Systematic Trade Structuring

Tristan Farran

BSc in International Economics and Finance

Bocconi University

Advisor: Prof. Fulvio Ortu

Contents

Chapter 1

Introduction

This thesis presents a proof of concept for a trade structuring system that maps an investor's distributional views on an underlying variable into an implementable derivatives position, assuming a sufficiently liquid listed options market exists for the underlying. For simplicity, the underlying variable is taken to be a security price throughout. To ensure generality, the framework is fully non-parametric and does not impose a parametric model for the underlying distribution. Many of its objects and steps can also be interpreted through the lens of Bayesian probability, so the resulting construction can be viewed as a non-parametric Bayesian trade structuring system.

In brief, the system infers a market-implied prior distribution from option prices and incorporates investor views as constraints to obtain a posterior distribution that deviates minimally from the prior while rendering the views true. The posterior then determines an optimal trade structure under a chosen objective, such as expected log return growth, Sharpe ratio, or any other function of the distribution. Several examples demonstrate that the system produces sensible recommendations in benchmark scenarios and remains flexible enough to accommodate practical constraints such as liquidity, and transaction costs. Extensions for future work are discussed, including multi-asset, multi-period structuring.

Chapter 2

The Market-Implied Distribution

The first essential notion that must be established is that, by the fundamental theorems of asset pricing, if no arbitrage opportunities exist in a market, then prices in that market must imply a risk-neutral probability distribution. This is most straightforwardly demonstrated by risk-neutral valuation, a common and important derivatives pricing technique.

2.1 Risk-Neutral Valuation

Risk-neutral valuation obtains the price of a derivative X by finding a replicating portfolio Y that has the same payoff in all states of the world and is composed of assets whose prices are observable in the market. Assuming there are no arbitrage opportunities, the price of X must be equal to the price of Y, giving us a way to price X without making reference to its underlying's real-world probability distribution P. This reasoning can be also extended to securities which cannot be replicated statically using one unchanging portfolio and must instead be replicated continuously using a portfolio whose composition changes over time in response to changes in a set of relevant variables. Options are one such case, with continuous replication underpinning the foundational Black-Scholes model (Black and Scholes 1973).

2.2 The Risk-Neutral Measure

Conveniently, risk-neutral valuation implies a probability measure Q, equivalent to P for pricing purposes, called the risk-neutral measure. This measure can be thought of as what the market's estimate of P would be if investors were risk-neutral.[1]

In general, investors are not risk neutral, and we should be careful not to confuse Q for P. However, their equivalence for pricing purposes means that Q can be used to structure trades, because as long as the posterior we will construct leads to different prices than our risk-neutral prior, and the views used to construct it are correct, the position can be justly considered as providing positive profits in expectation.

To see how prices imply risk-neutral probabilities, consider a simple binary bet with a payoff function $\pi_\gamma(A)$ based on an event $A \in \mathcal{A} = \{A, A'\}$ that occurs with probability p in the discrete probability space $(\Omega, \mathcal{A}, P)$:

$$\pi_\gamma(A) := \begin{cases} M & \text{if } A \\ 0 & \text{if } A' \end{cases} \tag{2.1}$$

If the bet is fairly priced such that expected profits are equal to zero for both sides, then its price V and payout M fully specify the discrete risk-neutral measure:

$$E(\Pi) = pM - V = 0 \iff Q = \left(\frac{V}{M}, 1 - \frac{V}{M}\right) \tag{2.2}$$

[1]We can extend the analogy slightly to the case where the risk-preferences of investors weighted by their market impact are neutral "on average".

This intuition is formalized and extended by the First Fundamental Theorem of Asset Pricing (see Pascucci 2011, Ch. 2), which says that a discrete market on a discrete probability space is arbitrage-free if and only if $\exists Q \sim P$ such that,[2] denoting the price S:

$$S_0 = E_Q(S_T)e^{-rT} \tag{2.3}$$

$$E_Q(S_{T+1} \mid S_1, \ldots, S_T) = S_T \tag{2.4}$$

$$P(A) = 0 \iff Q(A) = 0 \tag{2.5}$$

The First Fundamental Theorem can also be extended to continuous processes, guaranteeing the same properties as above but under a stronger notion of arbitrage known as "no free lunch with vanishing risk" or NFLVR (see Delbaen and Schachermayer 1994).

All security prices contain information about the risk-neutral measure under a strong enough definition of no-arbitrage, but not to the same extent. For example, assuming that the set of possible prices of a stock is an interval $\mathcal{S} \subseteq \mathbb{R}^+$, the current price S_0 only implies $E_Q(S_T)$ for some future time T by (2.3), which is far from enough to fully specify Q. In fact, to try to recover Q without making further assumptions, we'll need prices from the entire options chain for that date.

As a rule of thumb, it can be said that the more complex the payoff, the more information the price contains about the underlying variable's risk-neutral measure. To see why, consider the fact that expected option payoffs depend on volatility due to their non-linear exposure to the price (i.e. non-zero gamma) as a result of Jensen's inequality, while expected stock or future payoffs do not depend on volatility due to their linearity.

[2] Note the similarity between condition (2.2) and conditions (2.3) and (2.4).

2.3 Recovering Q

As a first foray into recovering the risk-neutral measure from option price, we can employ some useful properties of the butterfly structure, a linear combination of vanilla options that can be constructed the same way using either calls or puts. The payoff function and price of a butterfly with center strike K^* and strike spacing Δ, denoted by π_B and B respectively, can be expressed in terms of call payoffs and prices, denoted π_C and C, as:

$$B(K^*, \Delta, T) = C(K^* - \Delta, T) - 2C(K^*, T) + C(K^* + \Delta, T) \tag{2.6}$$

$$\pi_B(K^*, \Delta, S_T) = \pi_C(K^* - \Delta, S_T) - 2\pi_C(K^*, S_T) + \pi_C(K^* + \Delta, S_T) \tag{2.7}$$

Resulting in a concave payoff that takes on its maximum value Δ when $S_T = K^*$:

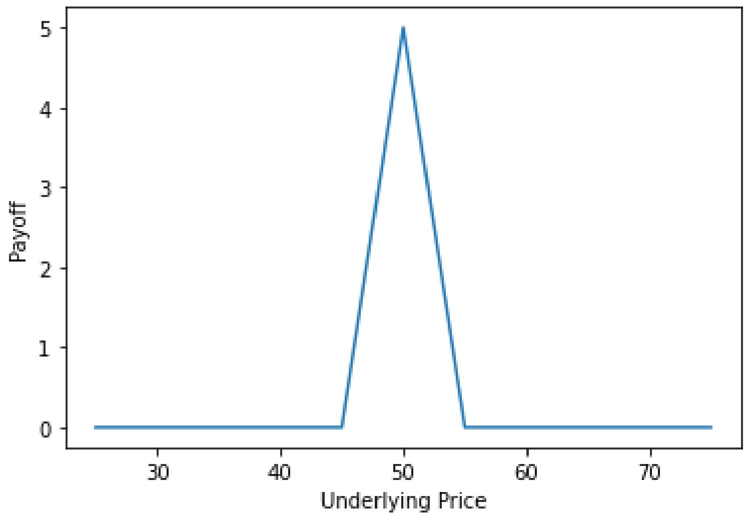

Figure 2.1: A butterfly payoff with $K^* = 50$ and $\Delta = 5$.

Assuming $S_T \in \{S \mid S = n\Delta,\, n \in \mathbb{Z}\}$, the connection between butterfly prices and probabilities is simple because $\pi_B = \pi_\gamma$, paying out $M = \Delta$ if and only if $S_T = K^*$ since S_T cannot take on another value in the region where the payoff is positive (i.e. between strikes), enabling the retrieval of a risk-neutral probability by (2.2) after discounting the payoff: $\frac{B(K^*, \Delta, T)}{\Delta e^{-rT}} = P_Q(S_T = K^*)$.

When we try to extend this argument to the case where S_T is a continuous random variable by taking the limit, we encounter the following indeterminate form:

$$\frac{\lim_{\Delta \to 0^+} B(K^*,\, \Delta,\, T)}{\lim_{\Delta \to 0^+} \Delta} = \frac{0}{0} \tag{2.8}$$

If the PDF $q(S_T)$ is continuous, then the existence of $\frac{\partial C(K,T)}{\partial K}$ is guaranteed by (2.3) for reasons that will become clear with the derivation of the Breeden-Litzenberger formula at the end of the section, enabling us to apply L'Hôpital's rule to show the following:

$$\begin{aligned} \lim_{\Delta \to 0^+} \frac{B(K^*,\, \Delta,\, T)}{e^{-rT}\Delta} &= e^{rT} \lim_{\Delta \to 0^+} \frac{\partial B(K^*,\, \Delta,\, T)}{\partial \Delta} \\ &= e^{rT} \lim_{\Delta \to 0^+} \frac{\partial C(K^* - \Delta,\, T)}{\partial \Delta} + \frac{\partial C(K^* + \Delta,\, T)}{\partial \Delta} \\ &= e^{rT} \lim_{K \to K^*} -\frac{\partial C(K,\, T)}{\partial K} + \frac{\partial C(K,\, T)}{\partial K} \\ &= 0 \end{aligned} \tag{2.9}$$

This is a sensible result since the probability of a continuous random variable taking on any single value is always zero:

$$\lim_{\Delta \to 0^+} \int_{K^*-\Delta}^{K^*+\Delta} f(x)\, dx = \lim_{\Delta \to 0^+} F(X + \Delta) - F(X - \Delta) = 0 \tag{2.10}$$

Given the above, the natural next step is to try to find probability density from Butterfly prices, which we can do by dividing by Δ once more and recognizing the expression as the second order central difference of $C(K,\, T)$ on K (see Breeden and Litzenberger 1978):

$$\begin{aligned} q(K) &= e^{rT} \lim_{\Delta \to 0^+} \frac{B(K^*,\, \Delta,\, T)}{\Delta^2} \\ &= e^{rT} \lim_{\Delta \to 0^+} \frac{C(K^* + \Delta,\, T) - 2C(K^*,\, T) + C(K^* - \Delta,\, T)}{\Delta^2} \\ &= e^{rT} \frac{\partial^2 C}{\partial K^2} \end{aligned} \tag{2.11}$$

An alternative intuition for this is to interpret (2.9) as the price of a derivative with a payoff equal to a butterfly payoff rescaled by Δ, which becomes a binary bet with $M = 1$ in the limit,[3] implying that its price is equal to $P_Q(S_T = K^*)$ as discussed earlier.

$$\lim_{\Delta \to 0^+} \frac{\pi_B(K^*, \Delta, T)}{\Delta} = \begin{cases} 1 & \text{if } S_T = K^* \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

Another, more direct derivation of the Breeden-Litzenberger formula (Zhou 2018) is:

$$\begin{aligned} C(K) &= e^{-rT} \int_0^\infty \max(x - K, 0)\, f(x)\; dx \\ &= e^{-rT} \int_K^\infty (x - K)\, f(x)\; dx \\ &= e^{-rT} \left(\int_K^\infty x\, f(x)\; dx - K \int_K^\infty f(x)\; dx \right) \\ \therefore\ \frac{\partial C(K)}{\partial K} &= -e^{-rT} \int_K^\infty f(x)\; dx \\ &= e^{-rT} \left(\int_{-\infty}^K f(x)\; dx - 1 \right) \\ &= e^{-rT} \left(F(K) - 1 \right) \\ \therefore\ \frac{\partial^2 C(K)}{\partial K^2} &= e^{-rT} \frac{dF}{dK} \\ f(K) &= e^{rT} \frac{\partial^2 C(K)}{\partial K^2} \end{aligned} \tag{2.13}$$

If Q is not continuous, there is no general method to evaluate (2.8), but we can establish some bounds on the risk-neutral probability of S_T being within the width of a butterfly. Since π_B is bounded from above by π_γ with $M = \Delta$, B provides the following lower bound (Bossu and P. Carr 2014, Ch. 3):

$$P(K^* - \Delta \leq S_T \leq K^* + \Delta) \geq e^{rT} \frac{B(K^*, \Delta, T)}{\Delta} \tag{2.14}$$

[3] This kind of derivative is also known as an Arrow-Debreu security, or a double digital option.

2.4 Real-World Considerations

In the real world, option chains have a limited number of strikes and only some of them have liquid enough options associated to them for their prices to contain useful information, this section will briefly cover some implications of this and ways to address them.

The discreteness of option strike listings means that almost all of the methods presented herein must be discretized when implemented, e.g. differentiation must be approximated using finite differences, and the reliability of the system therefore critically depends on the stability of underlying the numerical approximations.

Additionally, illiquid options must be filtered from any given chain before it is used to infer a risk-neutral measure. This is because their prices may reflect outdated information or even imply arbitrage opportunities where none exist (since the last price may be far from the price at which it is currently possible to trade), invalidating our central assumption. Illiquid options can be discerned according to a variety of metrics including bid-ask spread, volume, and market impact statistics (see Lybek and Sarr 2003).

To make use of all of the available information in the market, the put-call parity relationship, which follows from the no-arbitrage assumption (Stoll 1969), should be used to generate the prices of calls wherever puts are more liquid:

$$C(K,\, T) = P(K,\, T) + S_0 - Ke^{-rT} \tag{2.15}$$

Yielding prices for a vector of strikes with sufficient liquidity denoted $\mathbf{K} = \{K_1, \ldots, K_n\}$.

Furthermore, ask prices should be used as the reference for long positions, while bid prices should be used as the reference for short positions, as these are the prices at which it is actually possible to put each side of the trade on.[4]

The realities of option chains also imply that it is only possible to sample Q at certain points ($S_T \in \mathbf{K}$) and that there are always price intervals for which no probabilities can be recovered ($I \nsubseteq \mathbf{K}$). Two nonparametric methods to try to deal with this are:

1. Fitting a constrained spline to the volatility surface to generate additional prices (see Ait-Sahalia and Duarte 2003).

2. Extending Q by the maximum entropy (uniform) distribution taking the value $\frac{1-\int q(x)\,dx}{\alpha+K_0-K_n}$ on $[0,\, K_1)$ and $(K_n,\, \alpha]$, where α is what the investor believes to be the maximum possible price. This is useful to avoid inadvertently assuming that the market deems $S_T \notin [K_0,\, K_n]$ to be impossible.

2.5 Comparison

To get a better sense of the effectiveness of the Breeden-Litzenberger formula in practice, we can test how accurately it recovers the log-normal distribution from a simulated options chain priced via Black-Scholes with no implied volatility smile.

We will do this by plotting the recovered distribution and its Q-Q plot, as well as the percentage error of the recovered distribution's volatility against the coverage ratio[5] $Z = \frac{K_n - K_0}{\sigma S_0}$, the strike spacing Δ, and the true Black-Scholes volatility σ.

[4] Short positions entail additional complexity including borrow costs and margin dynamics, which will be addressed in a subsequent section.

[5] To replicate the fact that options tend to be more liquid the nearer they are to at-the-money, strikes are generated in evenly both directions from S_0 as Z grows.

It is important to note that the relationship of the errors with each variable may be sensitive to the specific distribution that is being recovered and the values of its parameters and that the recovered distribution has not been adjusted via either of the options presented at the end of Section 2.4.

Recovered Distributions

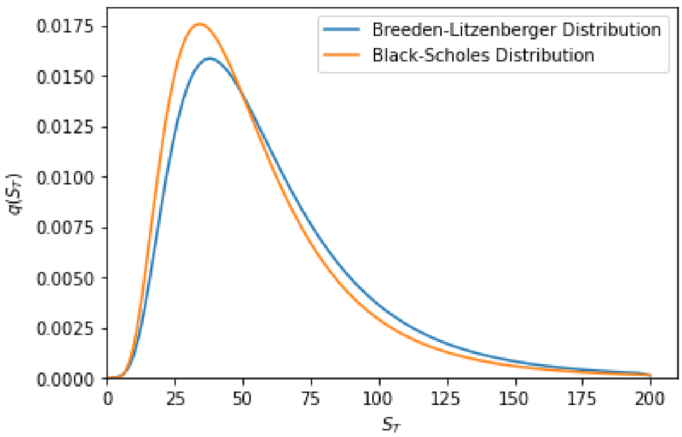

Figure 2.2: A distribution recovered from simulated prices.

$S_0 = 50,\ Z = 10,\ \Delta = 2,\ T = 2,\ \sigma = 0.4,\ r = 0.07$

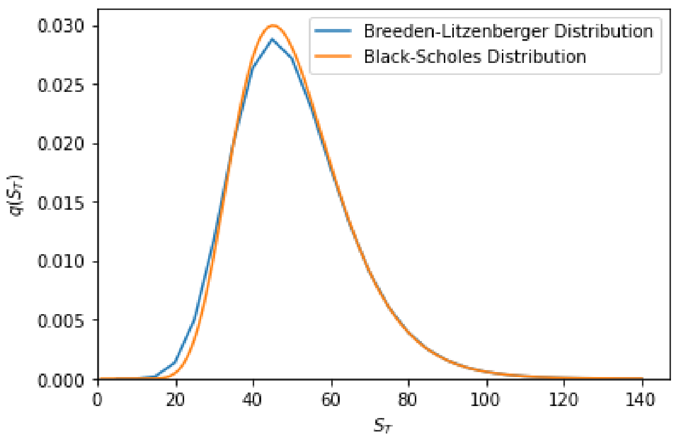

Figure 2.3: A distribution recovered from simulated prices.

$S_0 = 50,\ Z = 7,\ \Delta = 5,\ T = 0.5,\ \sigma = 0.4,\ r = 0.02$

Q-Q Plot

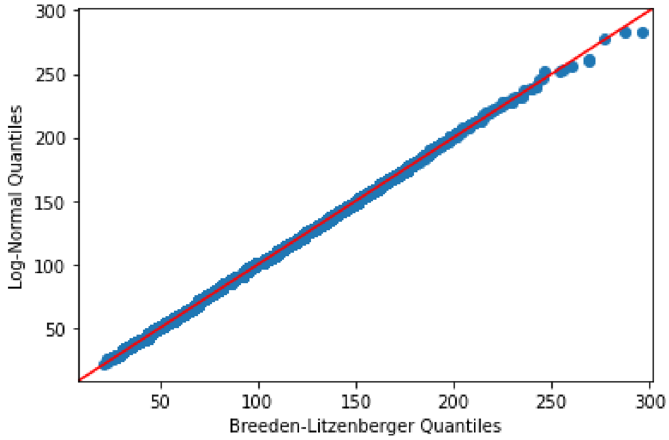

Figure 2.4: A Q-Q plot of the recovered distribution against the true distribution.

$$S_0 = 50,\ Z = 15,\ \Delta = 1,\ T = 1,\ \sigma = 0.3,\ r = 0.1$$

Errors

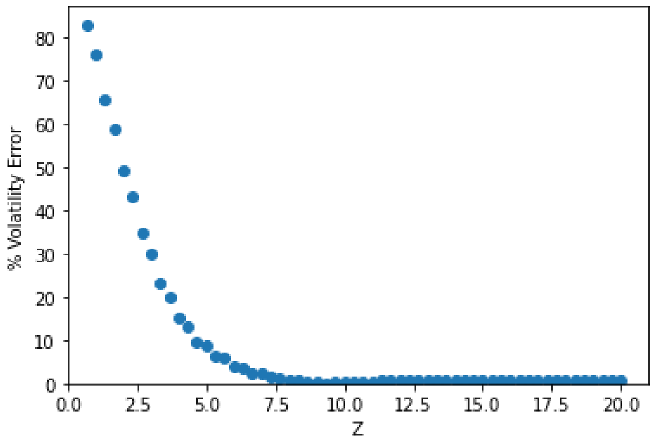

Figure 2.5: Percentage volatility error against Z.

$$S_0 = 75,\ \Delta = 2,\ T = 1,\ \sigma = 0.2,\ r = 0$$

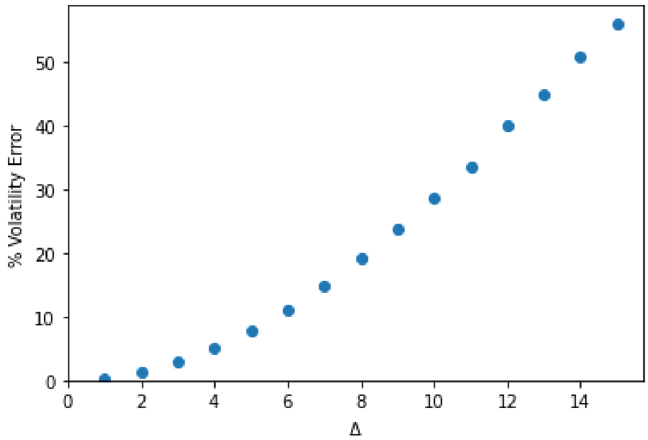

Figure 2.6: Percentage volatility error against Δ.

$S_0 = 50,\ Z = 15,\ T = 1,\ \sigma = 0.2,\ r = 0$

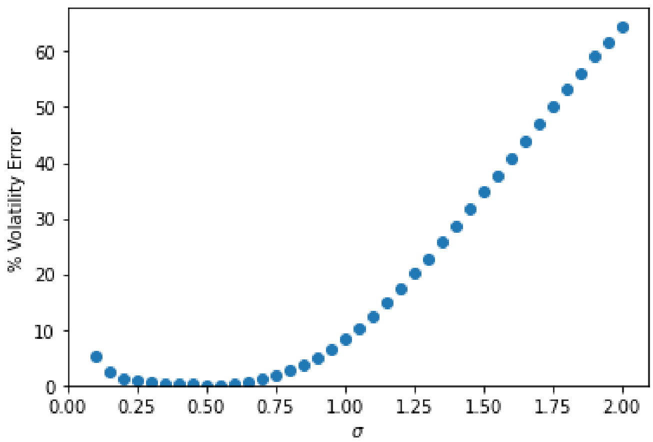

Figure 2.7: Percentage volatility error against σ.

$S_0 = 50,\ Z = 15,\ \Delta = 2,\ T = 1,\ r = 0$

The relationships presented in these plots are quite intuitive and show that the recovered distribution should be quite accurate for moderate values of the variables.

Chapter 3

The Fair Posterior

The next step is to construct a posterior probability distribution F, which we'll call the "fair posterior". Informally, our aim in constructing the fair posterior is to construct the distribution that renders Q the "least wrong" it could be while ensuring our views, denoted $\mathbb{V}$, are correct. Loosely, this is analogous to ensuring that derivatives on this market are the "least mispriced" they could be under F.

This is desirable philosophically, as a way of respecting the information baked into market prices, and practically because it ensures that expected profits are maximally conservative, which is useful for trade sizing. It is loosely analogous to assuming that anything that the investor does not have a view on is best estimated by the market.[1]

In practice, this means minimizing a measure of how different F is from the market-implied distribution subject to the view set, which we restrict to be a set of consistent constraints on statistics of F, consistency meaning the absence of contradictions that mean no solution can be found. Some examples of inconsistent $\mathbb{V}$s are:

[1] If this were not the case, then the fair posterior should simply be the maximum-entropy distribution corresponding to the view set.

1. $\mathbb{V} = \{E_F(S_T) > 10,\, E_F(S_T) < 10\}$ is inconsistent for obvious reasons.

2. $\mathbb{V} = \{\sigma_F^2 > E_F(S_T^2),\, E_F(S_T) > 0\}$ is inconsistent because $\sigma_F^2 = E_F(S_T^2) - E_F(S_T)^2$.

3. $\mathbb{V} = \{\sigma_F > 5\}$ with $\alpha = 15,\ \mu = 10,\ K_0 = 5$ is inconsistent because $\sigma_F \leq \sqrt{(\alpha - \mu)(\mu - K_0)}$ (Bhatia and Davis 2000).

3.1 Ground Rules

There are five conditions which must be ensured when transforming Q into F to ensure that the posterior is valid, allows the use of the measure of difference we'll use (the Kullback-Leibler divergence), and does not introduce arbitrage opportunities (Carr and Madan 2005). These can be implemented as constraints in a suitable optimizer, they are:

$$\int_{\mathrm{dom}(f)} f(x)\ dx = 1 \tag{3.1}$$

$$f(x) \geq 0,\ \forall x \in \mathrm{dom}(f) \tag{3.2}$$

$$f(x) = 0 \iff q(x) = 0 \tag{3.3}$$

$$B_F(K^*,\, \Delta,\, T) > 0,\ \forall K^* \in \{K_2, \ldots, K_{n-1}\} \subseteq \mathbf{K} \tag{3.4}$$

$$C_F(K_b,\, T) - C_F(K_a,\, T) > 0,\ \forall (K_a,\, K_b) \in \mathbf{K}^2,\ b > a \tag{3.5}$$

3.2 The Kullback-Leibler Divergence

Of the many functions that could be used to measure the difference between the fair posterior and the market-implied distribution, the Kullback-Leibler divergence of F from Q, denoted $D_{\mathrm{KL}}(F \parallel Q)$, is the most suitable for our purposes. Minimizing $D_{\mathrm{KL}}(F \parallel Q)$ will endow F with precisely the intended properties because it is the expected value of the log-likelihood ratio of F relative to Q, which is:

1. The most powerful statistic for determining whether two distributions are different[2] according to the Neyman-Pearson lemma (Kullback 1959).

2. Closely related to the measure of risk-adjusted expected returns due to mispricing which we'll use to assess payoff optimality, as we'll show.

3. Interpretable as both:

   (a) The excess surprise from believing $P_Q(A)$ is the probability of an event when it is actually $P_F(A)$, measuring how wrong believing in Q is if F is correct.

   (b) The Bayes factor Q is updated by to get F, a reflection of how much rendering the view set true changes the posterior relative to the prior.

The continuous Kullback-Leibler divergence of F from Q is:

$$D_{\text{KL}}(F \parallel Q) = \int_{-\infty}^{+\infty} f(x) \left(\log \frac{f(x)}{q(x)} \right) \, dx \tag{3.6}$$

The discrete Kullback-Leibler divergence of F from Q is:

$$D_{\text{KL}}(F \parallel Q) = \sum_{A \in \mathcal{A}} P_F(A) \, \log \left(\frac{P_F(A)}{P_Q(A)} \right) \tag{3.7}$$

3.3 Minimization

The simplest approach to minimization would be to assume a form for the posterior distribution and optimize its parameters, but we will continue to use a non-parametric approach to maximize generality, taking the probabilities themselves as the parameters to be optimized subject to the view set and ground rules. This will be done using a Python library that employs a conjugate gradient method.

[2] In the sense that it has the highest probability of detecting a true positive.

3.4 Examples

The Python library used for this implementation can only handle discrete distributions, so the prior and posterior in this example will be given by finding the lower bound for the interval corresponding to each non-overlapping butterfly width by (2.14), normalizing, obtaining the discrete fair posterior from the discrete prior, and then assuming the price to be uniformly distributed within each interval to get a probability density function.

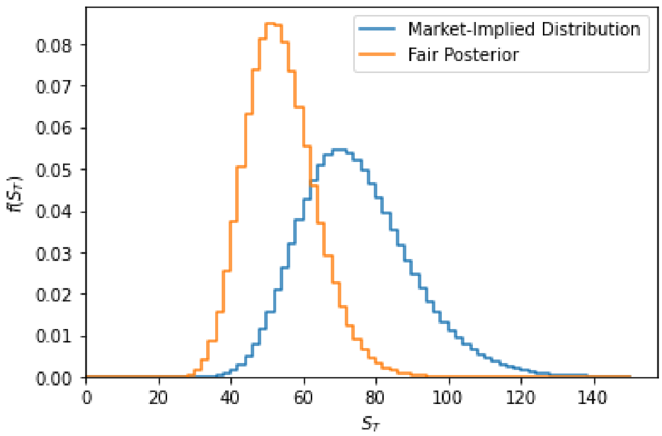

Figure 3.1: A fair posterior with reduced expected value.

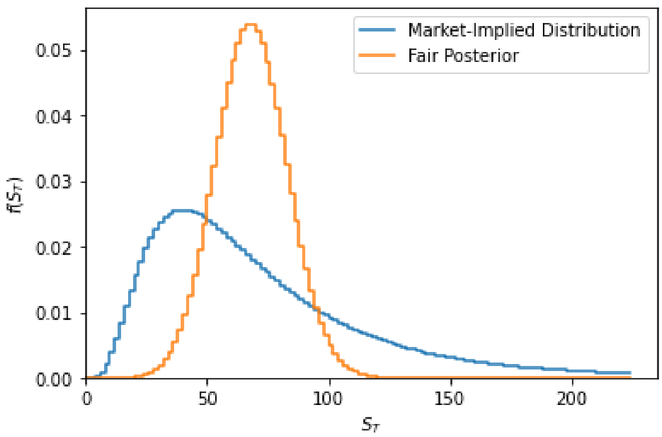

Figure 3.2: A fair posterior with reduced variance.

Chapter 4

The Optimal Trade

4.1 Measuring Optimality

Given F, we can obtain an optimal trade by finding the combination of securities which maximizes an objective function, some obvious candidates being expected returns or the Sharpe ratio. In this implementation, we will maximize expected log return à la Kelly, accounting for the volatility-drag associated with reinvestment across trials and maximizing terminal wealth in the limit (Kelly 1956).[1]

Denoting the log return function by:

$$g_x(S_T) = \log\left(\frac{\pi_x(S_T)}{V_x}\right) \tag{4.1}$$

The continuous and discrete expected log return of a long position are, respectively:

$$G(F,\, V,\, \pi) := \int_{-\infty}^{+\infty} f(x) \log g_x(S_T)\ dx \tag{4.2}$$

[1]It is worth noting that the fair posterior also provides estimates of variance and skewness, which can be useful in position sizing.

$$G(F,\,V,\,\pi) := \sum_{A\,\in\,\mathcal{A}} P_F(A)\,\log g_x(S_T) \tag{4.3}$$

Extending this to a long-short derivatives portfolio (i.e. an option structure) is not trivial however, due to the complexities introduced by leverage and the fact that the typical form of the short return $R = \frac{V_0}{V_T}$ diverges to infinity as the terminal price goes to zero, throwing off any attempts at return-based optimization.

To account for this, we must construct long and short position payoff and cost functions, denoted H and J respectively, that can avoid infinite short returns and account for each position's leverage at expiry based on p. 5-7 in the *CBOE Margin Manual* 2021, ignoring variation margin and any adjustments associated to particular structures.[2] Letting $\tilde{x}$ denote a put if x is a call and vice versa, and taking j to be an indicator variable taking the value 1 for a long position and -1 for a short position, these can be defined as follows:

$$H_x(S_T,\,j) := \begin{cases} \max(\pi_x(S_T) - 0.25V_x,\,0) & j = 1 \\ \max(V_x + J_x(S_T,\,j) - \pi_x(S_T),\,0) & j = -1 \end{cases} \tag{4.4}$$

$$J_x(S_T,\,j) := \begin{cases} \max(V_x - \pi_x(S_T),\,0.75V_x) & j = 1 \\ \max(\pi_x(K,\,S_T),\,0.2(S_0 - \pi_{\tilde{x}}(K,\,S_0)),\,0.1S_0) & j = -1 \end{cases} \tag{4.5}$$

Giving us the return of a long-short options portfolio X as:

$$R_X(S_T) = \frac{\sum_X H_x(S_T,\,j)}{\sum_X J_x(S_T,\,j)} \tag{4.6}$$

Note that since margin requirements are slightly less stringent for long positions, being long and short the same options structure actually leaves the portfolio slightly net long.

[2] It is also worth noting that trading on margin introduces path dependency into the cost function (the collateral required changes dynamically), but it would be too complex to account for this here.

4.2 The Optimal Trade

We could try to maximize G directly to find the optimal trade but since the log function is only defined for the positive real numbers and returns are often zero, we'd be forced to use an approximation. For example, a second-order Taylor expansion around $E(R_X)$:

$$E(\log(R_X)) = \log(E(R_X)) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{nE(R_X)^n} E((R_X - E(R_X))^n) \tag{4.7}$$

$$\therefore G(F,\, X) \approx \log(E(R_X)) - \frac{V(R_X)}{2E(R_X)^2} \tag{4.8}$$

To avoid using an approximation, we will first find the optimal return function as in Soklakov 2011, denoted $\aleph(S_T)$, and then find the trade whose return function best approximates it. To find $\aleph(S_T)$, we will re-frame the problem of optimal investment as the problem of finding the optimal weight vector $\mathbf{W}^* = \{w_1, \ldots, w_{|\mathcal{A}|}\}$ for a portfolio of binary bets β where each bet pays out M if and only if the corresponding event in $\mathcal{A} = \{S_T \in I_i \mid I_i \in I\}$ occurs, where I is a partition of $[0, \alpha]$ and all events are mutually exclusive. As mentioned in section 3.4, we will recover a distribution of this form directly in this implementation, but it can also be obtained from any distribution by integrating the PDF or summing the PMF (depending on whether it's continuous or discrete) over the subintervals of the partition.

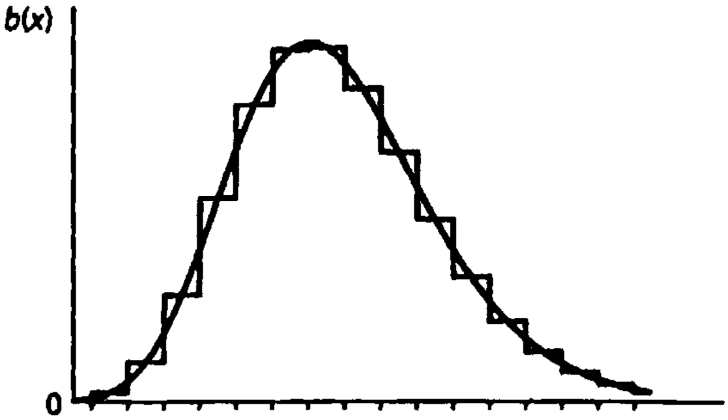

Figure 4.1: An intervalized continuous distribution.

Letting $\mathbf{V} = \{v_i, ..., v_{|\mathcal{A}|}\}$ and noting that $\pi_\beta(A) = w_i M$, we can show (Soklakov 2011):

$$\mathbf{W}^* = \arg\max_{\mathbf{W}}(G(F,\, \beta)) = \arg\max_{\mathbf{W}}(G(F,\, \mathbf{W} \cdot \mathbf{V},\, M \times \mathbf{W})) = F \tag{4.9}$$

Then, interpreting the index as a function of the underlying price $i = i(S_T)$ that indicates which interval it is in, we can show the following by (2.2) ibid.[3]

$$\aleph(S_T) = P_F(i(S_T)) \times \frac{M}{v_i} = \frac{P_F(i(S_T))}{P_Q(i(S_T))} = \frac{f(S_T)}{q(S_T)} \tag{4.10}$$

Implying that the Kullback-Leibler divergence is the maximum expected log return under F attainable by investing in binary bets priced according to Q, as alluded to earlier:

$$\max(G) = \sum_{A \in \mathcal{A}} P_F(A) \log \frac{P_F(A)}{P_Q(A)} = D_{\mathrm{KL}}(F \parallel Q) \tag{4.11}$$

Thereby showing that minimizing the Kullback-Leibler divergence of F from Q precisely minimizes the mispricing of derivatives and implying that $\aleph$ itself inherits the interpretation of $D_{\mathrm{KL}}(F \parallel Q)$ as the Bayes factor:

$$f(S_T) = \aleph(S_T) q(S_T)$$
$$P(S_T \mid \text{research}) = \frac{P(\text{research} \mid S_T)}{P(\text{research})} P(S_T)$$

The optimal trade of order n for a given underlier, denoted $\Psi_n^*(\Gamma)$, can then be defined as the sub-portfolio (n-combination with repetitions) of the set of payoffs attainable by taking long or short positions in the liquid options associated to the underlier, denoted Γ, that minimizes a metric $D(\Psi)$ which measures the difference between its return function $R_\Psi(S_T)$ and $\aleph(S_T)$.

[3] This is valid when the bet prices imply fair odds such that their sum is equal to one as in (2.2), but the same reasoning can be extended to unfair odds by normalizing returns.

Using the total square deviation:

$$D(\Psi) := \int (R_\Psi(S_T) - \aleph(S_T))^2 \, dS_T \tag{4.12}$$

$$\Psi_n^*(\Gamma) := \arg \min_{\Psi_n(\Gamma)} D(\Psi_n(\Gamma)) \tag{4.13}$$

This boils down to an integer programming problem, which is NP-complete. Though finding exact solutions via brute-force search is feasible in a matter of minutes for small structures with $n \leq 3$, for more complex structures optimization quickly becomes intractable due to combinatorial explosion, so heuristic methods such as simulated annealing or hill climbing must be used to find approximate solutions (see Chen, Batson, and Dang 2011).

4.3 Examples

One way of confirming that the system behaves as intended is to check that it produces sensible results in line with the typically recommended structures for simple view sets. To do so, we will supply common view sets for which there are commonly accepted solutions, and obtain the market-implied distribution from call prices generated via Black-Scholes with a set of parameters denoted χ, producing a log-normal distribution.

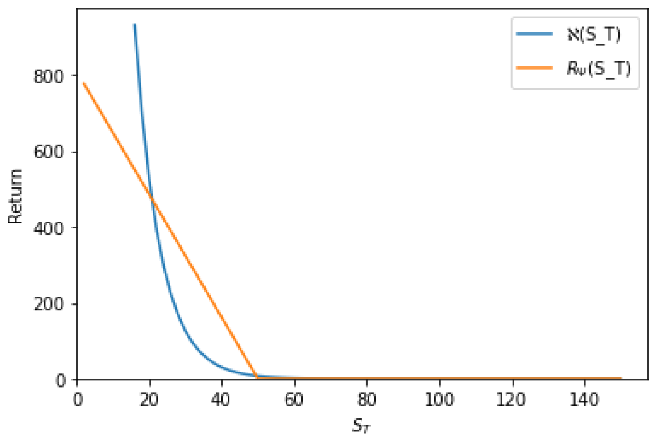

Figure 4.2: The return function of Ψ_1^* for a bearish view on $E(S_T)$.

For $\chi = (S_0 = 75,\ Z = 10,\ \Delta = 1, \sigma = 0.2,\ T = 1,\ r = 0.01)$ and $\mathbb{V} = \{E_F(S_T) = 55\}$, we obtain a long out of the money put position, which is one of two structures typically recommended for this view and $n = 1$ (long put or short call).

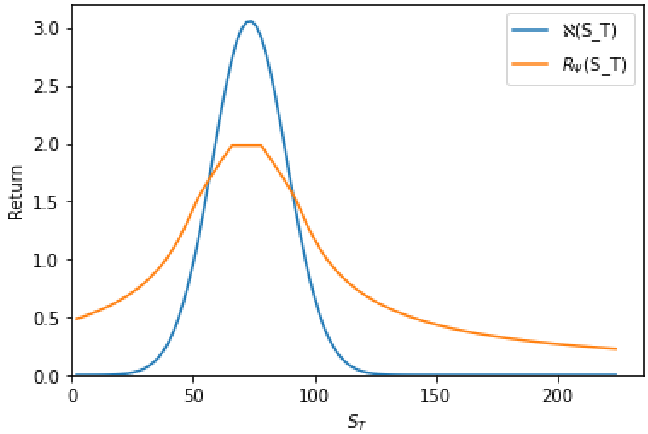

Figure 4.3: The return function of Ψ_2^* for a bearish view on σ. For $\chi = (S_0 = 75,\ Z = 10,\ \Delta = 1, \sigma = 0.2,\ T = 10,\ r = 0)$ and $\mathbb{V} = \{\sigma_F = 0.15\}$, we obtain a short strangle position, which is the structure that is typically recommended for this view and $n = 2$.

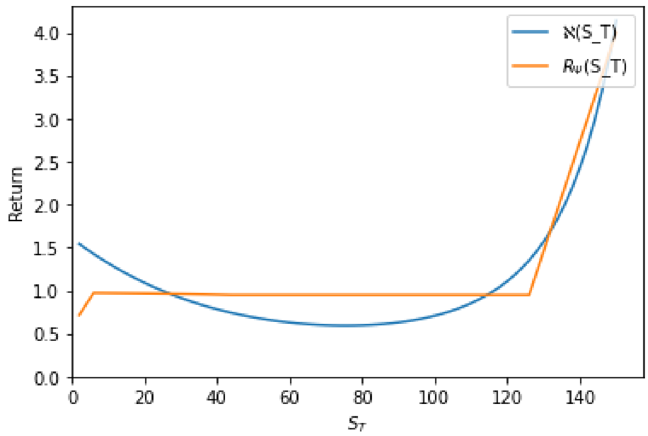

Figure 4.4: The return function of Ψ_2^* for a bullish view on $\text{Skew}(S_T)$. For $\chi = (S_0 = 75,\ Z = 10,\ \Delta = 1, \sigma = 0.2,\ T = 30,\ r = 0)$ and $\mathbb{V} = \{\text{skew}_F(S_T) = 3750000\}$, we obtain a long out of the money put position, which is one of two structures typically recommended for this view and $n = 2$ (risk reversal or bear spread).

Chapter 5

Extensions

5.1 Extending Optimality

In Chapter 3, we derived a method to find the optimal trade using n positions, leaving open the question of what value n should take on. Before sketching an answer to this, we will extend the framework to account for transaction costs and investors' budget constraint.

Each additional position in Ψ entails some amount of transaction costs (exchange fees, slippage etc.) that can be estimated according to the security's liquidity. Importantly, the transaction cost estimate for each position, denoted τ_x, should increase with the number of repetitions of any given position in the trade because the higher the volume of a position in the trade, the more market impact the investor will have when putting it on. τ_x can then be included in each position's cost function, yielding a more realistic trade return function.

To speed up the system and make it more useful for small investors who cannot afford to trade larger structures we can restrict the search space of possible trades to those which the investor can afford with the amount of capital at their disposal η, denoted Φ_η:

$$\Phi_\eta := \{\Psi \mid \max(\{J_\Psi(S_T) \mid S_T \in [0,\, \alpha]\}) \leq \eta\} \tag{5.1}$$

The introduction of τ_Ψ and Φ_η makes finding the optimal trade across all levels of n easier because:

1. If the search space isn't restricted to Φ_η it is infinite, rendering exhaustive search impossible and search in general more difficult. Though exhaustive search for a global minimum is still likely to be infeasible for large values of η, it is in principle possible and it should be achievable for small values of initial capital.

2. If τ_x for each position isn't included in $J_\Psi(S_T)$, $D(\Psi_n^*)$ is guaranteed to have no unique minimum across levels of n in an unrestricted search space. This is because supposing there exists a globally optimal structure Ψ_N^*:

$$\frac{n}{N} \in \mathbb{Z} \implies D(\Psi_n^*) = D(\Psi_N^*) \tag{5.2}$$

$$\frac{n}{N} \notin \mathbb{Z} \implies D(\Psi_n^*) > D(\Psi_N^*) \tag{5.3}$$

   Since Ψ_n^* must deviate from Ψ_N^* for values of n which are not multiples of N and simply be $\frac{n}{N}$ of repetitions of Ψ_N^* for those which are. In this case, Ψ_N^* could be identified by observing the repetition of a structure, but this would require computing the optimal trade up to $2N$. Including τ_Ψ remedies this because it is increasing on repetitions of each position, meaning that repeating the same structure does not yield the same return profile and $D(\Psi_n^*)$ would be monotonically increasing on n for $n > N$. Then, assuming $D(\Psi_n^*)$ is monotonically decreasing on n while $n < N$,[1] this allows us to find the optimal structure in $N+1$ iterations rather than in $2N$ by identifying N as the value of n after which $D(\Psi_n^*)$ begins to increase.

[1] This is intuitive but nonetheless an assumption, proof of this conjecture is also left as an extension.

Thus, Ψ_N^* is guaranteed to exist and can be found as soon as $D(\Psi_n^*)$ begins to increase or Φ_η is exhausted. An additional tactic that could be used to speed up this process is to implement some tolerance level for $D(\Psi_n^*)$ below which a trade is considered optimal.

5.2 Existing Positions

The system can straightforwardly be expanded to deal with existing positions in a given market by including the payoff and cost functions of existing positions in the return calculations, thereby finding the combination which most improves the portfolio position by the usual process.

5.3 Structure Preferences

Similarly, investor preferences regarding structures can easily be accommodated. If an investor wants to trade a particular structure and only wants advice on strike selection from the system, they can restrict Φ_η to be the set of affordable trades composed of the available variants of that structure. Conversely, if the investor doesn't want to trade a particular structure, they can simply exclude it from Φ_η. In both cases, the investor speeds up the system by reducing the search space.

5.4 The Optimal Portfolio

Given views on the distribution of several securities' prices at several option expirations, the framework could potentially be extended to incorporate relative views e.g. $\mathbb{V} = \{\sigma_X > \sigma_Y\}$ and find the optimal portfolio cross-sectionally and across expirations.

The fair posterior can be found independently for each security and expiration if no relative views are given. If relative views are given, the fair posteriors for the distribution pertaining to the views must be obtained jointly, this could be done by minimizing the sum of their Kullback-Leibler divergences, for example. It should also be noted that if relative views concern the same security at different points in time, the additional constraint that all calendar spread prices are non-negative must be implemented (Carr and Madan 2005). Then, the optimal portfolio could be found using the usual process but minimizing the sum of the difference metric across each security and expiration.

Another subtlety which must be accounted for in doing this is that different payoffs occur at different times, so trades with earlier expirations should be favored over trades with later expirations (since any proceeds can be reinvested at the risk free rate). This can be done crudely by discounting the distance metric for each trade before minimizing to find the optimal trade. Additionally, having positions that expire at different times opens the door to re-balancing, which can be achieved by re-evaluating the optimal trade given the portfolio, along the lines of what was described in Section 5.2.

5.5 Extending Views

Due to imperfections in the minimization process used in this implementation, it is possible that no solution will be found for a consistent view set. One way of dealing with this would be to dampen views (reducing their percentage variation from the statistic they regard according to the prior) until a solution can be found. This can be done using a confidence ranking, iterating over the view set from lowest to highest confidence, dampening each view somewhat until a solution can be found.

A more extensive treatment of confidence weighting and generating the posterior under extreme views can be found in Meucci, Ardia, and Keel 2010.

Chapter 6

Conclusion

In summary, this thesis demonstrates the feasibility of a fully non-parametric, Bayesian trade-structuring system that links market prices, investor views, and implementable option portfolios within a single coherent framework. Starting from a market-implied risk-neutral prior, the system incorporates subjective constraints via minimum-relative-entropy updating to obtain a "fair" posterior that preserves as much information as possible from observed prices while enforcing the view set. The resulting posterior yields an optimal target payoff profile, which can then be approximated using liquid listed options under realistic constraints. Across representative examples, the framework produces intuitive structures consistent with standard practitioner recommendations, while remaining flexible enough to accommodate practical considerations such as liquidity filtering, transaction costs, and existing portfolio exposures.

Bibliography

[21] *CBOE Margin Manual.* 2021. URL: `https://cdn.cboe.com/resources/membership/Margin_Manual.pdf`.

[AD03] Yacine Ait-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.

[BC14] S. Bossu and P. Carr. *Advanced Equity Derivatives: Volatility and Correlation.* Wiley Finance. Wiley, 2014.

[BD00] Rajendra Bhatia and Chandler Davis. "A Better Bound on the Variance". In: *The American Mathematical Monthly* 107.4 (2000), pp. 353–357.

[BL78] Douglas Breeden and Robert Litzenberger. "Prices of State-Contingent Claims Implicit in Option Prices". In: *The Wharton School* 51.4 (1978), pp. 621–651.

[BS73] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (1973), pp. 637–654.

[CBD11] Der-San Chen, Robert G Batson, and Yu Dang. *Applied Integer Programming.* John Wiley & Sons, 2011.

[CM05] Peter Carr and Dilip B. Madan. "A Note on Sufficient Conditions for No Arbitrage". In: *Finance Research Letters* 2 (2005), pp. 125–130.

[DS94] Freddy Delbaen and Walter Schachermayer. "A general version of the fundamental theorem of asset pricing". In: *Mathematische Annalen* 300 (1994), pp. 463–520.

[Kel56] John Kelly. "A new interpretation of information rate". In: *IEEE Transactions on Information Theory* 2 (1956), pp. 185–189.

[Kul59] Solomon Kullback. *Information Theory and Statistics.* Dover Publications, 1959, pp. 12–31.

[LS03] Tonny Lybek and Abdourahmane Sarr. "Measuring Liquidity in Financial Markets". In: *International Monetary Fund Working Papers* 2 (2003).

[MAK10] Attilio Meucci, David Ardia, and Simon Keel. "Fully Flexible Extreme Views". In: *Bloomberg Education & Quantitative Research Paper Series* (2010).

[Pas11] Andrea Pascucci. *PDE and Martingale Methods in Option Pricing.* Springer-Verlag, 2011, pp. 31–35.

[Sok11] Andrei Soklakov. "Bayesian Lessons for Payout Structuring". In: *Risk* 24.9 (2011), pp. 115–119.

[Sto69] Hans Stoll. "The Relationship Between Put and Call Option Prices". In: *The Journal of Finance* 24.5 (1969), pp. 801–824.

[Zho18] Sen Zhou. "Estimation of the risk-neutral density function from option prices". In: 2018.