Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

Mathias Barkhagen
Linköping Studies in Science and Technology.
Dissertations No. 1657

Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

Copyright © Mathias Barkhagen, 2015

Typeset by the author in \LaTeX documentation system.

ISSN 0345-7524
Printed by LiU-Tryck, Linköping, Sweden 2015
This dissertation is centered around two comprehensive themes: the extraction of information embedded in equity index option prices, and how to use this information in order to be able to make optimal decisions in the equity index option markets. These problems are important for decision makers in the equity index options markets, since they are continuously faced with making decisions under uncertainty given observed market prices. The methods developed in this dissertation provide robust tools that can be used by practitioners in order to improve the quality of the decisions that they make.

In order to be able to extract information embedded in option prices, the dissertation develops two different methods for estimation of stable option implied surfaces which are consistent with observed market prices. This is a difficult and ill-posed inverse problem which is complicated by the fact that observed option prices contain a large amount of noise stemming from market micro structure effects. Producing estimated surfaces that are stable over time is important since otherwise risk measurement of derivatives portfolios, pricing of exotic options and calculation of hedge parameters will be prone to include significant errors. The first method that we develop leads to an optimization problem which is formulated as a convex quadratic program with linear constraints which can be solved very efficiently. The second estimation method that we develop in the dissertation makes it possible to produce local volatility surfaces of high quality, which are consistent with market prices and stable over time. The high quality of the surfaces estimated with the second method is the crucial input to the research which has resulted in the last three papers of the dissertation.

The stability of the estimated local volatility surfaces makes it possible to build a realistic dynamic model for the equity index derivatives market. This model forms the basis for the stochastic programming (SP) model for option hedging that we develop in the dissertation. We show that the SP model, which uses
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

generated scenarios for the squared local volatility surface as input, outperforms the traditional hedging methods that are described in the literature. Apart from having an accurate view of the variance of relevant risk factors, it is when building a dynamic model also important to have a good estimate of the expected values, and thereby risk premia, of those factors. We use a result from recently published research which lets us recover the real-world density from only a cross-section of observed option prices via a local volatility model. The recovered real-world densities are then used in order to identify and estimate liquidity premia that are embedded in option prices.

We also use the recovered real-world densities in order to test how well the option market predicts the realized statistical characteristics of the underlying index. We compare the results with the performance of commonly used models for the underlying index. The results show that option prices contain a premium in the tails of the distribution. By removing the estimated premia from the tails, the resulting density predicts future realizations of the underlying index very well.
Sammanfattning

Avhandlingen fokuserar på två övergripande teman: att extrahera information som finns inbäddad i priser för aktieindexoptioner, samt hur man kan använda denna information för att fatta optimala beslut på aktieindexoptionsmarknader. Dessa problem är viktiga för beslutsfattare som agerar på aktieindexoptionsmarknader, eftersom de kontinuerligt ställs inför att fatta beslut under osäkerhet givet observerade marknadspriser. Metoderna som utvecklas i avhandlingen tillhandahåller robusta verktyg som kan användas av marknadsaktörer för att förbättra kvalitén på fattade beslut.


Stabiliteten för de estimerade lokala variansytorna gör det möjligt att bygga en realistisk dynamisk modell för aktieindexderivatmarknaden. Denna modell ligger till grund för den stokastiska programmerings-modell (SP-modell) för optionshedgning som vi utvecklar i avhandlingen. Vi visar att SP-modellen, vilken använder
genererade scenerier för den lokala variansytan som indata, presterar bättre än de traditionella hedgningsmetoder som finns beskrivna i litteraturen. Förutom att ha ett korrekt estimat av variansen för relevanta riskfaktorer, är det när man bygger en dynamisk modell också viktigt att ha ett bra estimat för väntevärdena, och därmed riskpremierna, för dessa faktorer. Vi använder ett resultat från nyligen publicerad forskning som låter oss extrahera den fysiska densiteten via en lokal volatilitetsmodell, genom att enbart använda ett tvärsnitt av observerade optionspriser. De extraherade fysiska tätheter används sedan för att identifiera och estimera likviditetspremier som finns inbäddade i optionspriser.

Vi använder också de extraherade fysiska tätheter för att testa hur väl optionsmarknaden predikterar de realiserade statistiska egenskaper för det underliggande indexet. Vi jämför resultaten med prediktionsförmågan för vanligen använda modeller för det underliggande indexet. Resultaten visar att optionspriser innehåller en premie i svansarna för fördelningen. Genom att ta bort de estimerade premierna från svansarna, predikterar den resulterande tätheten väl framtida realiseringar för det underliggande indexet.
Acknowledgements

Writing these sentences means that I am approaching the end of a long journey which has been immensely inspiring and exciting, and at times also frustrating and exhausting. Over the last five years as a Ph.D. student I have met a number of people to whom I owe my gratitude.

First of all I want to thank my supervisor Jörgen Blomvall for the excellent supervision, and for providing guidance and inspiration during all parts of my journey as a Ph.D. student. I have learnt a lot from you! I also want to thank my second supervisor Ou Tang for organizing the research seminars at the division and for always being able to give insightful comments. I would also like to thank Torbjörn Larsson at the Mathematics Department, for inviting me to give a presentation there in 2010. This eventually led to me being accepted as a Ph.D. student in finance.

To all colleagues at the division of Production Economics: thank you for providing such a friendly and intellectually stimulating work environment. I have really enjoyed my time at the division. I would especially like to thank Kicki for making all new, and old, people feel welcome at the division, and for providing warmth. I also especially thank my two fellow Ph.D. students in finance, Jonas Ekblom and Johan Gustafsson. Jonas, for being a companion on the "rough path" as a finance Ph.D. student over the past 5 years. I have really enjoyed our discussions and your passion for knowledge and research in finance. Johan, for your sense of humour and your fresh perspective on many issues.

Work is not everything, luckily. I would like to thank my friends and family for helping me take my mind off work and for giving me inspiration. Finally, I would like to thank Linnea, for your patience with me over the last 4 months when I was finishing this thesis. You have put up with me working late into the nights and neglecting most things outside of work. I take it as a sign of love.

Linköping, 9 April 2015

Mathias Barkhagen
List of Papers

The following papers are appended and will be referred to by their alphabetic letters.


D. M. Barkhagen, J. Blomvall and E. Platen, Recovering the real-world density and liquidity premia from option data, Submitted, 2015.

Contents

Abstract i

Sammanfattning iii

Acknowledgements v

List of Papers vii

1 Introduction 1

2 Optimal decisions under uncertainty in the equity index derivatives markets 11

3 Empirical properties of equity index options 17
   3.1 Empirical statistical properties of the underlying equity market returns 17
   3.2 Empirical properties of the equity index options market 21
   3.3 Risk premia 23

4 Option implied surfaces 29
   4.1 The implied volatility surface 34
      4.1.1 Static no-arbitrage conditions 35
      4.1.2 Asymptotic results 36
## Optimal Decisions in the Equity Index Derivatives Markets
### Using Option Implied Information

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.3 Estimation methods described in the literature</td>
<td>38</td>
</tr>
<tr>
<td>4.2 The local volatility surface</td>
<td>41</td>
</tr>
<tr>
<td>4.2.1 Static no-arbitrage conditions</td>
<td>43</td>
</tr>
<tr>
<td>4.2.2 Estimation methods described in the literature</td>
<td>51</td>
</tr>
<tr>
<td>5 Dynamic models of option implied surfaces</td>
<td>55</td>
</tr>
<tr>
<td>5.1 No-arbitrage approach</td>
<td>56</td>
</tr>
<tr>
<td>5.2 PCA based approach</td>
<td>58</td>
</tr>
<tr>
<td>6 Estimating the physical PDF from observed option prices</td>
<td>61</td>
</tr>
<tr>
<td>6.1 The growth optimal portfolio</td>
<td>62</td>
</tr>
<tr>
<td>6.2 Pricing of European-style derivatives under the benchmark approach</td>
<td>67</td>
</tr>
<tr>
<td>6.3 Recovery of the physical PDF from observed option prices</td>
<td>73</td>
</tr>
<tr>
<td>7 Summary of included papers</td>
<td>79</td>
</tr>
<tr>
<td>7.1 Answers to RQ3 and RQ4</td>
<td>82</td>
</tr>
<tr>
<td>7.2 Answer to RQ1</td>
<td>87</td>
</tr>
<tr>
<td>7.3 Answer to RQ2</td>
<td>89</td>
</tr>
<tr>
<td>7.4 Answer to RQ5</td>
<td>90</td>
</tr>
<tr>
<td>7.5 Answer to ORQ</td>
<td>92</td>
</tr>
</tbody>
</table>

## Appended Papers

### Paper A - An Improved Convex Model for Efficient Estimation of Option Implied Surfaces

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1 Introduction</td>
<td>107</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>A.2 No-arbitrage constraints on call option prices</td>
<td>111</td>
</tr>
<tr>
<td>A.2.1 Equivalent formulation of the no-arbitrage conditions in the strike direction</td>
<td>114</td>
</tr>
<tr>
<td>A.3 Estimating the option implied RND for a single maturity</td>
<td>115</td>
</tr>
<tr>
<td>A.3.1 Examples of methods for estimating the option implied RND</td>
<td>116</td>
</tr>
<tr>
<td>A.3.2 Optimization based estimation of the implied RND for a single maturity</td>
<td>118</td>
</tr>
<tr>
<td>A.3.3 Derivatives prices as functions of the implied RND</td>
<td>120</td>
</tr>
<tr>
<td>A.3.4 Discretization of the optimization problem</td>
<td>121</td>
</tr>
<tr>
<td>A.3.5 Choosing the weight vectors $a, a'$ and $a''$ and the penalty matrices $E_a$ and $E_b$</td>
<td>124</td>
</tr>
<tr>
<td>A.4 Arbitrage-free estimation of the option implied RND surface</td>
<td>127</td>
</tr>
<tr>
<td>A.4.1 Continuous formulation of the full surface problem</td>
<td>127</td>
</tr>
<tr>
<td>A.4.2 Discretization of the full surface problem</td>
<td>129</td>
</tr>
<tr>
<td>A.5 Empirical implementation and demonstrations</td>
<td>130</td>
</tr>
<tr>
<td>A.5.1 Description of the data set</td>
<td>131</td>
</tr>
<tr>
<td>A.5.2 Extracting the RND and implied volatility surfaces</td>
<td>133</td>
</tr>
<tr>
<td>A.5.3 Extracting the local volatility surface</td>
<td>139</td>
</tr>
<tr>
<td>A.5.4 Time series study of estimated surfaces</td>
<td>142</td>
</tr>
<tr>
<td>A.6 Conclusions</td>
<td>146</td>
</tr>
<tr>
<td>Paper B - Non-Parametric Estimation of Stable Local Volatility Surfaces</td>
<td>151</td>
</tr>
<tr>
<td>B.1 Introduction</td>
<td>154</td>
</tr>
<tr>
<td>B.2 Static no-arbitrage constraints for the local volatility surface</td>
<td>157</td>
</tr>
<tr>
<td>B.3 Formulation of the estimation problem</td>
<td>159</td>
</tr>
<tr>
<td>B.3.1 Continuous formulation of the estimation problem</td>
<td>160</td>
</tr>
<tr>
<td>B.3.2 Derivatives prices as functions of the implied RND</td>
<td>162</td>
</tr>
</tbody>
</table>
CONTENTS

C.4.1 Scenario generation ............................................. 222
C.4.2 The SP model for the hedging problem ....................... 227
C.4.3 The hedge based on a Taylor expansion ...................... 230
C.5 Empirical results .................................................... 231
C.6 Concluding remarks ................................................. 234

Paper D - Recovering the Real-World Density and Liquidity Premia from Option Data 239

D.1 Introduction .......................................................... 241
D.2 Main results .......................................................... 243
D.2.1 Empirical estimation under the risk-neutral measure ....... 243
D.2.2 Option implied MMM parameters under the real-world measure ......................................................... 247
D.2.3 Time series of option implied MMM parameters .......... 250
D.2.4 Discovering the liquidity premium in option prices ....... 257
D.3 General modeling framework ....................................... 259
D.3.1 Real-world pricing under the benchmark approach ...... 260
D.3.2 Pricing of S&P500 index derivatives under the benchmark approach ......................................................... 264
D.3.3 The generalized Minimal Market Model ..................... 266
D.4 Estimation technique ................................................. 269
D.4.1 RND recovery ....................................................... 270
D.4.2 Recovering the real-world density ............................ 273
D.4.3 Simulated Maximum Likelihood ............................... 274
D.5 Conclusions ........................................................... 278
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

Paper E - Option Market Prediction of the S&P 500
Index Return Distribution 283

E.1 Introduction ............................................. 285
E.2 Recovery of the option implied real-world density ............. 287
  E.2.1 Estimation of the risk-neutral and real-world densities ... 288
  E.2.2 Transformation to the option implied MMM density ....... 294
E.3 Historical equity market models ............................. 297
  E.3.1 GARCH models .................................. 297
  E.3.2 Maximum Likelihood estimation of the GARCH models . 299
  E.3.3 Stochastic volatility models ........................ 300
  E.3.4 MCMC estimation of the stochastic volatility models .... 302
E.4 Density forecast evaluation framework ........................ 304
  E.4.1 Evaluations based on the log-likelihood .................. 304
  E.4.2 Evaluations based on probability integral transform ...... 307
E.5 Evaluation of the option implied density forecasts ............. 309
  E.5.1 Simultaneous evaluation of option implied volatility esti-
  mates and stochastic volatility model specifications .......... 314
  E.5.2 Evaluation of the option implied MMM densities ........... 320
E.6 Conclusions .............................................. 324

Appendix

A-1 Description of the S&P 500 index option data ................. 328
A-2 Derivation of posterior distributions of parameters and latent vari-
  ables in the Log SVLE model ............................. 329
  A-2.1 Choosing the prior distributions of the model parameters . 330
A-2.2 Posterior distribution of $\mu$ ............................ 330
A-2.3 Posterior distribution of $\kappa$ ........................... 332
A-2.4 Posterior distribution of $\theta$ ........................... 334
A-2.5  Posterior distributions of $\sigma_h$ and $\rho$ ........................................... 336
A-2.6  Posterior distributions of $\{h_t\}_{t=1}^{T}$ ............................................. 341

A-3  The MCMC algorithm for the Log SVLE model .......................................... 346

A-4  Derivation of posterior distributions of parameters and latent variables in the Log SVJLE model ................................................................. 351
A-4.1  Choosing the prior distributions for the jump specific model parameters ................................................................. 353
A-4.2  Posterior distribution for the latent jump times $\{q_t\}_{t=2}^{T}$ ........ 353
A-4.3  Posterior distribution for the latent jump sizes $\{\xi_t\}_{t=2}^{T}$ .... 354
A-4.4  Posterior distribution of the jump intensity $\lambda_y$ ................. 355
A-4.5  Posterior distribution of the mean of the jumps $\mu_y$ ............ 356
A-4.6  Posterior distribution for the variance of the jumps $\sigma_y^2$ ........ 357

A-5  The MCMC algorithm for the Log SVJLE model ...................................... 358
Introduction

Decision makers in the financial markets are continuously faced with the problem of making decisions under uncertainty. From the perspective of a bank, making decisions under uncertainty could e.g. refer to how to optimally hedge a portfolio of options, which risk factors to be exposed to in order to achieve a high return relative to the risk that is associated with the exposure, or the level of risk that an operation such as equity derivatives trading should be allowed to be exposed to. In order to be able to make good decisions, it is necessary to have an accurate approximation of the multivariate distribution of the relevant risk factors. It is important that the distribution of the risk factors take rare extreme events into account, since these events can potentially have a large impact on the outcome from the decisions. Given that we only have access to a limited observed history of market prices, such events may not be present in the available history on which the models are built. Evidence from financial history shows that obtaining an adequate description of relevant risk factors is in most cases a very challenging task. There are many examples where an insufficient understanding and modelling of relevant risk factors has led to catastrophic consequences. One such example is the collapse of Lehman Brothers on 15 September 2008, which is widely seen as the trigger for the Great Recession. The Lehman Brothers collapse was caused by large losses in the bank’s portfolio of credit derivatives linked to the US housing market. There are of course many reasons why the bank’s exposure towards the US housing market were allowed to get so large, but it is safe to say that had the banks internal risk models captured the true risks associated with the exposure, the collapse could likely have been avoided.

One way to gain understanding of the characteristics of relevant market risk factors is to use historically observed asset price returns in order to estimate the
multivariate probability distribution of those factors. Since the asset price returns can be directly observed on the market, the estimation is for this case performed under the so-called physical, or real-world, measure $\mathbb{P}$. There is however also an alternative way to extract information regarding the distribution of asset prices. Breeden and Litzenberger (1978) showed that it is possible to extract the probability distribution function of future prices of the underlying asset under the so-called risk-neutral measure $\mathbb{Q}$ from observed European option prices. Since derivatives are priced under the risk-neutral measure, also called the pricing measure, the implied density can then be used in order to price illiquid or exotic derivatives consistently with the prices of the standardized options. Investors, central banks, and policy makers also use information implied in derivatives prices in order to analyze market participants’ aggregate beliefs about future movements in the price of the underlying assets of those derivatives. For example, the Bank of England estimates and publishes risk-neutral densities (RNDs) of a number of financial assets and commodities on a daily basis. As is pointed out in Bank of England (2015), these RNDs do not necessarily provide the actual, or physical, probabilities of an asset price realizing particular values in the future. One major difference between the RND and the physical density is that the expected values implied by the two densities differ. Another difference is that RNDs are more negatively skewed than physical densities.

When making decisions about how to e.g. optimally hedge a certain exposure or which risks to gain exposure to in order to achieve a high risk-adjusted expected return, we would of course need to have an accurate view of the distribution of future asset prices under the physical measure $\mathbb{P}$. The RND can be transformed into the physical density, where the transformation parameters are estimated from a set of RNDs together with realized values of the prices or index levels that are predicted by the densities. Remarkable recent research has shown that it is possible to also extract information of the distribution of the underlying asset under the physical measure using only observed option prices as input to the estimation (see Heath and Platen, 2006; Carr and Yu, 2012; Ross, 2013). This means that the derivatives markets can provide us with forward looking information of the distribution of future asset prices under both the risk-neutral measure $\mathbb{Q}$ and the physical measure $\mathbb{P}$, and this information can then be used in order to make decisions on the financial markets. There are numerous applications for estimated densities. They can be used to assess market beliefs about political and economic events, price exotic derivatives, measure risk or to estimate risk preferences.

Even though it is possible to extract information of aggregate market expectations of future asset prices directly from the derivatives market, it is a difficult task to separate the noise present in market prices from the true information. The noise
stems from market micro structure effects such as the bid-ask spread, discrete price ticks, non-synchronized trading or misprinted prices. The overall themes for this dissertation are the extraction of implied information embedded in equity index option prices, and how to use this information in order to be able to make optimal decisions in the equity index option markets. These two themes provide the necessary background in order to answer the overall research question (ORQ) for the dissertation, which can be formulated as

ORQ: How can a general framework for making optimal decisions in the equity index derivatives market be developed?

Practitioners in the derivatives markets make trading decisions under uncertainty on a regular basis. This is a difficult problem since there is a large number of traded assets in the derivatives markets, and the observed prices of these assets are affected by noise. In addition, there are significant transaction costs that cannot be disregarded. Thus, there is a need for robust tools that can aid practitioners to make good trading decisions in the derivatives markets. The framework should be able to handle arbitrary distributions for the relevant risk factors as well as taking into account the significant transaction costs present in the equity index derivatives markets. We use stochastic programming (SP) in order to model the decision making, since SP is well suited to handle the important characteristics that need to be captured when modeling the derivatives market. The SP model will only require as input the prices of the assets today, and a description of these prices at some future date, e.g. the next trading day. The quality of the solution to the stochastic program will crucially depend on that the generated scenarios that are used as input to the problem accurately describe the distribution of the relevant risk factors (see e.g. King and Wallace, 2012). Thus, the generated scenarios need to capture both the expected returns and the uncertainty of the index level as well as for a collection of options written on the index. A given set of option prices has to satisfy a number of different no-arbitrage constraints in order to preclude arbitrage in the model. Thus we need to be careful and take this into account in order to not introduce arbitrage in the model when we generate scenarios. An optimization algorithm which solves the stochastic program would immediately realize this, since this would give an unbounded optimal objective function value.

The importance of generating realistic scenarios for the relevant risk factors in the equity derivatives markets leads to the first research question for this dissertation.

RQ1: Which properties are important to capture in order to obtain optimal decisions in the equity index derivatives markets?
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

One general concept which is important for optimal decisions is the risks in the equity index derivatives markets. At each point in time when markets are open we can observe quoted derivatives prices, but they contain a lot of noise which makes it difficult to create a reasonable model of the true risks in market prices and not to a large extent capture the noise. The risks for an option portfolio are complex and can therefore not adequately be captured by one-factor stochastic volatility models. The model that we use for capturing the risks in the equity index derivatives markets will therefore be based on surfaces that are implied by a set of observed option prices. We will thus need to have a good method for estimating option implied surfaces that can be used in order to accurately approximate the true risks for equity index options. There are many possible ways to approach the problem of estimating option implied surfaces and there are numerous methods proposed in the literature. Apart from capturing the risks in the derivatives markets it is for many applications also important to include estimates of expected option returns in the model. In order to obtain an accurate estimate of the expected option returns, we also need to have a good understanding and estimation of the risk premia associated with relevant risk factors.

When we have a good static description of the relevant risk factors, the next question that arises is then how to build a realistic dynamic model for those risk factors. In order to obtain decisions of high quality it is vital to accurately approximate the true dynamics of the relevant risk factors. This leads to the second research question of this dissertation

RQ2: How should a dynamic model for the equity index derivatives markets be built?

The dynamic model should describe how the underlying index and a collection of options on that index simultaneously evolves under the physical probability measure \( P \). For pricing and risk measurement of complex option portfolios, possibly involving exotic options, we will also need to specify this model under the risk-neutral measure \( Q \). The research community has, through recently published papers (Heath and Platen, 2006; Carr and Yu, 2012; Ross, 2013), started to gain a better understanding of what observed market prices of options implies for the underlying asset under the physical probability measure. This means that the information embedded in option prices can be used in order to infer the distribution for the index under the physical measure for future points in time, and this information can in turn be used when building a dynamic model for the index option markets.

A common method used in the literature for building a dynamic model of deriva-
tives markets is to assume that the underlying index follows a one-factor stochastic volatility model such as e.g. the Heston model (Heston, 1993). This model will then implicitly describe the simultaneous dynamics of derivatives prices and the underlying index. The dynamics in the equity index options markets are far too complex to be adequately captured by a one-factor stochastic volatility model (see e.g. Christoffersen et al., 2009). The question then arises how the dynamics of the existing risk factors for equity index options should be described? We can start by noting that there exist a large number of derivatives traded in the market and that the prices of these derivatives are tightly linked through a dependence structure. If we have access to a time series of snapshots of option implied surfaces we can use principal component analysis (PCA) in order to describe the historical dynamics of the surface. Since market data contains a lot of noise it is important to have an estimation method for the option implied surface that is able to separate out the noise in order to estimate surfaces of high quality. Otherwise the PCA will identify risk factors that do not exist and only stem from the noise in the input data.

It is also important to make sure that the generated scenarios do not contain any arbitrage opportunities. Therefore, it is important to build a dynamic model which does not give rise to arbitrage in the generated scenarios. This will have important implications for which surface we choose to model when we build a dynamic model. The most common approach described in the literature is to build a dynamic model for the Black-Scholes (Black and Scholes, 1973) implied volatility surface (IVS). The static no-arbitrage conditions for the IVS are however very intricate which makes it difficult to build a dynamic model for the IVS that avoids arbitrage in the created scenarios. On the contrary, the static no-arbitrage conditions for the local volatility surface (LVS) are much simpler which makes the LVS a more suitable surface to use in a dynamic model of the option market. The LVS is however much more sensitive to noise in market data, which places high demands on the method we use for estimation of option implied surfaces. When building a dynamic model for the equity index derivatives market we also need to have accurate estimates of the time-dependent risk premia, since the optimal decisions for most problems to a large extent will be governed by expected values under the physical probability measure $\mathbb{P}$.

To be able to build a realistic dynamic model simultaneously for an underlying index and options written on that index, we first need to have a method that is able to estimate option implied surfaces of high quality. This especially means that the estimation method should produce surfaces that are stable over time which is a difficult problem. In order to be able to develop a method for estimation of option implied surfaces with desired properties, we first need to determine the important
properties that we must handle in the estimation. This leads to the third research question of the dissertation which can be formulated as:

RQ3: Which properties are essential when estimating option implied surfaces?

The estimation problem is complicated by the fact that market prices contain noise and possibly also arbitrage. Options are priced under the risk-neutral measure $\mathbb{Q}$. Breeden and Litzenberger (1978) showed that the risk-neutral density, $q$, of the underlying index, $S$, at time $T$, is proportional to the following second derivative of call option prices, $C(K,T)$,

$$
q(S_T) \sim \left. \frac{\partial^2 C(K,T)}{\partial K^2} \right|_{K=S_T}.
$$

(1.1)

This means that if market data contains a lot of noise, this noise will, through the second derivative in equation (1.1), be amplified for the option implied surface. The large amount of noise present in market data is visible e.g. when inspecting many of the examples from the literature of estimated local volatility surfaces. Market prices can also contain arbitrage which further complicates the estimation. As we have established earlier, in order to build a dynamic model for the option market we need to have access to arbitrage-free option implied surfaces of high quality. Determining the essential properties for the estimation gives us the necessary input to how to formulate the estimation problem in order to handle the identified difficulties. Thus, the answer to RQ3 serves as input to the fourth research question for this dissertation which we formulate as:

RQ4: How should option implied surfaces be estimated?

There are many examples of methods for estimation of option implied surfaces described in the literature, but most, if not all, of these methods are not able to produce estimated local volatility surfaces that are stable over time. As we have argued, it is vital to be able to estimate local volatility surfaces that are stable over time in order to be able to build a realistic dynamic model for the equity index options market. The estimation will provide snapshots of the state of the option market which will form the basis for the dynamic model, as well as directly give the current option prices each time the optimal decision problem is solved. In paper C of this dissertation we develop and implement a non-parametric estimation method which produces estimated surfaces of very high quality. The high quality of the estimated surfaces has enabled the research that has led to the last 3 papers of the dissertation. In paper C, we build and implement a dynamic model for the local volatility surface based on PCA. The implementation of the dynamic model
crucially relies on that we have estimated realistic local volatility surfaces that are stable over time. In papers D and E we use the estimated surfaces in order to estimate the option implied physical densities and liquidity premia embedded in equity index option prices. Also these results critically depends on the quality of the estimated surfaces that are used as input.

The information of the future index level distribution under the physical measure embedded in options can be used as input to the optimal decision problem. For many applications it is when solving the optimal decision problem important to adequately account for expected returns. Given that we have identified the relevant risk factors, we can include estimates of the expected returns for being exposed to those risk factors in the model. It will then be crucial to be able to estimate the distribution of the risk factors under the physical probability measure $\mathbb{P}$. Through the research on estimating option implied surfaces we have gained a very good understanding for the properties under the risk-neutral measure $\mathbb{Q}$. If we know the connection between the measure $\mathbb{Q}$ and $\mathbb{P}$, we can also obtain a very good understanding of these properties under $\mathbb{P}$. This leads to the fifth research question of the dissertation

RQ5: What does the information extracted under the risk-neutral measure $\mathbb{Q}$ imply under the physical measure $\mathbb{P}$?

The research community is only starting to understand how to extract information of the probability distribution under $\mathbb{P}$ using only observed option prices. This approach seems to offer a passable way forward in order to extend the model developed in paper C of this dissertation. When we have built up a better understanding for this approach it is also possible to build even better dynamic models of the equity index derivatives markets.

Figure 1.1 provides an overview of the research questions of the dissertation, as well as a simplified description of how the research questions are related. The figure can also be seen as an overview of the structure of my PhD project that has resulted in this dissertation. The overall research question was formulated early on in the project and has then gradually been broken down to the separate parts that can be summarized by the 5 research questions.

Now that the research questions of the dissertation have been formulated we will briefly outline the structure of the rest of the dissertation. The dissertation consists of two main parts - the first part provides an overview of the research areas covered by the dissertation, and the second part contains the five papers of the dissertation. In order to give the reader the sufficient background for the papers included in
this dissertation, and also to provide an overview of the relevant literature, this first part of the dissertation contains an exposition of the areas of the literature that are central for the dissertation.

The rest of the first part of the dissertation is organized as follows. Chapter 2 contains a short overview of three possible approaches for formulating the optimal decision problem that we consider for the equity index option market. This chapter also contains a motivation to why SP is suitable for this particular problem. In order to gain an understanding of the important properties that should be captured by our model, it is necessary to identify the empirical properties for both equity index options as well as the underlying equity market. These properties are summarized in Chapter 3 which also includes a short account of some of the risk premia that exist in the equity index option markets. In Chapter 4, the different surfaces that are implied by observed option prices are introduced, and it is shown how these surfaces are interconnected. This chapter also includes a full derivation of the necessary static no-arbitrage conditions for the local volatility surface. Dynamic models for option implied surfaces are considered in Chapter 5, and especially two different approaches for building dynamic models for the local volatility
surface. An important theme for the dissertation is how to extract information from observed option prices directly under the physical measure $\mathbb{P}$. Chapter 6 therefore contains a presentation of the benchmark pricing theory (Platen and Heath, 2006) of derivatives directly under the physical measure. This chapter also includes a derivation of the connection between call option prices and the physical PDF of the index for the case with deterministic dividends. The introductory part of the dissertation is then concluded in Chapter 7 with a short summary of the five papers in the dissertation pointing out the research contribution of each paper. This chapter also contains a summary of how the research questions of the dissertation are answered by the five papers. Finally, the second part of the dissertation contains the five appended papers.
Optimal decisions under uncertainty in the equity index derivatives markets

The overall research question for this dissertation was formulated in the previous chapter and is given by: *How can a general framework for making optimal decisions in the equity derivatives markets be developed?* Making optimal decisions in the equity index derivatives markets refers e.g. to how do we decide optimal hedging strategies for equity index option portfolios, or how should we make optimal investments on the equity index derivatives market. Irrespective of the method we choose for making these decisions we will need to have a mathematical description of the dynamics for how a collection of option prices as well as the underlying asset evolve with time under the physical measure $\mathbb{P}$. There are many possible approaches to modeling and solving optimal decision problems in the equity index derivatives markets. In paper C we develop a model for optimal hedging of an index option portfolio. When illustrating the advantages and disadvantages with different approaches for solving the optimal decision problem we will therefore study the problem from the perspective of optimal hedging of an option portfolio.

Traditional methods for the hedging problem described in the literature are e.g. delta-vega hedging or static hedging methods. One of the assumptions of the Black-Scholes model (Black and Scholes, 1973) is that an option can be perfectly replicated by continuous trading in the underlying asset and a risk-free bond. In practice we cannot perfectly replicate an option and the delta hedging strategy will lead to a hedging error. Since the value of a portfolio of options is very sensitive to the level of volatility, the delta hedging method can be expanded to a delta-vega
hedging method by also trading in options in order to hedge the vega exposure of the portfolio. One drawback with this method is that the delta-vega hedge need to be rebalanced frequently leading to large transaction costs. Another method when hedging exotic OTC options is to use a static hedging method. The static hedging method involves finding a buy-and-hold portfolio of exchange-traded options that aims to replicate the specific pay-off of the exotic option. In most cases though, it is not possible to find a static hedge that perfectly replicates the pay-off of the exotic option. Most methods that are described in the literature (see e.g. Hull, 2014) are based on a set of stochastic processes that describe how the underlying asset prices and prices of options on those assets evolve over time. When describing the simultaneous dynamics for market prices with the help of the standard models described in the literature one is faced with a dilemma – observed market prices will typically not evolve according to the model dynamics. In order to have a unified model it is required that the models that describe option price dynamics and the dynamics for the underlying asset are consistent. The most common extension to the Black-Scholes model described in the literature is to assume that the volatility of the underlying index follows a one-factor stochastic volatility model such as the Heston model (Heston, 1993). However, the true dynamics of a collection of option prices over time are far too complex in order to be adequately captured by a one-factor stochastic volatility model. This leads to the fact that traditional methods need to simplify the model dynamics so much that they cannot capture what is actually taking place on the markets.

A more general approach compared to the traditional methods described in the literature is to formulate the hedging problem as an optimization problem under uncertainty. The generic form for the optimization problem under uncertainty can be written as

$$\min_x \mathbb{E}[f(x)]$$

s.t. $x \in X.$

(2.1)

In problem (2.1), the expected value of $f$ is minimized given the decisions $x$ which have to belong to the feasible set $X$. There are different methods for solving the problem in (2.1) described in the literature. The choice of method depends on the properties of the specific problem instance. One alternative is to use dynamic programming where the focus normally is on long-term steady-state behaviour (see King and Wallace, 2012; Powell, 2012). The solution to the dynamic program is given in the form of a policy which maps states to decisions. Dynamic programming can be used if the dimension of the state space is sufficiently low so that it is possible to numerically solve for the policy, or alternatively find analytical solutions when possible. In order to be able to find analytical solutions it is however necessary to model the dynamics with a stochastic process which is
not too complex. If the model is simplified by e.g. assuming that the dynamics of the index level can be described by a geometric Brownian motion so that there exist analytical solutions to the subproblems in (2.1), then the problem can be solved with dynamic programming in discrete time (see e.g. Bertsekas, 1995). An advantage with dynamic programming is that it can be possible to find analytical solutions which can give valuable insights into the characteristics of the problem that we want to solve. The drawbacks are that we either have to use relatively few state variables, or use simplified models to describe the dynamics in order to be able to obtain analytical solutions. We want to solve (2.1) for the option market considering transaction costs. There exist a large number of options that can be traded even if there is only one underlying asset and hence the number of state variables will be large. Furthermore, the simultaneous dynamics of the index and the option implied surface cannot realistically be captured by a simplified process that allows analytical solutions. The combination of many state variables and the necessity to discretize the state pace makes the use of dynamic programming inappropriate for our problem due to the curse of dimensionality.

If the optimal decision problem (2.1) is solved with optimal control, it is necessary to specify stochastic processes that describe the market dynamics in continuous time. If we furthermore have a model that can be solved analytically, then we can use stochastic optimal control to find analytical solutions to (2.1) (see e.g. Bertsekas, 1995). If we want to have a realistic model for the market dynamics we need to specify a realistic continuous stochastic process for all risk factors that exist in the index option markets. Such a model will be far too complex in order to make it possible to solve the stochastic optimal control problem analytically. Since there are significant transaction costs in the index option markets the model needs to take the transaction costs into account. If we have a model with transaction costs and a more complicated model for the market dynamics, then (2.1) can be solved numerically with the help of stochastic optimal control. It is however complicated to include transaction costs in a stochastic optimal control model. Another issue with stochastic optimal control is that the solution is given as a continuous time controls which are typically of low dimensionality. Since there are many assets in the equity index option markets, the decision will typically be of high dimension making stochastic optimal control inappropriate as solution method in many cases.

A third alternative method to solve the problem (2.1) is to use stochastic programming (SP). By using SP to solve (2.1) we can have arbitrary models for the underlying asset price and derivatives prices. Furthermore there are no limitations in terms of how many options we can trade as was the case when solving (2.1) with dynamic programming or stochastic optimal control. SP is also well suited to handle transaction costs and other restrictions such as restrictions for shorting
and borrowing, as well as limitations for the risk level or for how much capital that is allowed to be tied in the options portfolio. This is achieved by discretizing time and thus we only determine the decision at a discrete set of time stages. SP can handle the restrictions with constraints in the optimization model and the only input that is needed is prices for the assets for each time stage. Thus, we have a high degree of freedom in choosing the distribution for the asset prices and are not restricted to specific models. SP solves (2.1) by only considering a set of scenarios. In Shapiro et al. (2009) it is shown that the problem (2.1) with a continuous sample space can be solved by discretizing the sample space. This leads to a so called deterministic equivalent problem which can be solved with standard methods from optimization. A drawback with using SP to solve (2.1) is that the scenario tree grows exponentially with the number of time steps, thus only a limited number of future decisions can be considered when the current decision is determined. Compared to dynamic programming and optimal control which give an optimal policy, which can be used for any state at any point in time, SP only gives the optimal decision right now. To get an optimal decision tomorrow the problem has to be resolved. SP can be solved efficiently by e.g. the method described in Blomvall and Lindberg (2002). Examples from the literature where SP has been applied to the problem of making optimal decisions in the equity derivatives markets are Blomvall and Lindberg (2003) and Gondzio et al. (2003).

As mentioned, the SP model will only require as input the prices of the assets today and a description of these prices at some future date e.g. the next trading day. As input to the SP model we thus need generated scenarios that describe the simultaneous distribution of the asset prices at the future date. This means that we need good methods for generating scenarios that capture the essential characteristics for the equity derivatives market and the market for the underlying asset. A given set of option prices has to satisfy certain no-arbitrage conditions. Thus, we need to be careful and take this into account in order to not introduce arbitrage in the model when we generate scenarios. Introducing arbitrage leads to an unbounded optimal objective function value which will be realized by an optimization algorithm which solves (2.1).

When building a framework for making optimal decisions in the equity index option markets we will use SP to solve the optimal decision problem. In order to obtain good decisions we need to use a dynamic model of all relevant risk factors that accurately captures the true market dynamics of the equity index option markets. As we have argued, stochastic optimal control is not well suited to handle complex models for the market dynamics. SP on the other hand can handle arbitrary distributions for the risk factors since the only input that is needed are prices for the assets for each time stage. There are a large number of assets that it is possible
to trade in on the equity index option markets. This means that the number of state variables is large for the problem that we want to solve, and as we have argued both dynamic programming and stochastic optimal control need to use relatively few state variables. As previously argued SP can handle the large number of state variables. SP, unlike stochastic optimal control, is also well suited to handle the significant transaction costs in the equity index option markets which cannot be disregarded. Given the characteristics of the equity index option markets and the problem we want to solve, we conclude that SP is best suited to solve the specific optimal decision problem that we consider.

SP is also intimately associated with the problem of pricing options which is illustrated by King (2002). There, the mathematical technique of conjugate duality is used to determine lower and upper boundaries that option prices must lie within in order to avoid arbitrage in incomplete markets. These boundaries should then also be included as constraints in an SP model involving options in an incomplete market.

We have argued why we use SP as solution method when building a framework for making optimal decisions in the equity index option markets. The optimal solution from the SP model will be a good decision only if we have accurately modeled the uncertainty in the equity index option markets. It is thus important that the modeled uncertainty accurately captures what has empirically been observed for the equity index option markets as well as the underlying equity markets. We will therefore in the next chapter go through empirically observed properties of both equity markets and associated equity index option markets.
Empirical properties of equity index options

To lay the foundation for the subsequent chapters of this introductory part of the dissertation, we will in this chapter list some empirical properties of both the underlying equity market and for equity index options. In order to be able to build a realistic dynamic model of the equity index option markets we need to capture the important empirical properties of both the underlying and the option market. When making optimal decisions in financial markets it is also vital to have a good understanding and modeling of empirically observed risk premia. We will therefore in this chapter also include a summary of findings in the literature regarding the volatility risk premium which affects expected returns of index options.

3.1 Empirical statistical properties of the underlying equity market returns

This section provides a summary of empirical statistical properties of equity returns which can also be found in Barkhagen (2013). The list of empirical statistical properties can be used as input for deciding which model to use for the underlying equity index when building a dynamic model for the equity index derivatives markets. Generally, time series of equity returns display a non-normal distribution which is negatively skewed and exhibits fat tails. A good overview of empirical properties of asset returns is given in Cont (2001) and most of the points on the list below are given in that paper. Other sources for the points on the list below
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

are Jondeau et al. (2007), Ghysels et al. (1996) and Bollerslev et al. (1994).

1. **Fat tails:** The unconditional distribution of returns (meaning that we do not take into account that volatility changes with time) displays fatter tails than that expected from a normal distribution. This means that if we use the normal distribution to model financial returns, we will underestimate the number and magnitude of large negative and positive returns. Leptokurtic distributions of financial return data have been observed since the early sixties, in e.g. Mandelbrot (1963) and Fama (1963, 1965) among others. Observation of the leptokurtic distribution of stock returns led to a large literature on modeling stock returns as i.i.d. draws from fat-tailed distributions (see e.g. Mandelbrot, 1963; Fama, 1963, 1965). There are a number of other fat-tailed parametric distributions proposed in the literature such as the Student’s t-distribution (see e.g. Blattberg and Gonnedes, 1974; Kon, 1984) and normal inverse Gaussian distributions (see e.g. Barndorff-Nielsen, 1997).

2. **Asymmetry:** Large negative returns are more frequent than large positive returns, and thus the unconditional distribution of returns is negatively skewed. The asymmetry and fat-tail phenomena persist even after adjusting for conditional heteroscedasticity (e.g. via GARCH-type models), meaning that the conditional distribution is also non-normal (see e.g. Jondeau et al., 2007).

3. **Volatility clustering:** Volatility of returns is serially correlated, suggesting that a large (positive or negative) return tends to be followed by another large return. There is a close relation between volatility clustering and fat tails of the distribution of equity returns. The fat tailedness is a static explanation whereas a key insight provided by ARCH models is a formal link between time varying (conditional) volatility behavior and (unconditional) heavy tails (see e.g. Ghysels et al., 1996). ARCH-type models, first introduced in Engle (1982), as well as stochastic volatility models are partly built to capture the volatility clustering. As stated in Ghysels et al. (1996), it is widely documented that ARCH effects disappear with temporal aggregation, where Drost and Nijman (1993) is given as an example.

4. **Jumps:** Stock prices display large sudden, discontinuous, and usually negative movements. There are several examples in the option pricing literature of models that aim at capturing jumps in equity returns (see e.g. Bates, 1996; Merton, 1976).

5. **Leverage effect:** Empirical evidence shows that volatility is negatively correlated with the returns, which was first mentioned in the literature by Black (1976). The term leverage effect stems from the fact that falling stock prices imply an increased leverage (debt to equity ratio) of firms, and it is believed that this results
Empirical statistical properties of the underlying equity market returns

in more uncertainty and hence volatility. As argued by Black and stated in Ghysels et al. (1996), the response of stock volatility to the direction of returns is too large to be explained by leverage alone however. Another, and perhaps more plausible, explanation of the leverage effect is that when there is a large negative return on the stock market, the fear of further losses creates a herd like behavior of market participants which in turn increases the market volatility.

6. Mean reversion of volatility: When volatility is disturbed, it tends to return to its normal level, which may itself vary over time (see e.g. Engle and Patton, 2001). This property is captured by most GARCH and stochastic volatility models.

7. Long memory and persistence of volatility: Most measures of volatility suggest that volatility is highly persistent. This property is closely related to the property of volatility clustering. As stated in Ghysels et al. (1996), Ding et al. (1993) studied the autocorrelations of $|r(t, t+1)|^c$ for positive values of $c$, where $r(t, t+1)$ is a one-period return. They found $|r(t, t+1)|^c$ to have quite strong correlations for long lags, while the strongest temporal dependence was for $c$ close to one. Similarly, the estimation of stochastic volatility models show similar patterns of persistence (see e.g. Jacquier et al., 1994).

8. Volume/volatility correlation: Trading volume is positively correlated with all measures of volatility. One example from the literature which documents the relationship between trading volume and volatility is Gallant et al. (1992).

9. Aggregated normality: As the time scale over which returns are calculated increases, the return distribution gets closer to the normal distribution (see e.g. Cont, 2001). Thus, e.g. weekly returns are closer to being normally distributed than daily returns.

10. Asymmetry in time scales: The shape of the return distribution is not the same at different time scales. In particular, coarse-grained (longer time scale) measures of volatility predict fine-scale (shorter time scale) volatility better than the other way around (see e.g. Cont, 2001).

11. Absence of linear autocorrelations of returns: Returns generally do not display significant linear autocorrelations, except for very small intraday time scales (less than 20 minutes) for which microstructure effects are present (see Cont, 2001). The absence of significant linear autocorrelations in asset returns have been widely documented (see e.g. Fama, 1971) and is often cited as support for the effi-
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

cient market hypothesis (Fama, 1991). If price changes exhibit significant negative linear autocorrelation, this correlation can be exploited by a simple strategy with positive expected returns. Such strategies will therefore tend to change correlations so that they become closer to zero except for very short time scales, which represents the time the market takes to react to new information (Cont, 2001).

In high-frequency return series of transaction prices, one actually observes negative autocorrelation at very short lags (typically, one or a few trades) which is traditionally attributed to the bid-ask bounce (Campbell et al., 1997). Transactions will typically take place either close to the ask or closer to the bid and tend to bounce between these two limits. However, one also observes negative autocorrelations at the first lag in bid or ask prices themselves, suggesting a fast mean reversion of the price at the tick level (Cont, 2001). This property can be explained by the action of a market maker (Goodhart and O’Hara, 1997).

12. Nontrading periods: Information that affect stock prices accumulates slower when markets are closed than when they are open (Bollerslev et al., 1994). Return volatilities tend to be higher following weekends and holidays compared to regular week days. In Fama (1965); French and Roll (1986) the authors found that although variances are higher following weekends and holidays than on other days, not nearly as high as would be expected if the news arrival rate were constant.

13. Forecastable events: Increased stock market volatility around dividend and earnings announcements, and macro-economic data releases is documented in numerous papers (Bollerslev et al., 1994). For example, in Cornell (1978) and Patell and Wolfson (1979, 1981) it is shown that individual firms’ stock return volatility is high around earnings announcements.

There are also important predictable changes in volatility across the trading day. For example, the volatility is typically much higher at the open and close of stock exchange trading than during the middle of the day (Bollerslev et al., 1994). This property is related to the characteristic that trading volume and volatility are positively correlated. The increase in volatility at the open can also be attributed to the fact that information accumulated while the market was closed.

14. Co-movements in volatilities: Volatilities within and across stock markets tend to move together in response to common underlying factors (Bollerslev et al., 1994). Typically, factor models are used to model the linkage of international volatilities, as in e.g. Engle and Susmel (1993).

15. Time-varying cross-correlation: Correlation between equity returns (as
Empirical properties of the equity index options market

well as with other asset classes) tends to increase during high-volatility periods, in particular during crashes (see e.g. Jondeau et al., 2007).

The models described in the literature for capturing the dynamics of equity returns are designed to capture some of the important empirical and statistical features which were presented above. These models can primarily be divided into two main types of models, namely GARCH-type models and stochastic volatility models. For the GARCH (Generalized Autoregressive Conditional Heteroscedasticity) models there is only one source of randomness which drives the log return process, and unlike for the stochastic volatility models GARCH models are only specified as processes in discrete time. For stochastic volatility models also the volatility process is driven by a source of randomness which will typically be negatively correlated with the innovations that drive log returns.

3.2 Empirical properties of the equity index options market

The 1987 stock market crash caused a large increase in the slope and curvature of the Black-Scholes implied volatility smirk for equity index options (see e.g. Rebonato, 2004). Prior to the 1987 crash the implied volatility curve for a given maturity was relatively flat, implying that the market participants believed that the Black-Scholes model with a constant volatility and a log-normally distributed index level was a fairly accurate description of the index dynamics. However, since the crash of 1987 the equity index option markets have persistently displayed a sharp negative slope for the implied volatility smirk. Since implied volatilities are merely a direct transformation of option prices to the space of implied volatilities, empirical properties of observed option prices can be illustrated by the empirical properties of the implied volatility surface. Below, we have summarized the most important empirical properties that are reported in the literature. The main sources for the summary below are Rebonato (2004), Cont et al. (2002), Fengler (2005) and Cont and da Fonseca (2002).

1. Non-flat profile of the implied volatility surface: Since the market crash of 1987, equity index option markets have displayed a non-flat profile of implied volatilities, both in the strike direction and in the time-to-maturity direction (Rebonato, 2004).
2. **The implied volatility smirk is more pronounced for short maturities:** For shorter maturity options the implied volatility smirk is much more pronounced than for options with longer time-to-maturity. The longer time-to-maturity the more shallow the smirk will be for that maturity (Rebonato, 2004; Fengler, 2005).

3. **The shape of the implied volatility surface changes over time:** The shape of the implied volatility surface is not constant but changes over time (Cont and da Fonseca, 2002; Cont *et al.*, 2002; Rebonato, 2004). The observed changes of the implied volatility surface are complicated, making it unlikely that the observed dynamics can be captured by a one-factor stochastic volatility model that only describes the randomness of the volatility of the underlying index.

4. **The implied volatility is lowest around ATM:** The implied volatility curve for a given maturity has a minimum which most of the time is located at a strike that is close to the current index level (Fengler, 2005).

5. **Asymmetry in the implied volatility curve:** The implied volatility smile is much more pronounced going from ATM in the direction of lower strikes than in the opposite direction. The smile is much less pronounced moving from ATM to higher strikes and may even be monotonically decreasing in that direction (Rebonato, 2004).

6. **Dependence between the slope of the smirk and the level of volatility:** The steepness of the slope of the implied volatility smirk tends to increase when there is high volatility in the market, i.e. during market turbulence (Rebonato, 2004). Other evidence from the literature suggests that the slope of the smirk changes little when volatility changes during normal market conditions (see e.g. Derman, 1999; Christoffersen *et al.*, 2009).

7. **The volatility of implied volatilities depends on strike and time-to-maturity:** The volatility of implied volatilities is higher for options with short time to maturity than for longer maturity options (Fengler, 2005). The volatility of implied volatility for options with low strikes will typically also be higher than for options close to ATM.

8. **Mean reversion of implied volatility:** Cont and da Fonseca (2002) and Fengler (2005) find that implied volatility exhibits mean reversion.

9. **Shocks to the implied volatility surface are highly correlated**
implied volatilities for different points on the implied volatility surface are highly correlated. This means that most of the dynamics of the implied volatility surface can be captured by a small number of components (Cont and da Fonseca, 2002; Fengler, 2005).

10. Negative correlation between the index level and the implied volatilities: Estimating the correlation between historical index returns and implied volatilities of individual options reveals a negative correlation (Fengler, 2005). This is consistent with the leverage effect of Black (1976) and with the observation that the estimated correlation parameter in a stochastic volatility models such as the Heston model is negative.

The empirical properties of the implied volatility surface will provide us with important guidelines when we build a framework for estimating realistic option implied surfaces. The empirically observed dynamic properties of the implied volatility surface will also be important to take into account when building a dynamic model for the equity index option markets. Now that we have listed some empirically observed properties of a collection of option prices we will in the next section illustrate the volatility risk premium through an example model.

3.3 Risk premia

As we saw in Chapter 3.2, the standard Black-Scholes model with constant volatility cannot explain what is empirically observed in the equity index option markets. The first attempts in the literature to take the observed volatility smirk into account assumed that the index level volatility is stochastic and hence is driven by its own source of randomness, which for most models is assumed to be correlated with the source of randomness for the index level. The extra source of randomness in a stochastic volatility model means that derivatives, contrary to the Black-Scholes model, cannot be perfectly hedged by taking positions in the underlying index future and a risk-less bond. The stochastic volatility gives rise to a volatility risk premium that will have an impact on expected returns of both naked and delta hedged index option positions.

In order to exemplify the emergence of the volatility risk premium in a stochastic volatility model, and investigate the effect of this premium on expected instantaneous excess returns of derivatives, we will use the Heston model as an illustration. In the Heston model the SDEs for the index level, $S_t$, and the stochastic variance
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

process, \( v_t \), under the physical probability measure \( \mathbb{P} \) are given by (see Heston, 1993)

\[
\begin{align*}
    dS_t &= \mu_t S_t dt + \sqrt{v_t} S_t dW_t^{P,s}, \\
    dv_t &= \kappa(\theta - v_t) dt + \sigma_v \sqrt{v_t} dW_t^{P,v},
\end{align*}
\]

where \( W_t^{P,s} \) and \( W_t^{P,v} \) are standard Wiener processes under \( \mathbb{P} \) and where

\[
E \left( dW_t^{P,s} dW_t^{P,v} \right) = \rho dt.
\]

From Heston (1993) we have that the pricing PDE that a derivative, \( f(t, S_t, v_t) \), under the Heston model have to fulfill is given by

\[
\begin{align*}
    \frac{\partial f}{\partial t} + \frac{1}{2} \sigma_v^2 v_t S_t^2 \frac{\partial^2 f}{\partial S^2} + \frac{1}{2} \sigma_v^2 v_t \frac{\partial^2 f}{\partial v^2} + \rho \sigma_v v_t S_t \frac{\partial^2 f}{\partial S \partial v} + (\kappa(\theta - v_t) - \lambda^v(t, S_t, v_t)) \frac{\partial f}{\partial v} + \\
    (r_t - \delta_t) S_t \frac{\partial f}{\partial S} - r_t f = 0,
\end{align*}
\]

where \( r_t \) and \( \delta_t \) are the instantaneous deterministic risk-free interest rates and dividend yield respectively, and where the term \( \lambda^v(t, S_t, v_t) \) is the volatility risk premium. For simplicity we have used the short hand notation \( f \) for the derivative price \( f(t, S_t, v_t) \). In the Heston model the volatility risk premium is assumed to be a linear function of the stochastic variance according to

\[
\lambda^v(t, S_t, v_t) = kv_t.
\]

The market price of equity risk is given by

\[
\Lambda^e_t = \frac{\mu_t - (r_t - \delta_t)}{\sqrt{v_t}} = \frac{\lambda^e_t}{\sqrt{v_t}}.
\]

where \( \lambda^e_t \) is the equity risk premium. Analogously we can define the market price of volatility risk according to (see e.g. Gatheral, 2006)

\[
\Lambda^v_t = \frac{\lambda^v_t(t, S_t, v_t)}{\sigma_v \sqrt{v_t}}.
\]

From Girsanov’s theorem we have that the processes \( W_t^{Q,s} \) and \( W_t^{Q,v} \) given by

\[
\begin{align*}
    W_t^{Q,s} &= W_t^{P,s} + \int_0^t \Lambda^e_u du, \quad \text{and} \\
    W_t^{Q,v} &= W_t^{P,v} + \int_0^t \Lambda^v_u du,
\end{align*}
\]

From Heston (1993) we have that the pricing PDE that a derivative, \( f(t, S_t, v_t) \), under the Heston model have to fulfill is given by

\[
\begin{align*}
    \frac{\partial f}{\partial t} + \frac{1}{2} \sigma_v^2 v_t S_t^2 \frac{\partial^2 f}{\partial S^2} + \frac{1}{2} \sigma_v^2 v_t \frac{\partial^2 f}{\partial v^2} + \rho \sigma_v v_t S_t \frac{\partial^2 f}{\partial S \partial v} + (\kappa(\theta - v_t) - \lambda^v(t, S_t, v_t)) \frac{\partial f}{\partial v} + \\
    (r_t - \delta_t) S_t \frac{\partial f}{\partial S} - r_t f = 0,
\end{align*}
\]

where \( r_t \) and \( \delta_t \) are the instantaneous deterministic risk-free interest rates and dividend yield respectively, and where the term \( \lambda^v(t, S_t, v_t) \) is the volatility risk premium. For simplicity we have used the short hand notation \( f \) for the derivative price \( f(t, S_t, v_t) \). In the Heston model the volatility risk premium is assumed to be a linear function of the stochastic variance according to

\[
\lambda^v(t, S_t, v_t) = kv_t.
\]

The market price of equity risk is given by

\[
\Lambda^e_t = \frac{\mu_t - (r_t - \delta_t)}{\sqrt{v_t}} = \frac{\lambda^e_t}{\sqrt{v_t}}.
\]

where \( \lambda^e_t \) is the equity risk premium. Analogously we can define the market price of volatility risk according to (see e.g. Gatheral, 2006)

\[
\Lambda^v_t = \frac{\lambda^v_t(t, S_t, v_t)}{\sigma_v \sqrt{v_t}}.
\]

From Girsanov’s theorem we have that the processes \( W_t^{Q,s} \) and \( W_t^{Q,v} \) given by

\[
\begin{align*}
    W_t^{Q,s} &= W_t^{P,s} + \int_0^t \Lambda^e_u du, \quad \text{and} \\
    W_t^{Q,v} &= W_t^{P,v} + \int_0^t \Lambda^v_u du,
\end{align*}
\]

24
Risk premia

are standard Wiener processes under the risk-neutral measure $Q$ (see e.g. Björk, 2011). Using these expressions for the $Q$-Wiener processes, we can express the SDEs for the index level and the stochastic variance in equation (3.1) according to

$$dS_t = (r_t - \delta_t)S_t dt + \sqrt{v_t}S_t dW_t^{Q,a},$$

$$dv_t = \kappa^Q(\theta^Q - v_t)dt + \sigma_v\sqrt{v_t}dW_t^{Q,v},$$

where the parameters in the drift term for the stochastic variance under the risk-neutral measure $Q$ are given by

$$\kappa^Q = \kappa + k,$$

and, $\theta^Q = \frac{\kappa \theta}{\kappa + k}.$

Now we will investigate how the equity and volatility risk premia affect expected instantaneous excess returns of options in the Heston model. A similar analysis can be found in Broadie et al. (2009) in a jump diffusion model setting. Using Itô’s lemma we have that the differential of the derivatives price, $f$, is given by

$$df = \left(\frac{\partial f}{\partial t} + \frac{1}{2}v_tS_t^2\frac{\partial^2 f}{\partial S^2} + \frac{1}{2}\sigma_v^2v_t\frac{\partial^2 f}{\partial v^2} + \rho \sigma_v v_t S_t \frac{\partial^2 f}{\partial S \partial v}\right) dt + \frac{\partial f}{\partial S} dS_t + \frac{\partial f}{\partial v} dv_t,$$

and using the SDEs for the index and the variance under the measure $\mathbb{P}$ we have that

$$df = \left(\frac{\partial f}{\partial t} + \frac{1}{2}v_tS_t^2\frac{\partial^2 f}{\partial S^2} + \frac{1}{2}\sigma_v^2v_t\frac{\partial^2 f}{\partial v^2} + \rho \sigma_v v_t S_t \frac{\partial^2 f}{\partial S \partial v}\right) dt +$$

$$\frac{\partial f}{\partial S} \left(\mu_t S_t dt + \sqrt{v_t}S_t dW_t^{P,a}\right) + \frac{\partial f}{\partial v} \left(\kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dW_t^{P,v}\right).$$

Using the shorthand notation, $\lambda_t^v$, to denote the volatility risk premium, we have from equation (3.4) that

$$\frac{\partial f}{\partial t} + \frac{1}{2}v_tS_t^2\frac{\partial^2 f}{\partial S^2} + \frac{1}{2}\sigma_v^2v_t\frac{\partial^2 f}{\partial v^2} + \rho \sigma_v v_t S_t \frac{\partial^2 f}{\partial S \partial v} =$$

$$- (\kappa(\theta - v_t) - \lambda_t^v) \frac{\partial f}{\partial v} - \left(r_t - \delta_t\right)S_t \frac{\partial f}{\partial S} + r_t f,$$

and plugging this into equation (3.15) we get

$$df = \frac{\partial f}{\partial S} \left(\mu_t S_t dt + \sqrt{v_t}S_t dW_t^{P,a}\right) + \frac{\partial f}{\partial v} \left(\kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dW_t^{P,v}\right) - \lambda_t^v \frac{\partial f}{\partial v} \left(\kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dW_t^{P,v}\right).$$
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

\[
\left( (\kappa(\theta - v_t) - \lambda_t^i) \frac{\partial f}{\partial v} + (r_t - \delta_t) S_t \frac{\partial f}{\partial S} - r_t f \right) dt. \tag{3.17}
\]

Cancelling terms in equation (3.17) yields the following expression for the differential

\[
df = \frac{\partial f}{\partial S} (\mu_t - r_t + \delta_t) dt + \frac{\partial f}{\partial v} \lambda_t^v dt + \frac{\partial f}{\partial S} \sqrt{v_t} S_t dW_t^{p,s} + \frac{1}{f} \frac{\partial f}{\partial v} \sigma_v \sqrt{v_t} dW_t^{p,v} + r_t f dt, \tag{3.18}
\]

which is rearranged into

\[
\frac{df}{f} - r_t dt = \frac{S_t}{f} \frac{\partial f}{\partial S} (\mu_t - r_t + \delta_t) dt + \frac{1}{f} \frac{\partial f}{\partial v} \lambda_t^v dt + \frac{S_t}{f} \frac{\partial f}{\partial S} \sqrt{v_t} dW_t^{p,s} + \frac{1}{f} \frac{\partial f}{\partial v} \sigma_v \sqrt{v_t} dW_t^{p,v}. \tag{3.19}
\]

Taking expectation under \( \mathbb{P} \), conditional on the information available at time \( t \), we have that the annualized expected instantaneous excess return of derivatives in the Heston model is given by

\[
\frac{1}{dt} \mathbb{E}^\mathbb{P} \left( \frac{df}{f} - r_t dt \middle| \mathcal{F}_t \right) = \frac{S_t}{f} \frac{\partial f}{\partial S} (\mu_t - r_t + \delta_t) + \frac{1}{f} \frac{\partial f}{\partial v} \lambda_t^v + \frac{S_t}{f} \frac{\partial f}{\partial S} \lambda_t^s + \frac{1}{f} \frac{\partial f}{\partial v} \lambda_t^v. \tag{3.20}
\]

As is apparent from equation (3.20) the expected instantaneous excess return is in the Heston model determined by one term that depends on the risk premia for the underlying equity index and one term for the stochastic variance. The risk premia are also scaled by the sensitivity of the derivative with respect to the risk factors. Since volatility is a constant in the Black-Scholes model, the second term in equation (3.20) disappears and the expected return of the derivative will only depend on the term involving the equity risk premium, i.e. we have that

\[
\frac{1}{dt} \mathbb{E}^\mathbb{P} \left( \frac{df}{f} - r_t dt \middle| \mathcal{F}_t \right) = \frac{S_t}{f} \frac{\partial f}{\partial S} \lambda_t^s. \tag{3.21}
\]

Since the equity risk premium is typically positive we have that European call options, since the delta of call options is positive, will have a positive expected excess return in the Black-Scholes model. By an analogous argument, since the delta of put options is negative, the expected excess return of put options is negative in the Black-Scholes model. Turning to the expected excess return of derivatives in the Heston model we have that the expected excess return in this setting also depends on the volatility risk premium. For equity index options the volatility risk
premium is in the literature found to be significantly negative (see e.g. Coval and Shumway, 2001; Gârleanu et al., 2009), implying that selling index options in the market is profitable on average. Since the vega is always positive for European call and put options, the second term in (3.20) will give a negative contribution to the expected return. Thus, taking the volatility risk premium into consideration, we have that the expected excess return of call options may even be negative, and makes the expected excess return of put options even more negative. If we include jumps in the underlying equity index level and in the stochastic volatility in the model, then also a jump risk premium will emerge in the expression for the expected excess return of derivatives (see Broadie et al., 2009).

Now that we have shown how the volatility risk premium affects the expected excess return of equity index options in the Heston model we will give a short overview of the findings in the literature regarding expected returns and embedded risk premia in index options. The literature review is by no means complete and we refer to e.g. Constantinides et al. (2013) or Gârleanu et al. (2009) for a more comprehensive review.

In Ilmanen (2011) it is demonstrated that selling equity index volatility has been profitable since the market crash of 1987, which can be explained by a negative volatility risk premium embedded in index options. The volatility risk premium can be approximated by the difference of implied volatility calculated from observed option prices and the realized volatility over the corresponding time period. According to Ilmanen (2011) the average of this difference is for S&P500 index options close to 4% over the time period January 1986 to December 2009. The volatility exposure can either be obtained through trading in variance swaps or delta hedged short straddles (a straddle consists of a long position in a put and a call option with the same strike).

The profitability of volatility selling strategies in the index options market is also supported by a number of empirical studies in the literature. Coval and Shumway (2001) found that selling zero delta ATM straddles for S&P500 options earned an average of 3% per week over the period January 1990 to October 1995, while selling straddles with moneyness between 85% and 90% earned weekly returns of 4.5%. Coval and Shumway (2001) also found that the average weekly returns of long positions in OTM puts was -14.6% over the same period whereas ATM puts had an average weekly return of -9.50%. For call options the average weekly returns were slightly positive with 1.85% for ATM options. In a more recent study Broadie et al. (2009) find that average monthly S&P500 put option returns over the period August 1987 to June 2005 is -56.8% for moneyness 0.94 and -29.9% for ATM options.
The overall evidence from the literature is that it is profitable to sell equity index volatility in the market and a number of possible explanations has been put forth for explaining this. Since options are in zero net supply in the market and end users are net long in index options, it means that market makers hold net short positions in index options. Since the sellers are providing insurance to the end users they should earn compensation for this which can explain some part of the negative volatility premium. Also, since options cannot be perfectly hedged due to trading at discrete points in time, stochastic volatility and jumps in the underlying asset and the volatility, market makers will charge a premium for exposure to these risks. Selling volatility has according to Ilmanen (2011) incurred large losses at mainly two occasions since 1987, namely the market crashes of 1987 and 2008. It can therefore be argued that selling volatility should earn a premium since large losses from this strategy tend to occur at the same time where the equity market has large drawdowns. Gårleanu et al. (2009) find through an empirical study using a demand based option pricing model that excess demand from end users can help explain the overall expensiveness of index options. Using their demand based model they find that one-third of the index option expensiveness can be explained by demand effects. In the literature there is also evidence that OTM puts are comparatively very expensive compared to e.g. ATM options, as manifested by the steep slope of the implied volatility smirk. Gårleanu et al. (2009) also find that end users are especially long OTM index puts, and that their demand based model can explain the relative expensiveness of OTM puts.
Option implied surfaces

In this chapter we will give an overview of the different surfaces that are implied by a collection of options, and that can be extracted with an appropriate method. The implied volatility surface has already been introduced when we listed the empirical properties of a collection of option prices in Chapter 3.2. The most obvious surface that is implied by a cross-section of options is the price surface. This surface could for example be the surface of European call prices on some specified domain in the strike and time-to-maturity plane. The Black-Scholes implied volatility surface is then simply obtained by inverting the Black-Scholes formula. Thus, the implied volatility surface only represents a direct transformation of the price surface. Under certain assumptions, it is also possible to infer the beliefs of market participants of the future states of the equity market through the implied risk-neutral density surface (RND). The RND surface contains information of the unconditional risk-neutral probabilities for different index levels in the future. It is however also possible to extract more local information from a collection of option prices. The local volatility surface contains information about the expected value, under the risk-neutral measure $\mathbb{Q}$, of the future stochastic variance conditional on different index levels in the future. Dupire (1994) showed how to extract the local volatility surface from a continuous surface of call option prices. In order for the algorithm to work one needs to start from a call price surface which is free of arbitrage, and where prices for all strikes and maturities on the domain have already been interpolated and extrapolated. Derman and Kani (1998) derived the following connection between the squared local volatilities, $\sigma_L^2(K,T)$, and the stochastic instantaneous variance

$$\sigma_L^2(K,T) = \mathbb{E}^Q (v_T | S_T = K, \mathcal{F}_t), \quad (4.1)$$
where \( \nu_T \) is the instantaneous stochastic variance at time \( T \), and where \( \mathcal{F}_t \) denotes the information available at time \( t \). From equation (4.1) we see that risk-neutral expectations of the stochastic variance are contained in the local volatility surface. Apart from the unconditional RND surface it is also possible to extract the conditional RND surface which contains local information of risk-neutral probabilities. Recent papers in the literature (Heath and Platen, 2006; Ross, 2013; Carr and Yu, 2012) have also shown that it is possible to extract the physical PDF from observed option prices. This information could potentially be very valuable when building a model for making optimal decisions in the equity index option markets. All of the surfaces are interconnected which means that given one surface, the other surfaces can be extracted by some numerical procedure. We will after the illustration below go through how we can transform one surface into another. The fact that all surfaces are interconnected means that we have full freedom in choosing which surface to estimate when we formulate the estimation problem. This means that we can focus on estimating a surface which will make it possible to formulate an optimization problem with good properties as we will see later.

We will now through a discretized example illustrate the connections between the unconditional RND surface and the two local surfaces – the local volatility surface and the conditional RND surface. There are two possible interpretations of local volatilities which we will explain in Chapter 4.2. In one of the two interpretations of the local volatility model (see e.g. Dupire, 1994), the risk-neutral dynamics of the underlying equity index is described by the SDE

\[
dS_t = (r_t - \delta_t)S_t dt + \sigma_L(t, S_t)S_t dW_t,
\]

where \( r_t \) is the deterministic instantaneous risk-free interest rate, \( \delta_t \) is the deterministic instantaneous dividend yield and where the local volatility, \( \sigma_L(t, S_t) \), is a deterministic function dependent on \( t \) and the index value, \( S_t \), at time \( t \). In order to gain intuition for the local volatility model we will use discretization in order to illustrate some properties of the model. Discretizing the SDE (4.2), yields the following difference equation

\[
S_{t_i} - S_{t_{i-1}} = (r_{t_{i-1}} - \delta_{t_{i-1}})S_{t_{i-1}} \Delta t + \sigma_L(t_{i-1}, S_{t_{i-1}})S_{t_{i-1}} \Delta W_{t_i},
\]

where \( \Delta W_{t_i} = W_{t_i} - W_{t_{i-1}} \sim N(0, \Delta t) \). Hence we have the following conditional distribution of \( S_{t_i} \) under the measure \( \mathbb{Q} \)

\[
S_{t_i} | S_{t_{i-1}} \sim N \left( S_{t_{i-1}} + (r_{t_{i-1}} - \delta_{t_{i-1}})S_{t_{i-1}} \Delta t, \sigma_L^2(t_{i-1}, S_{t_{i-1}})S_{t_{i-1}}^2 \Delta t \right).
\]

Given the conditional distribution arising from the discretized SDE, we interpret the local volatility model as giving rise to normally distributed conditional RNDs
which is illustrated in Figure 4.1. From the figure we see that the distribution of the index level in time step \( t_i \), given that the index level at time \( t_{i-1} \) is equal to \( S_j \), is given by the red distribution curve in the plot. For simplicity, the variances of the distributions in Figure 4.1 are illustrated as being equal. In reality the variances will typically be larger for points on the grid which are further away from the current index level. The true local volatility describing the dynamics of the index level will typically not be a deterministic function of the current time and index level, but will instead depend on some exogenous random source. Thus, the volatility is itself stochastic giving rise to a stochastic volatility model. As we saw in equation (4.1) the squared local volatility is the risk-neutral expectation of the future stochastic variance conditional on the index level at that time. Since volatility is stochastic, there should be one additional dimension in Figure 4.1, and the distributions in the figure can be seen as marginal distributions where the stochastic volatility has been integrated out.

Denoting the joint probability density function for the spot levels at time \( t_{i-1} \) and \( t_i \) by \( p(S_{t_i}, S_{t_{i-1}}) \), we have that

\[
p(S_{t_i}, S_{t_{i-1}}) = p(S_{t_{i-1}})p(S_{t_i}|S_{t_{i-1}}), \tag{4.5}
\]

where \( p(S_{t_{i-1}}) \) is the marginal density of the spot at time \( t_{i-1} \), and \( p(S_{t_i}|S_{t_{i-1}}) \) is the conditional density of the spot at time \( t_i \) given the value of the spot at time \( t_{i-1} \). Hence, we have that the marginal distribution of the spot at time \( t_i \) is given by

\[
p(S_{t_i}) = \int_0^\infty p(S_{t_{i-1}})p(S_{t_i}|S_{t_{i-1}})dS_{t_{i-1}}. \tag{4.6}
\]
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

The conditional density of $S_{t_i}$ given that the spot is $S_j$ at time $t_{i-1}$ can be discretized such that we obtain the following discrete set of probabilities

$$\left\{ q_{i,k}^{i-1,j} \right\}^{N+1}_{k=0}.$$  \hspace{1cm} (4.7)

Thus, $\left\{ q_{i,k}^{i-1,j} \right\}^{N+1}_{k=0}$ is the discrete version of the conditional RND given that the spot level is $S_j$ at time $t_{i-1}$. The unconditional RND value in node $(i,j)$ can then, analogously to the continuous case, be extracted from a collection of discrete local RNDs according to

$$q_{i,j} = \sum_{l=0}^{N+1} q_{i-1,l}^{i,j} q_{i-1,l},$$ \hspace{1cm} (4.8)

and the unconditional RND for time $t_i$ is given by

$$\left\{ q_{i,j} \right\}^{N+1}_{j=0}.$$ \hspace{1cm} (4.9)

Now that we have illustrated the connections between the local volatility surface and the conditional and unconditional RND surfaces we are ready to go through how one surface can be transformed into another. The interconnections between the different surfaces are illustrated in Figure 4.2. We will not explicitly go through the connections between all the surfaces, however given our description it is possible to transform every surface into the others either directly or indirectly. In the description above we have explained that the local volatility surface can be obtained from observed option prices through the method described in Dupire (1994). This method thus represents the arrow going from prices to local volatilities in connection (3) in Figure 4.2. Transforming local volatilities to prices can be achieved by numerical methods, but there does not exist any exact analytical expression in this case. There are however different analytical approximations that can be used. Obtaining European option prices from the unconditional RND can easily be achieved by a summation when the RND is discretized. This connection is represented by the arrow going from the RND to prices in connection (1) in the figure. For the opposite direction it is shown in e.g. Andersen and Brotherton-Ratcliffe (1997) that prices of Arrow-Debreau (A-D) securities can be extracted from a grid of call prices, and the prices of A-D securities are proportional to the RND. In Andersen and Brotherton-Ratcliffe (1997), the authors also derive a system of equations describing the relationship between the discretized local volatility surface and the discretized conditional RND surface. This paper also contains a derivation of the relationship between the discretized RND surface and the discretized local volatility surface which gives the connection (5).
discretized (unconditional) RND surface can be obtained from the discretized conditional RND surface by summation, which gives the arrow in one direction in connection (2).

Heath and Platen (2006) derive how the physical PDF can be extracted from the call price surface when the dynamics of the index is assumed to follow a local volatility model, and that the growth optimal portfolio (GOP) can be approximated by the index that the call options are written on. This connection is represented by the arrow going from the PDF to prices in connection (6). It is also possible to relate the physical PDF directly to the RND in this setting which is represented by connection (7) in the figure. Now that we have given an overview of how the different surfaces are connected we will present some properties of the two surfaces that are most commonly used as variable for the estimation methods used in the literature – the implied volatility surface and the local volatility surface. For the RND surface we provide a short overview of the estimation methods described in the literature in paper A of this dissertation. In Chapter 6 of this introductory part we will also provide a derivation of how the physical PDF can be backed out from observed option prices under specific assumptions.
4.1 The implied volatility surface

The implied volatility surface is the surface that is most commonly used by practitioners in the equity derivatives markets to extract information from observed option prices. Given the price of an European call or put option, the Black-Scholes implied volatility is defined as the parameter that when inserted in the Black-Scholes formula (Black and Scholes, 1973) recovers the option price. Hence, the implied volatility for a European call option with strike price $K$ and maturity date $T$, $\sigma(K,T)$, is the unique solution to the equation

$$C(K,T) = C^{BS}(K,T,\sigma(K,T)),$$  \hspace{1cm} (4.10)

where $C(K,T)$ is the market price of the option and where $C^{BS}(K,T,\sigma(K,T))$ is the price of the option given by the Black-Scholes formula. The implied volatility is suitable to use in order to express option prices, since it provides a metric under which option prices with different strikes, time-to-maturities, underlying assets and observation times can be compared. For these reasons implied volatilities are in the equity index option markets used to quote prices of options. The implied volatility surface is also the surface that is most commonly used for building dynamic models of option implied surfaces. Many results in the literature regarding properties of option implied surfaces are also expressed in terms of the implied volatility surface. In this section we will start to present the static no-arbitrage conditions that the call price function has to satisfy. Since implied volatilities represents a transformation of call option prices, these no-arbitrage conditions can be translated into no-arbitrage conditions for the implied volatility surface. We will then use these no-arbitrage constraints on the call price function as the starting point when we derive the static no-arbitrage conditions on the local volatility surface.

When building a framework for the estimation of realistic option implied surfaces important guide-lines regarding plausible shapes of the surface are given by theoretical asymptotic results from the literature. Most of the asymptotic results in the literature are reported for the implied volatility surface. We will therefore in this section also briefly present an important result from the literature regarding the slope of the implied volatility function for small and large strikes. We will then present some of the most commonly used methods described in the literature for estimating the implied volatility surface.
4.1.1 Static no-arbitrage conditions

In this section we will formulate the static no-arbitrage conditions that a call price function must satisfy. From put-call parity these constraints can be directly translated into no-arbitrage conditions for the put price function. The no-arbitrage conditions that we formulate for the call price function will then serve as a starting point when we derive the no-arbitrage conditions on the local volatility surface in Chapter 4.2.1. In order to formalize notation, let $S_t$ denote the level of an index with a continuous deterministic dividend yield $\delta_s$, where $s \in [t,T]$. Using the risk-neutral valuation of Cox and Ross (1976), the price of a European call option $C(K,S_t,t,T,r_s,\delta_s)$ with strike price $K$ is given by

$$C(K,S_t,t,T,r_s,\delta_s) = e^{-\int_t^T r_s ds} \int_K^\infty (S_T - K)q(S_T|S_t,t,T,r_s,\delta_s)dS_T,$$

(4.11)

where $T$ is the maturity of the option and $r_s$, where $s \in [t,T]$, is the deterministic risk free interest rate. In equation (4.11), $q(S_T|S_t,t,T,r_s,\delta_s)$ denotes the risk-neutral density (RND) of $S_T$ conditional on the information available at time $t$. Similarly, the price of an European put option $P(K,S_t,t,T,r_s,\delta_s)$ is given by

$$P(K,S_t,t,T,r_s,\delta_s) = e^{-\int_t^T r_s ds} \int_0^K (K - S_T)q(S_T|S_t,t,T,r_s,\delta_s)dS_T.$$

(4.12)

Since $r_t$ and $\delta_t$ are assumed to be deterministic the futures price and the forward price coincide and is given by

$$F_{t,T} = \mathbb{E}^Q[S_T] = S_t e^{\int_t^T (r_s - \delta_s) ds}.$$

(4.13)

From now on we will use the shorthand notation $C(K,T)$ to denote call option prices, and $q(S_T)$ to denote the RND. Differentiating equation (4.11) once with respect to the strike price yields

$$\frac{\partial C(K,T)}{\partial K} = -e^{-\int_t^T r_s ds} \int_K^\infty q(S_T)dS_T.$$

(4.14)

Since the RND is nonnegative we have that $0 \leq \int_K^\infty q(S_T)dS_T \leq 1$, which leads to the well-known bounds on the first derivative

$$-e^{-\int_t^T r_s ds} \leq \frac{\partial C(K,T)}{\partial K} \leq 0,$$

(4.15)

which implies monotonicity of the call price function with respect to the strike price.
Differentiating once more with respect to $K$ yields (as was first shown in Breeden and Litzenberger, 1978)

$$
\frac{\partial^2 C(K,T)}{\partial K^2} = e^{-\int_t^T r_sds}q(K) \geq 0, \quad (4.16)
$$

and hence the call option price is a convex function with respect to the strike price. From equation (4.16) we can recover the risk-neutral density from the call price function according to

$$
q(S_T) = e^{\int_t^T r_sds} \frac{\partial^2 C(K,T)}{\partial K^2} \bigg|_{K=S_T}. \quad (4.17)
$$

General no-arbitrage arguments give the following lower and upper bounds on call option prices

$$
\max(S_te^{-\int_t^T \delta_sds} - Ke^{-\int_t^T r_sds}, 0) \leq C(K,T) \leq S_te^{-\int_t^T \delta_sds}. \quad (4.18)
$$

If the call price function satisfies equations (4.15), (4.16) and (4.18) for a given time to maturity, $\tau = T - t$, then there exists no arbitrage opportunities for this maturity.

In the time-to-maturity direction it is in general not possible to state sufficient conditions which prevent calendar arbitrage. It is however possible to formulate a necessary condition which is given by

$$
e^{\int_t^{T_2} r_sds}C(K_2,T_2) \geq \frac{e^{\int_t^{T_1} r_sds}C(K_1,T_1)}{K_2}, \quad \text{where} \quad \frac{K_2}{F_{t,T_2}} = \frac{K_1}{F_{t,T_1}}. \quad (4.19)
$$

If we let $K/F_{t,\tau}$ denote the forward moneyness, where $F_{t,\tau}$ is the futures price with time to maturity $\tau$, we see that the condition that precludes calendar arbitrage is defined for a constant forward moneyness. The condition in equation (4.19) translates into the requirement that total implied variance defined by $\sigma_{BS}^2(x, \tau)\tau$, where $\sigma_{BS}$ is the Black-Scholes implied volatility, must be a non-decreasing function of $\tau$ for a constant forward moneyness $x$ (see Fengler (2009) for a proof).

### 4.1.2 Asymptotic results

In this section we will formulate two important asymptotic results for the slope of the implied volatility smile for extreme strikes derived by Lee (2004) and summarized in Lee (2005). The results are model independent which makes them
The implied volatility surface

particularly important. The results are formulated in terms of the forward log-
moneyness which is defined as

\[ x = \ln \left( \frac{K}{F_T} \right), \]  

(4.20)

where \( K \) is the strike price and where \( F_T \) is the forward price of the index for
maturity date \( T \). The first result derived by Lee (2004) states that for any given
expiry date \( T \), the implied volatility function for sufficiently small or large strikes
cannot grow faster than \( \sqrt{|x|} \) due to the strict monotonicity of call and put options.
Lee derives the following bound on the slope of the implied volatility function in
the right tail, i.e. for large strikes

\[ \sigma(x, T) \leq \frac{\sqrt{2x}}{T}, \]  

(4.21)

for some sufficiently large \( x > x^* \). If this condition is not satisfied call option
prices will not be monotonically decreasing for increasing strikes \( K \). There is a
similar result for the slope of the implied volatility function in the left tail. In the
left tail, i.e. for low strikes, for any \( \beta > 2 \) there exists \( x^* \) such that, for all \( x < x^* \),

\[ \sigma(x, T) \leq \sqrt{\frac{\beta |x|}{T}}. \]  

(4.22)

For \( \beta = 2 \), the same conclusion holds, if and only if \( S_T \) satisfies \( \mathbb{P}(S_T = 0) < 1/2 \). If
this condition is not satisfied put option prices will not be monotonically decreasing
for decreasing strikes \( K \). This result has implications when choosing the functional
form for a parametrization of the implied volatility smile. Specifically the results
in equations (4.21) and (4.22) restricts the slope of the implied volatility function
in the tails to not grow faster than \( \sqrt{|x|} \).

Lee (2004) also proves another result, named the moment formula, which replaces
the coefficients in equations (4.21) and (4.22) by a real number determined by \( \tilde{p} \)
in the right-hand tail, and \( \tilde{q} \) in the left-hand tail, where

\[ \tilde{p} = \sup\{ p : \mathbb{E}^Q(S_T^{1+p}) < \infty \}, \]  

(4.23)

\[ \tilde{q} = \sup\{ q : \mathbb{E}^Q(S_T^{-q}) < \infty \}. \]  

(4.24)

Thus, the coefficient in the right tail is, by abuse of language, determined by the "number" of finite moments of \( S_T \), whereas the coefficient in the left tail is determined by the "number" of finite moments of \( S_T^{-1} \). The coefficients in the left
and right tail is then in Lee (2004) showed to be given by

\[ \beta_R(T) = \limsup_{x \to \infty} \frac{\sigma^2(x, T)}{|x|/T}, \]  

(4.25)
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

\[ \beta_L(T) = \lim_{x \to -\infty} \sup \frac{\sigma^2(x, T)}{|x|/T}. \quad (4.26) \]

Lee (2004) shows that \( \beta_R, \beta_L \in [0, 2] \), and further proves that

\[ \hat{p} = \frac{1}{2\beta_R} + \frac{\beta_R}{8} - \frac{1}{2}, \quad (4.27) \]
\[ \hat{q} = \frac{1}{2\beta_L} + \frac{\beta_L}{8} - \frac{1}{2}. \quad (4.28) \]

Equivalently,

\[ \beta_R(T) = 2 - 4(\sqrt{\hat{p}^2 + \hat{p}} - \hat{p}), \quad (4.29) \]
\[ \beta_L(T) = 2 - 4(\sqrt{\hat{q}^2 + \hat{q}} - \hat{q}), \quad (4.30) \]

where the right-hand expressions are to be read as zero, if \( \hat{p}, \hat{q} \) are infinity. As stated in Lee (2005) one can think of \( \beta_R \) and \( \beta_L \) as the right-hand and absolute left-hand slopes of the linear asymptotes to the implied variance \( \sigma^2(x, T) \). As discussed in Lee (2005) some option pricing models admits explicit calculation of \( \hat{q} \) and \( \hat{p} \) in terms of the model parameters. Thus, in these cases the moment formula can be used directly to infer some of the model parameters. Therefore the moment formula can be used in order to provide an accurate guess for the starting values for the model parameters in these cases. Intuitively the result provided by the moment formula ties the asymptotic behaviour of the smile to the tail behaviour of the risk-neutral density. As formulated in Fengler (2005) the results in Lee (2005) provides important guidelines for how to extrapolate the smile function outside the region where we have observed option quotes. From equations \((4.21)\) and \((4.22)\) we have that the smile function should not grow faster than \( \sqrt{|x|} \). On the other hand the smile function should also not grow slower than \( \sqrt{|x|} \) unless we assume that \( S_T \) and \( S_T^{-1} \) have finite moments of all orders since then we have that \( \beta_R(T) = \beta_L(T) = 0 \).

There is a growing literature providing asymptotic results for the implied volatility surface in different model settings. We will not attempt to summarize these results here, but rather point out that these results can provide important guidelines when estimating the implied volatility surface under different model settings.

### 4.1.3 Estimation methods described in the literature

In this section we will go through some of the estimation methods commonly used in order to extract the implied volatility surface. Since all surfaces that we described in the beginning of this chapter are interconnected, we are referring to
The implied volatility surface

an explicit extraction of the implied volatility surface and not indirect estimation through some of the other surfaces. A common non-direct approach used by practitioners as well as academics is to fit a stochastic volatility model, such as the Heston model, to observed option prices and then infer the IVS from the option prices implied by these fitted Heston parameters. This approach can also be used to extract e.g. the implied RND surface or the local volatility surface.

The most commonly used method described in the literature for extracting the IVS is based on direct parametrizations of the implied volatility surface. A parametrization referred to as Practitioners Black-Scholes (PBS) in Christoffersen and Jacobs (2003) and Christoffersen et al. (2009) is given by the following polynomial in the strike $K$ and time-to-maturity $T - t$,

$$\sigma(K, T - t) = \theta_0 + \theta_1 K + \theta_2 K^2 + \theta_3 (T - t) + \theta_4 (T - t)^2 + \theta_5 K (T - t). \quad (4.31)$$

This parametrization was first used by Dumas et al. (1998) who concluded that the PBS produced lower out-of-sample pricing errors compared to the local volatility function models they consider in their paper. The PBS has been extensively used in the literature as a benchmark model when evaluating the performance in terms of out-of-sample pricing for competing models (see e.g. Dumas et al., 1998; Christoffersen and Jacobs, 2003; Christoffersen et al., 2009; Orosi, 2010). The model parameters of the PBS can be estimated with the help of ordinary least squares and since the implied volatility function is linear in the model parameters the resulting optimization problem is convex. However, the resulting implied volatility surface cannot be guaranteed to be arbitrage-free without explicitly adding no-arbitrage constraints on the parameters when solving the estimation problem. Adding no-arbitrage constraints on the parameters to the estimation problem would lead to a non-convex optimization problem. The convexity condition on call prices can be shown (see Brunner and Hafner, 2003) to imply the following no-arbitrage condition on the implied volatility function for a given maturity $T$ and strike price $K$

$$F_T \sqrt{T - t} \varphi \left( d_1(\bar{K}, T) \right) \left( \frac{1}{K^2 \sigma(\bar{K}, T)} + \frac{2d_1(\bar{K}, t)}{K \sigma(\bar{K}, t) \sqrt{T - t}} \frac{\partial \sigma(\bar{K}, T)}{\partial \bar{K}} + \right.$$

$$\left. \frac{d_1(\bar{K}, T) d_2(\bar{K}, T)}{\sigma(\bar{K}, T)} \left( \frac{\partial \sigma(\bar{K}, T)}{\partial \bar{K}} \right)^2 + \frac{\partial^2 \sigma(\bar{K}, T)}{\partial \bar{K}^2} \right) \geq 0, \quad (4.32)$$

where $F_T$ is the futures price for maturity $T$ and $\varphi(\cdot)$ is the standard normal density. In equation (4.32) $d_1(\bar{K}, T)$ and $d_2(\bar{K}, T)$ are given by

$$d_1(\bar{K}, T) = \frac{\ln(S_t / \bar{K}) + (r - \delta + \frac{1}{2} \sigma^2(\bar{K}, T))(T - t)}{\sigma(\bar{K}, T) \sqrt{T - t}}, \quad (4.33)$$
where $S_t$ is the current index level, and $r$ and $\delta$ is the constant interest rate and dividend yield respectively. Similarly the monotonicity condition on call prices can be transformed into conditions for the implied volatility function, and the no-arbitrage condition in the time-to-maturity direction demands that the total implied variance, $\sigma^2(K,T)(T-t)$, is a non-decreasing function of $T-t$ for a constant forward moneyness

$$x = \frac{K}{F_T},$$

(4.35)

As is apparent from the expression in equation (4.32) the static no-arbitrage conditions on the implied volatility surface are very involved. This is one of the reasons that we will not use the IVS as variable when we build a dynamic model for the index option market, since it will be difficult to generate scenarios for the IVS in the model which do not contain static arbitrage.

Another parametrization that is commonly used in the literature and by practitioners is the stochastic volatility inspired (SVI) parametrization (Gatheral, 2006). The parametrization is defined for each maturity, $T$, separately and is given in terms of the log-forward moneyness, $x = \ln(K/F_T)$, by

$$\sigma^2(x) = a + b \left( \rho(x - m) + \sqrt{(x - m)^2 + \gamma^2} \right),$$

(4.36)

where the model parameters depend on the expiration and $b > 1$, $\gamma > 0$ and $|\rho < 1|$. No-arbitrage in the strike direction in terms of the SVI parameters can be derived from equation (4.32). According to Gatheral (2006) one appealing property of this parametrization is that it is relatively easy to eliminate arbitrage in the time-to-maturity direction. This is accomplished by fitting the functional form (4.36) to all maturities simultaneously, with one set of model parameters for each maturity, and then interpolate the total implied variance in the time-to-maturity direction. According to Gatheral and Jacquier (2011) the SVI parametrization is commonly used by practitioners for mainly two reasons. These are the relative ease of ensuring no-arbitrage in the time-to-maturity direction, and that the behaviour in the wings as $|x| \to \infty$ is consistent with the moment formula of Lee (2004). However, as was the case for the PBS parametrization, estimating the SVI parameters from observed option prices leads to a non-convex optimization problem with several local optima. This means that estimated parameters will likely vary a lot when estimated over time, making it problematic to build a dynamic model based on SVI.
The local volatility surface

The PBS and SVI parametrizations are two of the most commonly used models for the IVS in the literature, but they represent only two examples from a vast literature. Other approaches are non-parametric estimation such as the non-parametric Nadaraya-Watson estimator which is used by e.g. Ait-Sahalia and Lo (2000); Cont and da Fonseca (2002); Benko et al. (2007). We will however not attempt to give a full overview here, but instead turn our attention to the local volatility surface.

4.2 The local volatility surface

The local volatility model was independently developed in three different papers (Dupire, 1994; Derman and Kani, 1994; Rubinstein, 1994) that appeared in the literature at around the same time. Contrary to the IVS the local volatility surface (LVS) is not directly observable from option prices, which makes the estimation much more difficult. Despite not being observable, there are several strong reasons to choose the LVS to simultaneously describe a collection of option prices when building a framework for making optimal decisions in the equity derivatives market. Since the LVS contains more local information than e.g. the IVS or the RND surface, it is much easier to determine if the LVS is realistic than it is for the IVS or the RND surface. The reason for this is that for e.g. the IVS the noise in the market data will be integrated out, making it difficult to determine if a particular IVS implies a realistic LVS. Furthermore, as we will see, it is much easier to ensure no-arbitrage for the LVS in the time-to-maturity direction than for the price, IV or RND surfaces, which will have important implications when we estimate the option implied surface as well as when we are building a dynamic model. To conclude, the two main reasons for choosing the LVS to simultaneously describe a collection of option prices is that it is a suitable variable to choose when estimating option implied surfaces, and that it is the most appropriate surface to use when building a dynamic model for the equity index option market.

There are two interpretations of local volatility which both assume the following SDE for the evolution of the index under the physical measure $\mathbb{P}$

$$dS_t = \mu_t S_t dt + \tilde{\sigma}_t S_t dW^P_t,$$

where $\tilde{\sigma}_t$ is the instantaneous volatility that may be driven by another exogenous stochastic source than $W^P_t$. Derman and Kani (1998) showed that

$$\mathbb{E}^Q (\tilde{\sigma}_T^2 | S_T = K, \mathcal{F}_t) = \frac{\partial C}{\partial T} + \left( r_T - \delta_T \right) K \frac{\partial C}{\partial K} + \delta_T C$$

$$\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}$$

(4.38)
where the squared local volatility is defined as

$$\sigma^2_{K,T}(t, S_t) = E^Q (\tilde{\sigma}^2_T | S_T = K, F_t).$$

(4.39)

Hence, the squared local volatility is defined as the expectation, under the risk-neutral measure $Q$, of the instantaneous variance at time $T$, conditional on that the index level $S_T$ is $K$ at time $T$, and on the information at time $t$. This means that the instantaneous volatility and the local volatility coincide for the point $(t, S_t)$, i.e. we have that

$$\tilde{\sigma}_t = \sigma_{S_t,t}(t, S_t).$$

(4.40)

For this interpretation of the local volatility, the local volatility surface changes stochastically over time since it is dependent on the information available at each point in time. If we assume that the instantaneous volatility, $\tilde{\sigma}_t$, only is a function of the time, $t$, and the index level, $S_t$, i.e. $\tilde{\sigma}_t = \tilde{\sigma}(t, S_t)$, and thereby does not depend on any additional exogenous random source, we have from equation (4.39) that

$$\sigma^2_{K,T}(t, S_t) = E^Q (\tilde{\sigma}^2(T, S_T) | S_T = K, F_t) = \tilde{\sigma}^2(T, K).$$

(4.41)

This means, since the squared local volatility in this case is independent of the information available at time $t$, that the local volatility surface is static over time. Derman and Kani (1998) calls this model an effective theory where all sources of randomness, except the part that stem from the index level, have been integrated out. In the effective theory for local volatility the dynamics of the index under the physical measure $P$ is then given by

$$dS_t = \mu_t S_t dt + \tilde{\sigma}(t, S_t) S_t dW^P_t.$$  

(4.42)

The difference between the two interpretations of local volatility is hence determined by how the surface evolves over time. For the effective theory the surface is static over time, whereas the surface evolves stochastically over time for the general theory. Pricing of European call options using the PDE in equation (4.38) yields identical prices for the both interpretations of local volatility. This means that each time the local volatility surface is calibrated to observed option prices the surface is identical for both interpretations.

In the illustrative example given in the beginning if this section we used the interpretation according to the effective theory in order to gain some intuition for the local volatility model. When we derive the static no-arbitrage conditions for the local volatility surface in Chapter 4.2.1, we will also start from the definition given by the effective theory. Since the local volatility surfaces momentarily are identical...
The local volatility surface

for both interpretations, it means that the static no-arbitrage conditions are the same for both the effective and the general theory. Now that we have introduced the local volatility model and the two interpretations of the model we will derive the static no-arbitrage conditions that must be satisfied by the local volatility surface. The derivation of the static no-arbitrage conditions for the local volatility surface is to my knowledge not collected anywhere in the literature, which is why the full derivation is included in Chapter 4.2.1.

4.2.1 Static no-arbitrage conditions

The no-arbitrage conditions on call option prices are summarized in Chapter (4.1.1). We will now transform the no-arbitrage conditions on call prices into conditions that the local volatility surface has to satisfy in order to be free of arbitrage. In the effective theory for local volatility the diffusion coefficient \( \tilde{\sigma}(t, S_t) \) which is consistent with observed European option prices is a function of time \( t \) and the spot value \( S_t \). Given the risk-neutral density function \( q(S_T) \), Dupire (1994) shows that there exists a unique risk-neutral diffusion process which generates this distribution. The diffusion process under the risk-neutral measure \( Q \) which is consistent with observed European option prices is given by

\[
dS_t = (r_t - \delta_t) S_t dt + \tilde{\sigma}(t, S_t) S_t dW_t,
\]

(4.43)

where the instantaneous interest rate \( r_t \) and the instantaneous dividend yield \( \delta_t \) are assumed to be deterministic functions, and where \( \tilde{\sigma}(t, S_t) \) denotes the local volatility. Starting from this diffusion process and the fact that the risk-neutral density must satisfy the Kolmogorov forward equation, the Dupire equation for the local volatility can be derived and is given by

\[
\tilde{\sigma}^2(T, K) = \frac{\frac{\partial C}{\partial T} + (r_T - \delta_T) K \frac{\partial C}{\partial K} + \delta_T C}{\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}}.
\]

(4.44)

The Dupire equation was originally derived in Dupire (1994) for the case where interest rates and dividends are assumed to be zero. From now on we will drop the tilde symbol in the local volatility term and, unless otherwise stated, let \( \sigma \) denote local volatility.

In order to state the necessary no-arbitrage conditions on the local volatility surface we need to establish conditions which state when there exist a unique solution to
the SDE given in equation (4.43). Using the notation in Björk (2011), we want to
determine when there exist a stochastic process \( X_t \) that satisfies the SDE
\[
\begin{align*}
    dX_t &= \bar{\mu}(t, X_t)\,dt + \bar{\sigma}(t, X_t)\,dW_t, \\
    X_0 &= x_0.
\end{align*}
\] (4.45) (4.46)

The theorem which states the conditions for existence and uniqueness of a solution
to the SDE (4.45) is given in Björk (2011) and below. We note that there exist
alternative ways to formulate the existence and uniqueness conditions and that
the conditions given below is one way to do this.

**Theorem 4.2.1. (Existence and uniqueness of a solution to the SDE)**

Suppose that there exists a constant \( K \) such that the following conditions are sat-
satisfied for all \( x, y \) and \( t \)
\[
\begin{align*}
    \|\bar{\mu}(t, x) - \bar{\mu}(t, y)\| &\leq K\|x - y\|, \quad (4.47) \\
    \|\bar{\sigma}(t, x) - \bar{\sigma}(t, y)\| &\leq K\|x - y\|, \quad (4.48) \\
    \|\bar{\mu}(t, x)\| + \|\bar{\sigma}(t, x)\| &\leq K(1 + \|x\|), \quad (4.49)
\end{align*}
\]

where \( \|\cdot\| \) denotes the Euclidean norm. Then there exists a unique solution to the
SDE (4.45). The solution has the following properties

1. \( X \) is \( \mathcal{F}^W_t \)-adapted.
2. \( X \) has continuous trajectories.
3. \( X \) is a Markov process.
4. There exist a constant \( C \) such that \( \mathbb{E}\left[\|X_t^2\|\right] \leq Ce^{Ct}(1 + \|x_0\|^2) \).

Now, since the solution \( X_t \) is a stochastic (Markov) process, the corresponding
probability density function must be a proper density function. This follows since a
stochastic process is a collection of random variables which represents the evolution
of some random variable over time. A random variable is defined on a probability
space \( (\Omega, \mathcal{F}, \mathbb{P}) \), where \( \mathbb{P} \) assigns probabilities to the events given in the set \( \mathcal{F} \). From Kolmogorov’s axioms we have that the probability of an event must be a
real number between zero and one. The definition of a probability density function
\( f_X(x) \) is such that
\[
\int_a^b f_X(x)\,dx = \mathbb{P}(X \in [a, b]),
\] (4.50)
for any interval \([a, b]\). Since probabilities cannot be negative we must have that
\[
\int_a^b f_X(x)\,dx \geq 0,
\] (4.51)
for any interval \([a, b]\). The integral in equation (4.51) is non-negative for all intervals \([a, b]\) only if the integrand function itself is non-negative, i.e. \(f_X(x) \geq 0\) for all \(x\). We also have that the probability of a sure event must equal one and hence we have that
\[
\int_{-\infty}^{\infty} f_X(x) \, dx = \mathbb{P} (X \in (-\infty, \infty)) = 1.
\] (4.52)

Since under the risk-neutral measure the probability \(Q(S_T < 0) = 0\), we have the following properties for the risk-neutral density function
\[
q(S_T) \geq 0 \quad \forall \ T, \ S_T,
\] (4.53)
\[
\int_0^\infty q(S_T) dS_T = 1.
\] (4.54)

The equations (4.47) and (4.48) in Theorem 4.2.1 state that the drift and diffusion coefficients of the SDE must be Lipschitz continuous. In our case we have that the drift and diffusion coefficients are given by
\[
\bar{\mu}(\cdot, \cdot) = (r_t - \delta_t) S_t,
\] (4.55)
\[
\bar{\sigma}(\cdot, \cdot) = \sigma(t, S_t) S_t,
\] (4.56)
and hence we demand that the local volatility surface satisfies the condition that \(\sigma(t, S_t) S_t\) is Lipschitz continuous.

In order to show that the local volatility surface is free of arbitrage, we need to show that the option prices implied by the local volatility surface satisfy the bounds given in equations (4.15) and (4.18)
\[
- e^{-\int_t^T r_s \, ds} \leq \frac{\partial C}{\partial K} \leq 0, \quad \text{and}
\] (4.57)
\[
\max \left( S_t e^{-\int_t^T \delta_s \, ds} - K e^{-\int_t^T r_s \, ds}, 0 \right) \leq C \leq S_t e^{-\int_t^T \delta_s \, ds}.
\] (4.58)

In the following we will assume that \(\sigma(t, S_t) S_t\) is Lipschitz continuous and hence that \(q(S_T) \geq 0\) for all \(S_T\). We have from equation (4.14) that the derivative of the call price with respect to \(K\) is given by
\[
\frac{\partial C}{\partial K} = - e^{-\int_t^T r_s \, ds} \int_{K}^{\infty} q(S_T) dS_T,
\] (4.59)
and since the risk-neutral density is positive for all \(S_T\) and \(e^{-\int_t^T r_s \, ds} \geq 0\) we have from equation (4.54) that the bounds in equation (4.57) are satisfied. For the upper bound in the equation (4.58) we have that
\[
S_t e^{-\int_t^T \delta_s \, ds} = e^{-\int_t^T r_s \, ds} F_{t,T} \overset{(4.13)}{=} e^{-\int_t^T r_s \, ds} \mathbb{E}^Q[S_T],
\] (4.60)
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

and

\[ C = e^{-\int_t^T r_s ds} \mathbb{E}^Q[(S_T - K)^+] = e^{-\int_t^T r_s ds} \int_K^\infty (S_T - K) q(S_T) dS_T \leq \]
\[ e^{-\int_t^T r_s ds} \int_K^\infty S_T q(S_T) dS_T \leq e^{-\int_t^T r_s ds} \mathbb{E}^Q[S_T] = e^{-\int_t^T r_s ds} F_{t,T} = e^{-\int_t^T \delta_s ds} S_t, \tag{4.61} \]

since \( q(S_T) \geq 0 \) for all \( S_T \) and \( S_T \geq 0 \).

For the lower bound in equation (4.58) we have that since \( q(S_T) \geq 0 \) for all \( S_T \) and \( e^{-\int_t^T r_s ds} \geq 0 \),

\[ C = e^{-\int_t^T r_s ds} \mathbb{E}^Q[(S_T - K)^+] = e^{-\int_t^T r_s ds} \int_K^\infty (S_T - K)^+ q(S_T) dS_T \tag{4.62} \]

will always be greater than zero. For the lower bound we also have that

\[ S_t e^{\int_t^T \delta_s ds} - K e^{-\int_t^T r_s ds} = e^{-\int_t^T r_s ds} (F_{t,T} - K) = e^{-\int_t^T r_s ds} \mathbb{E}^Q[(S_t - K)] = e^{-\int_t^T r_s ds} \int_0^\infty (S_T - K) q(S_T) dS_T \leq e^{-\int_t^T r_s ds} \int_K^\infty (S_T - K) q(S_T) dS_T = C, \tag{4.63} \]

and hence we have that

\[ S_t e^{\int_t^T \delta_s ds} - K e^{-\int_t^T r_s ds} \leq C, \tag{4.64} \]

and thus the lower bound in equation (4.58) is satisfied.

From equation (4.19) we have that in order to satisfy the necessary no-arbitrage in the time-to-maturity direction option prices with different maturities must satisfy the following inequality

\[ \frac{e^{\int_{t_2}^{T} r_s ds} C(K_2, T_2)}{K_2} \geq \frac{e^{\int_{t_1}^{T} r_s ds} C(K_1, T_1)}{K_1}, \tag{4.65} \]

where

\[ \frac{K_1}{F_{t,T_1}} = \frac{K_2}{F_{t,T_2}} = x. \tag{4.66} \]

Setting \( K_2 = K_1 F_{t,T_2} / F_{t,T_1} \) in equation (4.65) and simplifying yields the inequality

\[ \frac{e^{\int_{t_2}^{T} r_s ds} C(K_2, T_2)}{F_{t,T_2}} \geq \frac{e^{\int_{t_1}^{T} r_s ds} C(K_1, T_1)}{F_{t,T_1}}. \tag{4.67} \]
The local volatility surface

The quotients in equation (4.67) are for general \( K \) and \( T \) given by

\[
\frac{e^{\int_t^T r_s ds} C(K, T)}{F_{t,T}} = \mathbb{E}^Q \left[ \left( \frac{K}{F_{t,T}} - x \right)^+ \right] = \mathbb{E}^Q \left[ \left( \frac{S_T - K}{F_{t,T}} - x \right)^+ \right] = \hat{C}(x, T), \quad \text{where}
\]

\[
x = \frac{K}{F_{t,T}} .
\]  

(4.68)

(4.69)

Hence, the necessary no-arbitrage conditions in the time-to-maturity direction in terms of \( \hat{C}(x, T) \) are given by

\[
\hat{C}(x, T_2) \geq \hat{C}(x, T_1), \quad \text{where} \quad T_2 > T_1
\]

\[
\iff \hat{C}(x, T_2) - \hat{C}(x, T_1) \geq 0 \iff \frac{d\hat{C}}{dT} \geq 0, \quad \text{for} \ T_2 \to T_1 ,
\]

(4.70)

where \( \frac{d\hat{C}}{dT} \) denotes the total derivative of \( \hat{C} \) with respect to \( T \).

Given that the no-arbitrage conditions in the time-to-maturity direction can be conveniently formulated in terms of the modified prices \( \hat{C}(x, T) \), we want to express the equivalence of the Dupire equation in terms of \( \hat{C}(x, T) \). In order to derive the result we note that we can write \( \hat{C} \) as

\[
\hat{C}(x, T) = \frac{\int_t^T r_s ds C(K(x, T), T)}{F_{t,T}} = \frac{\int_t^T \delta_s ds C(K(x, T), T)}{S_t}, \quad \text{where} \quad K(x, T) = xF_{t,T} .
\]

(4.71)

Taking the partial derivative of equation (4.71) once with respect to \( x \) yields, since the only direct dependence on \( x \) is through the function \( K(x, T) \),

\[
\frac{\partial \hat{C}}{\partial x} = \frac{\partial \int_t^T \delta_s ds C(K(x, T), T)}{\partial x} = \frac{\int_t^T \delta_s ds \partial C}{S_t} \frac{\partial K}{\partial x} ,
\]

(4.72)

and since

\[
\frac{\partial K}{\partial x} = \frac{\partial}{\partial x} (xF_{t,T}) = F_{t,T}
\]

(4.73)

we have that

\[
\frac{\partial \hat{C}}{\partial x} = \frac{\int_t^T \delta_s ds}{S_t} F_{t,T} \frac{\partial C}{\partial K} = \frac{\int_t^T \delta_s ds}{F_{t,T}} F_{t,T} \frac{\partial C}{\partial K} = \frac{\int_t^T r_s ds}{\partial K} \frac{\partial C}{\partial K} .
\]

(4.74)
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

Differentiating once more with respect to $x$ gives

$$\frac{\partial^2 \hat{C}}{\partial x^2} = \frac{\partial}{\partial x} \left( \frac{\partial \hat{C}}{\partial x} \right) = \frac{\partial}{\partial x} \left( e^{\int_t^T r_s ds} \frac{\partial C}{\partial K} \right) \frac{\partial K}{\partial x} = e^{\int_t^T r_s ds} \frac{\partial^2 C}{\partial K^2} F_{t,T},$$

and hence we have that

$$\frac{1}{2} x^2 \frac{\partial^2 \hat{C}}{\partial x^2} = \frac{1}{2} K^2 e^{\int_t^T r_s ds} F_{t,T} \frac{\partial^2 C}{\partial K^2} = \frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}. \quad (4.75)$$

From equation (4.76) we deduce the implication

$$\frac{\partial^2 C}{\partial K^2} > 0 \iff \frac{\partial^2 \hat{C}}{\partial x^2} > 0, \quad (4.77)$$

since $e^{\int_t^T r_s ds}, F_{t,T}$ and $K \geq 0$.

For the total derivative in the time-to-maturity direction we again use that we can write $\hat{C}$ as

$$\hat{C} = e^{\int_t^T r_s ds} C(K,T) = e^{\int_t^T r_s ds} C(K,T). \quad (4.78)$$

The total derivative of the RHS of equation (4.78) is given by

$$\frac{d}{dT} \left( e^{\int_t^T r_s ds} C(K,T) \right) = \frac{\partial}{\partial T} \left( e^{\int_t^T r_s ds} C(K,T) \right) \frac{dT}{dT} + \frac{\partial}{\partial K} \left( e^{\int_t^T r_s ds} C(K,T) \right) \frac{dK}{dT} = \left( \frac{\delta_T}{S_t} e^{\int_t^T r_s ds} C(K,T) + \frac{\delta_T}{S_t} \frac{\partial C}{\partial T} \right) + \frac{\delta_T}{S_t} \frac{\partial C}{\partial K} \frac{dK}{dT}, \quad (4.79)$$

and hence the total derivative of $\hat{C}$ with respect to $T$ is given by

$$\frac{d\hat{C}}{dT} = \frac{e^{\int_t^T r_s ds}}{F_{t,T}} \left( \delta_T C(K,T) + \frac{\partial C}{\partial T} + \frac{\partial C}{\partial K} \frac{dK}{dT} \right). \quad (4.80)$$

We also have that since $K = x S_t e^{\int_t^T (r_s - \delta_s) ds}$, the $T$-derivative of $K$ is given by

$$\frac{dK}{dT} = (r_T - \delta_T) x S_t e^{\int_t^T (r_s - \delta_s) ds} = (r_T - \delta_T) K, \quad (4.81)$$

and hence the total $T$-derivative of $\hat{C}$ is given by

$$\frac{d\hat{C}}{dT} = \frac{e^{\int_t^T r_s ds}}{F_{t,T}} \left( \delta_T C(K,T) + \frac{\partial C}{\partial T} + (r_T - \delta_T) K \frac{\partial C}{\partial K} \right). \quad (4.82)$$
The local volatility surface

We are now ready to formulate the equivalence of the Dupire equation in terms of the modified prices $\hat{C}(x, T)$

\[
\frac{d\hat{C}}{dT} = \frac{e^{\int_t^T r_s ds}}{F_{t,T}} \left( \frac{\delta T C(K, T) + \frac{\partial C}{\partial T} + (r_T - \delta_T)K \frac{\partial C}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}} \right) = \frac{e^{\int_t^T r_s ds}}{F_{t,T}} \frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2} \delta_T C(K, T) + \frac{\partial C}{\partial T} + (r_T - \delta_T)K \frac{\partial C}{\partial K} \right) \overset{(4.44)}{=} \sigma^2(K, T). \tag{4.83}
\]

Now, since the SDE describing the evolution of the index level in the local volatility model is given by

\[
dS_t = (r_t - \delta_t)S_t dt + \sigma(t, S_t)S_t dW_t, \tag{4.84}
\]

we have from Theorem 4.2.1 that if $\sigma(t, S_t)S_t$ is Lipschitz continuous, then the risk-neutral densities given from equation (4.17)

\[
q(S_T) = e^{\int_t^T r_s ds} \frac{\partial^2 C}{\partial K^2} \bigg|_{K=S_T}, \tag{4.85}
\]

are positive for all $K$ and $T$ and hence that

\[
\frac{\partial^2 C}{\partial K^2} \geq 0. \tag{4.86}
\]

We have from equation (4.70) that the local volatility surface is free of calendar arbitrage if

\[
\frac{d\hat{C}}{dT} \geq 0. \tag{4.87}
\]

Theorem 4.2.1 implies that if $\sigma(t, S_t)S_t$ is Lipschitz continuous then $q(S_T) \geq 0$ for all $S_T$ which from equation (4.85) implies that

\[
\frac{\partial^2 C}{\partial K^2} \geq 0. \tag{4.88}
\]

Now we have from equation (4.83) that

\[
\frac{d\hat{C}}{dT} = \frac{e^{\int_t^T r_s ds}}{F_{t,T}} \frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2} \sigma^2(K, T), \tag{4.89}
\]
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

and since \( \frac{e^{\int_t^T r_s ds}}{F_{t,T}} \frac{1}{2} K^2 \geq 0 \) we have that if Theorem 4.2.1 holds then

\[
\frac{dC}{dT} \geq 0 \iff \sigma^2(K, T) \geq 0, \quad (4.90)
\]

and thus the necessary condition to avoid calendar arbitrage given in equation (4.70) is satisfied. We are now ready to formulate a theorem which states a necessary condition for avoiding calendar arbitrage for the local volatility surface.

\textbf{Theorem 4.2.2. (Necessary condition for absence of calendar arbitrage)}

A necessary conditions for the local volatility surface to be free of calendar arbitrage is that \( \sigma(t, S_t)S_t \) is Lipschitz continuous and that the squared local volatility is positive for all \( K \) and \( T \), i.e. that

\[
\sigma^2(K, T) \geq 0. \quad (4.91)
\]

Since Lipschitz continuity of \( \sigma(t, S_t)S_t \) implies that \( q(S_T) \geq 0 \) for all \( S_T \), which in turn implies that the no-arbitrage conditions in equations (4.15), (4.16) and (4.18) are satisfied, the following theorem, which states the necessary conditions for the local volatility surface to be free of arbitrage, immediately follows.

\textbf{Theorem 4.2.3. (Necessary condition for an arbitrage-free local volatility surface)}

A necessary condition for the local volatility surface to be free of arbitrage is that the squared local volatility is positive for all \( K \) and \( T \), i.e.

\[
\sigma^2(K, T) \geq 0, \quad (4.92)
\]

and that \( \sigma(t, S_t)S_t \) is Lipschitz continuous.

In the framework for estimation of local volatility surfaces that we develop in paper B, the Lipschitz continuity condition is satisfied by penalizing large second derivatives of the squared local volatility surface. Now that we have derived necessary no-arbitrage conditions on the local volatility surface, we we present some estimation methods described in the literature.
4.2.2 Estimation methods described in the literature

Estimating local volatility surfaces from a set of observed option prices is an inverse problem which is typically ill-posed. A well-posed problem is a problem that has the following properties:

1. A solution exists.
2. The solution is unique.
3. Stability of the solution, i.e. the solution’s behavior does not change much when there is a small change in the observed data.

A problem that is not well-posed is said to be ill-posed. Most of the estimation methods described in the literature do not adequately address the third condition in the list. Since we want to use the historically estimated local volatility surfaces in order to build a dynamic model for a collection of option prices, it is vital that the estimation method that we use is able to produce stable solutions over time. Since the instantaneous volatility truly is stochastic it means that the local volatility surface that we estimate will evolve stochastically over time. Thus it is important to capture the actual change in the local volatility surface and not include changes that stem from the noise in market prices. The estimated local volatility surface tends to contain a lot of noise if a suitable regularization procedure is not used in the estimation.

Many authors in the literature have suggested methods for estimating the local volatility surface from observed option prices. However, as we have indicated, most described methods do not produce estimated surfaces that are stable over time. We will now give some examples from the literature of suggested methods for estimation of the local volatility surface, and highlight some of the problems that are typical for many of the methods through two examples. We will however not attempt to give a broad overview of the relevant literature, but refer to e.g. Fengler (2005) for a more comprehensive literature review.

The simplest approach is to assume a parametric form for the local volatility surface or the squared local volatility surface. One example of a parametrization of the squared local volatility surface is given in Carmona and Nadtochiy (2009) who use 9 parameters in order to parameterize the surface. As pointed out in Bouchouev and Isakov (1999) the inverse problem when assuming a parametric form for the surface is generally over-determined which means that the optimal parameters cannot be found in a unique and stable way. Assuming a parametric form will typically give rise to non-convex optimization problems with many local minima. This will lead to problems when estimating the local volatility surface.
Another estimation procedure that has been suggested in the literature is to first interpolate and extrapolate market data to all strikes and maturities on the estimation grid. In the next step one could then extract the local volatility surface by discretizing either the Dupire equation (4.44) as in suggested in Dupire (1994), or by discretizing the Kolmogorov forward equation as in Andersen and Brotherton-Ratcliffe (1997). As pointed out in Bouchouev and Isakov (1999), the difficult part is to interpolate the data in order to estimate a smooth surface. Since market data contains a lot of noise the estimated local volatility surface may even include arbitrage, and will be very sensitive and unstable due to errors that arise from the interpolation. In order to illustrate the difficulty of obtaining estimated surfaces that are smooth and stable over time, we have included an estimated surface from the paper by Andersen and Brotherton-Ratcliffe (1997) which was published in the Journal of Computational Finance. The estimated surface is illustrated in Figure 4.3. As is apparent from Figure 4.3, the noise in the input data together with the interpolation method has produced an estimated local volatility surface which includes spikes and contains a lot of noise. Using this estimation method for estimation of a time series of local volatility surfaces would produce very unstable
The local volatility surface

results. Hence this method cannot be used in order to estimate surfaces that can serve as input to a dynamic model of the option market. The property of this method to produce estimated surfaces with irregular shape and sharp spikes is common for many of the methods described in the literature.

The most common approach used in the literature for estimation of the LVS is what Bouchouev and Isakov (1999) call optimization based methods. The optimization based methods are typically applied to under-determined problems where the number of parameters is larger than the number of observed market prices. Generally there will exist an infinite number of surfaces that are consistent with observed market prices, and thus it is necessary to include and additional criterion in order to choose between the possible surfaces. Due to the ill-posedness of the estimation problem, small perturbations of the input data will otherwise typically lead to very different solutions. In order to stabilize the estimation, regularization methods are used. For example Lagnado and Osher (1997) use the $L^2$-norm of the gradient of the local volatility function in order to regularize the problem. The optimization based methods optimize some objective function and make sure that the estimated LVS is consistent with observed market prices by solving the Dupire equation, the Kolmogorov forward equation or the generalized Black-Scholes PDE by finite difference methods. One such example from the literature is Coleman et al. (1999) who use a spline functional approach in order to approximate the local volatility function. They ensure that the estimated local volatility surface is consistent with market prices by solving the generalized Black-Scholes PDE by a finite difference method. Since they choose the number of spline knots to be no greater than the number of observed market prices, the problem that they solve is not over-determined. An estimated local volatility surface from Coleman et al. (1999) is illustrated in Figure 4.4. As is visible in Figure 4.4 the estimated local volatility surface does not contain any sharp spikes as was the case earlier. The spline functional form for the local volatility surface is clearly visible in the figure.

Given the definition of local volatility which we repeat here

$$
\sigma^2_{K,T}(t, S_t) = \mathbb{E}^Q \left( \tilde{\sigma}^2_T | S_T = K, \mathcal{F}_t \right),
$$

it is however difficult to argue that this conditional expected value of the instantaneous variance should vary for different strikes and maturities according to what is implied by the surface in Figure 4.4. This means that this surface is not a realistic local volatility surface. We also argue that the locations of the hills and valleys in the surface are likely due to noise in the observed market prices. This means that these locations will typically change from one day to the next making the estimated local volatility surfaces unstable over time. When we develop our non-parametric optimization based method for estimation of the local volatility

53
surface in paper B, we will explicitly formulate the optimization problem such that we produce realistic option implied surfaces that are stable over time. Now that we have given a short overview of some estimation methods described in the literature, and pointed out that most methods do not produce local volatility surfaces that are stable over time, we will in the next chapter describe some dynamic models for a collection of options described in the literature. The surface that is most commonly used in order to build a dynamic model of a collection of option prices is the implied volatility surface. One of the reasons for this is that this is the surface commonly used in the market for simultaneously expressing the prices for a collection of options. Since the implied volatility surface can be seen as an integrated version of the local volatility surface it is much easier to produce estimated implied volatility surfaces that are stable over time, compared to what is the case for the local volatility surface.
As we saw in the previous chapter it is vital that we have a method for estimating option implied surfaces that produce surfaces that are stable over time. Only then can these surfaces be used in order to build a dynamic model simultaneously describing the dynamics of a collection of option prices and the underlying index. In order to not complicate the notation we assume that we only trade in options with one underlying index. The models described in this chapter can however be expanded to include options on several different underlying equity indices. An alternative to the models described in this chapter would be to use e.g. a single factor stochastic volatility model, such as the Heston model, in order to provide a consistent model that simultaneously describes the dynamics of a collection of option prices and the underlying index. As we have argued earlier though, such a model will be too simplistic to be able to capture the true dynamics observed in the option markets. Another advantage of using a model that specifies the dynamics of an option implied surface is that the starting value for the dynamic model will simply be given by the current surface. This can be compared to e.g. a one-factor stochastic volatility model where the model parameters must first be calibrated to the current surface. We will now present two different modeling approaches for the dynamics of option prices described in the literature. In the first approach the goal is to describe the no-arbitrage dynamics of the option implied surface under the risk-neutral measure $\mathcal{Q}$. In the second approach, which is based on PCA, the dynamics of the option implied surface is specified directly under the physical measure $\mathcal{P}$, and for this approach the no-arbitrage constraints on the drift parameters are not explicitly formulated.
5.1 No-arbitrage approach

The first approaches for developing a dynamic model of option implied surfaces under the risk-neutral measure $Q$ used the implied volatility surface as variable. Early examples of such models described in the literature are Ledoit and Santa-Clara (1998) and Schönbucher (1999). However, as we have argued earlier, the local volatility surface is more suitable to use for a dynamic model than the implied volatility surface. This has to do with the fact that the no-arbitrage conditions are much more involved and difficult to enforce than what is the case for the local volatility surface. This will have implications when building a no-arbitrage model for the option implied surface under the measure $Q$ since also the dynamical no-arbitrage conditions will be more complicated for the implied volatility surface than for the local volatility surface. In this section we will summarize some of the results for a dynamic no-arbitrage model of the squared local volatility surface developed by Derman and Kani (1998).

Derman and Kani (1998) define the continuous dynamics of the squared local volatilities under the physical measure $P$ according to the SDE

$$d\sigma^2_{K,T}(t, S_t) = \alpha_{K,T}(t, S_t)dt + \sum_{i=1}^{n} \theta^i_{K,T}(t, S_t)dW^P_i,$$

which in integrated form results in

$$\sigma^2_{K,T}(t, S_t) = \sigma^2_{K,T}(0, S_0) + \int_0^t \alpha_{K,T}(u, S_u)du + \sum_{i=1}^{n} \int_0^t \theta^i_{K,T}(u, S_u)dW^P_i, \quad (5.2)$$

where $\sigma_{K,T}(0, S_0)$ denotes the initial fixed local volatility surface. The dynamics of the underlying index under $P$ is given by the SDE

$$dS_t = \mu_t S_t dt + \tilde{\sigma}_t S_t dW^P_0,$$  

where $\tilde{\sigma}_t = \sigma_{\tilde{S},t}(t, S_t)$ denotes the instantaneous volatility. If we let $T = t$ and $K = S_t$ in equation (5.2) we have that the instantaneous variance satisfies the following dynamics expressed in integrated form

$$\tilde{\sigma}^2_t = \sigma^2_{\tilde{S},t}(0, S_0) + \int_0^t \alpha_{\tilde{S},t}(u, S_u)du + \sum_{i=1}^{n} \int_0^t \theta^i_{\tilde{S},t}(u, S_u)dW^P_i. \quad (5.4)$$

Derman and Kani (1998) then derives the following no-arbitrage conditions on the drift parameters

$$\alpha_{K,T}(t, S_t) = -\sum_{i=1}^{n} \theta^i_{K,T}(t, S_t) \left( \frac{1}{q(t, S_t, T, K)} \int_t^T \int_0^\infty \theta^{i'}_{K',T'}(t, S_t)q(t, S_t, T', K') \right).$$
No-arbitrage approach

\[ K'^2 \frac{\partial^2}{\partial K'^2} q(T', K', T, K) dK' dT' - \Lambda^i, \]  

(5.5)

where \( \Lambda^i \) for \( i = 1, \ldots, n \) are the market prices of risk for the diffusion components which are assumed to be independent of \( K \) and \( T \). In equation (5.5) \( q(t, S_t, T, K) \) denotes the conditional risk-neutral probability density function which according to Derman and Kani (1998) can be extracted from the current local volatility surface. After a change of measure to the risk-neutral measure \( Q \) the following processes

\[ dW^Q,i_t = dW^P,i_t + \Lambda^i dt, \quad i = 1, \ldots, n, \]  

(5.6)

\[ dW^Q,0_t = dW^P,0_t + \frac{\mu_t - (r_t - \delta_t)}{\sigma_t} dt, \]  

(5.7)

are Wiener processes under \( Q \). This means that if the drift terms satisfy the drift conditions in equation (5.5), then discounted index levels and option prices are simultaneously martingales under the measure \( Q \). As is apparent from equation (5.5), the drift parameters are completely specified by the terms \( \theta_{i,K,T}(t, S_t) \) and the market prices of risk. After a change of measure to the risk-neutral measure the market prices of risk will cancel in the expression for the drift parameters which will then be completely specified by the terms \( \theta_{i,K,T}(t, S_t) \). This also means that in this setting the pricing of derivatives will not depend on the market prices of risk. Apart from the no-arbitrage condition on the drift parameters the squared local volatility surface must in this framework also satisfy the consistency condition for the instantaneous variance

\[ \tilde{\sigma}_t = \sigma_{S_t,t}(t, S_t). \]  

(5.8)

The no-arbitrage conditions on the drift parameters are very involved, which according to Derman and Kani (1998) makes it difficult to directly implement the continuous time specification. Instead they propose to implement the model in discrete time with the help of stochastic implied trees. When they formulate their model in discrete time they, in order guarantee positive variances, change the definition of the drift and diffusion functions according to

\[ \alpha_{K,T}(t, S_t) \rightarrow \alpha_{K,T}(t, S_t) \sigma_{K,T}(t, S_t), \text{ and} \]  

(5.9)

\[ \theta_{i,K,T}(t, S_t) \rightarrow \theta_{i,K,T}(t, S_t) \sigma_{K,T}(t, S_t), \quad i = 1, \ldots, n. \]  

(5.10)

The discretized version of the evolution of the squared local volatility surface is then given by

\[ \Delta \sigma^2_{m,n}(i, j) = \sigma^2_{m,n}(i, j) \left( \alpha_{m,n}(i, j) \Delta t_i + \sum_{l=1}^{n} \theta_{m,n}(i, j) \Delta W^l_i \right). \]  

(5.11)
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

Derman and Kani (1998) then goes on to derive formulas for calculating the drift parameters from the no-arbitrage conditions in the discretized setting. The discretized dynamic model can then be used in order to price complex option positions consistently with observed market prices of standardized options. The model can also be used in order to calculate the sensitivities of complex option positions to changes in market variables. Since volatility is fully stochastic in this model setting this type of model is well suited to price path dependent and forward starting options. It should however be noted that this is a very challenging model to implement.

In a recent paper Carmona and Nadtochiy (2009) provide a more rigorous mathematical treatment of the dynamic LVS model and derive the no-arbitrage conditions on the drift parameters through an alternative route compared to Derman and Kani (1998). Carmona and Nadtochiy (2009) also point out that one term seems to be missing in the drift condition given in equation (5.5). The details of the difference between the drift conditions given in the two papers is out of the scope for this exposition but the interested reader is referred to Carmona and Nadtochiy (2009).

5.2 PCA based approach

An alternative to the no-arbitrage approach described in Chapter 5.1 is the PCA based approach. Instead of demanding dynamic no-arbitrage in a continuous-time setting, the goal here is to describe the empirical movements of the option implied surface. The implied volatility surface (IVS) is the most commonly used surface when building a PCA based model (see Cont and da Fonseca, 2002; Cont et al., 2002; Fengler et al., 2007). One reason for this is that the IVS is directly observable on the market and used by market makers to quote prices of European options. Another reason for using the IVS is that it is much easier to produce implied volatility surfaces that are stable over time since a lot of the noise in market data is integrated out. As we have pointed out it is much more difficult to produce local volatility surfaces (LVSs) that are stable over time. However, the LVS is a more appropriate surface to use in a dynamic model since the necessary static arbitrage conditions is much easier to enforce for the LVS than for the IVS. In order to guarantee absence of arbitrage when using the IVS for the dynamic model it is necessary to check that each generated scenario is free of arbitrage and then use a method to remove the arbitrage for each such scenario. This process of checking and removing arbitrage from the scenarios for the IVS is very computationally burdensome. The main reason that the methods described in the current literature
Figure 5.1: Figures (a) and (b) show the mean and standard deviation of the squared local volatility surface for S&P500 options over the period January 2011 to October 2014.

use the IVS instead of the LVS is most likely that it is very difficult to produce estimated LVSs that are stable over time. In paper C of this dissertation we build a stochastic dynamic model of the equity index option market based on the LVS. Having developed a good method for estimating LVSs that are stable over time, lets us use the LVS as basis when building the dynamic model. To the best of our knowledge this is the first example in the literature that uses the LVS as surface when implementing a PCA based model for the dynamics of the derivatives market.

In order to illustrate some empirical properties of the LVS for S&P500 index options, Figure 5.1 shows the mean and daily standard deviation for the squared local volatility over the period January 2011 to October 2014. The plot of the daily standard for the squared LVS in Figure 5.1b, illustrates that the surface is not static but fluctuates around its average shape. In paper C we build a stochastic model for the squared local volatility surface based on PCA. In Cont and da Fonseca (2002) and Cont et al. (2002) the PCA is based on the covariance matrix of log differences of the IVS. As is visible in Figure 5.1b the magnitude of the standard deviation is very different for different points on the grid. This means that when using the covariance matrix for the PCA, the points on the grid with the largest variances would dominate the first few principal components. This implies that the points on the grid corresponding to low and high moneyness would be too dominating for the first principal components, while they have limited impact on the option prices. For this reason we use the correlation matrix for the PCA in paper C. The factors that drive the randomness of the LVS, corresponding to the diffusion terms \( \theta_m(i,j) \) in equation (5.11), are in this PCA based model proportional to the eigenvectors of the correlation matrix of daily log changes of
the squared local volatility. The details of this model can be found in paper C. The number of factors that drives the randomness of the LVS is chosen such that a sufficient amount of the correlation of the log difference of the squared LVS is explained.

In paper C we use the PCA based dynamic model of the IVS in order to hedge a portfolio of European options using stochastic programming. This means that the prices of these options are given directly from the scenarios of the LVS under the physical measure $P$. If the portfolio to be hedged also included exotic options that could only be priced using Monte-Carlo simulation, we would need to take the drift restriction in equation (5.5) into account since the simulation used for pricing is performed under the risk-neutral measure $Q$. For each scenario under $P$, the exotic options would be priced using the dynamics of the LVS and the index under $Q$, using the scenario outcomes for the LVS and the index as starting values for the Monte-Carlo simulations under $Q$. Implementing the no-arbitrage drift conditions into a model evaluated under the physical measure $P$ is to our knowledge not done in the literature and thus remains an open question. It should also be noted that we do not take the consistency condition given in equation (5.8) explicitly into account in the dynamic model that is developed in paper C. Instead we chose to model the dynamics of the index in discrete time with a GARCH model that includes a Poisson jump term for the index level. This model provides a very close fit to historically observed returns for the S&P500 index. Not satisfying the consistency condition can be justified by an empirically observed characteristic of the index option markets described in Chapter 3.2. The empirical literature on index options has found that the dynamics of the index option markets and the dynamics of the underlying index are partly separated (see e.g. Gärleanu et al., 2009).
Estimating the physical PDF from observed option prices

In this chapter we will derive a result which makes it possible to back out the physical, or real-world, PDF from observed option prices under a specific assumption. The result is derived under a slightly different setup in Heath and Platen (2006), and papers D and E of this dissertation we will rely on a generalized version of the result presented there. The critical assumption in Heath and Platen (2006) is that the Growth Optimal Portfolio (GOP), (see Kelly, 1956; Long, 1990), can be approximated by any well-diversified equity total return index. In order to provide the reader with the appropriate background for the results presented in papers D and E we will in this chapter go through the derivations of the results given in Heath and Platen (2006). The setup is here slightly altered since we assume the existence of a deterministic dividend yield, whereas the dividend yield is assumed to be zero in Heath and Platen (2006). Since the critical assumption relies on approximating the GOP, we will, before we derive the results for backing out the physical PDF, define and derive some properties of the GOP. The derivations for the physical PDF is possible since fair prices of derivatives, when benchmarked with the GOP, are martingales under the physical probability measure. In order to lay the foundation for the derivations we will in this chapter also present the theory of benchmark pricing of derivatives directly under the real-world probability measure.

The results for backing out the physical PDF from option prices presented in Heath and Platen (2006) are related to recent papers in the literature that describes the recovery of the physical PDF. Ross (2013) has gained a lot of attention in the research community recently. There the author shows how the physical PDF can
be recovered from observed option prices in a particular setting where markets are complete, the state variable is a time homogeneous Markov process with a finite discrete state space, and additional assumptions regarding the utility function of the representative investor. In Carr and Yu (2012) the authors, similarly to our setup, relies on properties of the numéraire portfolio in order to back out the physical PDF from option prices. They assume that there is only one driver for the dynamics of the the numéraire portfolio, and this driver, $X_t$, is modeled as a bounded time homogeneous diffusion process. Thus, unlike in Ross (2013), the model setup in Carr and Yu (2012) is cast is continuous time and for a continuous state space. Under additional assumptions Carr and Yu are able to show that it is possible to recover the physical PDF for the driver from derivatives written on that same driver $X_t$. Apart from these two papers additional articles are starting to emerge in the literature and this research area is very active at the moment.

6.1 The growth optimal portfolio

The critical assumption behind being able to back out the physical PDF from option prices on a diversified index, is that this index with dividends reinvested is a good approximation for the GOP. Thus, in order to give the reader an understanding of the theory behind this assumption we will in this section derive some results for the GOP. For the exposition in this section we will follow Platen (2005).

In the model setup for deriving the results for the GOP we assume that there are $d$ sources of traded uncertainty in the market, and these uncertainties are modeled by $d$ independent Wiener processes, $W^1, W^2, \ldots, W^d$. The market consists of $d$ traded securities and a savings account. The traded securities are held in primary security accounts where all dividends or interest rate payments, for stocks and bonds respectively, are reinvested. We let

$$S^0_t = e^{\int_0^t r_s ds},$$

(6.1)

denote the value of the savings account, where $\{r_t\}_{0 \leq t \leq T}$ is the short rate, and $S^j_t$, $j = 1, \ldots, d$, denote the value of the $j$th primary security account at time $t$. The dynamics for the primary security accounts are assumed to follow the following SDEs under the physical probability measure $\mathbb{P}$

$$dS^j_t = S^j_t \left( a^j_t dt + \sum_{k=1}^d b^j_k W^k_t \right), \quad j = 1, \ldots, d,$$

(6.2)
The growth optimal portfolio

where \( S_j^0 > 0, j = 1, \ldots, d \). It should be noted that the drift terms, \( a_j \), the volatility terms \( b_{t,k}^{j} \) and the short rate, \( r_t \), could depend on other sources of uncertainty than the Wiener processes \( W^1, W^2, \ldots, W^d \) and hence the market is not assumed to be complete. In the model setup we could hence have stochastic volatilities and short rate.

We now introduce the market price of risk vector,

\[
\theta_t = [\theta_1^t, \ldots, \theta_d^t]^T, \quad (6.3)
\]

and let

\[
a_t = [a_1^t, \ldots, a_d^t]^T, \quad \text{and} \quad (6.4)
\]

\[
b_t = \begin{bmatrix} b_{t,1,1} & \ldots & b_{t,1,d} \\ \vdots & \ddots & \vdots \\ b_{t,d,1} & \ldots & b_{t,d,d} \end{bmatrix}. \quad (6.5)
\]

Then the equation

\[
b_t \theta_t = (a_t - r_t 1), \quad (6.6)
\]

where \( 1 \in \mathbb{R}^{d \times 1} \) is a vector of ones, has a unique solution if the volatility matrix, \( b_t \), is invertible. Thus, using the natural assumption, as in Platen (2005), that \( b_t \) is invertible we have that the market price for risk vector is given by

\[
\theta_t = b_t^{-1}(a_t - r_t 1). \quad (6.7)
\]

Using this relationship in equation (6.2) we can express the dynamics of the primary security accounts according to the following SDE

\[
dS^j_t = S^j_t \left( r_t dt + \sum_{k=1}^d b_{t,k}^{j}(\theta^k_t dt + dW^k_t) \right), \quad j = 1, \ldots, d. \quad (6.8)
\]

Portfolios consisting of investments in the primary security accounts are formed and the corresponding value process is given by

\[
P^h_t = \sum_{j=0}^d h^j_t S^j_t, \quad (6.9)
\]
where $h_j^t$ is the number of units of security account $j$ that is held at time $t$. If $h_j^t$ is a negative number means that the portfolio is short $h_j^t$ units of the $j$th security account. We only consider self-financing portfolios and hence we have that

$$dP_t^h = \sum_{j=0}^d h_j^t dS_t^j,$$

(6.10)

(see e.g. Björk, 2011). We introduce the relative fractions of the portfolio held in the primary security accounts according to

$$\pi_j^t = h_j^t \frac{S_j^t}{P_t^h},$$

(6.11)

and thus we have that

$$\sum_{j=0}^d \pi_j^t = 1.$$  

(6.12)

Using the relative weights in equation (6.10) we have that the dynamics of the portfolio value process is given by

$$dP_t^h = \sum_{j=0}^d \pi_j^t P_t^h dS_t^j = P_t^h \left( r_t dt + \sum_{j=1}^d \sum_{k=1}^d \pi_j^t b_{j,k}^t (\theta_k^t dt + dW_k^t) \right).$$  

(6.13)

Now that we have defined the dynamics of the value process for an arbitrary portfolio, we will define the GOP and derive the relative weights for this portfolio. Before we define the property of the GOP we will first give the dynamics of the logarithm of the value process for an arbitrary portfolio. Applying Itô’s lemma on $\ln(P_t^h)$ we have that

$$d\ln(P_t^h) = \frac{1}{P_t^h} dP_t^h - \frac{1}{2} \frac{d}{(P_t^h)^2} d[P_t^h],$$

(6.14)

where $[P_t^h]$ denotes the quadratic variation of $P_t^h$ and where

$$d[P_t^h] = (P_t^h)^2 \left( \sum_{j=1}^d \sum_{k=1}^d \pi_j^t b_{j,k}^t dW_k^t \right) \left( \sum_{l=1}^d \sum_{m=1}^d \pi_l^t b_{l,m}^t dW_m^t \right) = (P_t^h)^2 \sum_{j=1}^d \sum_{k=1}^d \pi_j^t b_{j,k}^t \sum_{l=1}^d \pi_l^t b_{l,k}^t dt.$$  

(6.15)
The growth optimal portfolio

Thus we have that the dynamics of \( \ln(P_h^t) \) is given by

\[
d\ln(P_h^t) = r_t dt + \sum_{j=1}^{d} \sum_{k=1}^{d} \pi^j_t b^j_t (\theta^k_t dt + dW^k_t) - \frac{1}{2} \sum_{j=1}^{d} \sum_{k=1}^{d} \pi^j_t b^j_t \sum_{l=1}^{d} \pi^l_t b^l_t dt
\]

\[
\left( r_t + \sum_{j=1}^{d} \sum_{k=1}^{d} \pi^j_t b^j_t (\theta^k_t - \frac{1}{2} \sum_{l=1}^{d} \pi^l_t b^l_t) \right) dt + \sum_{j=1}^{d} \sum_{k=1}^{d} \pi^j_t b^j_t dW^k_t,
\]

(6.16)

and hence we have that the logarithmic growth rate of the portfolio is given by

\[
g^h_t = r_t + \sum_{j=1}^{d} \sum_{k=1}^{d} \pi^j_t b^j_t (\theta^k_t - \frac{1}{2} \sum_{l=1}^{d} \pi^l_t b^l_t).
\]

(6.17)

Using vector notation we can express the logarithmic growth rate of the portfolio more compactly as

\[
g^h_t = r_t + (\pi^h_t)^T b_t (\theta_t - \frac{1}{2} b^T_t \pi^h_t),
\]

(6.19)

where \( \pi^h_t = [\pi^1_t, \ldots, \pi^d_t]^T \). Now that we have defined the logarithmic growth rate of the portfolio we are ready to formulate the definition of the GOP.

**Definition 6.1.1.** The GOP is defined as the portfolio with maximum logarithmic growth rate such that

\[
g^*_t \geq g^h_t,
\]

\( \forall t \in [0, T] \) and all strictly positive processes \( \{P^h_t\}_{0 \leq t \leq T} \).

Thus the definition of the GOP is given in a pathwise sense and is not defined in terms of a conditional expectation or utility function. However, the GOP coincides with the portfolio that maximizes expected logarithmic utility of the terminal wealth (see Karatzas and Shreve, 1998).

Now that we have defined the GOP we will derive the relative weights of the GOP. The relative weights are given by the first order conditions for maximizing the logarithmic growth rate, \( g^h_t \), for each \( t \in [0, T] \),

\[
\nabla_{\pi^h_t} g^h_t = b_t (\theta_t - \frac{1}{2} b^T_t \pi^h_t) = 0,
\]

(6.20)
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

and hence we have that the first order optimality condition is satisfied for

$$\theta_t - b_t^T \pi_t^h = 0,$$  \hspace{1cm} (6.21)

from which we, since $b_t$ is invertible, can extract the relative weights for the GOP

$$\pi_t^{h^*} = (b_t^T)^{-1} \theta_t.$$ \hspace{1cm} (6.22)

Observing that

$$\sum_{j=1}^d \sum_{k=1}^d \pi_t^{j,k} (\theta_t^k dt + dW_t^k) = (\pi_t^{h^*})^T b_t \theta_t dt + (\pi_t^{h^*})^T b_t dW_t,$$ \hspace{1cm} (6.23)

where $dW_t = [dW_t^1, \ldots, dW_t^d]^T$, we plug these relative weights for the GOP into the SDE (6.13) describing the portfolio dynamics, and arrive at the following dynamics for the GOP

$$dP_t^{h^*} = P_t^{h^*} \left( r_t dt + |\theta_t|^2 b_t \theta_t dt + |\theta_t|^2 b_t dW_t \right) =$$

$$P_t^{h^*} \left( (r_t + \sum_{k=1}^d (\theta_t^k)^2) dt + \sum_{k=1}^d \theta_t^k dW_t^k \right).$$ \hspace{1cm} (6.24)

Thus the annualized conditional variance of the GOP, $P_t^{h^*}$, is given by

$$\text{Var} \left( \frac{dP_t^{h^*}}{P_t^{h^*}} \bigg| P_t^{h^*} \right) \frac{1}{dt} = \sum_{k=1}^d (\theta_t^k)^2.$$ \hspace{1cm} (6.25)

Equivalently, we have that the conditional volatility of the GOP is given by

$$|\theta_t| = \sqrt{\sum_{k=1}^d (\theta_t^k)^2},$$ \hspace{1cm} (6.26)

which is the total market price of risk. Thus we can describe the dynamics of the GOP with the following one-dimensional SDE

$$dP_t^{h^*} = P_t^{h^*} \left( (r_t + |\theta_t|^2) dt + |\theta_t| dW_t \right),$$ \hspace{1cm} (6.27)

where $W_t$ is a one-dimensional Wiener process. We will use this observation, for the case of a local volatility model setting with deterministic dividend yields, when we derive the results for extracting the physical PDF from option prices on a diversified equity index approximating the GOP. From equation (6.27) we also identify the risk premium of the GOP which is simply the square of its volatility.
When deriving the results for backing out the physical PDF from option prices we will assume that the GOP can be approximated by a well-diversified index, such as the MSCI world total return index when the investment universe represents the global equity market or the S&P500 total return index for the US equity market. This approximation is supported by results in Platen (2004) and Le and Platen (2006) where it is shown that the GOP is approximated by any well-diversified total return index. These results has later been developed further in Platen and Rendek (2012b) and Rendek (2013).

### 6.2 Pricing of European-style derivatives under the benchmark approach

In this section we will describe how to price derivatives using the real-world probability measure under the benchmark approach. We will summarize the benchmark approach below by following the results presented in Platen (2009) and Platen (2011). As we have described earlier, the GOP or numéraire portfolio can be approximated by a well-diversified portfolio, e.g. the S&P500 total return index when the investment universe represents the US equity market. Similarly the MSCI world total return index is a good proxy of the numéraire portfolio when the global equity market is considered. We will let \( \bar{D}_t \) denote the S&P500 total return index which we will use as an approximation for the numéraire portfolio for the US equity market. Thus, since we assume a deterministic dividend yield, we have that the S&P500 price index is given by

\[
D_t = e^{-\int_0^t \delta_s ds} \bar{D}_t, \tag{6.28}
\]

where \( \{\delta_t\}_{t \geq 0} \) is the dividend yield. In order to simplify the notation we let

\[
B_t = e^{\int_0^t r_s ds}, \tag{6.29}
\]

denote the value of the savings account at time \( t \).

As in Platen (2011) and the setting in the previous section we consider a financial market with \( d \) risky, non-negative, primary securities. As before we let \( P^h_t \) denote the portfolio value at time \( t \) of an arbitrary self-financing portfolio. The portfolio value at time \( t \) is given by

\[
P^h_t = \sum_{j=0}^{d} h^j_t S^j_t, \tag{6.30}
\]
where \( h_j \) is the number of units invested in the \( j \)th primary security account at time \( t \). We let \( \mathcal{V}^+ \) denote the set of all strictly positive, finite, self-financing portfolios with initial capital \( x > 0 \). We use the following definition, formulated by Platen (2009), for the numéraire portfolio, also called the growth optimal portfolio (GOP).

**Definition 6.2.1.** For a given \( x > 0 \), a strictly positive, finite, self-financing portfolio \( P_t^{h*} \in \mathcal{V}^+ \) is called a numéraire portfolio if all non-negative portfolios \( P_t^h \), when denominated in units of \( P_t^{h*} \), are supermartingales.

When we price derivatives under the benchmark approach we choose the GOP as benchmark. We let \( \hat{P}_t^h \) denote the benchmarked value of the portfolio \( P_t^h \) and hence we have that

\[
\hat{P}_t^h = \frac{P_t^h}{P_t^{h*}}. \tag{6.31}
\]

From Definition 6.2.1 we have that the benchmarked value \( \hat{P}_t^h \) of any non-negative portfolio \( P_t^h \) satisfies the supermartingale property

\[
\hat{P}_t^h \geq \mathbb{E}^P \left( \hat{P}_s^{h*} \mid \mathcal{F}_t \right), \tag{6.32}
\]

for all \( 0 \leq t \leq s \leq \infty \). For the case where we have equality in relation (6.32) the benchmarked portfolio price forms a martingale, in which case we call the price fair. There can be many supermartingales that approach the same future random value. The martingale that matches the future random value is the minimal possible supermartingale. This leads to the following **Law of the Minimal Price** given in Platen (2009).

**Theorem 6.2.1. (Law of the Minimal Price)** If a fair portfolio replicates a given non-negative pay-off at a stopping time, then this portfolio represents the minimal replicating portfolio among all non-negative portfolios that replicate this pay-off.

Since there may exist unfair self-financing portfolios in a benchmark setting, the classical Law of one Price does not hold in general. For a given hedgeable pay-off the corresponding fair hedge portfolio represents the least expensive hedge portfolio.

We now turn to the problem of pricing a particular pay-off. We let \( H_T \) define a contingent claim that delivers a non-negative pay-off at maturity \( T \). The benchmarked pay-off

\[
\hat{H}_T = \frac{H_T}{P_T^{h*}}, \tag{6.33}
\]
Pricing of European-style derivatives under the benchmark approach

is assumed to have finite expectation,

$$\mathbb{E}^P \left( \hat{H}_T \middle| \mathcal{F}_0 \right) < \infty. \quad (6.34)$$

As formulated in Platen (2009), the fair price of the contingent claim, $H_T$, is by the Law of the Minimal Price given by the real world pricing formula:

**Corollary 6.2.1.** If for a contingent claim $H_T$, there exists a fair portfolio $P^h_H$ that replicates this claim at maturity $T$ such that $H_T = P^h_H$, then its minimal replicating price a time $t \in [0, T]$ is given by the real world pricing formula

$$P^h_H = P^h_t \mathbb{E}^P \left( \frac{H_T}{P^h_T} \middle| \mathcal{F}_t \right). \quad (6.35)$$

Now, let the benchmarked savings account be given by

$$\hat{B}_t = \frac{B_t}{P^h_t}, \quad (6.36)$$

and let $\Lambda_t$ denote the normalized benchmarked savings account, where

$$\Lambda_t = \frac{\hat{B}_t}{B_0}. \quad (6.37)$$

We have that

$$\mathbb{E}^P (\Lambda_T \middle| \mathcal{F}_0) \leq \Lambda_0 = 1, \quad (6.38)$$

since $\Lambda_T$ is a supermartingale. Now, we rewrite equation (6.35) according to

$$P^h_H = P^h_t \mathbb{E}^P \left( \frac{H_T}{P^h_T} \middle| \mathcal{F}_t \right) = \mathbb{E}^P \left( \frac{B_T / P^h_T}{B_t / P^h_t} H_T \middle| \mathcal{F}_t \right) = \mathbb{E}^P \left( \frac{\Lambda_T B_t}{B_T} H_T \middle| \mathcal{F}_t \right). \quad (6.39)$$

Since $\mathbb{E} (\Lambda_T \middle| \mathcal{F}_t) \leq 1$ we have for $t = 0$ that

$$P^h_H \leq \frac{\mathbb{E}^P \left( \frac{\Lambda_T B_0}{B_T} H_T \middle| \mathcal{F}_0 \right)}{\mathbb{E}^P (\Lambda_T \middle| \mathcal{F}_0)}. \quad (6.40)$$

Let the formally taken risk-neutral price be given by (see e.g. Björk, 2011)

$$\mathbb{E}^Q \left( \frac{H_T}{B_T} \middle| \mathcal{F}_0 \right) = \frac{\mathbb{E}^P \left( \frac{\Lambda_T H_T}{B_T} \middle| \mathcal{F}_0 \right)}{\mathbb{E}^P (\Lambda_T \middle| \mathcal{F}_0)}. \quad (6.41)$$
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

Under risk-neutral pricing the Radon-Nikodym derivative, \( \Lambda_T = \frac{dQ}{dP} \), is assumed to be a \( \mathbb{P} \)-martingale, which yields \( \mathbb{E}^\mathbb{P}(\Lambda_T|\mathcal{F}_0) = \Lambda_0 = 1 \), and we have in this case that

\[
\mathbb{E}^\mathbb{Q}\left(\frac{H_T}{B_T}\bigg|\mathcal{F}_0\right) = \mathbb{E}^\mathbb{P}\left(\Lambda_T \frac{H_T}{B_T}\bigg|\mathcal{F}_0\right). \tag{6.42}
\]

Without assuming the existence of an equivalent risk-neutral probability measure we have from equations (6.40) and (6.41) that

\[
P_0^{bh} \leq \frac{\mathbb{E}^\mathbb{P}\left(\Lambda_T \frac{B_0 H_T}{B_T}\bigg|\mathcal{F}_0\right)}{\mathbb{E}^\mathbb{P}\left(\Lambda_T\bigg|\mathcal{F}_0\right)} = \mathbb{E}^\mathbb{Q}\left(\frac{H_T}{B_T}\bigg|\mathcal{F}_0\right). \tag{6.43}
\]

We note that only if \( \Lambda_t \) is a martingale, i.e. that \( \mathbb{E}(\Lambda_T|\mathcal{F}_0) = \Lambda_0 = 1 \), then \( P_0^{bh} \) equals

\[
\mathbb{E}^\mathbb{Q}\left(\frac{H_T}{B_T}\bigg|\mathcal{F}_0\right).
\]

Hence, only if the Radon-Nikodym derivative, \( \Lambda_t \), is a martingale the real world price and the risk-neutral price are guaranteed to be equivalent. Otherwise the benchmark price is lower than or equal to the risk-neutral price. In Platen (2009) it has been argued, through an empirical example for the S&P500 total return index, that the typical path of a normalized benchmarked savings account seems highly unlikely to be the realization of a martingale, but rather resembles the trajectory of a strict supermartingale. The conclusion from this is that the risk-neutral prices of derivatives can be substantially higher than the corresponding real world prices, especially for longer term contracts. Also, in Heath and Platen (2006) the authors refer to papers in the literature (Sin, 1998; Lewis, 2000; Loewenstein and Willard, 2000; Delbaen and Shirakawa, 2002; Heath and Platen, 2002; Platen, 2002; Cox and Hobson, 2005; Fernholz et al., 2005) which indicate that for realistic market models no equivalent risk-neutral measure exists. For the setup that we use here there is no assumption made regarding the existence of an equivalent martingale measure.

In our approximate setting, where the S&P500 total return index is assumed to approximate the GOP for the US equity market, we will now give the expressions for European call and put options, futures and fair zero coupon bond prices under the benchmark approach. The real world price of a European call option, with strike price \( K \) and maturity \( T \), on the GOP, is under the benchmark approach given by

\[
C = \tilde{D}_t \mathbb{E}^\mathbb{P}\left(\frac{(D_T - K)^+}{D_T}\bigg|\mathcal{F}_t\right) = \tilde{D}_t \mathbb{E}^\mathbb{P}\left(\left(e^{-\int_t^T \delta_s ds} - \frac{K}{D_T}\right)^+ \bigg|\mathcal{F}_t\right). \tag{6.44}
\]
Pricing of European-style derivatives under the benchmark approach

Similarly, the real world price of the put option with the same strike and maturity is under the benchmark approach given by

\[ P = \tilde{D}_t \mathbb{E}^P \left( \frac{(K - D_T)^+}{D_T} \mid \mathcal{F}_t \right) = \tilde{D}_t \mathbb{E}^P \left( \left( \frac{K}{D_T} - e^{-\int_t^T \delta_s \, ds} \right)^+ \mid \mathcal{F}_t \right). \]  

(6.45)

The fair or real world futures price under the benchmark approach is given by solving for the futures price \( F_{t,T} \) in the equation

\[ \tilde{D}_t \mathbb{E}^P \left( \frac{D_T - F_{t,T}}{D_T} \mid \mathcal{F}_t \right) = 0, \]  

(6.46)

and thus the fair futures price under the benchmark approach is for a deterministic dividend rate given by

\[ F_{t,T} = \frac{\mathbb{E}^P \left( \frac{D_T}{D_T} \mid \mathcal{F}_t \right)}{\mathbb{E}^P \left( \frac{1}{D_T} \mid \mathcal{F}_t \right)} = \frac{1}{\mathbb{E}^P \left( \frac{1}{D_T} \mid \mathcal{F}_t \right)}. \]  

(6.47)

The benchmarked fair price of a zero coupon bond, \( P_{t,T} \), is given by the conditional expectation of the benchmarked value of a contract that pays out one monetary unit at maturity, i.e. \( H_T = 1 \). Thus, we have

\[ P_{t,T} = \tilde{D}_t \mathbb{E}^P \left( \frac{1}{D_T} \mid \mathcal{F}_t \right). \]  

(6.48)

Using the payout function \( H_T = 1 \) in equation (6.41), and assuming evaluation at time 0, we have that

\[ \mathbb{E}^Q \left( \frac{1}{B_T} \mid \mathcal{F}_0 \right) = \frac{\mathbb{E}^P \left( \Lambda_T \frac{1}{B_T} \mid \mathcal{F}_0 \right)}{\mathbb{E}^P \left( \Lambda_T \mid \mathcal{F}_0 \right)}. \]  

(6.49)

Since the Radon-Nikodym derivative, \( \Lambda_T \), from equation (6.37) is defined by

\[ \Lambda_T = \frac{\hat{B}_T}{\hat{B}_0} = \frac{B_T}{B_0} = \frac{\tilde{D}_0 B_T}{D_T}, \]  

(6.50)

we have that

\[ \mathbb{E}^P \left( \Lambda_T \frac{1}{B_T} \mid \mathcal{F}_0 \right) = \mathbb{E}^P \left( \frac{\tilde{D}_0 B_T}{D_T} \frac{1}{B_T} \mid \mathcal{F}_0 \right) = \tilde{D}_0 \mathbb{E}^P \left( \frac{1}{D_T} \mid \mathcal{F}_0 \right) = P_{0,T}. \]  

(6.51)
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

Thus we have from equation (6.49) and (6.51) that

$$P_{0,T} = \mathbb{E}^Q \left( \frac{1}{B_T} \big| \mathcal{F}_0 \right) \mathbb{E}^F (\Lambda_T | \mathcal{F}_0),$$

(6.52)

and since $\mathbb{E}^F (\Lambda_T | \mathcal{F}_0) \leq 1$ this means that

$$P_{0,T} \leq \mathbb{E}^Q \left( \frac{1}{B_T} \big| \mathcal{F}_0 \right) = P_{0,T}^*.$$  

(6.53)

The right-hand-side of equation (6.53) is the definition of the risk-neutral price at time 0 of a zero coupon maturing at time $T$, and hence we have that the fair zero coupon bond price is smaller than or equal to the corresponding risk-neutral price. Since the fair futures price under the benchmark approach is given by

$$F_{0,T} = \frac{1}{\mathbb{E}^F \left( \frac{1}{D_T} \big| \mathcal{F}_0 \right)} = \frac{\tilde{D}_0 e^{-\int_0^T \delta_s ds} \mathbb{E}^P (\frac{1}{D_T} \big| \mathcal{F}_0)}{P_{0,T}} = \frac{\tilde{D}_0 e^{-\int_0^T \delta_s ds} \mathbb{E}^P (\frac{1}{D_T} \big| \mathcal{F}_0)}{\tilde{D}_0 e^{-\int_0^T \delta_s ds} \mathbb{E}^P \left( \frac{1}{\tilde{D}_0} \big| \mathcal{F}_0 \right)},$$

(6.54)

and the corresponding risk-neutral futures price can be expressed as

$$F_{0,T}^* = \frac{\tilde{D}_0 e^{-\int_0^T \delta_s ds}}{P_{0,T}^*},$$

(6.55)

we thus have that the fair futures price under the benchmark approach is greater than or equal to the corresponding risk-neutral price,

$$F_{0,T} \geq F_{0,T}^*.$$  

(6.56)

In order to derive put-call parity under the benchmark approach we start by studying the difference $C - P$,

$$C - P = \tilde{D}_t \mathbb{E}^P \left( \left( e^{-\int_0^T \delta_s ds} - \frac{K}{D_T} \right)^+ \big| \mathcal{F}_t \right) - \tilde{D}_t \mathbb{E}^P \left( \left( \frac{K}{D_T} - e^{-\int_0^T \delta_s ds} \right)^+ \big| \mathcal{F}_t \right) =$$

$$\tilde{D}_t \mathbb{E}^P \left( \left( e^{-\int_0^T \delta_s ds} - \frac{K}{D_T} \right)^+ \big| \mathcal{F}_t \right) = \tilde{D}_t \mathbb{E}^P \left( \left( e^{-\int_0^T \delta_s ds} \right)^+ \big| \mathcal{F}_t \right) =$$

$$\tilde{D}_t \mathbb{E}^P \left( \frac{1}{D_T} \big| \mathcal{F}_t \right) = \tilde{D}_t e^{-\int_0^T \delta_s ds} - K \tilde{D}_t \mathbb{E}^P \left( \frac{1}{D_T} \big| \mathcal{F}_t \right) =$$

$$\tilde{D}_t e^{-\int_0^T \delta_s ds} - KP_{t,T},$$

(6.57)
Recovery of the physical PDF from observed option prices

and thus this expression is analogous to the put-call parity under risk-neutral pricing but where the risk-neutral price of a zero coupon bond is replaced by the fair zero coupon bond price. Thus in order to get the same value of Black-Scholes implied volatilities for call and put option prices estimated under the benchmark approach, we need to use the risk-free rate implied by the fair zero-coupon bond rather than the risk-neutral zero-coupon bond price observed on the market.

6.3 Recovery of the physical PDF from observed option prices

We will now derive the result for extracting the implied physical PDF from option prices when we have non-zero but deterministic dividends. This is a generalization of the results presented in Heath and Platen (2006), where the results with zero dividends are derived. The results are derived under the setting where the short rate is assumed to be a continuous deterministic function of time. In the derivation of the implied PDF we assume that the broad market index with dividends reinvested, \( \tilde{D}_t = D_t e^{\int_0^t \delta_s ds} \), approximates the GOP for the given equity market. In our case \( \tilde{D}_t \) represents the S&P500 total return index which approximates the GOP for the US equity market. The dynamics of the GOP is given in equation (6.27), which under our setting with a local volatility model and continuous deterministic dividends translates into the SDE

\[
d\tilde{D}_t = \tilde{D}_t \left( (r_t + \sigma^2(t, D_t)) dt + \sigma(t, D_t) dW_t \right),
\]

where \( \sigma(t, D_t) \) is the local volatility function.

The results for extracting the physical PDF from option prices will now be derived in two steps. We will first derive a relation that must hold between the PDF for the benchmarked savings account and scaled benchmarked call prices. In the next step we will then use this relation in order to derive the result that gives the connection between the PDF for the broad market index and call options on that index. Under the assumption that the broad market total return index approximates the GOP, the fair price in units of the local currency of a European call option is from equation (6.44) given by

\[
C(t, D_t, T, K) = \tilde{D}_t \mathbb{E}^{\mathbb{P}} \left( \frac{(D_T - K)^+}{D_T} \right| \mathcal{F}_t),
\]

(6.59)
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

The fair benchmarked price of a European call option is then given by
\[ \hat{C}(t, D_t, T, K) = \frac{C(t, D_t, T, K)}{D_t} = \mathbb{E}^P \left( \frac{(D_T - K)^{+}}{D_T} \right| \mathcal{F}_t \right) \]. \quad (6.60)

We will now derive a relation between the PDF for the benchmarked savings account
\[ \hat{B}_t = \frac{B_t}{D_t}, \quad (6.61) \]
and a scaled version of the fair benchmarked call price \( \hat{C}(t, D_t, T, K) \). In order to do that we define
\[ \kappa = \frac{B_T}{K}, \quad \tilde{\kappa} = \frac{B_T}{Ke^{\int_0^T \delta_s ds}} = \frac{B_T}{K} = \kappa e^{-\int_0^T \delta_s ds}, \quad \tilde{K} = Ke^{\int_0^T \delta_s ds}, \quad (6.62) \]
and let the scaled benchmarked call prices be given by
\[ u(t, \hat{B}_t, T, \tilde{\kappa}) = \kappa \hat{C}(t, D_t, T, K) = \mathbb{E}^P \left( \left( \kappa e^{-\int_0^T \delta_s ds} - \frac{K}{\tilde{K}} \right)^{+} \left| \mathcal{F}_t \right) \right) = \mathbb{E}^P \left( (\tilde{\kappa} - \hat{B}_T)^{+} \left| \mathcal{F}_t \right) \right. \]. \quad (6.63)

Writing this expectation in integral form we have that
\[ u(t, \hat{B}_t, T, \tilde{\kappa}) = \int_0^{\tilde{\kappa}} (\tilde{\kappa} - y)p_{\hat{B}}(t, \hat{B}_t, T, y)dy. \quad (6.64) \]

Now, taking the derivative with respect to \( \tilde{\kappa} \) of this integral yields
\[ \frac{\partial}{\partial \tilde{\kappa}} \left( \tilde{\kappa} \int_0^{\tilde{\kappa}} p_{\hat{B}}(t, \hat{B}_t, T, y)dy - \int_0^{\tilde{\kappa}} yp_{\hat{B}}(t, \hat{B}_t, T, y)dy \right) = \int_0^{\tilde{\kappa}} p_{\hat{B}}(t, \hat{B}_t, T, y)dy + \tilde{\kappa}p_{\hat{B}}(t, \hat{B}_t, T, \tilde{\kappa}) - \tilde{\kappa}p_{\hat{B}}(t, \hat{B}_t, T, \tilde{\kappa}) = \int_0^{\tilde{\kappa}} p_{\hat{B}}(t, \hat{B}_t, T, y)dy. \quad (6.65) \]

Taking the derivative with respect to \( \tilde{\kappa} \) again yields
\[ \frac{\partial^2 u(t, \hat{B}_t, T, \tilde{\kappa})}{\partial \tilde{\kappa}^2} = p_{\hat{B}}(t, \hat{B}_t, T, \tilde{\kappa}). \quad (6.66) \]
Thus we have that the PDF of the benchmarked savings account is given by the second derivative with respect to $\tilde{\kappa}$ of the benchmarked call price scaled with $\tilde{\kappa}$. We will now use this result in order to derive a relationship between call prices and the PDF for the benchmarked savings account. In the next step we will then derive the final relationship between call prices and the PDF for the broad market index $D_t$.

We have that
\begin{equation}
    u(t, \hat{B}_t, T, \tilde{\kappa}) = \kappa \hat{C}(t, D_t, T, K) = e^{\int_0^T \delta_s ds} \tilde{\kappa} \hat{C}(t, D_t, T, K),
\end{equation}
and thus the first derivative with respect to $\tilde{\kappa}$ is given by
\begin{align}
    \frac{\partial u(t, \hat{B}_t, T, \tilde{\kappa})}{\partial \tilde{\kappa}} &= \frac{\partial}{\partial \tilde{\kappa}} \left( e^{\int_0^T \delta_s ds} \tilde{\kappa} \hat{C}(t, D_t, T, K) \right) = \\
    e^{\int_0^T \delta_s ds} \left( \hat{C}(t, D_t, T, K) + \kappa \frac{\partial \hat{C}(t, D_t, T, K)}{\partial \tilde{\kappa}} \right) &= \\
    e^{\int_0^T \delta_s ds} \left( \hat{C}(t, D_t, T, K) + \frac{\partial \hat{C}(t, D_t, T, K)}{\partial K} \frac{\partial K}{\partial \tilde{\kappa}} \right) = \left\{ \frac{\partial K}{\partial \tilde{\kappa}} = -\frac{B_T e^{-\int_0^T \delta_s ds}}{\tilde{\kappa}^2} \right\} = \\
    e^{\int_0^T \delta_s ds} \left( \hat{C}(t, D_t, T, K) - \frac{B_T e^{-\int_0^T \delta_s ds}}{\tilde{\kappa}} \frac{\partial \hat{C}(t, D_t, T, K)}{\partial K} \right). \tag{6.68}
\end{align}

Taking the derivative with respect to $\tilde{\kappa}$ once more yields
\begin{align}
    \frac{\partial}{\partial \tilde{\kappa}} \left( e^{\int_0^T \delta_s ds} \left( \hat{C}(t, D_t, T, K) - \frac{B_T e^{-\int_0^T \delta_s ds}}{\tilde{\kappa}} \frac{\partial \hat{C}(t, D_t, T, K)}{\partial K} \right) \right) = \\
    e^{\int_0^T \delta_s ds} \left( \frac{\partial \hat{C}(t, D_t, T, K)}{\partial K} \frac{\partial K}{\partial \tilde{\kappa}} + \frac{B_T e^{-\int_0^T \delta_s ds}}{\tilde{\kappa}^2} \frac{\partial \hat{C}(t, D_t, T, K)}{\partial K} \right) = \\
    \frac{B_T e^{-\int_0^T \delta_s ds}}{\tilde{\kappa}^2} \frac{\partial^2 \hat{C}(t, D_t, T, K)}{\partial K^2} + \frac{B_T e^{-\int_0^T \delta_s ds}}{\tilde{\kappa}} \frac{\partial \hat{C}(t, D_t, T, K)}{\partial K} = \frac{K^3 e^{\int_0^T \delta_s ds}}{B_T} \frac{\partial^2 \hat{C}(t, D_t, T, K)}{\partial K^2}. \tag{6.69}
\end{align}

Thus we have that the relationship between the two second derivatives is given by
\begin{equation}
    \frac{\partial^2 u(t, \hat{B}_t, T, \tilde{\kappa})}{\partial \tilde{\kappa}^2} = \frac{K^3 e^{\int_0^T \delta_s ds}}{B_T} \frac{\partial^2 \hat{C}(t, D_t, T, K)}{\partial K^2}. \tag{6.70}
\end{equation}
From this expression we can directly formulate the relationship between call prices and the PDF for the benchmarked savings account. We will then use a change of variable in order to derive the relationship between call prices and the PDF for the broad market index. Since the relationship between call prices and benchmarked call prices is given by

\[
C(t, D_t, T, K) = \hat{D}_t \hat{C}(t, D_t, T, K),
\]

we have that the following relationship holds

\[
p_B(t, \hat{B}_t, T, \tilde{\kappa}) = \frac{K^3 e^{2\int_0^T \delta_s ds}}{D_t B_T} \frac{\partial^2 C(t, D_t, T, K)}{\partial K^2}.
\]

We now let \( p_D(t, D_t, T, K) \) denote the PDF under \( \mathbb{P} \) of the broad market index \( D_t \). Using the relationship

\[
\tilde{\kappa} = \frac{B_T}{K e^{\int_0^T \delta_s ds}}
\]

and the expression in equation (6.72) we can derive an expression for \( p_D(t, D_t, T, K) \). Introducing the function \( g(\cdot) \) according to

\[
K = g(\kappa) = \frac{B_T}{\tilde{\kappa} e^{\int_0^T \delta_s ds}},
\]

or equivalently

\[
\tilde{\kappa} = g^{-1}(K) = \frac{B_T}{K e^{\int_0^T \delta_s ds}},
\]

we have, since \( g^{-1}(K) \) is strictly decreasing in \( K \), that

\[
p_D(t, D_t, T, K) = p_B(t, \hat{B}_t, T, \tilde{\kappa}) \left| \frac{\partial g^{-1}(K)}{\partial K} \right|,
\]

(see e.g. Gut, 1995). Since

\[
\frac{\partial g^{-1}(K)}{\partial K} = -\frac{B_T}{K^2 e^{\int_0^T \delta_s ds}}
\]

we have that

\[
p_D(t, D_t, T, K) = p_B(t, \hat{B}_t, T, \tilde{\kappa}) \frac{B_T}{K^2 e^{\int_0^T \delta_s ds}} =
\]

76
Recovery of the physical PDF from observed option prices

\[
\frac{K^3e^{\int_0^T \delta_s ds}}{D_tB_T} \partial^2 C(t, D_t, T, K) \quad \frac{B_T}{K^2 e^{\int_0^T \delta_s ds}} \partial^2 C(t, D_t, T, K) = \frac{K e^{\int_0^T \delta_s ds}}{D_t} \partial^2 C(t, D_t, T, K) \partial K^2. \quad (6.78)
\]

Thus we have, under the setting where the GOP is approximated by a broad market index with dividends reinvested and where the dynamics of the broad market index is given by a local volatility model, derived the following relationship between observed option prices and the physical PDF of the broad market index

\[
p_D(t, D_t, T, K) = \frac{K e^{\int_0^T \delta_s ds}}{D_t} \frac{\partial^2 C(t, D_t, T, K)}{\partial K^2}. \quad (6.79)
\]

This means that if we observed market prices of call options for a continuum of strike prices and time-to-maturities, we could use the relationship in equation (6.79) in order to back out the physical PDF from observed option prices. Of course we only observe market prices of call options for a discrete set of strike prices and time-to-maturities, and hence we would need to have a method for interpolating and extrapolating the call price function in order to be able to back out the physical PDF. When we back out the physical PDF from option prices in paper 4 and 5 we use the optimization based method for extracting the squared local volatility surface as a tool for interpolating and extrapolating the call price function in equation (6.79). Even though the results in this section are derived without the assumption of the existence of a equivalent risk-neutral measure we can use the RND surface implied by the squared local volatility surface for extracting the physical PDF. Comparing equation (6.79) to the expression for the risk-neutral PDF for the broad market index which is given by

\[
q_D(t, D_t, T, K) = e^{\int_T^t r_s ds} \frac{\partial^2 C(t, D_t, T, K)}{\partial K^2}, \quad (6.80)
\]

we have that

\[
p_D(t, D_t, T, K) = e^{-\int_T^t r_s ds} \frac{K e^{\int_0^T \delta_s ds}}{D_t} q_D(t, D_t, T, K) = e^{\int_T^t (\delta_s - r_s) ds} q_D(t, D_t, T, K) \frac{K}{D_t} q_D(t, D_t, T, K). \quad (6.81)
\]

Thus, assuming that the equivalent risk-neutral measure exists, equation (6.81) lets us conveniently change between the physical and the risk-neutral measure when the broad market index with dividends reinvested is assumed to approximate the GOP.

This ends the theoretical exposition in this introductory part of the dissertation, and the next chapter summarizes and highlights the 5 papers in the second part of the dissertation.
Summary of included papers

In this chapter we will first give a short summary of the 5 papers that are included in this dissertation and also highlight the contributions for each of the papers. We will also give an exposition to how the research questions that were presented in Chapter 1 are answered by the papers in this dissertation. The summary and contributions of the 5 papers are given below. We also include an account of the contribution of each author for the 5 appended papers.

**Paper A** (Barkhagen and Blomvall, 2015a)
In paper A we have developed a non-parametric optimization based framework for estimation of the RND surface implied by observed option prices, while satisfying no-arbitrage constraints. The inverse estimation problem is regularized by only considering realistic surfaces in the optimization problem. The regularization is accomplished by including a roughness measure in the objective function. The roughness is measured as a weighted sum of the squared difference between the discretized first and second derivative of the RND and a reference PDF. We demonstrate empirically that the method produces smooth and realistic RND, local volatility and implied volatility surfaces. Through a time series study using end of day data for S&P500 index options we show that the method produces surfaces that are stable over time. The main contribution of paper A is that we showed that when using the RND as variable in the optimization problem, we can, given that the futures prices are considered to be known, formulate a convex optimization problem for the full surface problem which produces smooth and stable surfaces.

The research need for paper A was identified by the first author. The second author proposed that it might be possible to apply a method earlier applied to the interest rate market also to the equity index option market. The necessary
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

The mathematics for the paper was derived by the first author. The first author did all computer implementations and wrote the whole paper.

**Paper B** (Barkhagen and Blomvall, 2015b)  
In paper B we extended the framework in paper A to the estimation of the squared local volatility surface implied by observed option prices, while satisfying no-arbitrage constraints for the surface. As in paper A, the regularization is accomplished by including a roughness measure in the objective function. We demonstrate that, compared to paper A, it is much easier to choose the weights in the discretized roughness measure when we use squared local volatilities as variables instead of the RND. As in paper A, we demonstrate empirically that the method produces smooth and realistic local volatility, RND and implied volatility surfaces. The main contributions of paper B are that we have extended the methodology in Andersen and Brotherton-Ratcliffe (1997) to a non-uniform grid, and shown how to integrate it into the optimization based framework for estimation of smooth and stable option implied surfaces.

The research idea for paper B grew out of the research for the first paper. The second author contributed with insights which guided the decisions during the early stages of the research. The first author developed the necessary research methodology for the paper. The necessary mathematics for the paper was derived by the first author. The first author did all computer implementations and wrote the whole paper.

**Paper C** (Barkhagen and Blomvall, 2014)  
In paper C we develop a stochastic programming model for hedging an option book in a realistic setting where all transactions take place at observed bid and ask prices. The SP model relies on scenarios that accurately capture the important risk factors of the equity index option market. In order to generate the scenarios that are used as input we develop a dynamic model for the local volatility surface based on PCA. In order for the PCA to capture the true dynamics of the equity index option market, and not to a large extent describe the noise in observed market data, we use the estimation method described in paper B for estimation of historical local volatility surfaces. The SP model is then tested out-of-sample for options on the Swedish OMXS30 index. We show that the SP model produces a hedge that has both lower realized risk and cost compared with traditional delta and delta-vega hedging strategies. The main contribution of paper C is that it is, to the best of our knowledge, the first in the literature that for option hedging incorporates an implementation of a dynamic model for the local volatility surface.
The second author proposed the basic research idea for paper C. The necessary mathematics for the paper was derived by the first author. The first author did most of the computer implementations and wrote the whole paper.

**Paper D (Barkhagen et al., 2015)**

In paper D we recover the real-world density for the underlying index from S&P500 Index options. The recovery of the real-world density relies in a first step on extracting high-quality estimates of the risk-neutral density surface from liquid options using the method developed in Barkhagen and Blomvall (2015b). In a second step, these risk-neutral densities are transformed into real-world densities using the results developed in Heath and Platen (2006). The Minimal Market Model (MMM) of Platen and Rendek (2012a) is in paper D used as a tool for extracting liquidity premia embedded in observed option prices. The option implied MMM parameters are estimated from observed S&P 500 index option prices over the period January 2011 to October 2014 via Simulated Maximum Likelihood. We show that the densities implied by the MMM parameters are consistent with the shape of the realized return distribution of the underlying index. This admits the identification of liquidity premia embedded in option price data. The main contribution of paper D is that we identify and estimate two separate liquidity premia embedded in S&P 500 index options – one for low strikes and one for increasing time-to-maturity. This result is consistent with previous findings in the literature.

The third author proposed the basic research idea for paper D. The first author developed the necessary research methodology for the paper. The necessary mathematics for the paper was derived by the first author. The first author did all computer implementations and wrote the whole paper.

**Paper E (Barkhagen and Blomvall, 2015c)**

In paper E we use the methodological framework developed in paper D in order to evaluate the real-world density forecasts recovered from observed S&P 500 index option prices. We use a density forecast evaluation framework which uses evaluations based on both the probability integral transform and out-of-sample log-likelihood values. Based on observed option prices and index levels over the period January 2011 to October 2014, we evaluate the forecast ability of the extracted option implied information in two different ways. First we evaluate how well the option market estimates the prevailing volatility level of the underlying index. We also evaluate the forecasting ability of the option implied densities for a two day horizon. The results are compared to those obtained using 6 different GARCH models estimated from observed index levels of S&P500. The main contribution of paper E is that we show that our recovered real-world densities
Optimal Decisions in the Equity Index Derivatives Markets  
Using Option Implied Information

predicts the realized return distribution of the underlying index for a two day horizon very well. This can be contrasted to two recent papers (Shackleton et al., 2010; Yun, 2014), who find that historical models estimated from intra-day index data predicts the realized return distribution better than option implied densities for a forecast horizon of one day.

The research idea for paper E grew out of the research for paper D and was then jointly developed by the two authors. The first author developed the necessary research methodology for the paper. The necessary mathematics for the paper was derived by the first author. The first author did all computer implementations and wrote the whole paper.

Now that we have summarized the papers we will give an exposition to how the research questions in Chapter 1 are answered by the papers in the dissertation. The structure of the exposition is organized such that the research questions are answered in the order in which they were dealt with during my PhD project. I will first give an exposition to how the research questions RQ3 and RQ4 given in Chapter 1 are answered in the first 2 papers in this dissertation. Thereafter I will provide answers to RQ1 and RQ2 which are mainly answered in paper C, before finishing this chapter by summarizing the preliminary answer to RQ5 which is addressed in papers D and E.

7.1 Answers to RQ3 and RQ4

RQ3 asks which properties that are essential to capture when estimating option implied surfaces. This research question was formulated in Chapter 1 and is reproduced here for convenience:

RQ3: Which properties are essential when estimating option implied surfaces?

We will below list the essential properties and then later explain how these properties are captured in the optimization based frameworks that are presented in papers A and B in this dissertation. The list of essential properties to be captured is given below.

A1. Estimation of option implied surfaces is complicated by the fact that a cross-section of observed option prices contain noise and possibly also arbitrage. In order for a set of option prices to be free of arbitrage, certain no-arbitrage con-
A2. Estimating surfaces from a set of observed option prices is an inverse problem which is typically ill-posed. The properties of a well-posed problem were listed in Chapter 4.2 and are repeated here for convenience

1. A solution exists.
2. The solution is unique.
3. Stability of the solution, i.e. the solution’s behavior does not change much when there is a small change in the observed data.

A problem that is not well-posed is said to be ill-posed. Many traditional methods described in the literature do not adequately address the issue of producing a solution that is stable over time. Most parametric methods, such as e.g. the Heston model (Heston, 1993) and SVI (Gatheral, 2006), lead to non-convex optimization problems that have many local optima, and hence a small change in the input data can lead to very different estimated values for the model parameters when the optimization problem is solved to optimality. Since the input data contains a lot of noise this means that traditional methods often have difficulty producing stable solutions. In paper A we demonstrate through a time series study that our optimization based framework produces RNDs that are stable over time. The stability issue for inverse problems can be handled by so called regularization techniques, which is used in the objective function in paper A and B.

A3. The problem that is solved should be formulated such that it is comparatively easy to estimate the surface. Thus, the optimization problem should not have any local optima, since this makes it very difficult to actually solve the problem to optimality, and it is also related to the stability issue of a well-posed inverse problem as was stated in A2. In paper A we show that the problem formulation that is used leads to a convex quadratic program where the solution is a global
Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information

optimum. The problem formulation that is used in paper B contains non-linear equality constraints and is a non-convex problem, however empirical tests indicate that the solution does not seem to get stuck in local optima.

A4. When we formulate the problem we must handle that the surface contains an infinite number of variables in an adequate way. In both paper A and paper B this is solved by discretization and by only considering realistic solutions through an appropriate regularization. Many of the traditional methods described in the literature solves this problem by assuming a parametric form for the solution, however as mentioned earlier this leads to, among other things, the drawback that the problem typically becomes non-convex with many local optima.

A5. The method should produce a surface which is successful in explaining how the surface changes with time, i.e. we want to use a method that is able to capture the changes that actually occur in the options markets. Since observed option prices contain a lot of noise it means that is easy to capture changes that does not stem from properties of the market but rather from the methods that inadequately handles the noise in the input data. As we have mentioned, many traditional parametric methods have difficulties dealing with the noise in the input data. As a consequence, when the evolution of the surface over time is studied, there is a high risk that dimension reduction methods such as principal component analysis (PCA) fail to capture the actual movements and to a large extent only capture the noise. This has been observed for the interest rate market in Blomvall and Ndengo (2012).

A6. When we formulate the optimization problem, we want to choose a variable in the problem that is not an integration (or summation) of other variables. Otherwise it is difficult to separate the noise in the estimation since the noise has been integrated out. We refer to variables that are created by integration as being non-local, and local variables refer to variables that contain local information that have not been integrated out. We can rank the surfaces with respect to how local the information is according to

\[ \text{price} \prec \text{implied volatility} \prec \text{RND} \prec \text{local volatility} \prec \text{conditional RND}, \]

where variables to the left are more integrated than variables to the right. In paper A we use the RND as variable when we solve the problem and in paper B we use the squared local volatility. The conditional RND volume contains the most local information but the high dimensionality conveys that it is probably difficult to use as variable in the optimization problem. One obvious problem with
using the local RND as variable is that we then need to specify a large number of reference RNDs when we are solving the problem.

The answers that we have identified to RQ3 serves as input to the fourth research question which reads:

RQ4: How should option implied surfaces be estimated?

The answers to this question are given in papers A and B and are extracted below.

In paper A and paper B we are using a non-parametric optimization based method which has previously been successfully applied to the interest rate markets in Manzano and Blomvall (2004), Blomvall (2011) and Blomvall and Ndengo (2012). Studies on the interest rate markets have shown that the method is very robust with respect to noise in input data (A1). Given the desired properties listed above we should have a method which is able to exclude unrealistic surfaces (A6). This means that it is less desirable to estimate e.g. the call price surface or the implied volatility surface since these surfaces contain integrated variables, which in turn means that irregularities in more local surfaces are integrated out. Thus for example an implied volatility surface that looks realistic can imply an unrealistic RND- and local volatility surface (see e.g. Rebonato, 2004). Since the conditional RND volume is intractable, it remains to choose between the RND and squared local volatilities as variables in the optimization problem. Out of the RND and squared local volatility surfaces, the easiest to estimate is the RND surface and therefore we started with this surface when solving the problem of estimating surfaces implied by observed option prices. We show in paper A that when we choose the RND as variable, the resulting optimization problem is convex and can be solved fast and efficiently. However, since we in the optimization also solve for implied futures prices, together with the fact that the no-arbitrage conditions in the time-to-maturity direction depend on futures prices, means that we cannot solve the full surface problem directly. Instead, we are first solving the problem one maturity at a time which results in an implied futures price for all maturities in the data set. These futures prices are then treated as constants and serve as input when we solve the problem for the whole surface simultaneously.

Another drawback with choosing the RND as variable in the optimization problem is that it is non-trivial to choose the parameters in the roughness measure in the objective function. This has to do with the fact that the natural relative variation is large for RND values at different points in the strike direction. Since we to a large degree measure the roughness as discrete second derivatives, it means that different points on the grid will have very different impact on the objective function
value which must be handled when we choose the parameters. Put in other words, when we seek realistic surfaces, we must handle the fact that it is reasonable with a wide range in the RND values.

In light of these drawbacks that are present when choosing the RND as variable in the optimization problem we could instead work with the local volatility surface. The local volatility surface is more complicated to work with than the RND surface but it results in several desirable properties for the optimization problem that we want to solve. In paper B we choose the squared local volatility surface as variable. The squared local volatility surface has the property that the range in values of squared local volatilities is approximately a factor 10, which is much smaller than the range for the RND surface. This means that it is much easier to choose the parameters in the discretized roughness measure in the objective function, and as the empirical tests in paper B demonstrates it is sufficient to choose the parameter vectors as a constant multiplied by a vector of ones.

Choosing the squared local volatility surface as variable also brings another desirable property – namely that it is much easier to ensure no-arbitrage in the time-to-maturity direction for the squared local volatility surface. This rectifies the problem that we had with also solving for the implied futures prices when using the RND surface as variable, and we can thus solve the full surface problem directly, without having to first solve the individual maturity problems when using the squared local volatilities as variables. The drawback with using the squared local volatilities instead of the RND as variables is that the optimization problem becomes non-linear and non-convex. However, empirical tests have shown that the solution does not seem to get stuck in local optima.

As an answer to RQ4 we find that an option implied surface should be estimated with a method that has properties equal to or better than the ones in paper B. The important properties from RQ3 that our proposed methods satisfy is that they define well-posed inverse problems (A2), they can deal with the huge dimensionality of the set of all feasible surfaces (A4) and that the decision variables are local (A6). The most promising method is the one presented in paper B. Since the optimization algorithm does not seem to get stuck in local optima, this method satisfies property (A3), and as a consequence also that the problem is well-posed (A2). By selecting squared local volatilities as variables also property (A6) is satisfied and it is easy to guarantee an arbitrage-free surface (A1), and in combination with the regularization which minimize the squared second derivative, the method also produces realistic surfaces (A4). Since each surface is realistic and arbitrage-free, this should provide a good opportunity to study the evolution of the surface over time (A5) as the noise in input data (A1) to a large extent.
has been removed. This is illustrated in paper B for the squared local volatility surface. This method therefore provides an accurate picture of the option surface at each point in time.

Now that we have answered RQ3 and RQ4 formulated in Chapter 1 we will turn our attention to RQ1 and RQ2 of the dissertation.

7.2 Answer to RQ1

RQ1 was formulated in Chapter 1 and is repeated below:

RQ1: Which properties are important to capture in order to obtain optimal decisions in the equity index derivatives markets?

The answers to RQ1 are partly given in Chapter 2 of this introductory part and partly by the question at issue when developing the optimal decision framework used in paper C. Partial answers to RQ1 are summarized in the list below:

B1. A crucial component for obtaining optimal decisions is inputs which realistically describes the simultaneous dynamics of the option surface and the underlying index under the physical measure $\mathbb{P}$, and, if complex derivatives needs to be priced in order to make the decision, also under the risk-neutral measure $\mathbb{Q}$.

B2. In order to adequately take the expected return relative to the risk exposure into account, we need to develop good methods for estimating the risk premia both for the underlying equity market and for the equity index option market.

B3. Traditional one-factor stochastic volatility models, with or without jumps, which are commonly used in the literature, are not sufficiently complex in order to capture the true simultaneous dynamics for the index and the associated index options. This means that it is necessary to develop a more realistic dynamic model for the equity index derivatives market.

B4. The equity index derivatives market contains many assets which are possible to trade in, and there are significant transaction costs that cannot be disregarded. Furthermore, as we have argued, the simultaneous distribution of the risk factors for the equity index option market is complex and cannot be described by the standard models described in the literature. When we formulate the model framework we should hence use a method that can handle these specific problem properties.
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

B5. In order to reach a good decision it is more important to create scenarios that as accurately as possible captures the true distribution of the risk factors, than to take into account that we will make decisions at different points in time over the time horizon of the problem.

Given the partial answers provided in the list above we will now answer RQ1 of the dissertation. When building a framework for making optimal decisions in the equity index option markets it is important to realistically model the market dynamics (B1). This means that it is necessary to use a model which includes many state variables, and as argued in Chapter 2 this implies that dynamic programming and stochastic optimal control are not suitable as solution methods when building the optimal decision framework. Furthermore, the equity index option markets contains significant transaction costs which cannot be disregarded (B4), which is another reason that stochastic optimal control is not appropriate to use when solving the optimal decision problem. Since one-factor stochastic volatility models are not sufficiently complex in order to capture the true dynamics of the index option markets, also the traditional methods described in the literature such as delta-vega hedging, are ruled out when we build the optimal decision framework. As we have argued in Chapter 2, stochastic programming (SP) on the other hand is well suited to handle the important properties that we must take into account when building the framework. The only input that is necessary in SP is the current asset prices and scenarios that describe the simultaneous distribution of the asset prices at every time step in the SP model. This means that SP can handle arbitrary simultaneous distributions for the risk factors (B3 and B4). SP can also handle transaction costs and other restrictions that must be taken into account when the optimal decision is to be made (B4). Since it is more important to capture the true distribution of the risk factors than to take into account that new decisions will be made at future points in time, we choose to create a large number of scenarios as input to the SP model and use a 2-stage model without recourse (B5). By creating a large number of scenarios as input, the generated scenarios will also include rare events, which is very important for capturing the risks of the derivatives markets. Using a large number of scenarios also means that the exponential growth in tree size in SP means that a multi-stage model will be too computationally costly. Thus, even though a 2-stage SP model without recourse decisions has drawbacks, this is the best option to determine optimal decisions for derivative hedging. In paper C we show that our SP model for option hedging performs better than the traditional methods described in the literature.
7.3 Answer to RQ2

Now that we have formulated the answers to RQ1 we are ready provide a preliminary answer to RQ2 which is repeated below for convenience:

RQ2: *How should a dynamic model for the equity index derivatives markets be built?*

**C1.** If the instruments that are included in the problem to be solved can be priced directly from an observed local volatility surface, i.e. if the options are European-style, it is sufficient to specify a consistent dynamic model of the squared local volatility surface under the physical measure $P$.

**C2.** If the problem however also includes exotic instruments for which the price depends on the stochastic dynamics for the volatility surface, such as e.g. forward starting options, we also need to specify the dynamic model of the squared local volatility surface under the risk-neutral measure $Q$. Under the risk-neutral measure $Q$ the drift terms of the squared local volatility process can be determined by the no-arbitrage drift restrictions given in Chapter 5. Under the model described in Chapter 5, the risk-neutral drift terms are completely determined by the diffusion factors of the process for the squared local volatility and are thus independent of the market prices of risk.

**C3.** In order to determine the diffusion coefficients for the dynamic model for the volatility surface it is suitable to use PCA. PCA is a good way to compactly capture the dynamics that has historically been observed on the equity index option market.

**C4.** When building a stochastic dynamic model for an option implied surface we should choose a surface for which it is easy to ensure no-arbitrage for generated scenarios. As we saw in Chapter 4, the no-arbitrage conditions for both the price and implied volatility surfaces are very involved, making it difficult, and computationally expensive, to generate arbitrage-free scenarios. We also saw in Chapter 4 that the static no-arbitrage conditions for the local volatility surface are much simpler making this a suitable surface to build a dynamic model for. This will demand that the estimation method used is able to produce surfaces that are stable over time. This is very difficult to achieve which is likely why the implied volatility surface, instead of the local volatility surface, has been used in earlier studies described in the literature.
C5. In order to capture the dynamics for the underlying equity market it is also necessary to use a dynamic model for the index which captures the most important properties of the equity market listed in Chapter 3.2. These include fat tails, volatility clustering and jumps in the index level.

C6. In order to adequately capture the risk-return trade-off in the dynamic model, it is also necessary to incorporate the dynamics for the time dependent risk premia that exist in the equity index derivatives markets.

The dynamic model that we have developed in paper C should not be seen as an ultimate model, but rather as a very good step in the right direction. The dynamic model is built for a surface for which it is easy to ensure no-arbitrage (C4), and the randomness of the implied surface is compactly captured by using PCA (C3). In order to be able to take rare extreme events in the underlying index level into account, the dynamics of the index in paper C is modeled by a GARCH model that takes jumps in the index level into account (C5). In order to take rare extreme events for the index option market into account we could also include the possibility of jumps in the volatility factors given by the PCA. The model used in paper C is much more realistic than most models described in the literature, but there are still some parts remaining in order to achieve a complete model. As we have mentioned the model needs to be further developed in order to take the risk premia that exist in the market into account (C6). The model also needs to be extended in order to incorporate also the dynamics under the risk-neutral measure \( \mathbb{Q} \), by taking the drift restrictions that guarantee no-arbitrage into account (C2).

7.4 Answer to RQ5

RQ5 was formulated in Chapter 1 and is repeated below:

RQ5: What does the information extracted under the risk-neutral measure \( \mathbb{Q} \) imply under the physical measure \( \mathbb{P} \)?

The answers to RQ5 are partly given by the results in paper D and paper E. Partial answers to RQ5 are summarized in the list below:
D1. The research community has only recently started to develop methods for extracting information of the return distribution of the underlying directly from option prices without including observations of the underlying in the estimation. Two such approaches are developed in Carr and Yu (2012) and Ross (2013). By application of these methods, it is possible to directly infer what the information extracted under the risk-neutral measure $Q$ implies under the physical measure $P$.

D2. In this dissertation we use a result developed by Heath and Platen (2006) which lets us recover the physical density from observed option prices under the assumption that the index follows a local volatility model. This method relies on the assumption that the numéraire portfolio can be approximated by a well diversified index which is supported by the results in Platen (2004) and Le and Platen (2006).

D3. In paper D we show that by transforming the recovered physical densities to densities implied by the Minimal Market Model (MMM) of Platen and Rendek (2012a), we are able to identify and estimate two liquidity premia embedded in S&P 500 index option prices – one for low strikes and one for increasing time-to-maturity. The results are consistent with the results earlier published in Bollen and Whaley (2004) and Gärleanu et al. (2009).

D4. In paper E we show that the option implied MMM densities predict the realized return distribution for the underlying index for a horizon of two days very well. We also show that the option implied densities perform better than 6 different GARCH models that are estimated from historically observed index levels. This result differs from what has recently been found in the papers by Shackleton et al. (2010) and Yun (2014). This is an indication that using information extracted only from observed option prices implies a better forecast than by also using historical index levels in the estimation.

Extraction of option implied information under the physical measure is currently an active area of research in finance (D1) and much remains to be done. We have in this dissertation investigated two issues regarding what the information extracted from option prices under $Q$ implies under the physical measure $P$. The results are based on using the method developed by Heath and Platen (2006) in order to recover the physical density by only using observed option prices in the estimation (D2). We have shown that it is possible to identify two separate liquidity premia embedded in observed prices of S&P 500 index options (D3). These results are consistent with earlier findings in the literature which use methods different to ours in order to extract similar estimates. In e.g. Gärleanu et al. (2009) the authors use a demand based option pricing model in order to estimate premia.
that are similar to our results in shape and magnitude. We have also evaluated the physical density forecasts and compared the performance to that obtained with 6 methods that are estimated from historically observed index levels. The results show that the option implied densities outperforms the historical densities, and are able to very well capture the realized return distribution of the S&P 500 index (D4) for a forecast horizon of two days. These results are very promising and it would be interesting to evaluate the option implied density forecasts also for longer horizons. Given our limited option data set this investigation is left for future research. Our research on option implied information recovered from option prices is at an early stadium and we have so far only provided an initial and partial answer to RQ5 in this dissertation. Much remains to be done but the methods developed and applied in this dissertation seem very promising for this line of research.

7.5 Answer to ORQ

We will now answer the overall research question (ORQ) of the dissertation. ORQ was formulated in Chapter 1 and is repeated below:

ORQ: How can a general framework for making optimal decisions in the equity index derivatives market be developed?

The answers to ORQ are given in Chapter 2 and by the results in the five papers of the dissertation and are summarized in the list below:

E1. In Chapter 2 we argue that stochastic programming (SP) is well suited for modelling and solving optimization problems under uncertainty in the equity index derivatives market. When building a general framework for making optimal decisions in the equity index derivatives market it will thus be based on SP.

E2. The only input data that is needed for the SP model is prices of instruments right now and scenarios for those prices at the future point in time for which the problem is solved. In order to obtain good decisions from the SP model, it is vital to generate scenarios that provide an accurate description of the simultaneous distribution of derivatives prices and the underlying at the future point in time.

E3. In order to have an accurate view of the simultaneous distribution for future option prices it is very important to use a method that estimates option implied surfaces of high quality. These estimated surfaces can then form the basis
for building a dynamic model of option prices.

**E4.** In order to generate the scenarios for the underlying index and the option prices written on the index we need a model that describes the simultaneous dynamics for the index and the option prices. This dynamic model is formulated as a dynamic model for an option implied surface. Since it is much easier to ensure no-arbitrage for the local volatility surface, than for e.g. the implied volatility surface, such a dynamic model should model the local volatility surface as the dynamic object.

**E5.** The dimensions of the dynamic model can be reduced by using Principal Component Analysis. The dynamic model can be built directly under the physical measure $\mathbb{P}$ using the estimated principal components. The principal components can also be used as input to a no-arbitrage dynamic model as described in Derman and Kani (1998) and Carmona and Nadtochiy (2009).

**E6.** Apart from having an accurate view of the variance for the relevant risk factors it is in many applications also important to have a good estimate of the expected values, and thereby risk premia, for option prices. This information can then be used in order to make decisions for optimal investments in the index options market.

**E7.** Apart from having a good dynamic model for option prices, we also need to have a model that accurately describes the dynamics for the underlying index. This model could be estimated from historically observed index levels, from observed index option prices or from a combination of the two.

There are several reasons for using an SP model, compared to dynamic programming or stochastic optimal control, as basis for an optimal decision framework for the equity index derivatives markets (**E1**). By choosing SP we can use arbitrary models for the underlying index and derivatives written on that index. Furthermore, there are no limitations regarding how many options that can be included in the model as was the case for dynamic programming and stochastic optimal control. SP is also well suited for handling transaction costs and other restrictions such as e.g. how much capital that is allowed to be allocated to the portfolio. One drawback with using SP for solving the optimization problem is that SP only gives the solution right now. In order to obtain a new optimal decision at a later point in time, the problem has to be solved again. Obtaining good decisions from the SP model crucially depends on using scenarios as input which accurately capture the simultaneous distribution of the index and derivatives written on the index.
Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information

(E2). When creating the scenarios used as input to the SP model it is also vital that we do not introduce arbitrage in the model. Otherwise, an optimization algorithm which solves the problem would immediately realize this, since the optimal objective function value in this case is unbounded.

When estimating a model which describes the simultaneous distribution of the relevant risk factors, we use option implied surfaces as input to the estimation. In order to obtain an accurate description of the simultaneous distribution it is therefore vital that we have a method that produces estimated option implied surfaces that are realistic and stable over time (E3). Otherwise, the estimated model will not give an accurate view of the distribution but to a large extent describe the noise present in the data. An alternative would be to use e.g. a stochastic volatility model which describes the simultaneous dynamics of the index and the index volatility. Such a model will however not be able to accurately capture the complicated dynamics which is empirically observed for equity index option markets. The dynamic model used for generating the scenarios that are used as input to the SP model should be formulated for an option implied surface that facilitates the creation of scenarios that are free of arbitrage. We therefore build a model for the local volatility surface (E4), since it is much easier to ensure no-arbitrage for the local volatility surface than for e.g. the implied volatility surface. The estimated surfaces are high dimensional and the dimensions of the dynamic model can be substantially reduced by using PCA (E5). A dynamic model for the local volatility surface can then be built directly under the physical measure \( \mathbb{P} \) by specifying stochastic processes for the principal components.

When building a dynamic model for the index and options written on the index it is for many applications also important to also include a good estimate of the risk premia for the options and the index (E6). The results in paper D provides an estimate of liquidity premia embedded in S&P 500 index options for low strike prices and for longer maturities. These estimated liquidity premia can be included in the dynamic model in order to facilitate the identification of options that are overpriced by the market. The methods applied in paper D can therefore be used in order to create input to a SP model for optimal investments in the equity index options market. It is also important to have a model which accurately captures the return distribution of the underlying index (E7). As we have demonstrated in paper E, the physical densities recovered from S&P 500 index option prices accurately capture the realized distribution for the underlying index. It thus seems promising to use the option implied densities as a basis for the dynamical model for the index as input to the optimal decision model.

This ends the exposition of how the research questions that was formulated in
Chapter 1 are answered by the papers of this dissertation. The second part of the dissertation consists of the 5 appended papers.

References


Barkhagen, M., Blomvall, J. and Platen, E., Recovering the real-world density and liquidity premia from option data. Submitted 2015.


Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information


REFERENCES


Optimal Decisions in the Equity Index Derivatives Markets
Using Option Implied Information


Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information


Optimal Decisions in the Equity Index Derivatives Markets Using Option Implied Information


Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-117106