Article Title Here In *Title Case (Alt+A)*(题目字体：Times New Roman， 字号20，加粗，居中)

First name Last name[1,*] First name Last name[2](Alt+A)（作者名字字体Times

ABSTRACT

Since Doug Breeden and Bob Litzenberger's seminal 1978 method of inferring future probabilities from traded options, risk-neutral probabilities have become instrumental in market pricing for discovering intrinsic values. This study delves into the concept of market-based probabilities within derivative markets, examining their utility in forecasting future asset prices. Our research begins by assessing the potency of risk-neutral probabilities in predicting future market returns and volatility, validated through the construction of confidence intervals. Upon establishing their predictive capability, we utilize Lasso regression analysis to anticipate market prices and trends. We then design and retrospectively evaluate a long-short trading strategy that capitalizes on these probabilities, aiming to profit from the inherent data in option prices. Our findings indicate that, despite potential inaccuracies arising from market constraints and financing issues, risk-neutral probabilities offer a viable avenue for insight and exploitation of future market movements. (This paper not only elucidates the theoretical underpinnings of market-based probabilities) but also presents empirical analysis of their practical application in market strategies, providing valuable insights for investors and policymakers alike.(这里看能不能多有点数值)

Keywords: *Market probability distribution, Lasso*

1. INTRODUCTION

Investors often tap into the insights embedded in financial asset prices as a cornerstone for making informed investment decisions. In this context, derivative markets provide them with a rich source of information for gauging market sentiment; due to their forward-looking nature, futures and options prices efficiently encapsulate market perceptions about underlying asset prices in the future[1]. Building on this foundation, an important concept that investors encounter within these markets is the notion of market-based probabilities.

Theoretically, market-based probabilities allow investors to view the expected values of assets through a lens that discounts the varying attitudes towards risk. Since 1978, when Doug Breeden and Bob Litzenberger[2] proposed a method of using traded options to infer those future probabilities, risk-neutral probabilities have been widely used in market pricing to find intrinsic value. It is worth observing that this research is generally portrayed as uncovering and exploiting mispricing in the market. Excess returns might simply be fair compensation for bearing hard-to-hedge risks that are not included in the pricing model. Or the mispricing may reflect the effects of market constraints, such as short sale restrictions and financing issues[3].

In this paper, our area of focus is on how the distribution of expected underlying asset returns, as outlined by the Minneapolis Fed's data, can be used to provide a more accurate and actionable framework for investment decision-making. To demonstrate the effectiveness of our data, we conducted experiments using the S&P 500 as a case study. Firstly, we evaluate the potency of risk-neutral probabilities in forecasting future market returns and volatility—a facet crucial for investors and policymakers alike by constructing confidence interval. Upon confirming their predictive capability, we employ these projections to anticipate market prices and trends through Lasso regression analysis. Subsequently, we devise and retrospectively evaluate a long-short trading strategy that exploits these probabilities, aiming to exploit and profit from the data inherent in option prices.

2. THE EFFECTIVENESS OF MPD

With Breeden-Litzenberger (1978) method, the Federal Reserve Bank of Minneapolis utilized the option price to compute the market probability distribution of the underlying asset. Subsequently, employing Shimko (1993)[4]'s approach, the bank calculated market probability density functions (MPDs) for a diverse array of asset classes. This methodology is detailed in the Federal Reserve Bank of Minneapolis's documentation on the methodology for current and historical market-based probabilities, published in 2024[5]. However, these foundational articles did not evaluate the predictive accuracy of these estimates, nor did they assess the effectiveness of the MPD. In this context, we introduce methods and approaches to evaluate their predictive efficacy and conduct empirical tests in Section 5.2.

The Minneapolis Fed estimated the distribution of expected returns for the underlying assets over the next six months based on option prices. Our analysis involves hypothesis testing to examine the accuracy of this distribution prediction. Concurrently, we assess the precision of significant price movements predictions for the underlying assets using metrics and methods such as Brier Score, ROC and AUC.

2.1 Hypothesis Testing for Prediction Accuracy

The Minneapolis Fed provides weekly forecasts of the probability distribution of log returns for assets over the subsequent six months. Generally, verifying whether a distribution accurately describes data characteristics is equivalent to exploring the degree to which data conforms to a specific distribution. Statistical methods addressing this issue require a sample set to test their derivation from the presumed distribution. However, for our target data, each distribution forecast corresponds to a single verifiable real data point, precluding the use of traditional distribution verification statistical methods.

Our adopted methodology involves analyzing a sequence of time points from t_0 to $t_n = T$, where at each time point t_i, $i = 0,1, \ldots, n$, there exists a forecast for the distribution of log returns of the underlying asset at $(t_i + 0.5)^1$, denoted as f_i, for $i = 0,1, \ldots, n$. The actual log return of the underlying asset at time t_i is denoted as x_i. For each forecasted distribution f_i, we identify its 10th and 90th percentiles, denoted as $P_{10,\, i}$ and $P_{90,\, i}$, respectively. This allows us to construct an 80% confidence interval $[P_{10,i}\ ,\ P_{90,i}]$. We then compare whether the actual value x_i falls within this 80% confidence interval. If x_i falls within the interval, we mark it as 1; otherwise, we mark it as 0. If the forecast distribution is accurate, the proportion of actual values that fall within the 80% confidence interval should approximate 80%, indicating:

$$\text{Accuracy} = \frac{\sum_i^n 1_{\{x_i \in [P_{10,i}\ ,\ P_{90,i}]\}}}{n} \approx 0.8 \quad (1)$$

where $1_{\{\}}$ is an indicator function that takes the value of 1 if x_i falls within $[P_{10,i}\ ,\ P_{90,i}]$, and 0 otherwise.

By examining the proportion of points that fall within the confidence interval, we can assess the rationality of the distribution forecasts.

2.2 Assessing Predictive Accuracy of Large Price Fluctuations

The Federal Reserve Bank of Minneapolis provides forecasts that include the probability of significant[2] upward or downward movements in the returns of underlying assets. To validate the accuracy of these probability predictions, the actual returns of the underlying assets are calculated to determine if they meet the criteria for substantial fluctuations. At time t_i, if the actual return x_i reaches or exceeds an upward threshold, we assign $I_i = 1$; otherwise, $I_i = 0$. This process creates a categorical variable used to indicate whether the actual returns have experienced a significant increase. The formulation is as follows:

$$I_i = \begin{cases} 1 & \text{if } x_i \geq \text{upward threshold} \\ 0 & otherwise \end{cases} \quad (2)$$

Similarly, we can construct a categorical variable D_i to indicate whether the returns have significantly decreased, with the following formula:

$$D_i = \begin{cases} 1 & \text{if } x_i \geq \text{downward threshold} \\ 0 & otherwise \end{cases} \quad (3)$$

The data analysis then focuses on the probability of these events occurring and the categorical variables indicating whether the events have occurred. To assess the accuracy of these predictions, we employ commonly used metrics to evaluate the performance of classification models. These include the Brier Score[6], ROC[7] (Receiver Operating Characteristic) Curve, and AUC[4] (Area Under the Curve).

The Brier Score is a measure of the accuracy of probabilistic predictions. It is used when predictions must assign probabilities to a set of mutually exclusive outcomes. The score is calculated as the mean squared difference between the predicted probabilities and the actual outcomes, where the outcomes are coded as 1 for the event happening and 0 for it not happening. The formula for the Brier Score is:

$$BS = \frac{1}{n}\sum_{i=1}^{n} (F_i - R_i)^2 \quad (4)$$

Here, n is the number of forecasting instances, F_i is the forecasted probability for the event at instance t_i, and R_i is the actual outcome of the event at instance t_i. The Brier Score should range between [0, 1], with a lower Brier Score indicating a better model fit.[8]

Typically, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) are performance metrics widely adopted in machine learning for evaluating binary classifiers [9]. The Receiver Operating Characteristic (ROC) curve is a graphical plot. The ROC curve is created by plotting the True Positive Rate (TPR)[3] against the False Positive Rate (FPR)[4] at

[1] This paper assumes T = 1 as one cycle, meaning T = 0.5 represents half a year.

[2] An increase or decrease of 20% within six months is considered a significant price movement.

[3] True Positive Rate (TPR): It measures the proportion of actual positives that are correctly identified by the classifier. It is calculated as:

various threshold settings.

To construct a Receiver Operating Characteristic (ROC) curve, the decision threshold of the classifier is varied, with both the True Positive Rate (TPR) and False Positive Rate (FPR) calculated at each threshold level. These rates are subsequently plotted on a graph, positioning the FPR on the x-axis and the TPR on the y-axis.

The area under the ROC curve (AUC) serves as a comprehensive measure for assessing the overall performance of the classifier. An AUC value of 1 indicates an ideal classifier with perfect accuracy, whereas an AUC value of 0.5 denotes a classifier lacking any discriminative power, equivalent to random guessing, and thus unable to differentiate between positive and negative classes effectively.

3. PREDICTIONS BASED ON MPD

Following the validation of the MPD effectiveness, our objective is to predict the returns, volatilities, and reversals of the underlying assets using MPD. We selected the Lasso model to achieve this goal. Section 3.1 will provide a brief overview of the principles behind Lasso, while Section 3.2 will detail how we specifically construct models for different regression targets.

3.1 Method in predictions :lasso

Lasso[10] (The Least Absolute Shrinkage and Selection Operator) stands as an efficient technique for variable selection across various application scenarios. Fundamentally, it incorporates L1 regularization as a penalty term to linear regression. During model training, Lasso applies this penalty to reduce the coefficients of independent variables with minor influence on the dependent variable to zero. This process effectively conducts variable selection within the fitting procedure. The general formula for Lasso is presented as follows:

$$\hat{\beta}^{\text{lasso}} = \arg\min_{\beta} \left(\frac{1}{2n}\|Y - X\beta\|_2^2 + \lambda\|\beta\|_1\right) (5)$$

Here:

Y represents the dependent variable.

X represents the matrix of independent variables.

β is the vector of coefficients.

n is the number of observations.

This formulation allows Lasso to simultaneously perform coefficient estimation and model complexity reduction, making it a valuable tool for achieving sparse models in high-dimensional data settings.

3.2 Predicting Returns, Volatilities, and Reversals

These estimators reflect market expectations for the future. Due to the characteristics of time series data, there is generally a higher correlation between data points that are closer in time. Therefore, we utilize both the estimators and lagged data of the target asset for prediction purposes. Moreover, our focus is on forecasting the changes in the asset's value one week ahead.

In the context of predicting returns, we designate the complete set of estimates data as matrix $E_{n\times m}$, where i and j index the rows and columns, respectively, denoting the observation at the i^{th} time point for the j^{th} feature. Similarly, the data pertaining to the target asset are represented by matrix $A_{p\times q}$. It is important to note that n and p may not necessarily be equal, owing to differences in the frequency of data generation, which results in varying amounts of data. Unless specified otherwise, the data structure within this paper is organized such that each row corresponds to observations of different features at a given time point.

Following the construction methodology typical of time series analysis, we apply lags to all variables intended to serve as independent predictors, incorporating these lagged variables into our predictive model. The dependent variable, or the target of our prediction, is constructed accordingly to facilitate a Lasso regression model. Specifically, only those estimands related to the target asset are selected as predictors. For instance, in predicting the returns of the S&P 500, we utilize the forecast data from options on the S&P 500 that expire in 6 months and 12 months as our independent variables.

To align the frequency of the independent variables, we first adjust the dimensions of $A_{p\times q}$ to match those of $E_{n\times m}$, resulting in a modified matrix $A'_{n\times q}$. This adjustment ensures that all independent variables are based on the same time intervals, facilitating a coherent analysis. Subsequently, we concatenate $E_{n\times m}$ and $A'_{n\times q}$ to form a new matrix, denoted as $[E_{n\times m}, A'_{n\times q}]$. This matrix is then subjected to lag operations to incorporate historical data into the predictive model.

The final matrix of independent variables, X, is composed of the original matrices and their lagged versions, symbolically represented as

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

where TP is the number of true positives, and FN is the number of false negatives.

[4] False Positive Rate (FPR): It measures the proportion of actual negatives that are incorrectly classified as positives. It is calculated as:

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

Where FP is the number of false positives, and TN is the number of true negatives.

$$X = [E_{n\times m}, A'_{n\times q}, \mathrm{lag}_1, ..., \mathrm{lag}_\ell] \quad (6)$$

where lag_ℓ denotes the operation of applying ℓ lags to the matrix $[E_{n\times m}, A'_{n\times q}]$. Symbolically represented as $\mathrm{lag}_\ell[E_{n\times m}, A'_{n\times q}]$. This structured approach captures both the current and historical information deemed relevant for prediction.

Utilizing the Lasso regression model, we engage in modeling for predicting both returns and volatilities. The distinction between these two outcomes lies solely in the dependent variable Y, with all other aspects of the modeling process remaining consistent across both predictions. The Lasso method's incorporation of regularization facilitates the selection of the most informative variables, thereby enhancing the model's predictive performance and mitigating the risk of overfitting.

这里还有reversals的没有写。

4. STRATEGIES

4.1 Baseline Strategy

4.2 Improved Strategy

To more robustly assess the reversibility of stock trends, we need to construct a more stable reversal factor. Traditional reversal factors assume that individual stocks will regress towards the average level of the market:

$$\mathrm{Reverse_i} = \mathrm{mean(Ret_{20})} - \mathrm{Ret_{i,20}} \quad (6)$$

However, this method of taking the average return of the entire market's stocks as a uniform benchmark for all stocks is overly simplistic and crude, as it neglects the specific qualitative information among different industries and companies. Therefore, finding a suitable expected return benchmark for each stock is key to determining the performance of the reversal factor. Tsinaslanidis and Kugiumtzis[2014] posit that similar stock patterns are underpinned by similar background conditions, and by identifying the future performance of stocks with similar historical patterns, one can predict the future returns of a stock[11]. In this context, we require a more extensive set of stock price data and provide a unified framework for constructing reversal factors：

$$\mathrm{Reverse_i} = \sum_{j=1}^{N} \mathrm{w_{ij} Ret_{j,20}} - \mathrm{Ret_{i,20}} \quad (7)$$

$$\sum_{j=1}^{N} \mathrm{w_{ij}} = 1 \quad (8)$$

where N represents the total number of stocks in the market, and w_{ij} is the correlation function regarding stock i and stock j, representing the weighted weight of stock j.

This general form implies that the weighted return of similar stocks serves as the benchmark for their mean reversion. What we need to do is to find the most similar stocks for each individual stock and assign their respective weighted weights. Taking Bank of America as an example, to construct a reversal factor indicator, we can select the 10 stocks with the highest price correlation from the entire banking sector (JPMorgan Chase&Co., Wells Farfo&Co. and so on) and construct the reversal factor according to the previous formula. By constructing reversal factors, investors can discover the true intrinsic value of stocks for reasonable investment strategies, and avoid downside risks when the market is overheated to achieve better risk management.

5. EMPIRICAL ANALYSIS

The above section has already introduced the method we use to solve problems. Next, we will explain how we tested our ideas on real data using the aforementioned method.

5.1 Data

We first introduce the data utilized in our study. The estimands under evaluation are derived from the Federal Reserve Bank of Minneapolis[5], which provides Statistics for Market-Based Probability Densities (MPDs) calculated from fifty-two different options or forward prices across six asset categories[6]. The original dataset comprises 12,249 observations and 14 independent variables. The data dictionary for these Statistics is as follows:

Table 1. Data Dictionary for Statistics for (MPDs)

Variables	Description
market	Represents the category of the option or forward, encompassing a total of 52 types.
idt	Indicates the date of the forecast, including year, month, and day.
maturity_target	The time-to-expiry target for options selection
mu	The mean of the MPD; mu = sum(probability at ln-return x * ln-

[5] The data originate from : https://www.minneapolisfed.org/banking/current-and-historical-market--based-probabilities
[6] The six asset categories are: Rates, Inflation, Equity, Currency, Commodity, Bank.

	return x)
sd	The standard deviation of the MPD; sd = sqrt(sum(probability at ln-return x * (ln-return x - mu)^2))
skew	The skew of the MPD; skew = sum(probability at ln-return x * (ln -return x - mu)^3)/(sd^3)
kurt	The kurtosis of the MPD; kurt = sum(probability at ln-return x * (ln-return x - mu)^4)/(sd^4)-3
p10	The 10th percentile of the MPD
p50	The 50th percentile of the MPD
p90	The 90th percentile of the MPD
lg_change _prob	The change in the expected return in percentage terms defined as "large"
prDec	The probability of a "large decline" in return as defined by "lg_change_prob"
prInc	The probability of a "large increase" in return as defined by "lg_change_prob"

For different underlying assets, the Federal Reserve Bank of Minneapolis provides forecasts weekly or bi-weekly. The data coverage period varies across different markets. We selected markets that have complete data available from January 10, 2013, to February 7, 2024, and subsequently removed any data prior to January 10, 2013. Before September 4, 2014, the Federal Reserve Bank of Minneapolis issued data forecasts bi-weekly. To ensure consistency in data interval frequency, we used forward filling to address missing data, resulting in a weekly data frequency after the adjustment. Through such filtering, the processed data includes 8,080 forecasts for 14 markets from January 10, 2013, to February 7, 2024, with a weekly data interval frequency. We refer to the processed data as mpd_stats_cleaned.

When forecasting returns, volatilities, and reversals, our analysis primarily focuses on S&P 500 data. Therefore, we obtained all price information for the S&P 500[7] and VIX[8] for the target period from Yahoo Finance.

理论上应该说一下标普500和VIX滚动平均的处理，我再看看。

B-S model假设的底层资产是log normal的，但是给出的估计量都是偏态的（肥尾），那么分析波动性是很有意义的。[12]

这里其实最好再有一个数据预测的时间对应的图，不然语言总是说不清楚。

[7] The data originate from : https://finance.yahoo.com/quote/%5EGSPC

[8] The data originate from : https://finance.yahoo.com/quote/%5EVIX?.tsrc=fin-srch

5.2 Implementation

Building on the previously outlined data processing and methodological framework, we now provide a detailed examination of the performance of the aforementioned methods on actual data sets.

Given that the Federal Reserve Bank of Minneapolis has conducted MPD calculations using options on the S&P 500 index with maturities of 6 months and 12 months, and the data have been adjusted to reflect the distribution of expected values six months after each observation date, our analysis will incorporate both the 6-month (sp6m) and 12-month (sp12m) option data.

5.2.1 Validating the Effectiveness of MPDs through S&P 500 Data

In line with the approach described in Section 2.1, there are 578 observations each for sp6m and sp12m over the entire period. Given that both datasets pertain to the same underlying asset, we analyze the data from both markets. Our calculations reveal that the within_interval_accuracy for sp6m predictions is 0.8616, while for sp12m predictions, it is 0.9446. This indicates that the predictions for sp6m are relatively close to the 80% mark, though there are reasons for the discrepancies observed: Firstly, the limited number of available data points for validation (578) does not fully attest to the accuracy of the predictions, suggesting potential data biases; moreover, the predictions appear to be conservatively biased, as more actual data fell within the predicted intervals than expected, thus reducing the probability of Type I errors while increasing the probability of Type II errors. Overall, we consider these estimates to be usable, albeit with a possibly underestimated variance.

The accompanying figure illustrates that the majority of the data points fall within the 80% confidence interval.

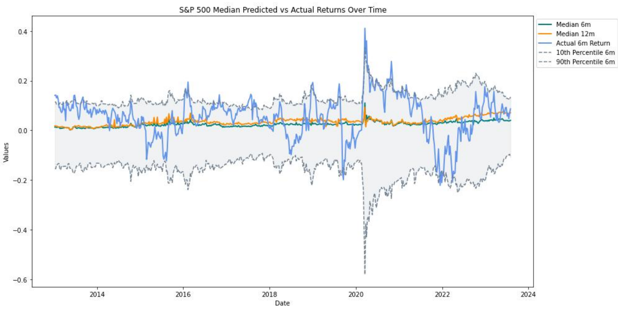

Figure 1 S&P 500 Median Predicted vs Actual Returns Over Time

(这个图在这里的排版还要再调，我是想的放大，在两栏中间)

Next, we analyze the predictions of significant price movements in the S&P 500 from 2013 to 2023. Initially, we calculate the actual price changes within a six-month period and then determine whether these changes meet or exceed an upward or downward threshold, which we have set at 20%. If met, we mark it as 1; if not, as 0. Subsequently, we calculate the Brier Score and AUC for both sp6m and sp12m predictions. The results are presented in Table 2.

Table 2. Brier Score and AUC for Significant Price Movement Predictions by sp6m and sp12m

	Brier Score	AUC
sp6m Increase	0.0124	0.9888
sp6m Decrease	0.0138	0.6867
sp12m Increase	0.0192	0.9560
sp12m Decrease	0.0240	0.6841

The data reveal that both sp6m and sp12m perform exceptionally well in predicting significant upward movements, with sp6m yielding slightly better results, achieving an AUC of 0.9888. This superiority can be attributed to the Breeden-Litzenberger approach, which forecasts the expiration day prices for futures, options, and forwards. The Federal Reserve Bank of Minneapolis builds on this by calculating predicted distributions of returns, adjusting all data to an expected value distribution within a six-month period following each observation date. Since sp6m directly predicts the returns after six months, its predictions tend to be more accurate.

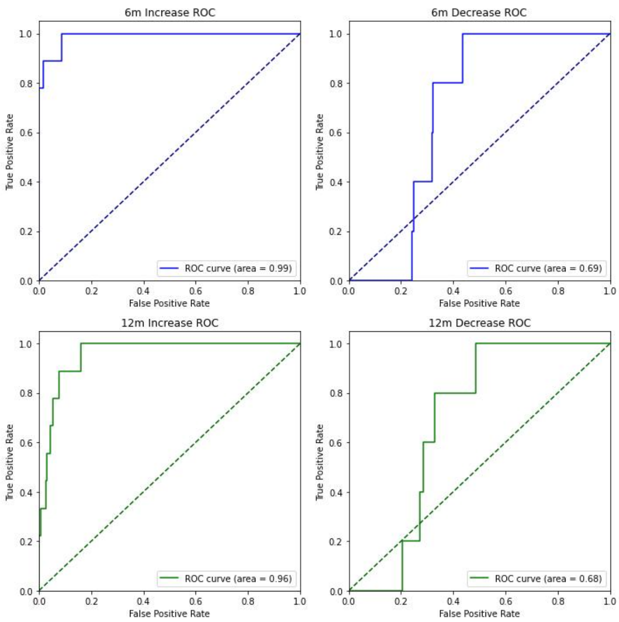

Figure 2 ROC Curve for Large Rise and Fall Predictions

5.2.2. 通过MPD预测

分别写对returns，volatilities，reversals的预测。

5.2.3. 策略的实施

这里放各种结果的图。

6. Conclusions and discussions

REFERENCE

[1] P. E. Tsinaslanidis and D. Kugiumtzis, "A prediction scheme using perceptually important points and dynamic time warping," Expert Systems with Applications, vol. 41, no. 15, pp. 6848-6860, 2014.

[2] Breeden, D. T., & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. The Journal of Business, 51(4), 621–651. http://www.jstor.org/stable/2352653

[3] Figlewski, S. (2018). Risk-Neutral Densities: A Review. Annual Review of Financial Economics, 10, 329-359. https://doi.org/10.1146/annurev-financial-110217-022944

[4] Shimko, David. (1993). Bounds of probability. Risk. 6. 33-37.

[5] Federal Reserve Bank of Minneapolis. (2024). Methodology for current and historical market-based probabilities. Retrieved from https://www.minneapolisfed.org/-/media/assets/banking/current-and-historical-market-based-probabilities/methodology.pdf

[6] BRIER, G. W. (1950). VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY. Monthly Weather Review, 78(1), 1-3. https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

[7] Zou, K.H., O'Malley, A.J., Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation, 6;115(5):654–7.

[8] Wikipedia contributors. (2024, February 25). Brier score. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Brier_score

[9] Wikipedia contributors. (2024, February 25). Receiver operating characteristic. In Wikipedia, The Free Encyclopedia. Retrieved [Your Date of Access], from https://en.wikipedia.org/wiki/Receiver_operating_characteristic#Curves_in_ROC_space

[10] Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological), 58(1), 267–288. http://www.jstor.org/stable/2346178

[11] Bahra, B. (1997). Implied Risk-Neutral Probability Density Functions from Option Prices: Theory and Application (Bank of England Working Paper No 66). Bank of England. Posted: 19 Apr 1998.

[12] Figlewski, Stephen, Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio (July 30, 2008). VOLATILITY AND TIME SERIES ECONOMETRICS: ESSAYS IN HONOR OF ROBERT F. ENGLE, Tim Bollerslev, Jeffrey R. Russell and Mark Watson, eds., Oxford, UK: Oxford University Press, 2008, Available at SSRN: https://ssrn.com/abstract=1256783