

# Calibration of the Ross Recovery Theorem to Real-world Data, and Tests of its Practical Value

Ling Lan                      Zhengxu (Andrew) Li  
ll3178@nyu.edu              zhengxu.li@nyu.edu  
New York University        New York University

Advisor: Robert V. Kohn  
kohn@cims.nyu.edu  
New York University

March 14, 2018

## Abstract

The Ross Recovery Theorem [1] challenges the commonly held thought that derivative prices do not contain useful predictive information about the distribution of financial variables (such as an equity index). The theorem recovers, under certain hypotheses on the character of the market, the subjective probability distribution of an equity index from current derivative prices. In this paper, building on the method of Backwell [2] for extracting state prices from option prices, we develop a strategy for combining option data with the Recovery Theorem to estimate the subjective distribution. Using real-world data, we then investigate whether the Recovery Theorem yields predictive information and has practical value, concluding that the answer might be no.

## 1 Introduction

It is generally believed that derivative prices do not contain useful information about the subjective distribution of financial variables. In the oldest model, people believe that, by the combination of derivatives with the maturity  $T$ , they can get any payoff. In this model, with the assumption that the market is free from arbitrage and the linearity relationship between derivative prices and the payoff, derivative prices determine a unique pricing measure. In this model involving derivative prices, there is no reference to subjective probability or its expectation.

The world-famous Black-Scholes model [7] also involves derivative prices. In this model, it is believed that if the market is complete, then derivative prices are determined by hedging. In a Black-Scholes world, derivative prices only carry information about volatilities, not about the drift, which indicates the future trend. Therefore, we cannot derive any information about subjective distribution of the future from it. However, the real world is not the Black-Scholes world: in reality, derivative prices are set by supply and demand. The Black-Scholes model, without considering the supply-and-demand relationship, could miss some important information when modeling the real-world situation.

Recently, Steve Ross came out with the Recovery Theorem, which challenges the views of the previous two models. In his opinion, since derivative prices come from the aggregated actions of investors with their own opinions about the future, a model involving derivative prices should also involve the investors' subjective probability distributions about the future. However, it would be impractical to include in a model each investor's subjective distribution, which could be very different from one another. Therefore, Ross makes a simple assumption: there is a central planner optimizing a representative agent's expected utility over consumption. The probability distribution involved in this optimization problem is the representative agent's subjective distribution. In this model, with the help of Perron-Frobenius Theorem, one can deduce the representative agent's subjective probability distribution from market data.

After the subjective probability distribution is obtained, some natural questions emerge: is the subjective distribution meaningful and useful? How well is the Recovery Theorem satisfied

by real-world prices? And what is the degree of match to market data? The goal of our project is to investigate how the Recovery Theorem can be calibrated to real-world data and to test its significance and practical value. In the end, we were able to calibrate the theorem, but we were not able to find any usefulness of it. This conclusion is different from some other people's, include that presented in Audrino et al's paper [8]. We will compare and contrast the results in this paper.

This paper is divided into four sections. Section 2 introduces the Recovery Theorem, and offers our own analysis of it. Section 3 summarizes our method to calibrate the theorem, explains in detail how we arrived at our specific method, and tests the robustness of the method. In section 4, by analyzing the probability distributions we get from our implementation and by applying the distribution to real-world trading, we explore the meaningfulness and usefulness of the Recovery Theorem. Section 5 concludes our paper.

It is worthwhile to mention other researchers' investigations into the Ross Recovery Theorem here. Walden [9] changes the assumption of a finite state space, and analyzes the possibility of recovering the probability distribution when the state variable is an unbounded diffusion process, concluding that in many cases the recovery is possible. Hansen and Scheinkman [10] factorize the Markovian pricing kernel, and discuss the limitation of Ross' assumption of transition independence of the Markovian pricing kernel, showing that transition independence is not always true in some cases. Qin and Vadim extend Hansen and Scheinkman's results on factorization, and thus extending the Recovery Theorem to continuous time. Audrino et al [8] conduct an empirical analysis of the Ross Recovery Theorem, which we will discuss in the conclusion section.

## 2 A Summary of the Recovery Theorem

The goal of this section is to introduce readers to some important results of the Recovery Theorem, upon which the subsequent sections are built. The theorem is due to Ross [1]. We include below a brief summary of the derivation of the theorem by Backwell, whose treatment of the theorem is mathematically more rigorous than Ross'. We present the most relevant results, with assumptions written in bold. For a complete derivation, please refer to section 2 of [2].

Notice that our final goal is to recover the subjective probability distribution of an equity index, which is a random variable in our context. Usually, an equity index can take any value, and therefore the probability problem should be in a continuous setting. But to simplify the problem, we assume that **the equity index is discrete and bounded**, and can only take values in  $\Psi := \{w_1, w_2, \dots, w_n\}$ . We also consider the problem as **a one-step discrete Markov chain**, where we only focus on the movement of the equity index from the current time 0 to the next time  $T$  (and do not consider any future beyond  $T$ ).

In this setting,  $\Psi := \{w_1, w_2, \dots, w_n\}$  becomes the state space of the Markov chain, and we want to get the transition matrix  $\mathbb{F}$  (an  $n \times n$  matrix with  $\mathbb{F}_{ij}$  defined as the probability to be in state  $w_j$  in the next time period if the current state is  $w_i$ ). Notice that  $\mathbb{F}$  contains all the information needed to describe the probabilities of how the system evolves, if we know the initial state.

Now we assume that **there exists a representative agent with time-additive intertemporal Expected Utility Theory (EUT) preferences over consumption**. The agent has an initial wealth of  $W$ , and he or she can either consume his or her wealth now or invest it to consume later. The problem he or she faces in the EUT model is to decide how much to consume now, and how much to invest now (to consume later) to maximize the expected utility over consumption.

The investment tools available in our model are Arrow-Debreu securities. (In our context, an Arrow-Debreu security for  $w_j$  is an asset that pays off \$1 if the index is  $w_j$  in the next time step, and pays off \$0 otherwise.) We define  $P := (p_{ij})$  as the Arrow-Debreu price matrix. This is an  $n \times n$  matrix where  $p_{ij}$  is the Arrow-Debreu security price, or state price, available in the initial state  $w_i$  and expiring in-the-money at  $T$  if the state variable takes on some particular  $w_j$ . Backwell assumes **a complete market free from arbitrage**: a full set of Arrow-Debreu securities is available in the beginning. We treat  $P$  as exogenous in this section.

Assume that the index is  $w_i$  in the beginning. Now define  $C := \{c_i^0, c_{i1}^T, \dots, c_{in}^T\}$ , where the initial consumption is  $c_i^0$ , and the consumption at  $T$  is  $c_{ij}^T$  (contingent on future state  $w_j$ ). If the agent wants to consume  $c_{ij}^T$  at  $T$  in the case where the index is  $w_j$ , he needs to invest in  $c_{ij}^T$  Arrow-Debreu securities for  $w_j$ , the price of which in the initial state is  $p_{ij}$ . Therefore

$$W = c_i^0 + \sum_{j=1}^n p_{ij} c_{ij}^T. \quad (1)$$

Now we continue to formalize the expected utility of consumption. We assume a simple time-additive extension. Therefore, we want to maximize  $E[U(c_i^0) + \delta U(c^T)]$  where  $c^T$  is the random variable representing the amount of actual consumption at  $T$ ,  $U$  is the utility function, and  $\delta$  is the utility discount factor: people generally prefer consumption now to consumption later. Notice that  $E[U(c_i^0)] = U(c_i^0)$  and  $E[\delta U(c^T)] = \delta E[U(c^T)] = \delta \sum_{j=1}^n \mathbb{F}_{ij} U(c_{ij}^T)$ . Therefore, our optimization problem is

$$\max_C U(c_i^0) + \delta \sum_{j=1}^n \mathbb{F}_{ij} U(c_{ij}^T) \quad (2)$$

such that (1) holds.

By solving the Lagrangian of the optimization problem and by assuming that **consumption is decided by the current state only, not by time or by the previous state**, that is  $c_i^0 = c_{ij}^T =: c_i$  (an equilibrium condition on consumption), we get

$$\frac{p_{ij}}{\mathbb{F}_{ij}} = \frac{\delta U'(c_j)}{U'(c_i)}.$$

Defining  $D := \text{diag}(U'(c_j))$ , it can be easily derived that

$$\delta F D = D P. \quad (3)$$

In his paper, Backwell proves that  $F$  can be uniquely determined by

$$F = \frac{1}{\delta} D P D^{-1}$$

where  $\delta$  is proved to be the eigenvalue corresponding to a unique positive eigenvector  $z$  of  $P$  (uniqueness being up to scaling), and  $D$  satisfies  $z := D^{-1}e$ . (Therefore, the choice of the utility function is irrelevant to our problem.) The existence and uniqueness of  $F$  is given by the Perron-Frobenius Theorem. Backwell formalizes the theorem as follows.

**Theorem 1:**

*Assuming*

- *a discrete and bounded state variable that makes one (Markovian) transition in an arbitrage-free and complete market,*
- *the equilibrium condition on consumption,*
- *the existence of a representative agent with time-additive intertemporal EUT preferences over consumption,*

*then, given an irreducible state-price matrix  $P$ , we can uniquely recover the real-world transition matrix  $F$ .*

Notice that Backwell's derivation only focuses on one time step: the Markov chain has only one transition, and the expected utility function to be maximized only concerns consumption at the initial time and the next time. Backwell himself has remarked that though such a modeling choice aids tractability and it is sufficient to consider only one step for a Markov chain, it simplifies the preference aspect of the model and makes decision-making myopic: the agent doesn't consider the future beyond  $T$ .

We have also considered a model based on Backwell's by adding one more step to the Markov chain. Our two-step analysis arrives at the same result as Backwell's one-step analysis. Carr and Yu [6] have also examined the Recovery Theorem and obtained an analogue of the result in a continuous setting.

### 3 Finding the Subjective Probability from Market Data

The previous section explains how to derive  $F$  from the state price matrix  $P$ . The remaining problem is how to estimate  $P$ , which was treated as exogenous in the previous section. The financial literature provides many methods to derive state prices from option prices. The method used by Backwell is built on [4] and [5]. We modified Backwell's method to build a more robust algorithm to obtain  $P$ .

In this section, we will first introduce our method to estimate  $P$ . Then, we will go back to explain our motivation for modifying Backwell's method. Finally, we test the robustness of our method.

#### 3.1 Method

First, we consider a subproblem: how to obtain the prices of Arrow-Debreu securities which expire at time  $t$  (we will call them  $t$ -dated prices). Suppose, as in the previous section, there are  $n$  states (that is  $n$  index levels), and then there are  $n$  state prices. We use  $n \times 1$  vectors  $p^{(t)}$  and  $S$  to represent  $t$ -dated state prices and index levels respectively.

Notice that, in a complete market where an Arrow-Debreu security exists for each possible state, we can obtain the price of any asset once we know the payoff of the asset at each state. For example, assume that there are 3 states  $s_1, s_2$ , and  $s_3$ , and  $AD1, AD2$ , and  $AD3$  are Arrow-Debreu securities for states  $s_1, s_2$ , and  $s_3$  respectively. Their prices are  $S_1, S_2$ , and  $S_3$ . If there is an Asset  $X$  which pays off  $X_1, X_2$ , and  $X_3$  at state  $s_1, s_2$ , and  $s_3$  respectively. Then buying one  $X$  is equivalent to buying  $X_1$  numbers of  $AD1, X_2$  of  $AD2$ , and  $X_3$  of  $AD3$ . Then the price of  $X$  must be  $X_1S_1 + X_2S_2 + X_3S_3$ , otherwise arbitrage is possible.

With this knowledge, we can come out with many equations involving state prices, asset payoffs at each state, and an asset price. In our model, we choose three types of assets (or investment tools):  $t$ -dated bonds (with continuous yield  $r^{(t)}$ ),  $t$ -dated futures (with price  $F^{(t)}$ ), and  $m$   $t$ -dated options (each with the price  $H_i^{(t)}$  and the strike  $X_i$ ). Then the following equations must hold

$$\sum_{j=1}^n p_j^{(t)} = e^{-r^{(t)}t} \quad (4)$$

$$\sum_{j=1}^n p_j^{(t)} S_j = F^{(t)} e^{-r^{(t)}t} \quad (5)$$

$$\sum_{j=1}^n p_j^{(t)} \max(S_j - X_k, 0) = H_k^{(t)} \quad (6)$$

$$k = 1, 2, \dots, m$$

In equation (4), the bond has a one-dollar payoff no matter what state we will be in at time  $t$ , and the current price of such a bond is just the bond's current discounted value. In equation (5), the payoff of a futures will just be the value of the stock at time  $t$ , and the current price of the futures should be discounted. Equation (6) follows the same idea, with  $\max(S_j - X_k, 0)$  representing the payoff of the option at each state  $H_k^{(t)}$  representing the price.

In his paper, Backwell obtained  $p^{(t)}$  by solving an optimization problem

$$p^{(t)} = \arg \min_{p \geq 0} \sum_{j=1}^n (p_{j-1} - 2p_j + p_{j+1})^2$$

such that (4), (5), and (6) hold.

In his optimization problem, equations (4), (5), and (6) are used as constraints. The optimization objective introduces a bias toward the smoothness of the discrete distribution of  $p^{(t)}$ . When we applied real-world data to his method, we found that equations (4), (5), and (6) were inconsistent when  $m$  was large. This is due, presumably, to the noises in the data. The real-world situation cannot be ideal and an unreasonable price could occur sometimes. Therefore, the equations could break down when we have too many inconsistent data. (More details about this will be discussed

in a subsequent subsection.) Therefore, we only keep (4) and (5) as the constraints, and include (6) in the function to be optimized. Our optimization problem is

$$p^{(t)} = \arg \min_{p \geq 0} \sum_{j=1}^n (p_{j-1} - 2p_j + p_{j+1})^2 + \sum_{k=1}^m \lambda_k \left( \sum_{j=1}^n p_j^{(t)} \max(S_j - X_k, 0) - H_k^{(t)} \right)^2 \quad (7)$$

$$k = 1, 2, \dots, m$$

such that equations (4) and (5) hold.

The function to be optimized (equation (7)) consists of a term featuring smoothness (the left term, which we will call the smoothness term) and a term featuring match to the market data (the right term, which we will call the error term). The smoothness term ensures the smoothness of the discrete distribution of  $p^{(t)}$ , and the error term is adapted from equation (6), making sure  $p^{(t)}$  does not deviate too much from the original constraint. Notice the introduction of  $\lambda_k$  before each error. In a subsequent subsection, we will test different values of  $\lambda$  to get the one which results in great smoothness and small error.

Define  $T$  as a fixed time horizon. Provided that we know the corresponding asset information, we can obtain  $p^{(T)}$ ,  $p^{(2T)}$ ,  $p^{(3T)}$ ,  $p^{(4T)}$ ,  $p^{(5T)}$  by solving the optimization problem above. (In application, we use Matlab's quadratic programming function *quadprog* to solve the optimization problem.)

Knowing  $p^{(T)}$ ,  $p^{(2T)}$ ,  $p^{(3T)}$ ,  $p^{(4T)}$ ,  $p^{(5T)}$ , our problem now becomes how to obtain state price matrix  $P$  from them. Ross suggests that, assuming that  $P$  is time-homogeneous, then

$$(p^{(i+1)T})^{tr} = (p^{iT})^{tr} P \quad (8)$$

$$i = 0, 1, 2, \dots$$

where  $p^{(0)}$  is simply a column of zeros with a one in the middle entry.

Also notice that we can write analogues to equations (4) and (5) to put more restrictions on  $P$

$$\sum_{j=1}^n p_{ij} \leq 1 \quad (9)$$

$$i = 1, 2, \dots, n$$

$$\sum_{j=1}^n p_{ij} S_j = S_i e^{-q^{(T)} T} \quad (10)$$

$$i = 1, 2, \dots, n$$

Equation (9) is an analogue to equation (4), which involves the bond: as we can only see the bond price relating to the current state, we cannot write a direct analogue, but we enforce a positive bond interest rate. Equation (10) is an analogue to equation (5), which involves futures: since we only know the futures price relating to the current state, we use the implied dividend yield  $q^{(T)}$  times the initial state  $S_i$  to estimate the futures price.

With equations (9) and (10) as constraints, we now define an optimization problem over  $P$ . The function to be optimized consists of

$$S(P) := \sum_{i=1}^n \sum_{j=1}^n (p_{i,j-1} - 2p_{i,j} + p_{i,j+1})^2 \quad (11)$$

which is a smoothness function similar to the one in the previous subproblem, and

$$D(P) := \sum_{k=1}^4 (p^{((k+1)T)} - p^{kT} P)(p^{((k+1)T)} - p^{kT} P)^{tr} \quad (12)$$

which is adapted from equation (8). We solve

$$P^{(t)} = \arg \min_{P \geq 0} (S(P) + D(P)) \quad (13)$$

such that equations (9) and (10) hold.

The optimization objective ensures that the smoothness is improved as much as possible while  $P$  does not deviate too much from equation (8). Again, we use the Matlab function *quadprog* to solve this optimization problem in our implementation.

Now that we get  $P$ , we can calculate  $F$  by the formula given in the previous section

$$F = \frac{1}{\delta}DPD^{-1}$$

where  $\delta$  is the eigenvalue corresponding to a unique positive eigenvector  $z$  of  $P$  (uniqueness being up to scaling), and  $D$  satisfies  $z := D^{-1}e$ .

### 3.2 Data and Variable Choices

This subsection lists what real-world data and variable values we use in our application, and briefly explains why we have made such data and variable choices.

- **State space**  $S = \{S_1, S_2, \dots, S_n\}$ : we are interested in the subjective probability distribution of S&P500; therefore,  $S$  is an  $n \times 1$  vector consisting of  $n$  S&P500 index levels. We let  $n = 21$ . At each of the 35 time points, we capture the current index level (**data resource: Wharton Research Data Services, WRDS**), and let it be the middle entry of  $S$  ( $S_{11}$ ). The interval between each entry is 5% of  $S_{11}$ , so  $S_1 = \frac{S_{11}}{2}$  and  $S_{21} = \frac{3S_{11}}{2}$ . We choose a 5% interval and let  $n = 21$  to make the state space  $S$  include all the possible index levels the SPX might end up in in the real world at the next time point: it is highly unlikely that the SPX index will fluctuate more than 50% for a short period of time.
- **Time range**  $T$ : we let the time range  $T = 3$  months, or 0.25 year. This choice comes from the fact that the expiration time of many assets (such as futures and options) is a multiple of 3 months.
- **Option data: strike price  $X_i$  and option price  $H_i^{(t)}$  (data resource: WRDS)**: there are two types of options in the market: American options, which can be exercised on or before the expiration date, and European options, which can only be exercised on the expiration date. Equation (6) assumes that the options can only be exercised on the expiration date. Therefore, European options on S&P500 must be used as our option data. To obtain the price of an option, we take the average of its lowest ask price and highest bid price—a common way to calculate an option price.
- **Futures price  $F^{(t)}$  (data resource: the Bloomberg Terminal)**: we use the futures on the SPX.
- **Treasury yield  $r^{(t)}$  (data resource: WRDS)**: our method requires the Treasury yields for the maturity time of  $T, 2T, 3T, 4T$ , and  $5T$  ( $T = 3$  months), but there is no Treasury yield data for 9-month and 15-month maturities. To determine the rates, we linearly interpolate between quoted maturities.
- **Dividend yield  $q^{(T)}$  (data resource: WRDS)**: we use the dividend yields of the SPX.
- Our method requires getting  $p^{(T)}, p^{(2T)}, p^{(3T)}, p^{(4T)}$ , and  $p^{(5T)}$ , but sometimes there is no option data available for  $5T$ . In this case, we simply ignore  $p^{(5T)}$  and remove the term involving  $p^{(5T)}$  in  $D(P)$  (equation (12)).

### 3.3 Improvements of Backwell's Method

When introducing the optimization problems, we mention that we have modified Backwell's method to get ours; however, we do not spend much time explaining our motivation for doing so. In this subsection, we will go back to Backwell's method, examine its defects, and explain how we start from Backwell's method and end up at our own.

For all the tests in this subsection and the next subsection, we use the data on 6/15/2001 as an example. The S&P 500 index at that time is 1214.36. Table 1 shows other relevant data information.

Table 1: Information of derivative prices starting at 2001/06/15

	$p^T$	$p^{2T}$	$p^{3T}$	$p^{4T}$	$p^{5T}$
Date	2001/09/17	2001/12/17	2002/03/15	2002/06/17	2002/09/16
Futures price	1226	1234.5	1243.5	1255	1267
Interest Rate	0.0352	0.0355	0.0354	0.0353	0.0441
# of Options available	46	36	26	27	0

• **Backwell’s optimization problem and its defects**

As mentioned in 3.1, Backwell uses the relationship of state prices and prices of investment tools (i.e. the price of the Treasury bond, futures, and options) as constraints, and the remaining freedom to maximize the smoothness of the discrete distribution of  $p^t$ . His optimization problem is (with  $m$  options)

$$p^{(t)} = \arg \min_{p \geq 0} \sum_{j=1}^n (p_{j-1} - 2p_j + p_{j+1})^2 \text{ such that } \begin{cases} \sum_{j=1}^n p_j^{(t)} = e^{-r^{(t)}t} \\ \sum_{j=1}^n p_j^{(t)} S_j = F^{(t)} e^{-r^{(t)}t} \\ \sum_{j=1}^n p_j^{(t)} \max(S_j - X_k, 0) = H_k^{(t)} \end{cases}$$

However, we notice two flaws in Backwell’s optimization problem:

- The state space  $S$  creates  $n + 1$  intervals. Our analysis has shown that if the system uses more than one option in the same interval, there is a high chance that the constraints in the optimization problem will contradict each other. Therefore, to obtain a feasible solution, **a large amount of option data has to be discarded**. In this case, **the solution  $p^{(t)}$  is very likely to be biased** by the selection of data. Thus, even if a combination of options is found, for which the system yields a feasible solution, it cannot be fully trusted.
- Additionally, real-world data inevitably contain noises. Therefore, using many option data as constraints makes **the solution not very smooth**, since the freedom left for optimization of the smoothness is largely reduced.

The above observations are consistent with the results we got when implementing Backwell’s method.

• **Implementation of Backwell’s method**

In order to confirm our analysis on the defects of Backwell’s design, we experiment with his method of calculating  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$ , and  $p^{4T}$  using data on 6/15/2001. Among our tests with all combinations of option data, there is no feasible solution if one interval contains more than one option. Beside this, even when no interval is filled with more than one option, there is often no solution with each entry positive. To find an ideal data set with feasible solutions which is as smooth as possible, we designed a simple "greedy algorithm" to do so, which

- selects the option in the middle state space interval, with which a feasible solution exists and the smoothness term is minimized compared to the conditions using other options in this interval. Fix this option as given in future steps.
- repeats the above process from middle intervals to intervals on the edge.
- skips the interval, if adding any option in this interval does not provide feasible solutions.

Figure 1 plots the solution of Backwell’s optimization problem on 6/15/2001 using the greedy algorithm. The optimized data combination chosen by the algorithm contains 11 options out of 46 available options expiring at 2001/09/22 when estimating  $p^T$ . That is, more than 75% of data are wasted. However, even if using this selective data set, the optimization problem offers a solution with poor smoothness, whose distribution oscillates drastically as shown in the figure.

It is surprised that our result of Backwell's optimization problem is less smooth than the result he provided in his paper, even though we have thrown out the noisy data. We doubt that it is due to the difference between the South African market (where Backwell has obtained his data) and the U.S. market (where we have obtained the data). But further research is needed to confirm our speculation.

Also, our greedy algorithm is designed in such a way that it might be hard to get a result with small noises. Our greedy algorithm regards the task of choosing as many option data as possible as its first priority, and then it maximizes the smoothness term. Therefore, even if a set of data is found, the smoothness of the result could be not good. Additionally, the robustness test of the greedy algorithm (illustrated in Section 3.5) shows that this algorithm is very sensitive to the change of the interval length of the current SPX index. Therefore, we intend to improve Backwell's optimization problem to take advantage of the data more efficiently without too much expense on the smoothness of the solution.

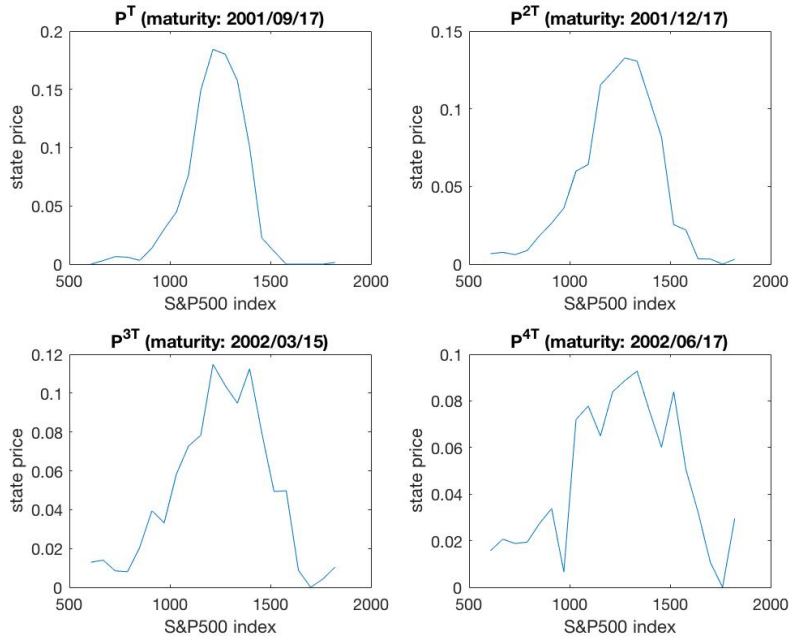


Figure 1:  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$ ,  $p^{4T}$  estimated by Backwell's method using the greedy algorithm. (# of state space: 21; interval length: 5% of current SPX index)

- **Improvements of Backwell's method**

The previous analysis, as well as the tests, demonstrates that it might not be a good idea to include option data as a constraint. Therefore, we improve Blackwell's optimization problem by treating options as part of the "goal" instead of the "restriction". Our updated optimization problem with  $m$  options is

$$p^{(t)} = \arg \min_{p \geq 0} \sum_{j=1}^n (p_{j-1} - 2p_j + p_{j+1})^2 + \sum_{k=1}^m \lambda_k \left( \sum_{j=1}^n p_j^{(t)} \max(S_j - X_k, 0) - H_k^{(t)} \right)^2 \quad (14)$$

$$\text{such that } \begin{cases} \sum_{j=1}^n p_j^{(t)} = e^{-r^{(t)}t} \\ \sum_{j=1}^n p_j^{(t)} S_j = F^{(t)} e^{-r^{(t)}t} \end{cases}$$

This method minimizes the smoothness of the solution and the error of the solution's deviation from option data at the same time. Each option  $k$  has a multiplier  $\lambda_k$ , which means our confidence in it. The higher  $\lambda_k$  is, the more influential the option  $k$  will be in this optimization problem. Therefore, besides the advantage that the solution to this updated



method always exists, compared to Backwell's method, this strategy is also able to utilize all option data and balance the idealization of the state price model and the significance of the market information. In Figure 2, which shows the estimated  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$ , and  $p^{4T}$  by the improved optimization problem with  $\lambda_k = 10^{-5}$ , we can see that the smoothness of the result is improved compared to the Backwell implementation in Figure 1. In this case, the improved method solves the predicament of data selection and transforms the influence of data noise into the calibration of multipliers  $\lambda$ . Details of the calibration are discussed in the following parts.

### 3.4 Calibration of the Multipliers

When we explain our method in 3.1 and 3.3, we introduced the multiplier  $\lambda_k$  of each option data  $k$  to balance the smoothness of the solution and its accuracy to real-world data, but did not discuss the calibration of multipliers. In this section, we explain how we determine the coefficients that best fit the data and analyze our decision by testing the outliers. A good selection of  $\lambda_k$ s should provide satisfying smoothness and not deviate too much from the market data. Our research of  $\lambda$  contains two aspects, which are finding the appropriate range and the distribution of "good"  $\lambda_k$ s. The range of the multipliers is a constant, around which the  $\lambda_k$ s balance the smoothness and the accuracy of the solution. The distribution of  $\lambda_k$  shows our different confidence in different option with the same maturity date.

#### 3.4.1 The Range of the Multipliers

In order to balance the relative importance of the smoothness term and the error term in equation (14), the range of the multipliers should depend on the range of both terms. The smoothness term is normally around  $10^{-3}$ , and the sum of errors is around  $10^2$ . Therefore, a constant  $\lambda$  of  $10^{-5}$  (no difference between options), which treats the two goals with equal attention, seems a reasonable start. Figure 2 shows the estimated  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$ , and  $p^{4T}$  by the improved optimization problem with  $\lambda_k = 10^{-5}$  ( $k = 1 \dots m$ ).

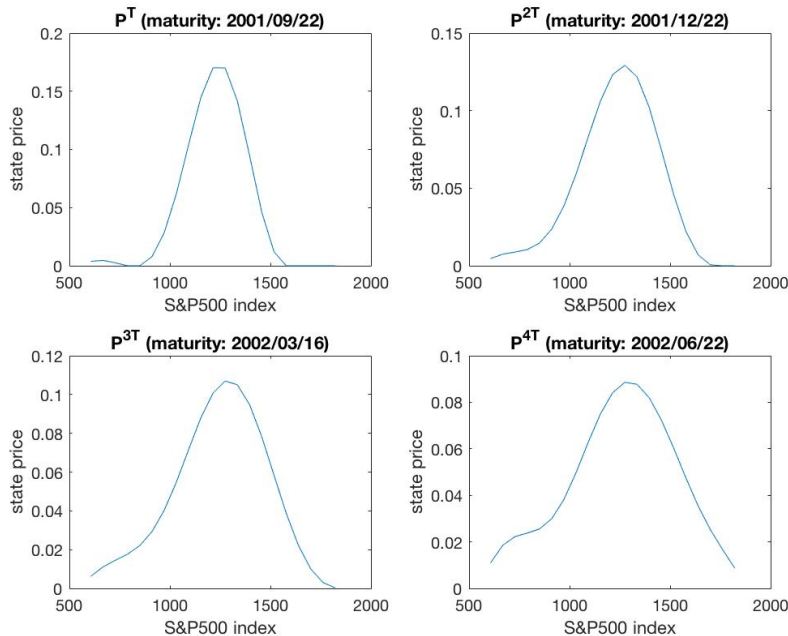


Figure 2:  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$ , and  $p^{4T}$  estimated by improved method with constant  $\lambda = 10^{-5}$  (# of state space: 21; interval length: 5% of current SPX index)

The maximization of the smoothness using  $\lambda = 10^{-5}$  turns out to work well. As we set  $p_{i,0} = p_{i,22} = 0$ , the algorithm drags the distribution down on both sides. However, the left tails of  $p^T$  and  $p^{4T}$  have peculiar structures of "going up" and  $p^{4T}$  seems to provide a distribution

worse than the rest. This is because the influence of options with further maturity and with strike prices further from current index is more remarkable than the influence of other options. This fact could reflect a phenomenon existing in the market, if investors expect extreme events (up or down). Therefore, it is natural to expect worse performance of the smoothness term on the tails of distributions and the distribution with further maturity dates.

Figure 2 confirms that  $\lambda = 10^{-5}$  offers satisfying smoothness of the distributions, but does not carry information about the error term. Table 2 shows the value of the smoothness term and the error of the estimation, with different  $\lambda$ s. And together with the numbers, Figure 3 shows the distributions of  $p^{4T}$  using  $\lambda = 10^{-6}, 10^{-4}, 10^{-3}$ , and  $10^{-2}$ . We found that the influence of  $\lambda$  is strongest in  $p^{4T}$ ; the reason, we think, is similar to what we wrote above concerning the tails.

Table 2: The smoothness term v.s. the error term with different constant  $\lambda$  (using improved method; # of state space: 21; interval length: 5% of current SPX index)

$\lambda = 10^{-3}$	$p^T$	$p^{2T}$	$p^{3T}$	$p^{4T}$
Smoothness	$6.18 \times 10^{-3}$	$1.35 \times 10^{-3}$	$4.94 \times 10^{-4}$	$1.25 \times 10^{-3}$
Error	3.58	1.94	1.07	2.67
$\lambda = 10^{-4}$	$p^T$	$p^{2T}$	$p^{3T}$	$p^{4T}$
Smoothness	$4.95 \times 10^{-3}$	$1.17 \times 10^{-3}$	$4.34 \times 10^{-4}$	$3.28 \times 10^{-4}$
Error	9.03	2.68	1.28	5.21
$\lambda = 10^{-5}$	$p^T$	$p^{2T}$	$p^{3T}$	$p^{4T}$
Smoothness	$3.37 \times 10^{-3}$	$8.76 \times 10^{-4}$	$3.67 \times 10^{-4}$	$1.80 \times 10^{-4}$
Error	70.9	14.6	4.11	10.0
$\lambda = 10^{-6}$	$p^T$	$p^{2T}$	$p^{3T}$	$p^{4T}$
Smoothness	$1.86 \times 10^{-3}$	$5.48 \times 10^{-4}$	$2.67 \times 10^{-4}$	$8.60 \times 10^{-5}$
Error	$5.98 \times 10^2$	$1.52 \times 10^2$	46.0	43.1

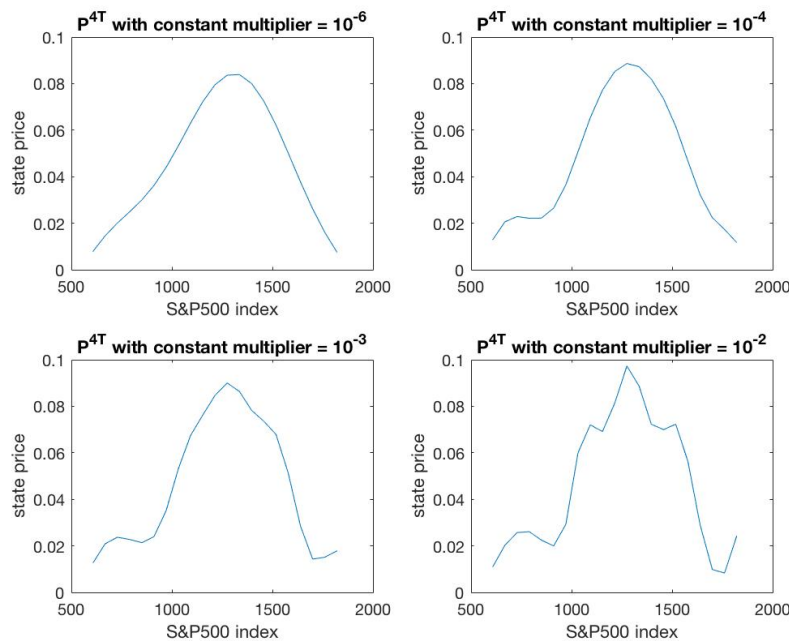


Figure 3:  $p^{4T}$  estimated by improved method with constant  $\lambda = 10^{-6}, 10^{-4}, 10^{-3}$ , and  $10^{-2}$  (# of state space: 21; interval length: 5% of current SPX index)

In Figure 3, as  $\lambda$  increases from  $10^{-6}$  to  $10^{-2}$ , the behavior of the state price distribution  $p$  is more unexpected, in that the distribution has larger tails and becomes more bumpy. However, as shown in Table 2, as  $\lambda$  decreases, the trade off between the improvement of smoothness and the diminishing accuracy becomes more remarkable. For example, when  $\lambda$  decreases from  $10^{-3}$  to  $10^{-4}$ , smoothness term of  $p^{4T}$  improved by 73.8% and errors increased by 94.9%. When  $\lambda$  drops from  $10^{-5}$  to  $10^{-6}$ , the smoothness term of  $p^{4T}$  decreased by 52.1% but the sum of errors increased by 330.3%. Therefore, a small  $\lambda$  provides distributions that are not smooth, and a large  $\lambda$  is inefficient by discarding too much information in exchange for little improvement of smoothness.

From Figure 2 and Figure 3, we observe that a  $\lambda$  less than or equal to  $10^{-4}$  provides tails that do not violate the set up of the model. And in Table 2, a  $\lambda$  greater than or equal to  $10^{-4}$  saves enough information to estimate accurate state price distributions (with error terms less than 10). Based on the above observations, we choose  $\lambda$  to be  $10^{-4}$  and keep this choice for experiments shown in future sections.

### 3.4.2 The Distribution of the Multipliers

In previous parts, we took  $\lambda_k$  to be constant for all  $k$  and focus on finding the appropriate range of the multipliers. However, we should expect more noise for long-dated options and therefore the multipliers should not be constant. Investors' opinions usually vary when they expect extreme events; therefore, the further the maturities are and the further the strike prices are away from the current *S&P 500* index, the more noises these option data should contain. Besides, practically speaking, deep-out-of-money options are not reliable for predicting investors' expectation of the future, since their low entrance value and high leverage may lure traders away from rational judgment in a real market. In Figures 2 and 3, most distributions show large tails, which reveal the unexpectedly large value of prices of options with strike prices far from the current index, especially of deep-out-of-money options. Therefore, in order to test the influence of different options with the same maturity, we design two distributions of  $\lambda_k$  around  $10^{-4}$  — a concave distribution and a convex one:

- Concave distribution: the multiplier of option  $k$  depends on the state space interval it belongs to. The options falling into the same state space interval have the same multipliers.  $\lambda$  of intervals on the edge is  $0.5 \times 10^{-4}$  and of central intervals is  $1 \times 10^{-4}$ , where the  $\lambda$  of different intervals varies linearly.
- Convex distribution: The convex distribution of  $\lambda$  is the same as the concave one, except that  $\lambda$  of intervals on the edge is  $1 \times 10^{-4}$  and the one of central intervals is  $0.5 \times 10^{-4}$ .

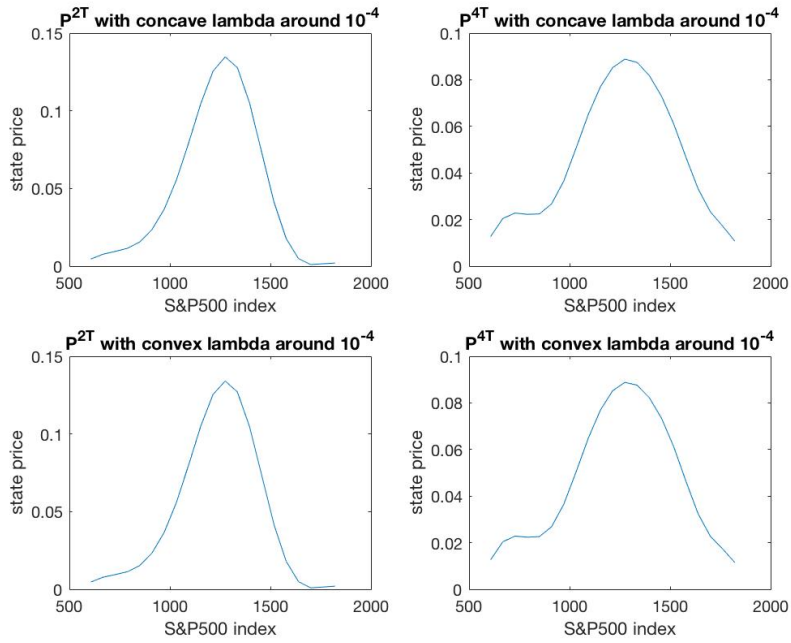


Figure 4:  $p^{2T}$  and  $p^{4T}$  estimated by improved method with concave and convex around  $\lambda = 10^{-4}$  (# of state space: 21; interval length: 5% of current SPX index)

Figure 4 plots  $p^{2T}$  and  $p^{4T}$  using the concave and convex distributions of multipliers. However, the results using concave and convex distributions are almost identical, meaning that noises do not seem to have much relationship with the strike prices as we expected. To test if there's truly no relationship between them, we compare the results without the top 10%, 20%, and 40% outliers. The outliers in our context are the options with high errors and the error of each option  $k$  is calculated by

$$Error_k = \sum_{j=1}^n p_j^{(t)} \max(S_j - X_k, 0) - H_k^{(t)} \quad (15)$$

Figure 5 plots the  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$  and  $p^{4T}$  dropping the top 10%, 20%, and 40% outliers.

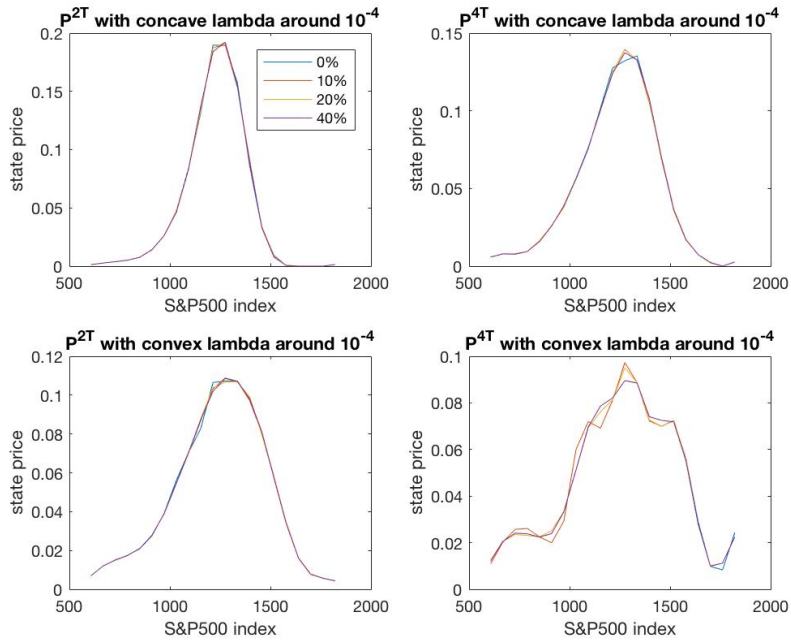


Figure 5:  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$  and  $p^{4T}$  estimated by improved method with  $\lambda = 10^{-2}$  and outlier drop rate = 0%, 10%, 20%, 40%. (# of state space: 21; interval length: 5% of current SPX index)

In Figure 5, no significant change is observed after we drop the outliers, not even after dropping 40% of the data. In fact, the distribution of outliers is quite arbitrary and the outliers calculated by equation (15) are not outliers. In this case, the option data are consistent and the noises they introduce are acceptable. We have also tested with other ranges of  $\lambda$ , and there's no clue that dropping outliers or using different distributions of the multipliers make much difference on the result. Therefore, we choose  $\lambda$  to be constant and equals to  $10^{-4}$ .

### 3.5 Robustness Test

So far, we have determined the method for recovering the subjective distributions, the value of coefficients in our method, and the real-world data for implementation. Before analyzing the results, we want to make sure that the method is robust. Notice that a robust algorithm should provide similar results despite small changes in the implementation. Therefore, we designed a robustness test by changing the length of state space intervals,  $\Delta S$ . The test compares two implementations:

- one with interval length is 5% of current SPX index and the number of state space is 21;
- one with interval length is 4% of current SPX index and the number of state space is 27;

The number of states changes in order to let the two discrete state models cover the same range.

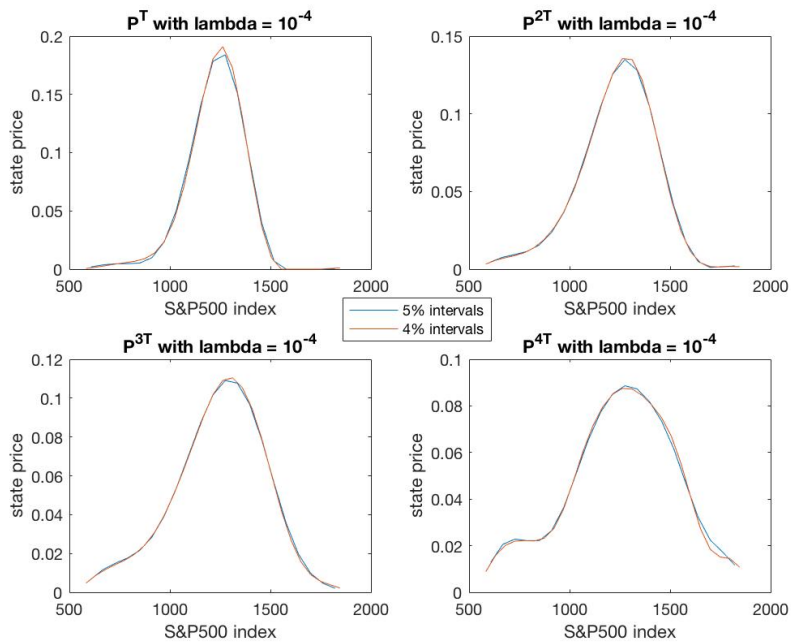


Figure 6:  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$  and  $p^{4T}$  estimated by improved method of different interval lengths: 5% with 21 intervals v.s. 4% with 27 intervals ( $\lambda = 10^{-4}$ ; outlier drop rate = 0%)

Figure 6 shows the four state price distributions under the two conditions, after normalizing the two conditions by multiplying  $5\%/4\% = 1.25$  to the distributions with 4% intervals. The distributions under different conditions coincide with each other, and this result gives us confidence on the robustness of the algorithm.

To complete the test, we include a similar robustness test on the original Backwell method. Figure 7 shows the four state price distributions obtained by the greedy algorithm under the same two conditions after normalization. The distributions under different conditions are not consistent with each other, especially at indexes that are far from the current index. This result shows that the algorithm is very sensitive to the change of the length of the index interval and confirms that our improvement of the method increases the robustness of the model. Further tests might be useful, but might not be necessary. In the next section, we will move on to the analysis of the subjective distributions using our method.

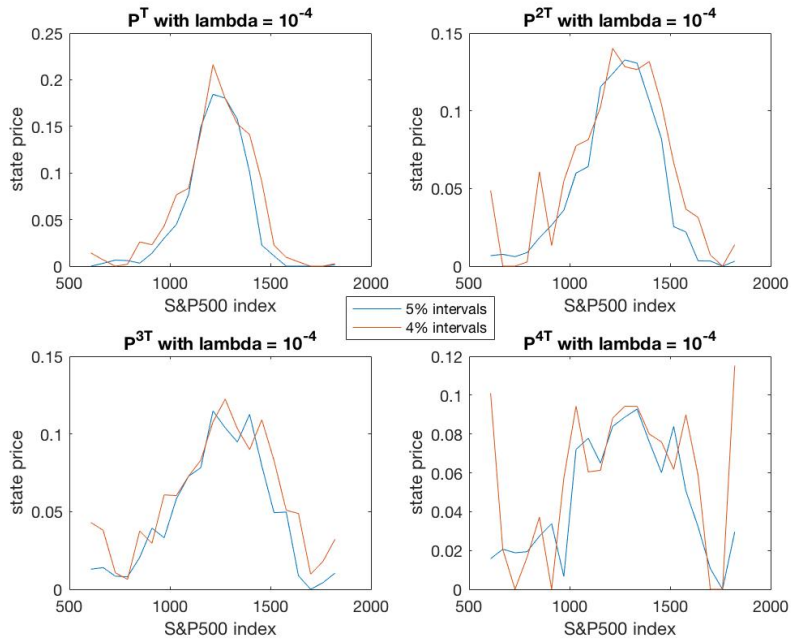


Figure 7:  $p^T$ ,  $p^{2T}$ ,  $p^{3T}$  and  $p^{4T}$  estimated by non-improved Backwell method of different interval lengths: 5% with 21 intervals v.s. 4% with 27 intervals ( $\lambda = 10^{-4}$ ; outlier drop rate = 0%)

## 4 Analysis of the Subjective Distributions

The Ross Recovery Theorem, introduced in Sections 1 and 2, has been controversial since it was published. It challenges the conventional view that derivative prices contain at best only partial information about a subjective probability distribution describing the market's views of the future. Instead, according to the Ross's theorem (and under his assumptions on the structure of the market), derivative prices actually determine the entire subjective probability distribution. This leads us to ask: does the subjective distribution have predictive value? And does the Recovery Theorem have practical implications for investors? To explain whether the theorem is satisfied by real market data and has practical value, in this section, we apply the Ross Recovery Theorem on data between 2000 and 2008 using the method summarized in section 3. First, we analyze the properties of the distributions, such as the expectations and skewness, to show what the subjective probability distributions look like. Then, we design a mock-trading strategy to test whether the theorem can be useful to an investor.

In the application, we choose a time horizon from 3/15/2000 to 9/15/2008, select 35 dates (time points) and capture the related data on these dates for implementation. These dates are evenly distributed, with the interval between each date approximately 3 months, meaning that we collect the data on 3/15, 6/15, 9/15, and 12/15 of each year. If the market is closed on any of the dates, we will try the next date until all the relevant data on this date are available.

### 4.1 Properties of the Distributions

The goal of this subsection is to explore the properties (expectation and skewness) of the 35 recovered subjective distributions, to analyze the trends of the distributions, and to explain the reasons behind the facts. We also explore whether the subjective distribution has predictive value.

#### 4.1.1 Expectation of the subjective distributions

Expectation is an effective index that measures the investors' prediction of the future. Figure 8 shows the weighted mean of the current subjective distribution, together with the *S&P500* index at the current time and the next time. We can see the expectations and the current indexes are closely correlated and the current index is always slightly higher than the expectation. Real-world information reflects current market knowledge and uncertainty. Therefore, according to how the

subjective probability distributions are recovered, the distributions appear to reveal information at current time. In fact, comparing the two graphs, we can see that the expectations of the distributions reflect only today's price and knowledge, and have no value in future prediction.

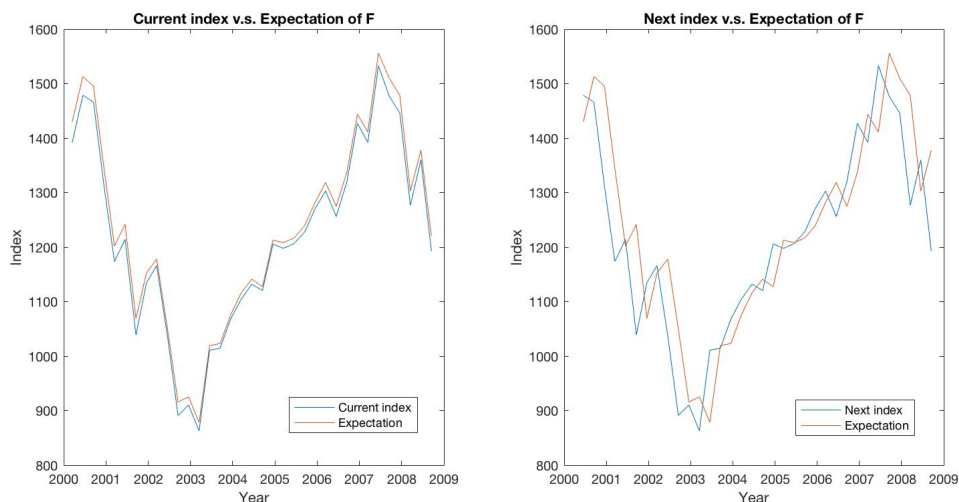


Figure 8: the current(next) indexes v.s.  $\mathbb{E}$  [price at next step] (estimated by improved method with  $\lambda = 10^{-4}$ ; no dropped outliers; # of state space: 21; interval length: 5% of current SPX index)

Notice that there is a systematic gap between the current indexes and the expectations, which we think results from the fact that there is a premium for risk. Variances of the subjective distributions represents the extent of risk contained within the prediction and Figure 9 plots the relationship between the systematic gap and variance. The correlation coefficient of the two variables is 0.8461, which confirms a strong uphill (positive) linear relationship. The distribution with a higher risk tends to provide a higher return, and therefore enlarges the gap between the current index and the mean of the subjective distribution.

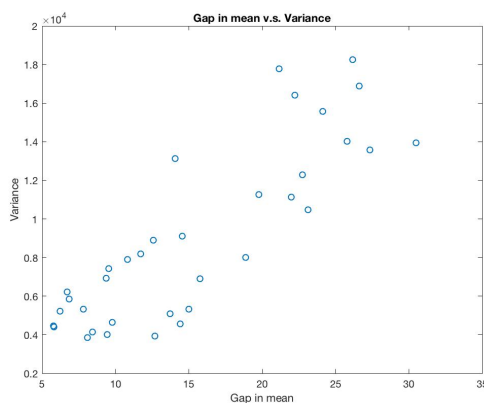


Figure 9: variance of distributions v.s. difference between current indexes and the expectation (estimated by improved method with  $\lambda = 10^{-4}$ ; no dropped outliers; # of state space: 21; interval length: 5% of current SPX index)

#### 4.1.2 Skewness of the subjective distributions

Besides the expectation of the distributions, skewness is another important property of a probability distribution. After recovering the subjective distribution, the unexpected structure in the tails of state space distribution  $p$  disappear, and the inclusion of a smoothness term in our optimizations is successful in keeping the subjective distribution smooth. The subjective probability distribution on 2001/06/15 is shown in Figure 10, whose shape is typical among the 35 distributions. We will



analyze more about the shape of subjective distributions in the following paragraph.

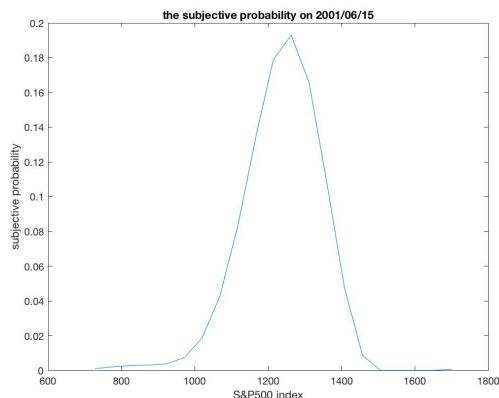


Figure 10: The subjective distribution on 2001/06/15 (estimated by improved method with  $\lambda = 10^{-4}$ ; no dropped outliers; # of state space: 21; interval length: 5% of current SPX index)

As shown in Figure 11, the subjective distributions are symmetric and concentrate in the middle. Also, the distributions are a little left-skewed, since all probabilities of larger than the means are greater than 50%, and the probabilities of larger than mean plus one standard deviation are greater than the ones of smaller than mean minus one standard deviation. It is natural since the stock price never goes negative. Additionally, tails represent option data in longer maturity, and therefore it is not surprising that the recovered distributions have fatter tails than Gaussian. However, the feature of skewness does not have much use for practical purpose.

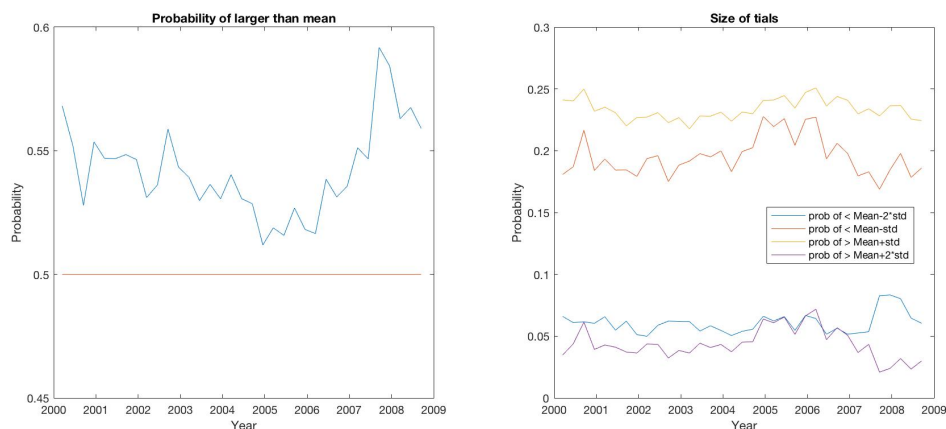


Figure 11: The probability of larger than the means of subjective distributions & the probability of larger(smaller) than the means  $\pm 1\sigma(2\sigma)$  (estimated by improved method with  $\lambda = 10^{-4}$ ; no dropped outliers; # of state space: 21; interval length: 5% of current SPX index)

## 4.2 A Trading Strategy Making Use of the Subjective Distribution

To test the practical value of the Ross Recovery Theorem, we design a simple trading strategy, which maximizes the expectation of log-return using the recovered probability distributions. We will first introduce our strategy below.

- **Log-Return Optimization**

First, we explain why we should maximize the expectation of the log return rate in a trading strategy.

Let  $W_0$  be the initial wealth of the investor, and  $W_n$  be the wealth after  $n$  time units.  $R_k$  is the random return variable, *i.e.*  $W_k = R_k \times W_{k-1}$ . Assume that all  $R_k$ 's are i.i.d. random

variables, we can see that  $W_n = R_n \times R_{n-1} \times \dots \times R_1 \times W_0$ . Then taking the logarithm to both sides, we get

$$\log W_n = \log W_0 + \sum_{k=1}^n \log R_k$$

By Law of Large Numbers,

$$\log \left( \frac{W_n}{W_0} \right)^{1/n} \rightarrow \mathbf{E}(\log R_1) =: m$$

Therefore, to maximize the final wealth, we only need to maximize the expectation of the log return rate, that is,  $m$ .

- **Strategy Design**

With the above knowledge, we design a trading strategy based on the maximization of log return. In each time period, we recover the subjective probability distribution  $F$ , and let the probability of going to state  $i$  in the next time period be  $F_i$ .

Our portfolio contains two assets: the stock index S&P500 and bonds. Let the investor have  $W_0$  at current time period. Then after investing  $\Theta$  ( $0 \leq \Theta \leq 1$ ) of the current wealth in stock, and the others in bonds, the wealth of the investor in next time period will be  $(1 - \Theta)(1 + r)W_0 + \frac{W_0 S_i}{S_0} \Theta$ , with probability  $F_i$ . Therefore, the expectation of the log return rate can be calculated by

$$m = \mathbf{E}(\log R_1) = \sum_{i=1}^n F_i \log[(1 - \Theta)(1 + r) + \Theta(S_i/S_0)]$$

To maximize  $m$ , the first derivative of  $\Theta$  should be zero, Let  $\hat{\Theta} \in [0, 1]$  satisfy

$$\frac{\partial m}{\partial \hat{\Theta}} = \sum_{i=1}^n F_i \frac{-(1 + r) + S_i/S_0}{(1 - \hat{\Theta})(1 + r) + \hat{\Theta}(S_i/S_0)} = 0$$

(In application, we use Matlab's equation and system solver with syntax *solve* to evaluate  $\hat{\Theta}$ .) Since  $\Theta$  should be larger than zero and less than one, we choose  $\Theta$ , among 0, 1, and  $\hat{\Theta}$ s, to be the one with the largest  $m$ . (If  $\hat{\Theta}$  does not exist, we choose  $\Theta$  between 0 and 1.)

The strategy repeats the above process using the 35 subjective distributions, and balance the portfolio each time period, that is trade with the optimized  $\Theta$ .

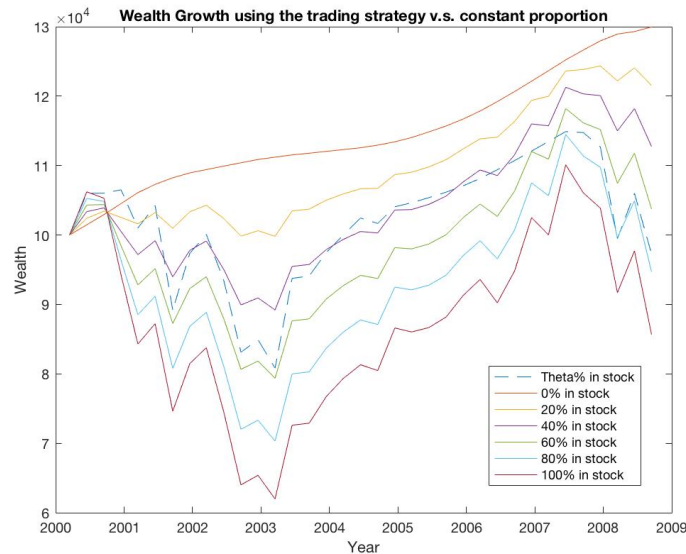


Figure 12: Long-term Wealth Growth with different portfolios (distribution estimated with  $\lambda = 10^{-4}$ ; 21 intervals with length of 5% of current SPX index)

Figure 12 compares the wealth evolution using the trading strategy that maximizes the log-return to the evolution using a constant proportional portfolio between 2000 and 2008. A constant proportional portfolio is a trading strategy using fixed  $\Theta$  every time period in our context. When *S&P500* index grows, the optimized strategy performs like the 40% portfolio; when *S&P500* index drops, it performs like a portfolio investing 100% of the wealth in stock every time period. The trading algorithm does not do much better in practice than a constant-proportion investing policy. Thus, it is not clear that the subjective probabilities obtained using Ross recovery are useful for portfolio optimization.

## 5 Conclusion

The Ross Recovery Theorem is remarkable. It challenges the traditional views that derivative prices do not carry useful information about the future. However, it is also controversial. Its assumptions, especially the preference metastructure, simplify the real-world situation. Furthermore, it is almost impossible to test how much the subjective probability distribution obtained from the theorem coincide with the real-world one: there is no way to test the correctness of the subjective distribution.

In our paper, we introduced the Recovery Theorem, calibrated it to real-world data, and tested its practical value. The obstacles to calibration were identified and overcome. The significance and practical value of the theorem were discussed. Audrino et al. [8] conduct an empirical analysis of the recovery theorem, but their conclusion is opposite to ours – they suggest that a useful trading strategy based on the theorem indeed exists. We speculate that the difference of the conclusions results from (1) how Audrino et al. and we handle the ill-posed system of (4), (5), and (6) differently, and (2) different trading strategies. Faced with the ill-posed optimization problem, while we choose to relax certain constraints in the problem, Audrino et al. choose Tykhonov Regularization. When conducting the empirical analysis, they design a trading strategy which makes a new decision every week based on the changes in risk-neutral volatility and recovered skewness. We believe that these two differences lead to different results.

In our analysis, we found that the subjective distribution seems to only present the current market situation and barely contains any future information. Also, the performance of a trading strategy based on the subjective distribution was mediocre.

In fact, the result is not surprising. The assumption that there is only one representative agent whose behavior gives rise to unique prices might be an oversimplification. Real-world prices come from the aggregated actions of different types of investors. This simplification might prevent the theorem from achieving a good distribution. Furthermore, current information reflects current knowledge and attitudes of the financial market, and it might barely contain useful information about the future, which is unpredictable. For instance, before the financial crisis, investors actually held optimistic attitudes toward the market: based on their attitudes, one could hardly conclude that a financial crisis would happen. Even if Ross Recovery does obtain the market's subjective probabilities, it is unreasonable to expect this to provide useful information about the future.

We conclude by saying that the Ross Recovery Theorem is mathematically correct, but we doubt its significance and value in the real financial world.

## References

- [1] Ross, Steve. The Recovery Theorem. *J. Financ.* **2015**, *70*, 615–648. doi:10.1111/jofi.12092
- [2] Backwell, Alex. State Prices and Implementation of the Recovery Theorem. *Journal of Risk and Financial Management*, **2015**, *8*, 2-16; doi: 10.3390/jrfm8010002
- [3] Samuelson, P.A. A Note on Measurement of Utility. *Rev. Econ. Stud.* **1937**, *4*, 155-161
- [4] Rubinstein, M. Implied Binomial Trees. *J. Financ.* **1994**, *49*, 771-818.
- [5] Jackwerth, J.C.; Rubinstein, M. Recovering Probability Distributions from Option Prices. *J. Financ.* **1996**, *51*, 1611-1631.
- [6] Carr, P.; Yu, J. Risk, Return, and Ross Recovery. *J. Deriv.* **2012**, *20*, 38.

- [7] Black, Fischer; Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*. **1973**, *81 (3)*, 637–654. doi:10.1086/260062.
- [8] Audrino, F.; Huitema, R.; Ludwig, M. An Empirical Analysis of the Ross Recovery Theorem. Available online: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2433170](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2433170) (accessed on 27 January 2018)
- [9] Walden, Johan. Recovery with Unbounded Diffusion Processes. *Review of Finance*, Forthcoming. Available online: <https://ssrn.com/abstract=2508414>
- [10] Hansen, Lars Peter.; Scheinkman, Jose A. Stochastic Compounding and Uncertain Valuation. Available online: <https://ssrn.com/abstract=2256246> or <http://dx.doi.org/10.2139/ssrn.2256246>
- [11] Qin, Likuan.; Linetsky, Vadim. Positive Eigenfunctions of Markovian Pricing Operators: Hansen-Scheinkman Factorization, Ross Recovery and Long-Term Pricing (February 2, 2015). Available online: <https://ssrn.com/abstract=2439002> or <http://dx.doi.org/10.2139/ssrn.2439002>