
_cons	0.013	0.100	0.004	0.044
	-0.54	-4.10	-0.19	-1.95
Ratio	0.899	0.509	0.910	0.779
	1.18	6.77	1.20	2.86
adj. R-sq	0.542	0.351	0.618	0.528
N	94	93	94	93

Appendix 10: Realized volatility prediction (2)

Below are provided results for regression specification of estimated predictability of realized volatility using GC and MF IV measures. Regression specification includes realized volatility as independent variable. *IV* is coefficient on option implied volatility and *hist* is coefficient on historic (alternative) forecast. Table 1 includes Newey-West HAC standard errors and corresponding t-stat value is provided. Estimated t-stats provide hypothesis that estimated *constant* $H_0 = 0$ and *ratio* $H_0 = 1$. Table 2 includes only non-overlapping observations for short-term options, with sampling period as 1st day of period when the options were issued. Estimated t-stats provide hypothesis that estimated *constant* $H_0 = 0$ and *ratio* $H_0 = 1$ with standard OLS errors.

Table 1: Predictability of realized volatility (overlapping observations)

	Calls						Puts					
	Short		Medium		Long		Short		Medium		Long	
	GC (1)	MFIV (2)	GC (3)	MFIV (4)	GC (5)	MFIV (6)	GC (7)	MFIV (8)	GC (9)	MFIV (10)	GC (11)	MFIV (12)
cons	0.031	0.034	0.054	0.054	0.121	0.106	0.017	0.035	0.038	0.039	0.148	0.135
	-1.39	-2.43	-3.63	-3.41	-10.81	-7.45	-0.95	-2.66	-3.08	-3.06	-9.69	-7.75
IV	0.887	0.364	0.858	0.512	0.621	0.651	0.927	0.380	0.918	0.622	0.460	0.331
	1.30	7.61	2.71	4.35	8.23	2.92	1.07	8.09	1.71	3.59	7.99	4.72
hist	-0.010	0.487	-0.002	0.358	0.001	0.075	-0.014	0.454	-0.011	0.292	0.012	0.223
	-3.22	7.97	-0.49	4.38	0.50	0.82	-2.94	7.37	-1.77	3.59	3.12	2.13
adj. R-sq	0.44	0.50	0.48	0.49	0.35	0.32	0.47	0.50	0.55	0.53	0.28	0.23
N	1907	1867	1950	1836	1505	1219	1904	1883	1902	1856	1427	1204

Table 2: Predictability of realized volatility (non-overlapping observations)

	Short-Calls		Short-Puts	
	GC (1)	MFIV (2)	GC (3)	MFIV (4)
cons	0.013	0.052	0.005	0.032
	-0.55	-2.36	-0.22	-1.50
IV	0.540	0.176	0.675	0.448
	3.61	10.15	2.64	4.83
hist	0.600	0.545	0.640	0.591
	-3.64	-6.21	-2.37	-3.74
adj. R-sq	0.60	0.55	0.64	0.59
N	94	93	94	93