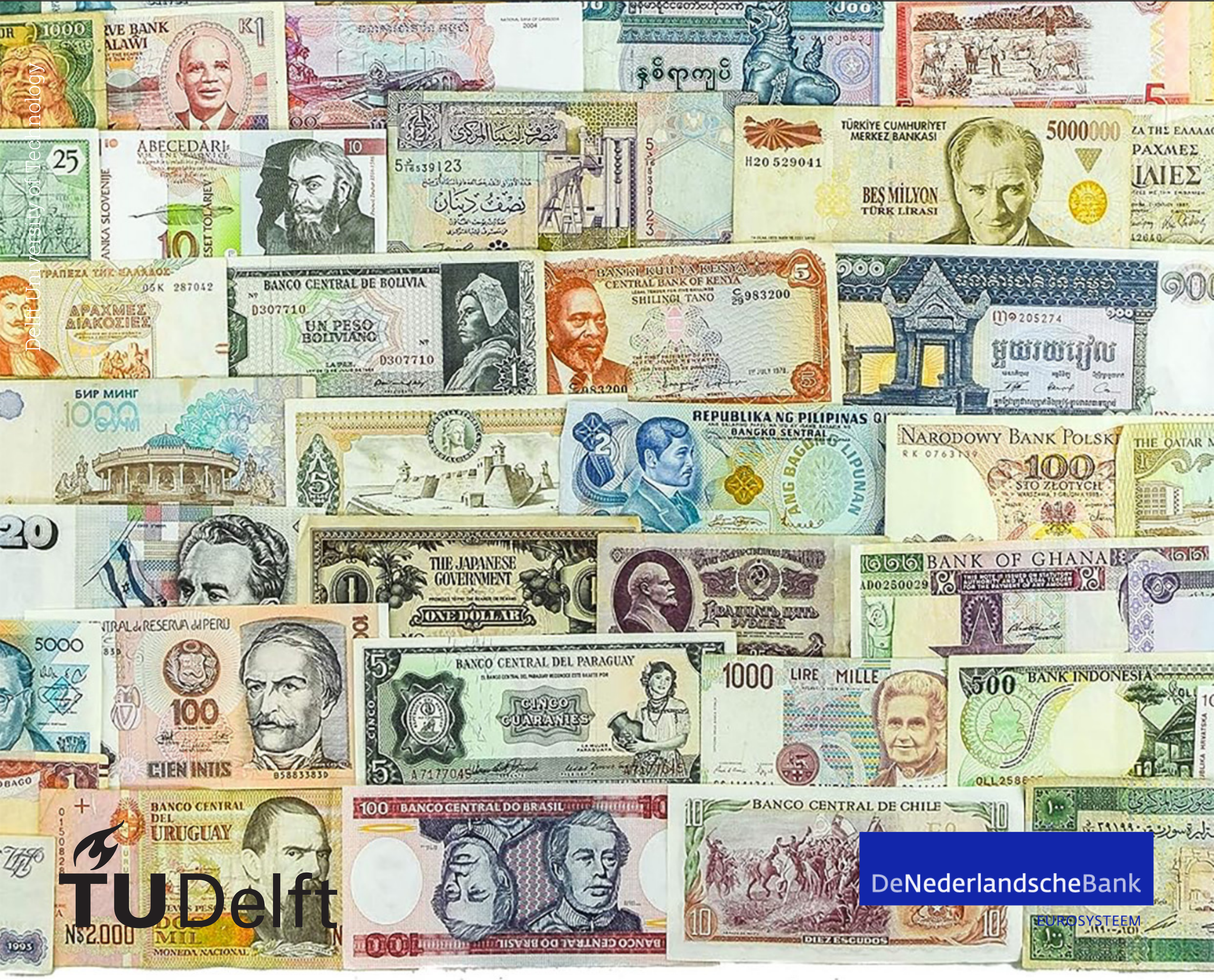
Estimating Option Implied Probability Distributions for Inflation

MSc Applied Mathematics

Theodorus Johannes Martinus Schuttenbeld



Delft University of Technology

TU Delft

De Nederlandsche Bank
EUROSYSTEEM

Estimating Option Implied Probability Distributions for Inflation

by

Theodorus Johannes Martinus Schuttenbeld

to obtain the degree of Master of Science
at the Delft University of Technology, faculty EEMCS,
to be defended publicly on Wednesday December 20th 2023 at 11:00.

Student number:	4271793	
Specialisation:	Financial Engineering	
Project duration:	April 1, 2023 – December 20, 2023	
Thesis committee:	Dr. L. E. Meester,	TU Delft, supervisor
	Prof. Dr. A. Papapantoleon,	TU Delft
	Dr. F. Fang,	TU Delft
	Dr. D. R. Rijsbergen,	De Nederlandsche Bank

Summary

This thesis investigates the estimation of option-implied probability density functions for inflation using inflation options, focusing not only on the expectated value but the whole distribution. The aim is to identify the most effective method for measuring the market expectation of future inflation.

The research explores both parametric and non-parametric approaches for deriving these density functions from inflation option prices. Methodologies include parametric models such as expansion, generalised distribution, and mixture methods, alongside non-parametric techniques using Breeden and Litzenberger's result, such as curve-fitting and kernel methods.

Implementing these methods involved analysing inflation option data sourced from the BVOL Bloomberg database, specifically for Harmonised Index of Consumer Prices excluding Tobacco (HICPxT) options from January 1, 2013. The study employed Shimko's method, various spline methods, the Delta method, and Kernel method, assessing their effectiveness and challenges.

Results reveal diverse implications for each method. Visual comparisons showcase the varying outcomes of the implemented methods, Likelihood-based assessments present a more numerical approach benefiting the Delta and Kernel methods due to higher scores and fewer negative likelihoods.

Conclusions suggest that while multiple methods offer insights into inflation prediction, the Kernel method shows promise in its reliability while the Delta method scores highest in the numerical methods. However, challenges in accurately modelling extreme values and tail behaviours persist across methodologies. Recommendations for further research involve addressing these limitations and exploring enhancements to refine inflation prediction models.

Keywords: *Inflation, Market Expectations, Option Pricing, Probability Density Functions, Risk Neutrality*

Preface

The last nine months have been a dedicated journey towards completing this thesis, a significant milestone in my pursuit of a master's degree in Applied Mathematics at Delft University of Technology. I owe thanks to several individuals who supported me along this rewarding yet challenging path.

I'm grateful for the guidance and support from De Nederlandsche Bank (DNB), particularly under the supervision of Meilina Hoogland, my daily supervisor, whose mentorship has been integral to this thesis. I am immensely appreciative of her guidance and support throughout this journey.

Additionally, I want to express my appreciation to Ludolf Meester, my university supervisor, whose guidance has been instrumental in shaping this thesis. I'm also thankful to Gibran Watfe for his contributions, and to David Rijsbergen for his managerial support and guidance.

The collaborative spirit within the Market Intelligence team at DNB enriched my understanding of economics and financial markets, and I'm grateful for their openness and willingness to share insights.

Lastly, my deepest appreciation goes to my friends and family for their unwavering support and encouragement throughout this enriching and exhausting journey.

With sincere thanks,
Theodorus Johannes Martinus Schuttenbeld
Delft, December 2023

Contents

List of Figures

List of Tables

1

Deriving inflation forecasts from the market

Inflation, the rise in the general price level of goods and services over time, has long been a focal point of economic discussions. As a phenomenon impacting economies on a global scale, inflation carries significant implications for financial markets, investment strategies, and policy formulation. The relevance of understanding and predicting inflation has become even more pronounced in current times, where economic stability is intricately tied to the ability to anticipate and respond to fluctuations in the inflationary environment.

The importance of inflation is underscored by its profound effects on purchasing power, interest rates, and investment returns. Whether it is the individual consumer grappling with increased living costs or the institutional investor hedging against the eroding value of assets, the dynamics of inflation are always present in financial decision-making. To navigate the uncertainties posed by inflationary pressures, the financial industry has developed sophisticated instruments like inflation options.

Deriving information about the distribution function from option prices involves using the prices of options to infer market expectations about the future movements of an underlying asset. Options are financial instruments that give the holder the right, but not the obligation, to buy or sell an asset at a predetermined strike price at the option's expiration. An intuitive explanation of the ability to derive a density function from option prices stems from the information encoded in the price differences between options with different strikes. These differences provide valuable insights into the market's perception of the likelihood that the underlying asset's future prices will fall within the range defined by those specific strikes. These density functions are not merely speculative, but are deeply intertwined with the actions and strategies of investors seeking to protect themselves from the economic consequences of inflation. Understanding and extracting valuable insights from market expectations has become crucial for financial professionals, policymakers, and investors.

This significance is particularly amplified in the context of recent high levels of inflation. Notably, these expectations hold considerable significance for central banks tasked with the mandate of maintaining inflation at a stable level, often targeted around 2%. Understanding these market-based inflation expectations has become crucial for central banks in fulfilling their mandates and steering monetary policies effectively.

While this study focuses on the comparison between methods to derive the option implied risk-neutral density (RND) for inflation, there have been previous comparison studies into these methods applied to data of other underlying assets. Particularly the review by Figlewski [9] gives a great overview of various methods. Furthermore, the book by Jackwerth [12] can be of great help for anyone interested in an accessible summary of these methods.
Inflation options are a recent invention, therefore the scientific literature does not date back very far.

The idea of risk-neutral densities was introduced by Black and Scholes [5] and Merton [15] with their famous closed form formula for option pricing. Since then there has been an interest in deriving the market expectation from asset prices. A very important result in this subject is by Breeden and Litzenberger [6], they showed how the risk-neutral density can be obtained from a large set of option prices. Built upon this result are the methods by Shimko[18], Aït-Sahalia and Duarte [1], Malz [13], and Grith, Härdle and Schienle [10], among many others.
The brief outline of these methods is that the RND can be found by differentiating the option price function twice with respect to the strike. Which means these methods try to find the best option price function that is differentiable and fits the data.

This thesis embarks on a comprehensive exploration of various methods to obtain option implied probability density functions for inflation. The aim is to deepen the understanding of the applicability of these methodologies on inflation data. Correspondingly, the following research question is devised: What is the best method of estimating option implied probability density functions for inflation?
To aid in answering that question the following will be answered first: Are models that are derived to estimate option implied densities for stocks also applicable to inflation data?
It is impossible to judge whether one of the implemented methods is the "best" method. However, in chapter 4 a few methods will be explained that can be used to compare the found methods.

In the first chapter the groundwork is laid through an exploration of the core concepts: risk neutrality and option pricing. Unveiling the essence of the risk-neutral density and its role in option pricing sets the stage for investigating option implied probability density functions tailored for inflation. Then, a literature review will be presented. This review will explore existing literature surrounding market-based inflation expectations and methodologies used to estimate option implied probability density functions. Next, in the methodology chapter, the detailed approach undertaken in this research to implement these models is explained. Subsequently, the findings are presented in the results chapter. Finally, this thesis will end in a discussion that sheds light on the implications of this research. Moreover, this final chapter also lays the groundwork for potential future research directions.

2 Theoretical Framework

This chapter explains a few concepts that are either required for understanding the results of this thesis, or are concepts that transcends all methods. We explain what the Risk-Neutral Density (RND) is and what its role is in pricing options, to be able to examine the methods to derive the option implied probability density functions for inflation. And secondly, some concepts concerning option pricing are discussed that are used in every method in the later chapters.

2.1. Risk-Neutrality

Note to readers, I am going to replace the example below, I am looking for a good source. I tried to show in a simple way what risk-neutral prices are, but Ludolf commented that this example is confusing.

The objective of this research is to derive the risk-neutral density from option prices, but to understand the results it is necessary to understand the meaning of the risk neutral density. It is assumed that, under the risk-neutral measure $\mathbb{Q}$, the expected return of an asset is equal to the risk-free rate r. This assumption simplifies complex financial models, making them more tractable and facilitating the valuation of derivative securities. In mathematical notation, the assumption of risk neutrality means the option price at time t is equal to the discounted expectation of the payoff, for example a call option with price C, maturity T and strike K:

$$C(T,K) = e^{-rT}\mathbb{E}_{\mathbb{Q}}(\max(S_T - K, 0)). \tag{2.1}$$

This means that, by using the integral representation of the expectation, the RND is the density function g that solves the following:

$$C(T,K) = e^{-rT}\int_K^\infty (s-K)g(s)ds. \tag{2.2}$$

In the case of inflation options the prices of the options are often higher than the expected payoff. This is due to their use as protection against extreme inflation scenarios, resulting in higher market-based inflation expectations than actual anticipated inflation. There have been attempts to decompose the prices of inflation linked products into their real expectation and risk premia [11], but that falls outside the scope of this thesis.

2.2. Inflation Options

There are two kinds of inflation options, caps and floors. Caps pay money if the realised inflation is above a predetermined strike rate K, while floors pay out if the realised inflation is below K. In such a contract it is also necessary to determine the maturity of the contract, denoted in T years. The payoff for a cap option with strike K and a duration of T years at maturity is therefore:

$$C_T(T,K) = \max(I_T - K^T, 0), \tag{2.3}$$

where I_T denotes the inflation from the beginning of the contract until maturity. Note that since both the inflation and the strike are rates the payoff is also a rate. In practice this means that for every contract

a principal amount is agreed upon. Neither party pays or receives this principal amount at any time, at the beginning a percentage of this principle is paid to purchase the option and at the end a percentage of this principal is paid out depending on the realised inflation. For the purpose of determining the RND the size of the principal does not matter, therefore it is most convenient to always use €1.
Note that a lot of rates are used which can be denoted as $k = 4\% = 0.04$ or $K = 1.04$, in this thesis this capitalisation will be used to tell which value is being used.

There are also two kinds of payoff structures, zero-coupon and year-on-year (Y-o-Y). A zero-coupon option only pays out money at the end of the contract, comparing the realised inflation over all T years with K^T. While a Y-o-Y option pays off every year comparing the realised inflation of that year with K. This means that there are potentially T payoffs, resulting in T cash flows which brings difficulties with discounting. It was therefore chosen to only use zero-coupon options in this research.

It is also necessary to determine which inflation index is used to measure I_T, since there are multiple measures of inflation such as national indices or indices excluding certain products. For this research the Harmonised Index of Consumer Prices (HICP) is used. The HICP is used in the euro area to measure the consumer price inflation. The version of the HICP that is used for options is the HICP excluding Tobacco, this is because of French legislation which forbids using tobacco as an underlying for certain financial instruments. Using option prices on different underlying indices should work in the same manner, it was decided to work with euro area data for this thesis because it is the main interest for the DNB and because the options on European inflation are traded more frequently than their US or UK counterparts.

2.2.1. Arbitrage

Given the highly illiquid nature of this market, the dataset poses a significant risk of containing arbitrage opportunities. Simply put, arbitrage is a free possibility of earning money without the risk of losing money. A lot of financial models make the assumption that arbitrage does not occur in the market. This benefits a lot of pricing problems in financial markets because it means all financial products have one unique arbitrage free price. However, in practice arbitrage is common, especially in illiquid markets. In very active markets arbitrage resolves itself because market participants will use the arbitrage opportunity to make money, and by doing so increasing demand of the portfolio of the arbitrage which increases its price until the arbitrage opportunity is removed.

For instance, in the case of inflation options it is possible to create a portfolio that is an arbitrage opportunity if its initial price is not positive. Consider a scenario where a 2% and a 3% cap are bought while two 2.5% caps are (short-)sold. Then if the realised inflation is x the payoff of this portfolio is as in Table 2.1:

Inflation	payoff	
$x \leq 2$	0	none of the options are in the money
$2 < x \leq 2.5$	$x - 2 > 0$	only the 2% is in the money
$2.5 < x \leq 3$	$x - 2 - 2(x - 2.5) = 3 - x \geq 0$	the 2% and 2.5% are in the money
$x > 3$	$x - 2 - 2(x - 2.5) + x - 3 = 0$	all the options are in the money

Table 2.1: Example of the payoff of an arbitrage portfolio.

This exemplifies that the initial price of this portfolio should be positive. Let C_t denote the price of a cap option at time t, and P_t the price of a floor option at time t. Then the constraint of the price of the portfolio can be represented mathematically as: $C_t(T, 1.02) - 2C_t(T, 1.025) + C_t(T, 1.03) > 0$. This is similar to the central difference estimator of the second order derivative of C_t in the strike rate, which means that $\frac{\partial^2 C_t(T,K)}{\partial K^2} > 0$ for all K in $[0, \infty)$.
This example shows that always for equidistant grid with N strikes K_i, $i = 1 \ldots N$, it must hold that:

$$C_t(T, K_i) - 2C_t(T, K_{i+1}) + C_t(T, K_{i+2}) > 0 \qquad \forall i = 1 \ldots N - 2. \tag{2.4}$$

Similarly one can reason for caps that since a higher strike rate always means a lower payoff, the price should be lower. Or in mathematical terms:

$$C_t(T, K_i) > C_t(T, K_j) \qquad \text{if } K_i < K_j. \tag{2.5}$$

And in the same manner for floor options, but reversed:

$$P_t(T, K_i) < P_t(T, K_j) \qquad \text{if } K_i < K_j. \tag{2.6}$$

This can be interpreted again as a restriction on the derivative of the option pricing functions, from Equation 2.5 it results that $\frac{\partial C_t(T,K)}{\partial K} < 0$. From Equation 2.6 it follows that $\frac{\partial P_t(T,K)}{\partial K} > 0$. Then using the Put-Call parity gives the conclusion that the first derivative of the cap price function is bounded:

$$-e^{-rT} < \frac{\partial C_t(T, K)}{\partial K} < 0. \tag{2.7}$$

Unfortunately in the data we find examples of days with option triplets that violate Equation 2.4, and option pairs that violate Equation 2.5 or Equation 2.6.

The problem with arbitrage in the dataset is that it could cause improper RNDs, for example if Equation 2.4 is not satisfied then using a curve-fitting method could lead to negative probabilities. To remove arbitrage from the dataset several things have been suggested. To prevent issues caused by arbitrage multiple things can be done. The first option is to either remove the days with arbitrage from the dataset, or to adjust those datapoints slightly such that there is no longer any arbitrage. The latter is done in the paper by Aït-Sahalia and Duarte [2], they suggest to apply constrained least squares regression to find the vector m that solves the following for price data vector y and strike data vector x of size n:

$$\min_{m \in \mathbf{R}^n} \sum_{i=1}^{n} (m_i - y_i)^2, \tag{2.8}$$

subject to the slope and convexity constraints of Equation 2.7 and Equation 2.4:

$$-e^{-rT} \leq \frac{m_{i+1} - m_i}{x_{i+1} - x_i} \leq 0 \quad \text{for all } i = 1, \dots, n-1, \tag{2.9}$$

$$\frac{m_{i+2} - m_{i+1}}{x_{i+2} - x_{i+1}} \geq \frac{m_{i+1} - m_i}{x_{i+1} - x_i} \quad \text{for all } i = 1, \dots, n-2. \tag{2.10}$$

However, Aït-Sahalia and Duarte also note that estimators of the call pricing function will satisfy the restrictions in the first and second derivatives only if the sample is large enough [1]. Our sample size of only 13 option prices is not particularly large.

Another option is to enforce the arbitrage restrictions in the method, for example when using a curve-fitting method there could be additional constraints on the method itself such that Equation 2.4, Equation 2.5, and Equation 2.6 are satisfied.

2.2.2. Inflation option market

This project entailed collaborating with industry experts, including analysts and traders across various commercial and central banks, to explore the dynamics of the markets of inflation-linked products. The goal of these conversations was to get a better understanding of how reliable the price data is. Regrettably, the conclusion of most of the traders was that they trade options on inflation very rarely. One of the traders told us that he trades the caps and floors only ”by appointment” and that he mostly trades these options once a month and only very occasionally multiple times a week in more volatile periods.

On the other hand, inflation linked bonds (ILB) and inflation swaps are traded in a much larger volume. In multiple conversations it was explained that the most liquid inflation option is the 0% floor option. This is because it is used in replicating an ILB, providing protection against deflation by capping the coupon rate at zero and preventing the principal to decrease. The trader we spoke from Crédit Agricole estimated that around 90% of the inflation options he has traded in the past year were 0% floors.

Unfortunately, this means that the daily data we receive from Bloomberg are most likely not prices of actually traded options. Given the over-the-counter nature of this market, Bloomberg relies on market participants for option prices. Most likely, participants will give their asking price of that day if they have not traded the options. However, when Bloomberg was asked about this, they could not give a complete answer. They mentioned that they pooled the price data from multiple market participants to give one price for that day. The low liquidity of these options means that the data is prone to having errors, including potential arbitrage opportunities.

2.3. Option pricing

As explained before, there is a connection between vanilla European options and the inflation options that are used in this thesis. Without loss of generality we can set the notional amount to 1 unit of currency. This simplifies the calculations since without it we have the following, where x denotes the notional and I_t the realised inflation at time t:

$$\begin{aligned} x \cdot C_t(T,K) &= e^{-rT}\mathbb{E}_{\mathbb{Q}}(x \cdot \max(I_T - K, 0)|\mathcal{F}_t), \\ &= x \cdot e^{-rT}\mathbb{E}_{\mathbb{Q}}(\max(I_T - K, 0)|\mathcal{F}_t), \\ \Leftrightarrow C_t(T,K) &= e^{-rT}\mathbb{E}_{\mathbb{Q}}(\max(I_T - K, 0)|\mathcal{F}_t). \end{aligned} \tag{2.11}$$

This shows that the price for an inflation cap is of a similar form as the price for a European call option. Thus, available option pricing models, like the renowned Black-Scholes-Merton model, can be applied to inflation caps and floors.

The Black-Scholes-Merton model offers a closed-form solution for European call option pricing:

$$C_{BS}(S_t, K, T, \sigma, r) = S_t \Phi\left(\frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right) - Ke^{-rT}\Phi\left(\frac{\ln\left(\frac{S_t}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right). \tag{2.12}$$

This closed form solution for the option price of a European call option was introduced by Black and Scholes [5] and Merton [15]. The downside of using Black-Scholes for the purpose of inflation options is the fact that it uses the current value of the stock in determining the price of the option. This is because for stocks the current price is equal to the discounted expected value of the stock on maturity of the option under the risk neutral measure. This is not the case for inflation, therefore it is more accurate to use Black's model [4]. In 1976 Black presented a variant on the Black-Scholes model where instead of the spot price the discounted futures price, F, is used.

$$C_B(r,T,F,K,\sigma) = e^{-rT}\left(F\Phi(d_1) - K\Phi(d_2)\right), \tag{2.13}$$
$$P_B(r,T,F,K,\sigma) = e^{-rT}\left(-F\Phi(-d_1) + K\Phi(-d_2)\right), \tag{2.14}$$

where

$$d_1 = \frac{\ln\left(\frac{F}{K}\right) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \tag{2.15}$$
$$d_2 = \frac{\ln\left(\frac{F}{K}\right) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}. \tag{2.16}$$

This model is more applicable for options on interest rates or inflation due to its reliance on the discounted future price, and will therefore be used in this thesis. However, several assumptions are necessary in Black's model. Besides it being based on risk neutral valuation, the key assumptions are a constant risk-free interest rate, a perfect market without transaction costs or taxes in which it is always possible to buy or sell any quantity of the underlying asset and option, and most importantly it assumes that the changes in the price of the underlying futures contract are log-normally distributed. These assumptions are also necessary for the Black-Scholes model.

2.3.1. Implied volatility

Implied volatility is a concept in option pricing and it can be derived from the option price and observable variables, assuming the option price is arbitrage-free. It corresponds to a value $\sigma^{imp}(T, K)$, known as the implied volatility of the market price $C_t(T, K)$, such that [17]:

$$C_t(T, K) = C_B(r, T, F, K, \sigma^{imp}(T, K)). \tag{2.17}$$

This implies that the implied volatility is the value required in Black's formula to align the model price with the option price. The volatility is the standard deviation of the log-returns of the underlying. Which means it describes the spread of the changes in the futures price, and it is the only variable that is not observable in the market. Implied volatility is often interpreted as the market's expectation of future volatility. Traders and investors use implied volatility to gauge how much the market anticipates the underlying commodity's price to fluctuate over the life of the option. High implied volatility may suggest uncertainty or anticipated significant price movements, while low implied volatility may indicate a more stable market outlook.
In Equation 2.13 and Equation 2.14 the parameters are the interest rate r, time to maturity T, futures price of inflation F, strike rate K, and the volatility σ. To compute the implied volatility various root-finding algorithms can be used, such as bisection or Newton's method.

Because the volatility is defined as the standard deviation, the implied volatility must always be non-negative. A different property of the implied volatility is the ”smile”. While Black's model assumes a constant implied volatility for options with the same maturity, real-world financial dynamics often reveal a more intricate picture. The implied volatility smile refers to the pattern where implied volatility varies with different strikes but for the same expiration date. This departure from a flat volatility surface is akin to a subtle smile on a graph, hence the evocative term ”smile”. At its core, the implied volatility smile suggests that market participants believe there are different levels of risk or uncertainty associated with various strike prices, even for options expiring at the same time.

3
Literature Study

The goal of this thesis is to compare methods of deriving risk neutral densities implied by inflation option prices. To be able to make the comparison first the different methods need to be studied. To do that a lot of scientific papers and working papers written by researchers working at central banks were read. The first techniques to derive a RND from option prices were not invented to be used on inflation data. In the early days these techniques were mainly used for equity, later for interest rates and recently also for inflation. This chapter delves into the mathematical foundations of various methods to derive the risk-neutral densities, exploring the origins and key contributors to each method.

Following the book by Jackwerth [12] these methods can be split up in two main categories, parametric methods and non-parametric methods. In short, parametric models assume that the inflation is distributed following a certain probability distribution, to then attempt to estimate its parameters. On the other hand, non-parametric methods fit the risk-neutral probability distribution either pointwise or build it up from linear segments.

3.1. Parametric Methods: Flexibility and drawbacks

The earliest research into this topic tried to find a density from a known family that best fit empirical RNDs [9]. This is called a parametric method. In general, it assumes a known distribution a priori. For example the Gamma distribution, to then try to find the parameters, α and β, such that the estimated prices of the options are as close as possible to the market prices. The prices would be estimated using the discounted expected payoff under the assumed distribution. To fit the parameters multiple constraints can be used, most frequently used is minimising the sum of squares:

$$\min_{\alpha,\beta} \sum_{i=1}^{m} (\hat{C}_i - C_i)^2 + \sum_{j=1}^{n} (\hat{P}_j - P_j)^2. \tag{3.1}$$

Where $\hat{C}_i$ and $\hat{P}_j$ are the estimated prices, and C_i and P_j the market prices.
This method has drawbacks if we use a parametric probability distribution that is not flexible enough for matching the observed option prices [12]. Within the parametric methods there are three groups identifiable, for a large overview I would like to refer the reader to Jackwerth (2004) [12].
The first group of methods is the Expansion Methods. They start with a simple known probability distribution and then add correction terms to it. These correction terms are often not guaranteed to preserve the integrity of the probability distribution, when using this method it is sensible to always check that the resulting distribution integrates to one and is strictly positive [12].
A second group is called the Generalised Distribution Methods, these methods use more elaborate distribution functions which have more parameters, they often add skewness and kurtosis parameters. Skewness is a measure of asymmetry in a distribution, putting more probability on low values than high values for example. Kurtosis measures the "tailedness" of a probability distribution, indicating the degree to which it deviates from a normal distribution by assessing the concentration of data in the tails.

The third group are the Mixture Methods. In these methods a probability distribution is created by adding several simple probability distributions with different mixing probabilities. This adds flexibility but comes at the cost of increasing the number of parameters that need to be fitted. For example, a mixture of three log-normal distributions requires 8 parameters, 2 for each of the 3 log-normals and another 2 to determine the weights. The weights need to sum to 1 so there only need to be 2 parameters for that. In using these mixture methods one should be careful that these methods are prone to over-fitting to the data[12].
A widely used example of these parametric methods is the sum of two log-normal distributions [16][7][3][20]. The mixture of two log-normals is given by:

$$g_{t,T}(x) = \lambda L(\alpha_1, \beta_1) + (1-\lambda)L(\alpha_2, \beta_2), \tag{3.2}$$

where $\lambda, \alpha_1, \beta_1, \alpha_2,$ and β_2 need to be estimated. The estimated call and put prices are then given by[7]:

$$\hat{C}_{i,t}(K_i, T) = e^{-r(T-t)} \int_{K_i}^{\infty} (s - K_i)(\lambda L(\alpha_1, \beta_1) + (1-\lambda)L(\alpha_2, \beta_2))ds, \tag{3.3}$$

$$\hat{P}_{j,t}(K_j, T) = e^{-r(T-t)} \int_{0}^{K_j} (K_j - s)(\lambda L(\alpha_1, \beta_1) + (1-\lambda)L(\alpha_2, \beta_2))ds \tag{3.4}$$

To fit the parameters multiple constraints can be used, most frequently used is minimising the following:

$$\min_{\lambda,\alpha_1,\beta_1,\alpha_2,\beta_2} \sum_{i=1}^{m} (\hat{C}_{i,t} - C_{i,t})^2 + \sum_{j=1}^{n} (\hat{P}_{j,t} - P_{j,t})^2. \tag{3.5}$$

However, in the same paper it is concluded that a non-parametric curve-fitting technique for estimating pdfs is an improvement upon this method [7].

3.2. Non-parametric Methods

The second approach to derive a RND is to use non-parametric methods. Although they are called non-parametric, in reality they require a lot more parameters. This is because instead of picking a few parameters for an assumed RND, these methods fit the RND either point-wise through the observed option prices or builds it up from line segments. Fitting the RND is rarely done however, since it is difficult to constrain the probability distribution to be positive and to integrate to 1. In the few cases that it is done maximum entropy methods are most often used[12].
Two other groups of Non-Parametric methods build on the result of Breeden and Litzenberger: Kernel methods and Curve-fitting methods. Breeden and Litzenberger found that the risk-neutral density is obtainable from option prices if a continuum of strikes were available. They showed that the RND can be computed by differentiating the option price function, $C_t(T, K)$, with respect to the strike K [6]. Since under the risk neutral measure it must hold that at time t the price of a European call option is equal to its discounted expected value [8]:

$$C(T, K) = e^{-rT} \mathbb{E}_{\mathbb{Q}}(\max(S_T - K, 0) | \mathcal{F}_t). \tag{3.6}$$

Rewriting this gives a more workable expression, where g_T is the risk neutral distribution function of the value of the underlying at time T:

$$\begin{aligned} C(T, K) &= e^{-rT} \mathbb{E}_{\mathbb{Q}}(\max(S_T - K, 0) | S_t), \\ &= e^{-rT} \int_0^{\infty} \max(s - K, 0) g_T(s) ds, \\ &= e^{-rT} \int_K^{\infty} (s - K) g_T(s) ds. \end{aligned} \tag{3.7}$$

Breeden and Litzenberger found that by differentiating both sides of Equation 3.7 with respect to K the RND can be computed [6][13]:

$$\frac{\partial}{\partial K} C(T, K) = e^{-rT} \left(\int_0^K g_T(s) ds - 1 \right). \tag{3.8}$$

And since from the definition of probability densities it follows that the cumulative distribution function G_T is equal to the integral of g_T this can be used to obtain:

$$G_T(K) \equiv \int_0^K g_T(s)ds = 1 + e^{rT}\frac{\partial}{\partial K}C(T,K). \tag{3.9}$$

Then by differentiating again with respect to K the following result of Breeden and Litzenberger is obtained:

$$g_T(K) = e^{rT}\frac{\partial^2}{\partial K^2}C(T,K). \tag{3.10}$$

The derivation of the RND from the put option price function works similarly, for that derivation I refer the reader to [6].

This result means that it is possible to retrieve the RND from the option price function, if the price function is twice differentiable with respect to the strike. This leads to the question what the correct function is for $C(T,K)$, as all that is known are the prices of the options traded in the market. Since this is a discrete data set of only a few strikes it is required to determine $C(T,K)$ by means of estimation or inter- and extrapolation. The most used non-parametric methods therefore focus on finding the function $C(T,K)$ instead of the RND directly.

3.2.1. Curve fitting

The most obvious way of using the result of Breeden and Litzenberger would be to fit a curve through the available option prices. However, there are two drawbacks to this method. The first drawback is that since the pricing function is different for puts and calls one needs to fit a curve through two seperate sets of data. Secondly, option prices vary greatly across strike prices, deep in-the-money calls are valued as high as the unerlying itself, while out-of-the-money options are valued close to zero [12]. To deal with both these drawbacks simultaneously Shimko suggested to use the implied volatilities across all strikes [18]. He fitted a second order polynomial through the implied volatilities, and then calculated the function of the call option prices using Black's model. This means that he used Black's model from Equation 2.13 to first transform the price-strike pairs into implied-volatility pairs, then fitting a second order polynomial, to then use the same formula to return to a function $C(T,K)$ which is twice differentiable.
Another way is to calculate the second derivative of Black's formula, and then input the fitted function directly. To do that first differentiate Equation 2.13 with respect to the strike K which results in the following:

$$\begin{aligned} e^{rT}\frac{\partial C_B}{\partial K} &= F\phi(d_1)\frac{\partial d_1}{\partial K} - \Phi(d_2) - K\phi(d_2)\frac{\partial d_2}{\partial K}, \\ &= F\phi(d_1)\frac{\partial d_1}{\partial K} - \Phi(d_2) - F\phi(d_1)\frac{\partial}{\partial K}(d_1 - \sigma\sqrt{T}), \\ &= F\phi(d_1)\sigma'\sqrt{T} - \Phi(d_2). \end{aligned} \tag{3.11}$$

Where $\phi(x)$ denotes the standard normal probability density function, and where the following result was used:

$$\begin{aligned} \phi(d_2) &= \phi(d_1 - \sigma\sqrt{T}), \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(d_1-\sigma\sqrt{T})^2}, \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}d_1^2 + d_1\sigma\sqrt{T} - \frac{1}{2}\sigma^2 T}, \\ &= \phi(d_1)e^{\ln\left(\frac{F}{K}\right) + \frac{1}{2}\sigma^2 T - \frac{1}{2}\sigma^2 T}, \\ &= \frac{F}{K}\phi(d_1). \end{aligned}$$

Then if Equation 3.11 is differentiated again it gives the following result:

$$\begin{aligned}
e^{rT}\frac{\partial^2 C_B}{\partial K^2} &= F\sqrt{T}\left(-d_1\phi(d_1)\sigma'\frac{\partial d_1}{\partial K}+\phi(d_1)\sigma''\right)-\phi(d_2)\frac{\partial d_2}{\partial K},\\
&= F\sqrt{T}\phi(d_1)\left(\frac{d_1\sigma'}{K\sigma\sqrt{T}}+\frac{(\sigma')^2 d_1 d_2}{\sigma}+\sigma''\right)-\frac{F}{K}\phi(d_1)\left(\frac{-1}{K\sigma\sqrt{T}}-\frac{\sigma' d_1}{\sigma}\right),\\
&= F\phi(d_1)\left(\frac{1}{K^2\sigma\sqrt{T}}+\frac{2d_1\sigma'}{K\sigma}+\frac{(\sigma')^2 d_1 d_2\sqrt{T}}{\sigma}+\sigma''\sqrt{T}\right).
\end{aligned} \tag{3.12}$$

Here the property of the standard normal PDF was used that:

$$\frac{\mathrm{d}}{\mathrm{d}x}\phi(x) = -x\phi(x).$$

In the result of Equation 3.12 the choice of how to fit the volatility is not yet made, it does show that both the first and second order derivative of σ with respect to K are required.

Besides transforming the option price data into the implied volatility it is also possible to use another transformation of variables. It is also useful to transform the price-strike pairs into implied volatility-Delta space [14]. In this context the definition of Δ is that of the Greeks in option pricing, which usually means $\frac{\partial C}{\partial S_t}$, where S_t is the underlying stock. Because in the case of Black's model the price of the underlying is replaced by the discounted forward price the definition of Δ becomes $\frac{\partial C}{\partial e^{-rT}F} = e^{rT}\frac{\partial C}{\partial F}$. Which gives the following:

$$\begin{aligned}
\Delta = e^{rT}\frac{\partial C_B}{\partial F} &= \Phi(d_1)+F\phi(d_1)\frac{\partial d_1}{\partial F}-K\phi(d_2)\frac{\partial}{\partial F}(d_1-\sigma\sqrt{T}),\\
&= \Phi(d_1)+F\phi(d_1)\frac{\partial d_1}{\partial F}-K\frac{F}{K}\phi(d_1)\frac{\partial d_1}{\partial F},\\
&= \Phi(d_1).
\end{aligned} \tag{3.13}$$

In some cases Δ is defined as $\frac{\partial C}{\partial F}$, which gives a slightly different result of $\Delta = e^{-rT}\Phi(d_1)$. However since e^{-rT} is constant for every individual day, this altered definition of Δ makes no difference in this context since it is a constant translation.
From Equation 3.13 it follows that Δ is a function of d_1 which is a function of K, σ, T and F. This means that it is first required to find the implied volatility. Then, since K, T, and F are observable the value for Δ can be directly computed. A curve can than be fitted through these newfound implied volatility-delta pairs.

This translation is purely used for fitting convenience [19], since the delta is bounded on the interval $[0, 1]$ there is less of an issue with tails and extrapolation.

3.2.2. Kernel Methods

Kernel based methods are local techniques for estimating the option price function at any strike in its domain, they use a weighted average of the observable prices, Y_i, to yield fitted values via [10]:

$$\hat{C}(K) = \sum_{i=1}^{n} w_i(K)Y_i, \tag{3.14}$$

where the weights w_i are functions of the strike that decline for increasing distance from the data points K_i, and $\frac{1}{n}\sum_{i=1}^{n} w_i(K) = 1$. These weights are constructed using kernel functions. There are a multitude of different kernel functions [10], for example the Quartic kernel function $\mathcal{K}(u) = \frac{15}{16}(1-u^2)\cdot\mathbb{1}(|u|\leq 1)$. Often these kernel functions are adjusted by using a bandwidth h such that $\mathcal{K}_h(u) = \frac{1}{h}\mathcal{K}(\frac{u}{h})$. Then, the simplest way of choosing the weights is by using the Nadaraya-Watson weights

$$w_i(K) = \frac{\mathcal{K}_h(K-K_i)}{\sum_{i=1}^{n}\mathcal{K}_h(K-K_i)} \tag{3.15}$$

Then by using Equation 3.14 a function for the option prices is found, which can then be differentiated numerically to obtain the RND. In a similar way to the previous section it is possible to first transform the data to strike-implied volatility space, to ease the fitting of the function. This way estimators of the second derivative function are constructed by twice differentiating the estimator of the function. However, these estimators have inferior statistical properties [10]. To improve this method it is possible to find a better estimator by minimising a locally weighted least squares regression of order p:

$$\min_{\beta} \sum_{i=1}^{n} \left(Y_i - \sum_{j=0}^{p} \beta_j(K)(K - K_i)^j \right)^2 \mathcal{K}_h(K - K_i). \tag{3.16}$$

The solution $\hat{\beta}_0(K)$ is an estimator of the option price at point K, while $j!\hat{\beta}_j(K)$, with $j = 1, \ldots, p$, are the estimated derivatives at point K. This method can therefore be used to find the second derivative of the option price function. Moreover, it can again also be used on the transformed implied volatility data. Then, using Equation 3.12 the RND can be obtained.
The choice of the bandwidth h influences the result greatly, other parameters such as the chosen kernel function have less influence on the final result in practice [10].

3.3. Tails

4 Methodology

In this chapter all choices made in the implementation of the various methods will be explained, as well as some additional explanations and calculations that were required to implement the methods explained in the previous chapter. All implementations were done in Matlab.

Because time was a limiting factor it was not possible to implement every method. It was therefore chosen to implement only the most promising methods: Shimko's method, $\sigma(K)$ curve-fitting using various splines, the Delta method, and the Kernel method.

4.1. Data

To be able to derive the RND from inflation option prices each method relies on daily data of option prices. The data that was used for this thesis was retrieved from the BVOL database of Bloomberg on 11-05-2023, it contains prices of inflation options on the Harmonised Index of Consumer Prices excluding Tobacco (HICPxT) for every business day since 01-01-2013. This pricing data consists of eight floor options with equidistant strikes from -1% to 2.5%, and eight cap options with equidistant strikes ranging from 1.5% to 5%. This means that there are both cap and floor options for the strikes 1.5, 2, 2.5. In this data set all options have the same maturity of $T = 5$ years. All prices are denoted in basis-points. These options are zero-coupon options, meaning that they do not pay out anything during the term of the contract. Only at the very end of the contract will it pay out the difference between the realised inflation during the contract and the strike rate, for the cap only if the inflation is higher than the strike rate and for the floor only if it is lower than the strike rate.

For most of the methods it is also required to use the inflation swap rate and the risk free rate. These two data sets were also retrieved from Bloomberg, for the same days as the option prices. For the risk free rate the euro area overnight index swap was used. For every PDF only one day of data is required, thus the input for every method is a single vector of size 18.
Since there are three overlapping strike rates for which there exist both a cap and floor option, it was decided to work only with out-of-the-money (OTM) options. Since most of the value for deep in-the-money (ITM) options is intrinsic value, it is common to eliminate ITM options and to extract RNDs using only OTM calls and puts [9]. Thus, of the 16 total option prices per day, only 13 are effectively used to derive the RND.
Additionally, to be able to compare the methods the data on realised inflation was required. This was also retrieved from Bloomberg as monthly data, and interpolated linearly to get a daily value for the inflation index.
The HICP is used in the euro area to measure the consumer price inflation. The version of the HICP that is used for options is the HICP excluding Tobacco, this is because of French legislation which forbids using tobacco as an underlying for certain financial instruments. Using option prices on different underlying indices should work in the same manner, it was decided to work with euro area data for this thesis because the options on European inflation are traded more frequently than their US or UK counterparts.

As explained in chapter 2, the market for these options is not perfect. The interviewed market participants told that trades are not made every day. Yet, the prices Bloomberg supplies for every strike rate change daily. Therefore it is uncertain whether the supplied data is from actually traded options or from a. When Bloomberg was asked about this they said that they collect their data from market participants daily, then taking a weighted average of the contributed market sources to get to a final price. The Bloomberg help desk could not disclose whether the supplied data was from actually traded options or a asking price from the pricing algorithm used by the market participants. This indicates that the dataset is not entirely reliable. This is an unfortunate consequence of the age and relative inactivity of the inflation option market.

The implication of the low liquidity of the options is that the data is prone to having errors due to for example arbitrage opportunities. Unfortunately the examples of arbitrage as explained in subsection 2.2.1 are present in the dataset. The problem with arbitrage in the dataset is that it could cause improper RNDs, to prevent this it was decided to apply the method of Aït and Sahalia explained in the previous chapter. Yet, to show the difference it all methods were also applied to the original data.

4.2. Implemented methods

It was decided early on in this project to implement some of the methods found in the literature. The decision of which methods to implement was based on the findings of the literature review, selecting the most promising and the most used methods. It was decided to implement various curve fitting methods, since they were praised the most in various scientific papers. The theory behind these methods is already explained in the previous chapter. In the following sections the various decisions that were required will be justified. Furthermore, certain methods posed challenges in terms of implementation, as there was ambiguity regarding the differentiation. In specific instances, method sources provided guidance limited to recommending a particular fitting procedure without explicit instructions.

The same method to obtain the implied volatility from option prices was used for all of the following methods. A simple bisection algorithm was applied to find the volatility such that the option price data was equal to the price from Black's model Equation 2.13.
Next, the OTM options were selected by comparing the strike rates with the swap rate on that day. This way every method is given the same dataset of 13 implied volatilities and corresponding strike rates per day.

4.2.1. Shimko's method

The first few implemented methods apply the curve-fitting method from subsection 3.2.1 to find a function for the relation between the implied volatility and the strike rate. Firstly, the method by Shimko was implemented as a benchmark. This model was already in use by DNB and therefore it was possible to reuse a lot of their implementation. It is the most simple method since it finds a second order polynomial to fit the data, and thus only requires three coefficients a, b, c such that:

$$\hat{\sigma}(K) = aK^2 + bK + c. \tag{4.1}$$

As a result the derivatives are also very straightforward:

$$\hat{\sigma}'(K) = 2aK + b, \tag{4.2}$$
$$\hat{\sigma}''(K) = 2a. \tag{4.3}$$

In Matlab the values for a, b and c were found by applying the *regress* function which finds the optimal values for the coefficients such that the sum of squares is minimised:

$$\sum_{i=1}^{n} (\hat{\sigma}(K_i) - \sigma_i)^2$$

Then if the coefficients are found, all three functions Equation 4.1, Equation 4.2 and Equation 4.3 can be used in Equation 3.12 to find the RND.

4.2.2. Splines

Secondly, to fit a function for the relation between the implied volatility and the strike rate splines were used.

4.2.3. Spline methods

4.2.4. Delta method

4.2.5. Kernel

4.3. Tails of the distributions

4.4. Comparison Methods

4.4.1. Likelihood method

4.4.2. Visual method

4.4.3. Moments

4.4.4. Swaprate

5 Results

In this chapter the results of the research will be presented, the next chapter will discuss the meanings of these results. Firstly, a few examples of the risk-neutral densities resulting from the various implemented methods will be shown. Which will already display some weaknesses. Secondly, the results of the likelihood method will be presented.

5.1. Visual comparison

Firstly, a lot of the characteristics can be seen in plots of the implied volatility and RND. However, since the methods produce a RND for every day of the dataset, and since there are multiple different implemented methods it is impossible to show figures of all of the resulting RNDs. Therefore only a few figures will be shown here to display and explain the differences between the methods. Since it is impossible to show the resulting RND of every day, six dates were selected to show the results.

The resulting fit of the implied volatility and the corresponding RND of Shimko's method can be found in Figure 5.1 and Figure 5.2. In both figures the upper tail does not go to zero. This problem is because the implied volatility exhibits quadratic growth by definition. Other than that it is a very robust method since the arbitrage problems in the data are less influential. The found function for the implied volatility does not necessarily go through all the given data points exactly, but instead allows for more flexibility in capturing the overall trend. Thereby ignoring if one of the datapoints is causing arbitrage.

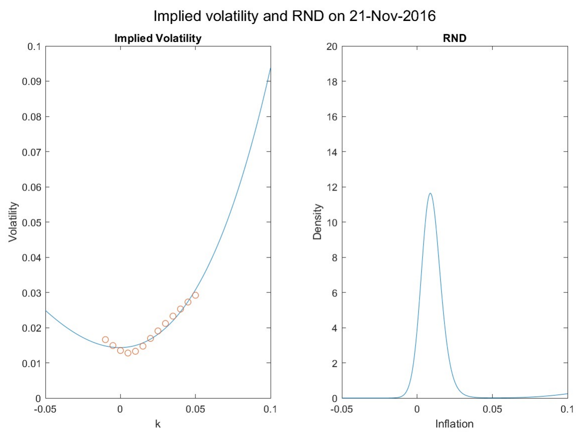


Figure 5.1: Results from Shimko's method on 21-11-16 with tails.

Secondly, an example of the results of the three different cubic spline methods are shown in Figure 5.3, Figure 5.4 and Figure 5.5. The differences between the methods is clearly visible in the figures of the implied volatility, while the RNDs are fairly similar. The only obvious thing is that with the constant tails the RND is negative close to the lowest and highest strike rates. This is a consequence of the

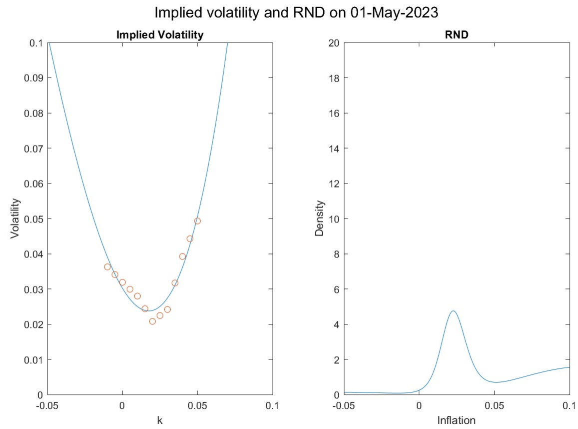


Figure 5.2: Results from Shimko's method on 01-05-23 with tails.

constraint that was put on the spline that the first derivative on the outer knot points should be 0. This was done such that the transition from the interpolated data to the tails is smooth. However, because of this the second derivative has to become negative close to the outer knot points. This is less of a problem for the linear tails because there the constraint on the outer knot points was that the second derivative should be zero there.

Another thing that should be noted from Figure 5.3 is that the implied volatility would become negative if the lower tail was extended further. This is a problem since by definition the implied volatility should always be positive. On certain days this did occur. Moreover, similarly to the issue that was encountered with Shimko's method, letting the cubic spline extrapolate the upper tail can also lead to issues because the implied volatility grows too large. This can be seen in Figure 5.6. On that day there are also issues with arbitrage, and since the cubic spline uses the data points as knot points it goes straight through all data points. This means that the implied volatility is not a smile as expected, and thus the second derivative of the implied volatility with respect to the strike rate becomes negative enough to make Equation 3.12 negative. Using linear tails, such as in Figure 5.7 solves the issue with the upper tail, but not the arbitrage yet.
The fact that some RNDs, such as Figure 5.7, seem to be built up of line segments is a consequence of the cubic property of the splines. Since the spline consists of segments of cubic polynomials, the second derivative consists of straight line segments. The cubic spline is defined such that the function and its first and second derivatives are continuous. However, this means that the cubic spline is not always smooth, and it follows from Equation 3.12 that the RND is also not smooth. Moreover, if the value of the second derivative differs greatly between two consecutive segments the transition between these segments can be very influential in the resulting RND.

In an attempt to solve the arbitrage issues smoothing splines are used, as explained in subsection 4.2.2 the difference is that this makes the function a lot smoother and it doesn't always go through the datapoints. As can be seen in Figure 5.9, this does help in avoiding certain issues caused by arbitrage. However, one should be careful in trusting this method, it does not guarantee that issues caused by arbitrage are prevented in all situations. It merely makes the function smoother, and thus it follows the general trend better. The smoothing also prevents the pointiness encountered in cubic splines.

The last spline method that was implemented applied a linear tail to the smoothing spline. This did not always lead to large differences, but a comparison can be made between Figure 5.10 and Figure 5.11.

The Delta method showed very differently looking figures. As can be seen in Figure 5.12 and Figure 5.13, the implied volatility is now fit as a function of Δ. For this fit splines were used again, making the fitted function $f(\Delta)$ continuous as well as its first and second derivative with respect to Δ. However, since the pricing function is differentiated with respect to the strike rate to retrieve the RND, f will also be differentiated with respect to the strike rate. These derivatives are not continuous, resulting in wrong looking RNDs. However, the shape of both RNDs are comparable to the results of the spline method in Figure 5.8 and Figure 5.7 respectively.

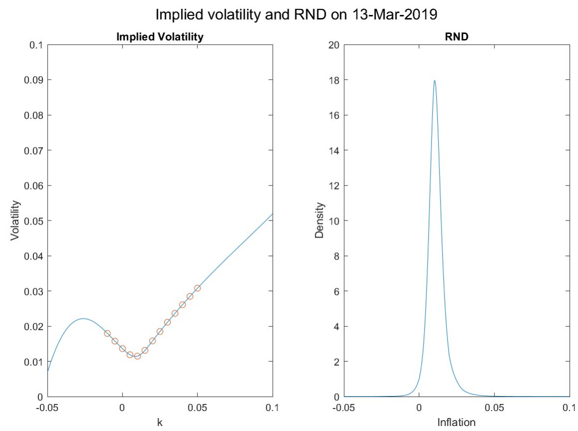


Figure 5.3: Results from the cubic spline method on 13-03-19 with regular tails.

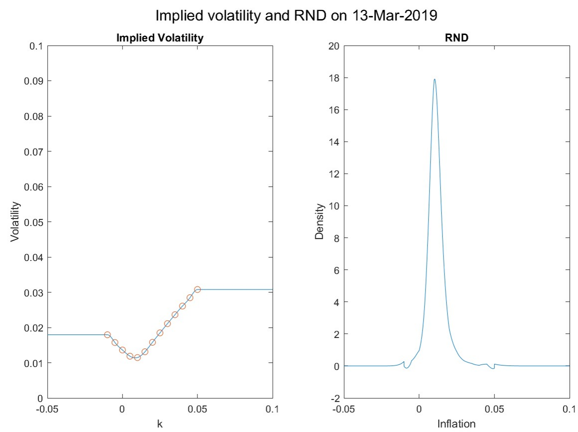


Figure 5.4: Results from the cubic spline method on 13-03-19 with constant tails.

It should also be noted that since the method defines the range of Δ, the length of the tails cannot be determined directly. As can be seen in Figure 5.13 in some cases the fitted tail is very long. There were no problems encountered with the tails using the Delta method, this property was also the reason for which most of the literature sources praised this method.
However, as is clear from Figure 5.13, this method is also heavily influenced by problems introduced by arbitrage. Another problem that was encountered with the Delta method was that sometimes when the data was transformed into Δ, the values for Δ were very close together. As can be seen in Figure 5.14, this results in very weird fits by the cubic splines. Without getting ahead of the findings explained in the next section, these issues led to extremely high values for the skewness and kurtosis.

A different kind of fitting method might help prevent these issues, but the problems introduced by transforming back and forth from Δ cannot be resolved by a different fit.
Lastly, the kernel method was very slow in implementation. In comparison with the previous methods, which all took around 90 seconds to compute the RND for all 2647 days, the kernel method took a lot longer with approximately 20 minutes. That being said, the resulting RNDs look a lot smoother, as can be seen in Figure 5.15 and Figure 5.16. Moreover, issues resulting from arbitrage are less of a problem with this method because the fitted line does not go exactly through the datapoints. This does depend on the size of the bandwidth, smaller size bandwidths follow the data more closely.
The only issue with this method is that it is impossible to set a condition on the slope of the endpoints such as in the spline method. This means that when adding the tails to the implied volatility, the slope of the function at the outer datapoints is used. As can be seen in Figure 5.17, this can lead to jumps in the RND. However, since this method finds a value for the function by taking a weighted average of

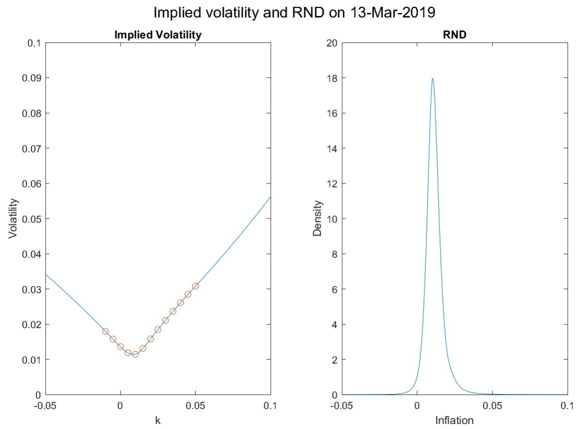


Figure 5.5: Results from the cubic spline method on 13-03-19 with linear tails.

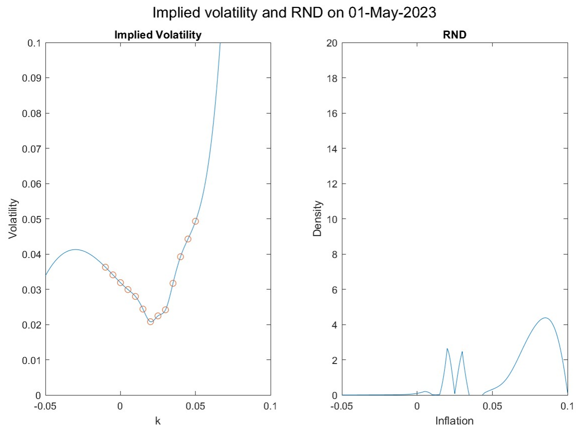


Figure 5.6: Results from the cubic spline method on 01-05-23 with regular tails.

the data that is close by, it is very bad at extrapolating. Thus it is required to fit a tail.

5.2. Likelihood method

The results of the likelihood method as explained in section 4.4 can be found in Table 5.1. There are four results per method. A distinction has been made between the original dataset and the preprocessed dataset as explained in subsection 2.2.1. Moreover, the results of including or excluding the tails of the distribution are compared. For the tails the Delta method used the grid $\Delta \in (\frac{1}{1500}, \frac{1499}{1500})$, all other methods used $K \in (0.95^T, 1.10^T)$. When the tails were excluded the grid ranged from the lowest strike rate or Δ to the highest strike rate or Δ.

On first glance, it is clear that the Delta method scores best except in the case of tails with preprocessed data. In that case there was at least one negative likelihood value. After that the kernel methods score best, with the best scores for bandwidth $h = 0.2$ and $h = 0.1$. And lastly, the smoothing cubic splines work the best out of the $\sigma(K)$ methods.

Overall, the Kernel methods seem the most reliable, because of the relatively high scores and no negative likelihoods.

5.3. Moments

The moments can be plotted together in the same graph to show differences between the methods. This time for the Kernel method only the bandwidth $h = 0.2$ was used, and for the Spline method the smoothing spline with linear tails and $p = 0.9999$ was used. When looking at the expectation in Figure 5.18, the methods seem to be following the same trend. Moreover, Shimko's method is generally

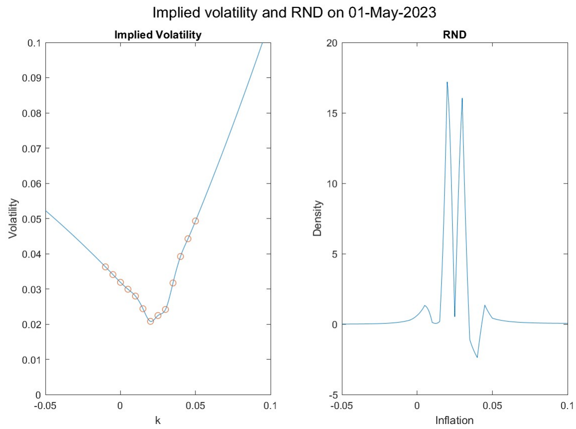


Figure 5.7: Results from the cubic spline method on 01-05-23 with linear tails.

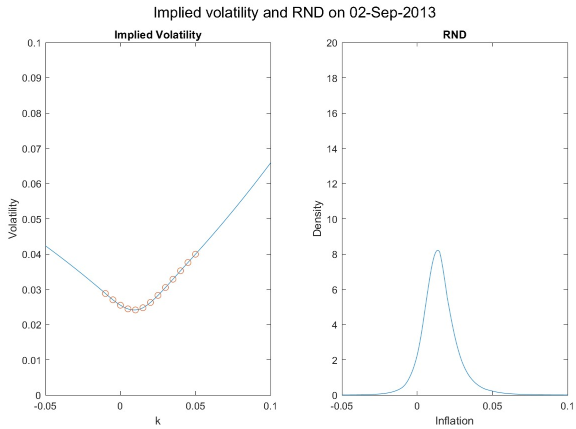


Figure 5.8: Results from the cubic spline method on 02-09-13 with linear tails.

higher than the others followed by the Kernel method, and the expectation of the Spline and Delta methods are close to equal for most days. The same goes for the variance, Shimko's is highest followed by Kernel and then spline and Delta close together the lowest. The later dates in the data set do show different behaviour than the earlier part of the dataset. The reason for this is unclear, it might be linked to the higher levels of inflation in recent times. As explained in subsection 2.2.2, the increase in inflation led to more activity in this market. The data also shows most arbitrage in the later days of the dataset.

There is not a lot to tell from the graphs of the skewness and kurtosis in Figure 5.18, since the Delta method has very large outliers. This is already shown in Figure 5.14, therefore the days where the skewness and kurtosis were too large were removed as shown in Figure 5.19. To interpret the results of the skewness and kurtosis it should be noted that these are heavily influenced by the tails of the distributions. Since the tails are often estimated in this thesis this can give a distorted view. However, it is still possible to compare the various methods. From Figure 5.19 it becomes clear that the skewness was positive for the largest time, when the inflation expectation was very low in the beginning of 2020 the skewness was negative for the Delta method. And again, in the last days of the dataset the skewness was negative for multiple methods. This comparison also shows that the skewness was close to zero for the Kernel and Delta method for most days, while it was a lot larger for Shimko's and the spline method. The positive skewness means that the upper tail is longer or fatter compared to the lower tail, meaning there is higher expected probability of large inflation than of deflation.
The Kurtosis is also positive most of the time. Meaning that the tails are fatter than that of the normal distribution. This is especially the case for the Spline method, which was to be expected since in chapter 3 it was explained that a linear extrapolation of the implied volatility leads to log-normal tails.

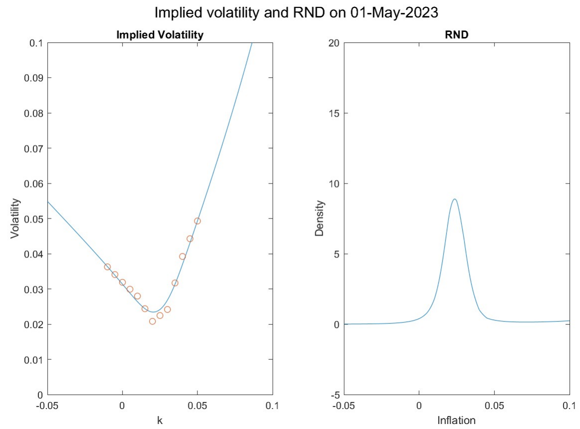


Figure 5.9: Results from the smoothing spline method on 01-05-23 with regular tails and $p = 0.9999$.

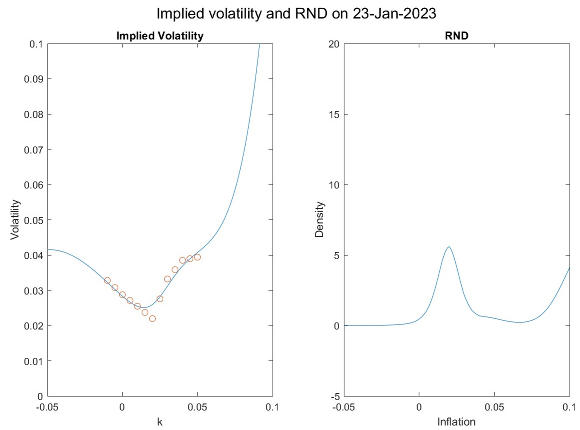


Figure 5.10: Results from the smoothing spline method on 23-01-23 with regular tails and $p = 0.9999$.

As suggested by economists at DNB the swap rate can also be used to measure inflation expectation. To compare the results of the methods with the swap rate Figure 5.20 was created. In this graph it is clear that the Spline and Delta method follow the swap rate very closely, and the other methods overestimate it.

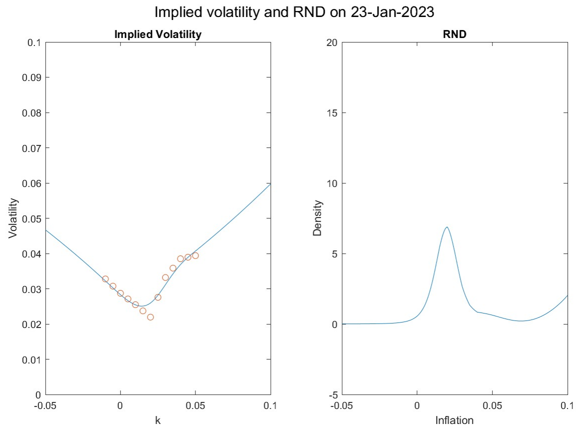


Figure 5.11: Results from the smoothing spline method on 23-01-23 with linear tails and $p = 0.9999$.

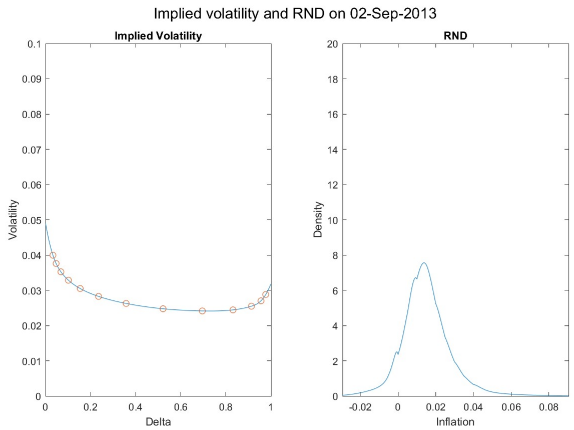


Figure 5.12: Results from the delta method on 02-09-13.

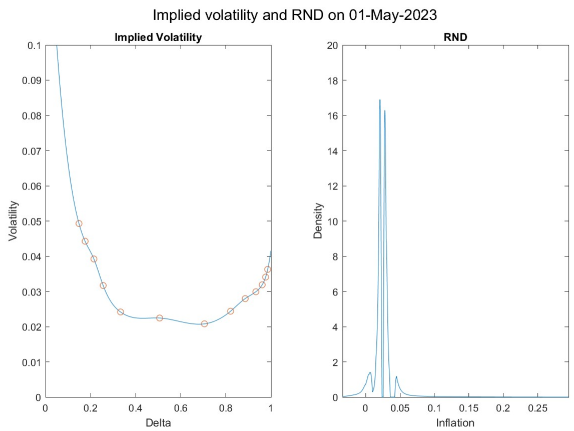


Figure 5.13: Results from the Delta method on 01-05-23.

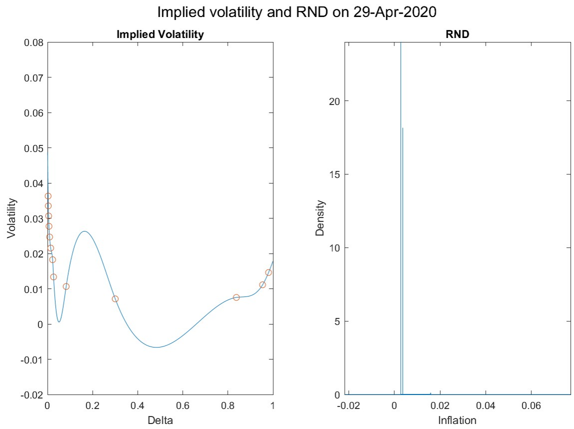


Figure 5.14: Results from the Delta method on 29-04-20.

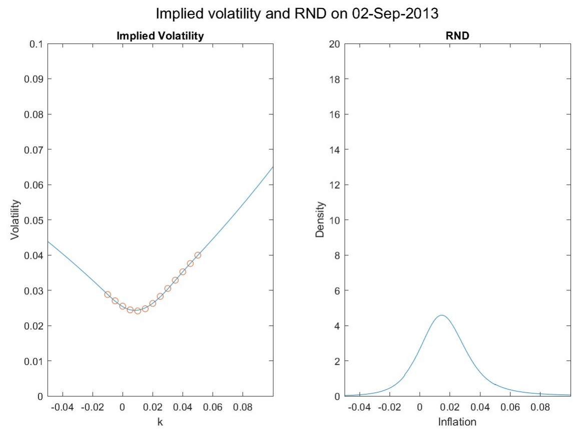


Figure 5.15: Results from the Kernel method with bandwidth $h = 0.2$ on 02-09-13 with linear tails.

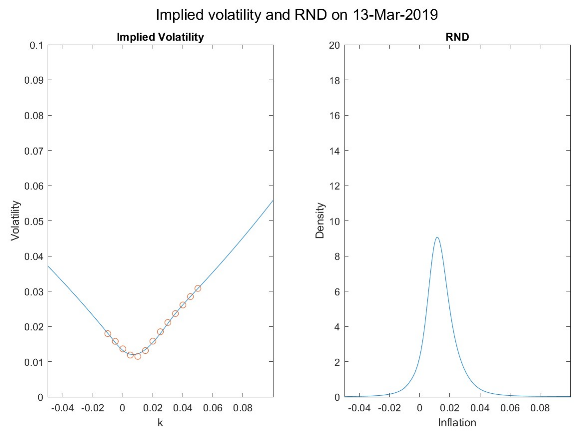


Figure 5.16: Results from the Kernel method with bandwidth $h = 0.2$ on 13-03-19 with linear tails.

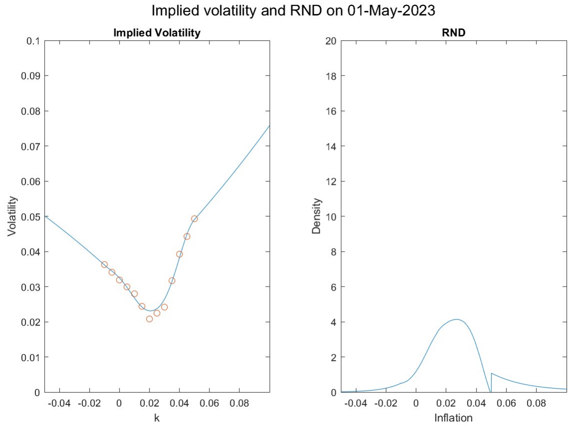


Figure 5.17: Results from the Kernel method with bandwidth $h = 0.2$ on 01-05-23 with linear tails.

	original data		preprocessed data	
	no tail	tail	no tail	tail
Shimko	53.78	50.22	54.56	48.06
Cubic Splines	56.21	$-\infty$	50.43	$-\infty$
Cubic Splines constant tail		55.91		50.39
Cubic Splines linear tail		55.51		49.58
Smoothing Cubic Splines	57.10	56.55	58.17	57.57
Smoothing Cubic Splines linear tail		50.22		57.55
Delta	72.63	69.72	78.60	$-\infty$
Kernel $h = 0.1$	63.74	60.95	63.68	59.84
Kernel $h = 0.2$	64.26	61.03	64.12	59.21
Kernel $h = 0.3$	64.10	60.09	63.86	57.19
Kernel $h = 0.4$	64.05	59.99	63.67	56.20
Kernel $h = 0.5$	64.03	60.00	63.62	56.03

Table 5.1: Results of the Likelihood method.

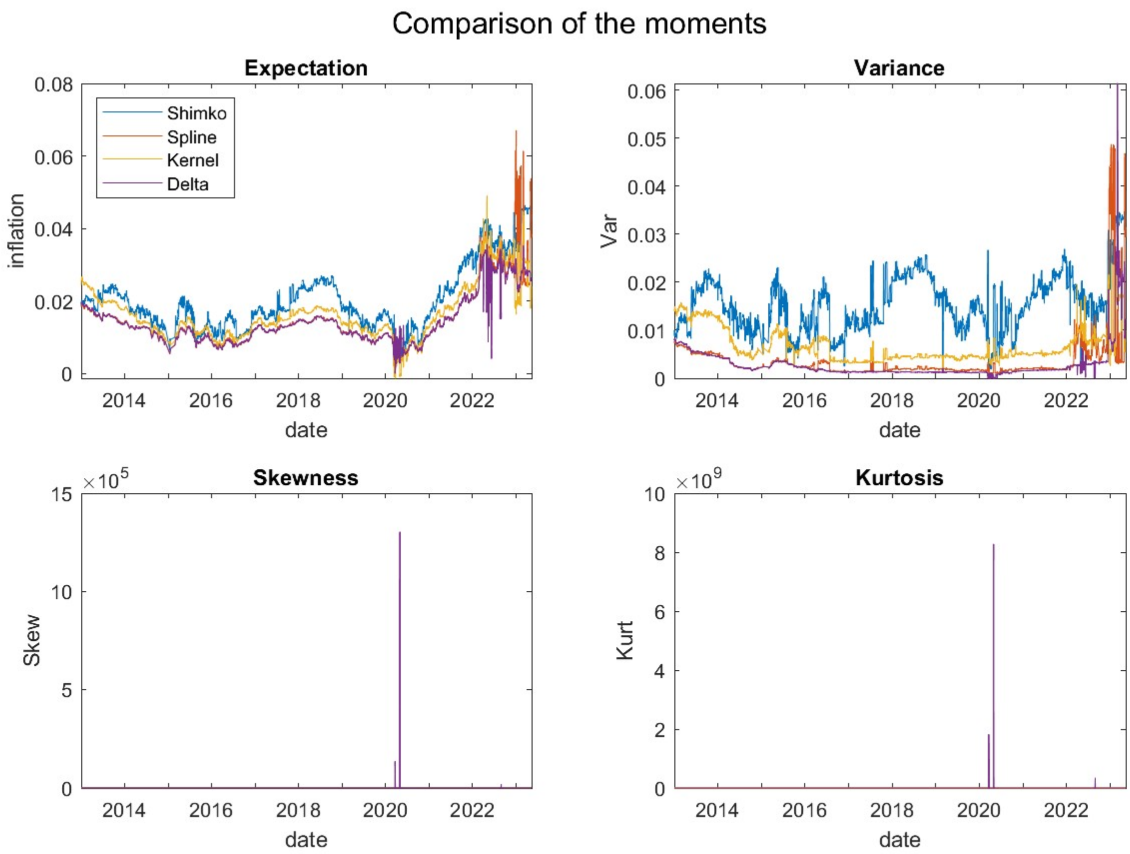


Figure 5.18: Comparison of the moments.

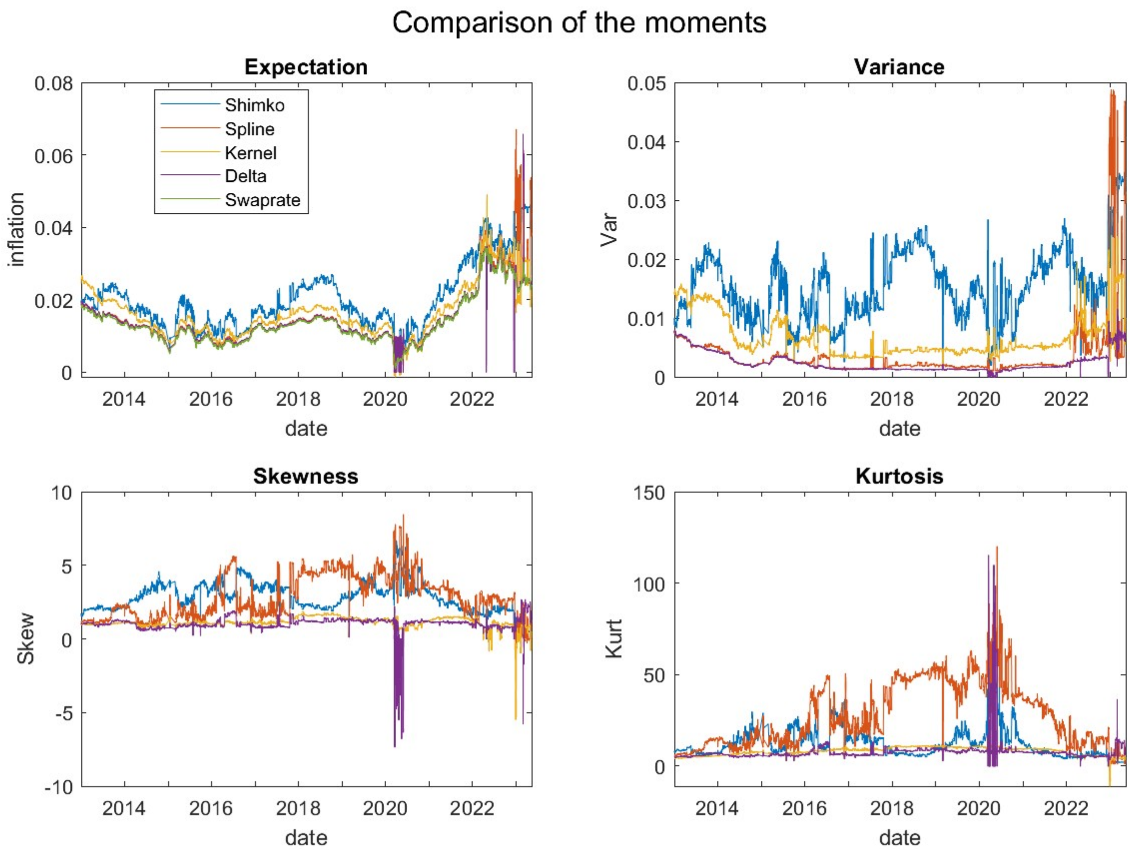


Figure 5.19: Comparison of the moments, where the outliers of the Delta method have been removed.

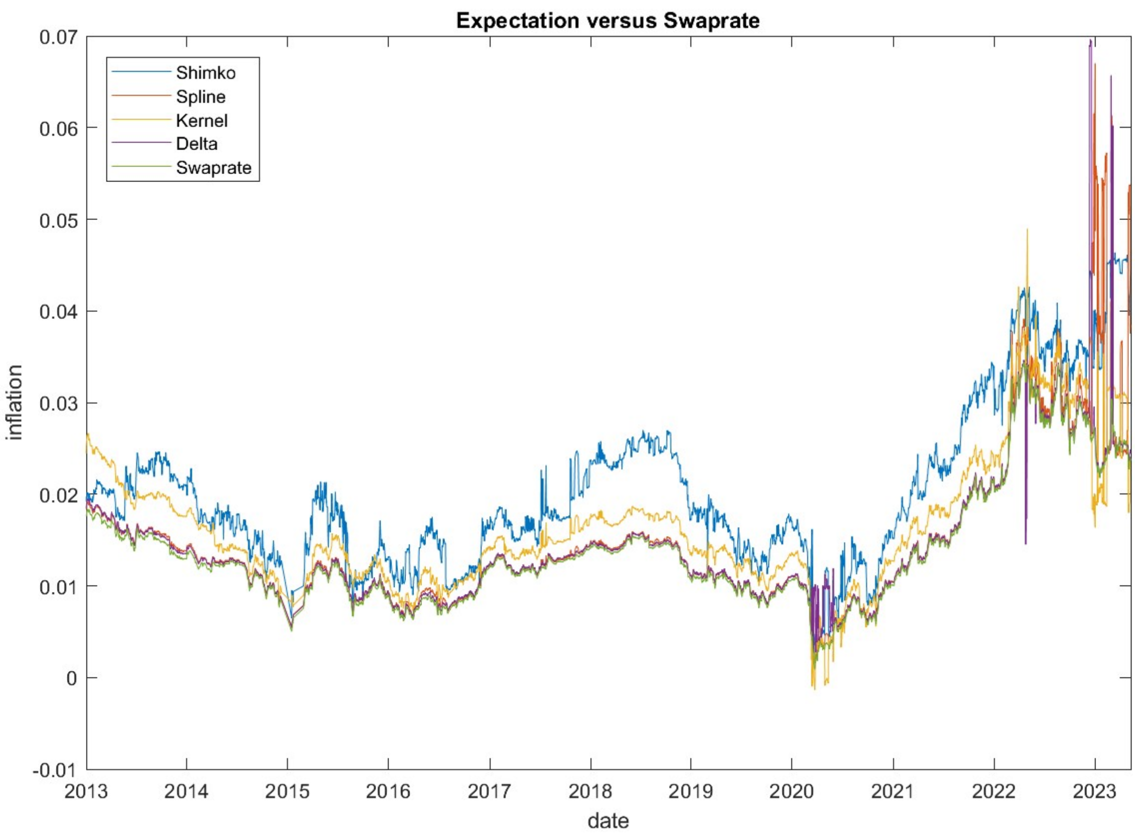


Figure 5.20: Comparison of the expectation of the methods with the swap rate.

6 Discussion and Conclusion

To be able to determine what the best method is of estimating option implied probability density functions for inflation, various methods were found in the existing scientific literature. A selection of these methods were implemented and compared. The short answer to the research question is that in theory all these methods can be applied to inflation option prices. However, in the current state of the inflation option market the data is not reliable enough. The problems introduced by arbitrage mean that some methods perform a lot better than others. In general it was found that the more closely a method follows the data, the poorer its performance in case of arbitrage. Adjusting the data to remove arbitrage from the dataset did help on some days, but it is not a certainty that this works on all days. Even if the constraints to prevent arbitrage are imposed on the data set, the fitting of a function for the intermediate points can introduce arbitrage opportunities again. The arbitrage opportunities in the dataset are a consequence of the current state of the inflation option market. Since only options exist on every half percentage-point, the data set is too small. If in the future more strike rates will be traded then this adjusting will work better.
Nevertheless, it was found that the Kernel method is the most robust and therefore performs best in the current state of the inflation option market. It does have two drawbacks, it is a slow method and it does not extrapolate well. An assumption must be made for the tails of the distribution. In contrast, the Delta method scores higher on the likelihood comparison method and it is able to extrapolate a lot better. Moreover, when compared to the swap rate the Delta method also scores as one of the best. The drawback of the Delta method is that the resulting RNDs do not look like proper density functions, this is not only because of arbitrage but also because of the transformation of variables it applies.

The interpretation of these results is that it is not possible to appoint the best method that works the best in every possible case. The advise for those who will use these methods in the future for inflation, would therefore be to always use multiple methods to derive multiple RNDs. To then use the one that gives the most plausible RND on that particular day. And the Kernel and Delta methods are most likely to perform the best given the comparisons done in this thesis. Another conclusion that can be drawn from the comparisons conducted in this thesis is that Shimko's method underperforms. Although it is a robust method in the case of arbitrage, using a second order polynomial does not work well for extrapolation. Moreover, it scored lower on the likelihood comparison and was the furthest away from the swap rate.
The various implemented spline methods scored relatively mediocre, especially the regular cubic splines could not handle the arbitrage well. The smoothing splines with linear tails scored relatively the best out of all the spline methods. It was as close to the swap rate as the Delta method, on the likelihood comparison it scored lower.

This research was heavily limited by the reliability of the data. Another conclusion of this thesis is therefore that the inflation option market is not yet evolved enough to apply these methods to derive a trustworthy risk-neutral density. On the other hand it is possible to say something on which method performs best with the data that is available, such as the conclusions above.
Another limiting factor is the fact that we try to find the market expectation of inflation, and we have no

better option than to test the found densities with the realised inflation. Therefore, it is more like testing which method predicts inflation the best, instead of finding the method that is the best at measuring the market expectation of inflation. Additionally, it is very difficult to test what is the best way to treat the tails of the distribution. As in a lot of extreme value theory there is limited data on hyperinflation or deflation, especially since the inflation option market is relatively young.

An avenue for further studies on this subject could therefore try to investigate the possibility to use generalised extreme value distributions (GEV) to fit tails to the distributions in a different way than extrapolating. Another suggestion for future research would be to investigate different ways of finding the relation between the implied volatility and the strike rate, such as using a stochastic volatility model or a fitting method that allows for restrictions on its derivatives. During this research I thought it would be a good idea to investigate a mean reverting stochastic volatility model. Since the central banks aim to keep inflation around 2% per year it could be hypothesised that inflation is mean reverting to that level. However, this idea was not implemented due to time restrictions. A second subject that was let go due to time restrictions was using the option-implied Fourier-cosine (iCOS) method to estimate the RNDs as suggested by Vladimirov in September of this year [21].
Furthermore, it could be interesting to see if it is possible to decompose the inflation option prices into an expected payoff part and various different risk premiums. This is already done for different inflation linked products, but I have not found any sources that did this for inflation options. This could help getting a better view on the error margin of option prices, which in turn could help in better dealing with arbitrage. Another way to reduce the influence of arbitrage would be to investigate aggregating the pricing data of various days to reduce price fluctuations. However, since contracts starting on different days are effectively defined on different time periods, they are difficult to aggregate.

References

[1] Yacine Aït-Sahalia and Andrew W. Lo. “Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices”. In: *The Journal of Finance* 53.2 (Apr. 1998), pp. 499–547. ISSN: 00221082. DOI: 10.1111/0022-1082.215228. URL: http://doi.wiley.com/10.1111/0022-1082.215228 (visited on 05/21/2023).

[2] Yacine Aït-Sahalia and Jefferson Duarte. “Nonparametric option pricing under shape restrictions”. In: *Journal of Econometrics* 116.1 (2003). Frontiers of financial econometrics and financial engineering, pp. 9–47. ISSN: 0304-4076. DOI: https://doi.org/10.1016/S0304-4076(03)00102-7. URL: https://www.sciencedirect.com/science/article/pii/S0304407603001027.

[3] Bhupinder Bahra. *Implied Risk-Neutral Probability Density Functions from Option Prices: Theory and Application*. Rochester, NY, July 1, 1997. DOI: 10.2139/ssrn.77429. URL: https://papers.ssrn.com/abstract=77429 (visited on 05/21/2023).

[4] Fischer Black. “The pricing of commodity contracts”. In: *Journal of Financial Economics* 3.1 (1976), pp. 167–179. ISSN: 0304-405X. DOI: https://doi.org/10.1016/0304-405X(76)90024-6. URL: https://www.sciencedirect.com/science/article/pii/0304405X76900246.

[5] Fischer Black and Myron Scholes. “The pricing of options and corporate liabilities”. In: *Journal of political economy* 81.3 (1973), pp. 637–654.

[6] Douglas T. Breeden and Robert H. Litzenberger. “Prices of State-Contingent Claims Implicit in Option Prices”. In: *The Journal of Business* 51.4 (1978), pp. 621–651. ISSN: 00219398, 15375374. URL: http://www.jstor.org/stable/2352653 (visited on 05/21/2023).

[7] Roger Clews, Nikolaos Panigirtzoglou, and James Proudman. “Recent Developments in Extracting Information from Options Markets”. In: *Bank of England Quarterly Bulletin* (2000).

[8] John C. Cox and Stephen A. Ross. “The valuation of options for alternative stochastic processes”. In: *Journal of Financial Economics* 3.1 (1976), pp. 145–166. ISSN: 0304-405X. DOI: https://doi.org/10.1016/0304-405X(76)90023-4. URL: https://www.sciencedirect.com/science/article/pii/0304405X76900234.

[9] Stephen Figlewski. “Risk Neutral Densities: A Review”. en. In: 3120028 (Mar. 2018). DOI: 10.2139/ssrn.3120028. URL: https://papers.ssrn.com/abstract=3120028.

[10] Maria Grith, Wolfgang Karl Härdle, and Melanie Schienle. *Nonparametric estimation of risk-neutral densities*. Springer, 2012.

[11] Jens Hilscher, Alon Raviv, and Ricardo Reis. *How Likely Is an Inflation Disaster?* Rochester, NY, Apr. 26, 2022. DOI: 10.2139/ssrn.4083404. URL: https://papers.ssrn.com/abstract=4083404 (visited on 05/09/2023).

[12] Jens Carsten Jackwerth. *Option-implied risk-neutral distributions and risk aversion*. In collab. with Research Foundation of AIMR. Charlottesville, Va: Research Foundation of AIMR, 2004. 86 pp. ISBN: 9780943205663.

[13] Allan M. Malz. “A Simple and Reliable Way to Compute Option-Based Risk-Neutral Distributions”. In: *SSRN Electronic Journal* (2014). ISSN: 1556-5068. DOI: 10.2139/ssrn.2449692. URL: http://www.ssrn.com/abstract=2449692 (visited on 02/28/2023).

[14] Allan M. Malz. *Option-Implied Probability Distributions and Currency Excess Returns*. Rochester, NY, Nov. 1, 1997. DOI: 10.2139/ssrn.943500. URL: https://papers.ssrn.com/abstract=943500 (visited on 05/21/2023).

[15] Robert C Merton. “Theory of rational option pricing”. In: *The Bell Journal of economics and management science* (1973), pp. 141–183.

[16] Ana Margarida Monteiro, Reha H. Tütüncü, and Luís N. Vicente. “Recovering risk-neutral probability density functions from options prices using cubic splines and ensuring nonnegativity”. In: *European Journal of Operational Research* 187.2 (2008), pp. 525–542. ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2007.02.041. URL: https://www.sciencedirect.com/science/article/pii/S0377221707002871.

[17] Antonis Papapantoleon. *Lecture Notes ”Introduction to Financial Mathematics”*. Jan. 22, 2022.

[18] David Shimko. “Bounds of probability”. In: *Risk* 6 (Apr. 1993), pp. 33–37.

[19] Tom Smith. *Option-implied probability distributions for future inflation*. 2012. URL: https://www.bankofengland.co.uk/-/media/boe/files/quarterly-bulletin/2012/option-implied-probability-distributions-for-future-inflation.pdf?la=en&hash=43C67B29B21E3279A0F74654721BB2FF67CEF087 (visited on 12/04/2023).

[20] Sami Vahamaa. “OPTION-IMPLIED ASYMMETRIES IN BOND MARKET EXPECTATIONS AROUND MONETARY POLICY ACTIONS OF THE ECB”. In: *ECB Working paper series* (Mar. 2004).

[21] Evgenii Vladimirov. “iCOS: Option-Implied COS Method”. In: (2023). DOI: 10.48550/ARXIV.2309.00943. URL: https://arxiv.org/abs/2309.00943.