

A dissertation submitted for the award of MSc

On estimating the risk-neutral and real-world probability measures

Author:
Trent Jon Spears

Supervisor:
Dr Johannes Ruf

June 21, 2013

Acknowledgements

Thank you to Dr Johannes Ruf for posing such an interesting topic, for organising workshops and correspondence with some of the thought leaders developing Ross Recovery, for providing me access to data, and for offering his time, support, ideas and enthusiasm as I developed this dissertation.

To those who like to get their hands dirty.

Contents

1 Introduction **3**

2 Some theory **6**

2.1 Matrices, Markov chains and ordinary least squares 6

2.2 Key financial theory 12

3 The Ross Recovery theorem **22**

3.1 Step 1. From option price data to S 22

3.2 Step 3. The Recovery theorem: F from P 24

4 Econometric issues in the application of Ross Recovery **26**

4.1 Step 2. Moving from S to P, Ross methodology. 27

4.2 A comparison of nine ways to estimate P from S. 30

4.2.1 Results 31

4.3 Remark: testing F 37

5 Conclusion **39**

A **41**

A.1 Estimating S 41

A.2 Ross P and F 41

A.3 Estimated P's and corresponding F's 41

A.4 Selected Matlab code 41

Bibliography **54**

Part 1

Introduction

This dissertation is about inferring market beliefs about the real-world probability distribution describing the future financial returns of an underlying asset from option price data.

It is a triumph of modern financial theory that the fair value for a financial option is not determined by a real-world probability distribution, or any expectation taken thereunder, but rather by the initial value of constructing a self-financing portfolio of assets replicating the terminal option payoff. Further, it is remarkable that almost for convenience sake, we also have a theory for finding such a value that does amount to expectation pricing, subject to technical conditions. Such an expectation is taken under the risk-neutral distribution, the distribution which induces the measure for which the discounted stock price process is a martingale. Conversely, it is well known that given knowledge of option values we can reverse-engineer and solve for the risk-neutral distribution [3]. However, theory has left us without a way to recover the real-world probability distribution describing the future financial returns of an underlying asset, effectively a crystal ball from which we could glean the future price path of an asset. A recent paper by Stephen Ross presents theory that offers a substitute [15]. Ross outlines his initial thoughts on a method for recovering the real-world probability distribution *as implied by collective market beliefs* indicative in option prices. Theoretically, the Ross Recovery idea provides multiple interesting directions for future research, while practically, the accurate recovery of such a distribution could have a profound effect on modern finance.

The primary goal of this dissertation is to present a replication of the econometrics and results of Ross Recovery as applied to S&P500 index option price data from 27 April, 2011, precisely as in Ross's paper. The process is outlined in Figure 1.1 shown below. Starting with

market data, we estimate the state price by tenor[1] matrix S, before using this estimate to estimate the state price transition matrix P. Normalising the rows of P yields the risk-neutral transition matrix $\mathbb{Q}$, and applying the Ross Recovery calculus to P yields us the real-world transition matrix F. We will see that this goal can be achieved to some degree of accuracy, despite the fact that we use data compiled from Bloomberg where Ross uses a proprietary – and hence to some extent differing – investment bank data set. We will also see that the method of Ross is somewhat ad hoc and illustrative, which is most obvious in his estimation of P from S. We offer further detail than is provided in [15], and span the entire estimation procedure . Also, we show why it is not possible to construct Ross's P from S using only the process Ross has presented, and further show that on implementing the likely method that Ross used, we do not estimate P to within a suitable margin of error of Ross's result. This provides us an opportunity to compare and contrast nine different robust non-parametric methods for estimating P from S, and we show that one method stands out as superior for estimating P. This is the chief contribution of this dissertation to the literature. Further, we have initial evidence to suggest that the real-world transition matrix estimated from our preferred P may be useful for forecasting returns in the short term.

[1] 'Tenor' is synonymous with 'time to maturity'.

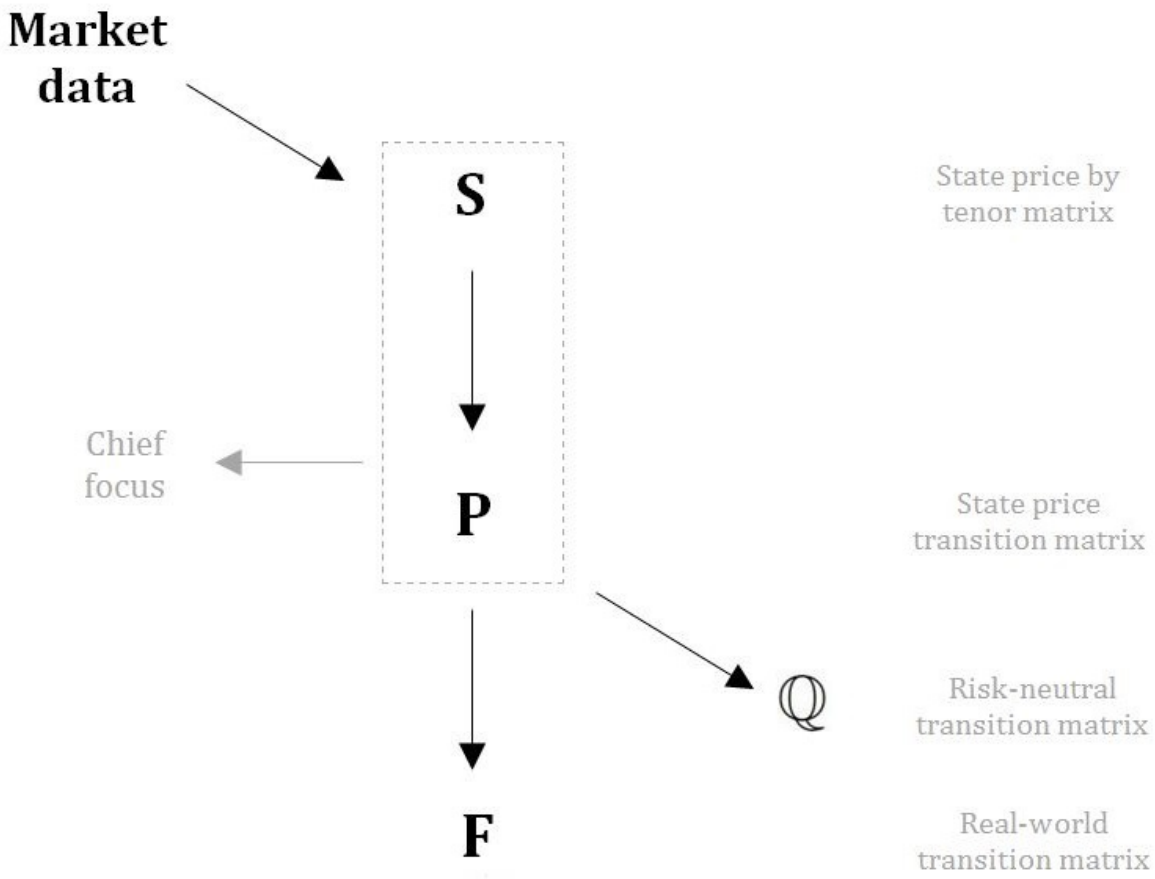


Figure 1.1: This image depicts a simple flow diagram showing the links between the important objects of the Ross Recovery process, which are the subject of this dissertation.

Part 2

Some theory

The ideas of Ross [15] are presented in a discrete-time setting on a discrete, bounded state space. In this setting, we make use of the following theory.

2.1 Matrices, Markov chains and ordinary least squares

Basic Markov chain ideas

A random process $\mathbf{X}$ is a collection of random variables $(X_t : t \in T)$ indexed by a set T. Such a set is often considered as accounting for the evolution of time. When T is a countable set, we call $\mathbf{X}$ a discrete-time process. Each X_t takes values in a state space $\mathbb{S}$, which we will also consider countable, and we say that $\mathbf{X}$ is a collection of discrete random variables each taking values in $\mathbb{S}$. We call the state space bounded (so that each X_t can only take a finite number of values) if $|\mathbb{S}| := \#\{\text{distinct elements of } \mathbb{S}\} < \infty$.

We seek to impose some structure on the elements of $\mathbf{X}$ so that we can build interesting models of real-life random phenomena. It is simple to assume each X_t is independent of the other, but then the evolution of $\mathbf{X}$ may lack the richness to be useful as an adequate model. A next step is to suppose the Markov condition,

$$\mathbb{P}(X_{t+1} = x_{t+1} | \{X_i = x_i : 1 \leq i \leq t\}) = \mathbb{P}(X_{t+1} = x_{t+1} | X_t = x_t), \qquad \forall t \in T,$$

in which case we call $\mathbf{X}$ a Markov chain. Call the transition probability the quantity on the right-hand side of the above equation.

If $\mathbb{P}(X_{t+1} = j | X_t = i) = \mathbb{P}(X_2 = j | X_1 = i) =: \mathfrak{p}_{ij}$ $\forall t \in T$ then $\mathbf{X}$ is a time-homogenous Markov chain[1]. In this case, we have a convenient form in which to describe the evolution of

[1]For simplicity here we have supposed T is the set of natural numbers $\{1, 2, \dots\}$.

X: as an $|\mathbb{S}| \times |\mathbb{S}|$ transition matrix $\mathscr{P}$ such that

$$\mathscr{P} = [\mathfrak{p}_{ij}].$$

Recall that a matrix $\mathbf{M}$ is called non-negative when the entry of the i^{th} row and j^{th} column $m_{ij} \geq 0$, for each i, j. Further, $\mathbf{M}$ is called positive if strict equality holds: $m_{ij} > 0$. Clearly then, $\mathscr{P}$ is a non-negativite matrix.

An important property of $\mathscr{P}$ is that by the law of total probability, its row sums are equal to unity: $\sum_j \mathfrak{p}_{ij} = 1 \ \forall i$. Also, unless otherwise stated, we will assume that all such matrices are irreducible, so that any one state can always be reached from any other at some time in the future, with positive probability:

$$\forall i \ \forall j \ \exists t_m \text{ such that } \mathbb{P}(X_{t_m} = x_j | X_1 = x_i) > 0.$$

Perron-Frobenius theory

In the early 20th century German mathematicians Oskar Perron and Georg Frobenius proved a suite of basic but useful results regarding the spectral structure of classes of non-negative matrices. Such results will prove indispensable to achieving Ross Recovery in our setting. Of central importance is the Perron-Frobenius theorem.

Let $\sigma(\mathbf{M})$ denote the spectrum of $\mathbf{M}$, that is, the set of distinct eigenvalues for $\mathbf{M}$. Define the spectral radius of $\mathbf{M}$, $\rho(\mathbf{M}) = \max_{\lambda \in \sigma(\mathbf{M})} |\lambda|$. Then we have:

Perron-Frobenius theorem (partial) [14]. For a non-negative, irreducible $n \times n$ matrix $\mathbf{M}$

- The Perron root $r := \rho(\mathbf{M}) \in \sigma(\mathbf{M})$, and $r > 0$.
- There exists a unique eigenvector p such that $p > 0$, $\mathbf{M}p = rp$, and $||p||_2 = 1$, called the Perron vector. The only non-negative eigenvectors for $\mathbf{M}$ are positive multiples of p, regardless of the eigenvalue.

We can extend the theorem as it relates to a reducible matrix; it is made use of exactly once in this study:

Theorem [14]. For a non-negative $n \times n$ matrix $\mathbf{M}$

- The value $r = \rho(\mathbf{M}) \in \sigma(\mathbf{M})$, and $r \geq 0$.
- There exists an eigenvector $p \in \{x | x \geq 0 \text{ but } ||x||_1 \neq 0\}$ such that $\mathbf{M}p = rp$.

Ordinary least squares

Suppose we have the matrix equation

$$A = BP \tag{2.1}$$

for A a $j \times k$ matrix of known entries, B a $j \times l$ matrix of known entries, and P a $l \times k$ matrix of unknown entries for which we wish to solve. We can write the system in the form

$$\tilde{\tilde{A}} = \tilde{\tilde{B}}\tilde{\tilde{P}} \tag{2.2}$$

for $\tilde{\tilde{A}}$ a $j \cdot k \times 1$ vector, $\tilde{\tilde{B}}$ a $j \cdot k \times l \cdot k$ matrix and $\tilde{\tilde{P}}$ a $l \cdot k \times 1$ vector. More precisely, $\tilde{\tilde{A}} = \{\{a_{1i} : 1 \leq i \leq j\}; \{a_{2i} : 1 \leq i \leq j\}; \ldots; \{a_{ki} : 1 \leq i \leq j\}\}$, $\tilde{\tilde{B}} = diag\{B, B, \ldots, B\}$ is a diagonal matrix with k copies of B across its diagonal, and $\tilde{\tilde{P}} = \{\{p_{1i} : 1 \leq i \leq l\}; \{p_{2i} : 1 \leq i \leq l\}; \ldots; \{p_{ki} : 1 \leq i \leq l\}\}$. In this form, our system presents as a classical linear regression problem, and we can solve for $\tilde{\tilde{P}}$ straightforwardly using the principle of ordinary least squares. Recall our estimate is given by the vector $\tilde{\tilde{P}}$ that minimises the sum of the squared residuals, that is,

$$\min_{\tilde{\tilde{P}}} \; SSR(\tilde{\tilde{P}}) := \left(\tilde{\tilde{A}} - \tilde{\tilde{B}}\tilde{\tilde{P}}\right)^T \left(\tilde{\tilde{A}} - \tilde{\tilde{B}}\tilde{\tilde{P}}\right). \tag{2.3}$$

This system is easily solved in many good, standard software packages.

Suppose now we have the same problem as equation (2.1), but that we wish to restrict row l_i to be strictly fixed, for some $1 \leq l_i \leq l$. In this case, we can write

$$\begin{aligned} A &= BP \\ \therefore A &= \tilde{B}\tilde{P} + B(:, l_i) \cdot P(l_i, :) \\ \therefore \tilde{A} &= \tilde{B}\tilde{P}, \end{aligned} \tag{2.4}$$

where

$$B(:, l_i) \cdot P(l_i, :) = \{\{B_{1l_i} P(l_i, :)\}; \ldots; \{B_{jl_i} P(l_i, :)\}; \}$$

is a $j \times k$ matrix, $\tilde{A} = A - B(:, l_i) \cdot P(l_i, :)$, $\tilde{B}$ is matrix B with the l_i^{th} column removed, and $\tilde{P}$ is matrix P with the l_i^{th} row removed. We call this reduced matrix system system tilde. It can also be written in the form (2.2) in the obvious way, which we will call the reduced equation,

$$\tilde{\tilde{A}}_R = \tilde{\tilde{B}}_R \tilde{\tilde{P}}_R, \tag{2.5}$$

and which can also be solved by ordinary least squares.

Constraints for our systems

In this section we consider methods for solving the equation (2.1) via versions of equations (2.2), (2.4) and (2.5), with additional constraints imposed on the system.

Firstly, consider equation (2.2) and (2.5). Recall that a general linear optimisation problem can be expressed in the following normal form:

$$\min_{\mathbf{x}} f(\mathbf{x}) = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to the constraints

$$x_i \geq 0, \qquad \forall i, \qquad \text{and} \tag{C1}$$

$$B\mathbf{x} = a, \tag{2.6}$$

where B is a $m \times n$ matrix, and a is a $m \times 1$ matrix. This problem is well known, and an optimal solution can be found, for instance, by the equally well known simplex method [12], easily implemented in good mathematical software packages. We have seen that equations (2.2) and (2.5) are solved as solutions to the above minimisation problem written in the form (2.3), and we also have that we can impose restrictions on our system as in (C1) and (2.6). We have flexibility to adapt the latter condition to impose interesting qualities on our solution. For instance, we may wish to set particular values of x_i to 0. Suppose we wish to impose $\{x_{p_k} = 0 : 1 \leq k \leq M \leq n, p_1 < p_2 < \cdots < p_M\}$. To do this, we could write

$$C\mathbf{x} = d \tag{C2}$$

where C is an $M \times n$ matrix such that the k^{th} row of C has a 1 in the p_k^{th} position and 0's elsewhere, and a is an $M \times 1$ matrix of 0's.

Another interesting constraint is to require that particular values of x_i are equal to one another. Suppose we wish to impose that $\{x_{p_k}$ are equal $: 1 \leq k \leq M \leq n, p_1 < p_2 < \cdots < p_M\}$. To do this, we could write

$$C\mathbf{x} = d \tag{C3}$$

where C is a $\binom{p_M}{2} \times n$ matrix such that for each unique (up to ordering) combination of pair (x_{p_i}, x_{p_j}) chosen from $\{x_{p_k}\} \times \{x_{p_k}\}$, there is a row of C with a 1 in the $p_{k_i}^{th}$ position and a -1 in the $p_{k_j}^{th}$, and 0's elsewhere, and a is a $\binom{p_M}{2} \times 1$ matrix of 0's. An alternative to imposing this constraint is to write our original equations (2.2) or (2.5) in a further reduced form. To do this, we substitute our requirement that $\{p_k$ are equal $\}$ into our $\tilde{\tilde{P}}$ vector. Next, we delete all but the first of each repeated entry in the vector. Suppose the first of a repeated entry

was in row k, and the replicates were in $k_1, \ldots, k_m$. We then add columns $k_1, \ldots k_m$ of $\tilde{\tilde{B}}$ into column k, before deleting them from $\tilde{\tilde{B}}$. We can proceed to solve this new system directly by unconstrained least squares.

Next we consider constraints on our system that can not be written in the form (2.6). In particular, we now consider the constraint

$$\text{the rows of } P \text{ are unimodal.} \tag{C4}$$

We call a matrix unimodal if for each of its rows the entries change from increasing to decreasing in either direction at most once. We focus on systems of the form (2.4), and we will also consider the alternative constraint that

$$\text{the rows of } P \text{ are unimodal, with the modes lying on the main diagonal of } P. \tag{C5}$$

This is in contrast to (C4), where our estimation algorithm must also optimise over mode location in each row of P. The definitive reference paper for implementing these constraints given our system is [**?**]. We will briefly review the basic method.

Firstly, recall the idea of alternating least squares (ALS). Given a matrix system that we wish to solve, such as (2.4), we can break the system into smaller parts, solving for the parameter estimates in each part in turn by ordinary least squares, in an iterative fashion, and also by conditioning on the estimates of previous iterates at each step. We terminate the scheme when significant changes in parameter updates cease to occur, up to some level of error tolerance. Since successive iterations improve parameter estimates in the least squares sense, or at worst leave them unchanged, the scheme will eventually converge.

An ALS algorithm for solving (2.4) is (i) make some initial guess for P; (ii) solve for P row-wise subject to unimodality in an iterative fashion; then (iii) update P and repeat till the changes in P are sufficiently small. It is clear that it is step (ii) carries the weight, and we refer the interested reader to [**?**] for further details of its implementation.

We note here that a limitation of ALS is a system may have multiple local minima, one of which the scheme will converge to, that we can not, however, guarantee to be globally optimal. Further, note that a unimodal regression on [1 0 0 0 1] yields two globally optimal solutions, [0.25 0.25 0.25 0.25 1] and [1 0.25 0.25 0.25 0.25], so that the globally optimal solution to (2.4) is not necessarily unique. Hence solutions to our system can vary significantly, and depend both on our data and the initial guess for P.

In summary, and to put what we have learned in perspective, we define and explain nine methods for finding solutions to (2.1) that we will implement in our numerical work later in

this dissertation. We also list the number of parameters that are required to be set for each method. The methods are henceforth labelled $M1 - M9$. Further, we suppose P is a $k \times k$ square matrix, and for some methods that we are restricted to have the middle-most row of P fixed by given constants, indicated by the words 'fixed row'.

M1: Ordinary least squares, fixed row

Solves equation (2.5) directly with no constraints for k(k-1) unique parameters.

M2: Ordinary least squares, fixed row, non-negative

Solves equation (2.5) with constraint (C1) for k(k-1) unique parameters.

M3: Unimodal, fixed row, non-negative

Solves equation (2.4) with constraints (C1), (C4) for k(k-1) unique parameters.

M4: Unimodal, fixed mode on diagonal, non-negative

Solves equation (2.4) with constraints (C1), (C5) for k(k-1) unique parameters.

M5: 0's placement, fixed row, non-negative

Solves equation (2.5) with constraints (C1), (C2) for $k-1+\sum_{i=1}^{u}(k-1-i)+\sum_{j=1}^{d}(k-1-j)$ unique parameters. Here u denotes the maximum number of non-zero entries permitted above the main diagonal of P, and d is the corresponding number below the diagonal.

M6: Sliding window, non-negative

Solves equation (2.2) with constraints (C1), (C3) for 2k-1 unique parameters. Entries of each diagonal for P are equal.

M7: Sliding window, fixed row, non-negative

Solves equation (2.5) with constraints (C1), (C3) for k-1 unique parameters. Entries of each diagonal for P are equal, and must include the given middle-most row.

M8: Sliding window, fixed row, non-restricted, non-negative

Solves equation (2.5) with constraints (C1), (C3) for 2(k-1) unique parameters. Entries of each diagonal for P are equal, with the exception of the given middle-most row.

M9: Sliding window, fixed row, two state, non-negative

Solves equation (2.5) with constraints (C1), (C3) for 3(k-1) unique parameters. The upper and lower matrices separated by the middle-most row both satisfy that their diagonals are equal, independently of each other.

2.2 Key financial theory

Some classic Black-Scholes-Merton

Since the influential work of Black and Scholes [2] and Merton [13], who modelled the relationship between option values and the underlying non-dividend paying asset (which we will call a stock), a commonplace assumption in quantitative finance has been that the distribution of future stock prices returns on a finite interval is normal.

Recall that for a stock with value S_0 now, that grows to S_T in a year, we have an annualised rate of return of

$$\frac{S_T - S_0}{S_0} \approx \log\Big(\frac{S_T}{S_0}\Big),$$

where we have made use of the Taylor series expansion for $\log x$, for x about 1. Indeed, for practical applications, the right-hand side term has become the accepted quantity for describing returns. Of course, if S_T is a normal random variable (and S_0 is a known constant), then $\log(S_T/S_0)$ is a log-normal random variable[2].

Further assuming no transaction costs, that the short-term risk-free interest rate r is a known constant, that we are considering European options, that short selling is allowed, and that there exists a money-market account within which we can lend or borrow at the risk-free rate [2, 13], then the Black-Scholes-Merton option pricing formula is given by

$$\frac{V_t}{B_t} = \mathbb{E}^{\mathbb{Q}}\Big[\frac{V_T}{B_T}\Big|\mathscr{F}_t\Big].$$

Here, the index $\mathbb{Q}$ represents that the expectation is taken with respect to the risk-neutral distribution. This distribution induces a probability measure equivalent to $\mathbb{P}$ – the natural probability measure – such that the discounted stock price process $(S/B)_{t\geq 0}$ is a martingale. Also, T is the terminal payoff time, V_t is the option value at time $t < T$, $\mathscr{F}_t$ is the information generated by the underlying up till time t, and $(B)_{t\geq 0}$ represents the numeraire process, typically taken as the money-market account, so that $B_t = e^{rt}$.

In particular, a European call option with strike K and terminal payoff $V_T = \max(S_T - K, 0)$ has value given by

$$\begin{aligned} V_t &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\big[\max(S_T - K, 0)|\mathscr{F}_t\big] \\ &= S_t N(d_1) - e^{-r(T-t)} K N(d_2) \end{aligned}$$

[2]Modern financial theory has drawn light to the fact that these distributional assumptions are, for practical purposes, inadequate [17]. Indeed, the distribution of stock price returns has historically proved to have 'heavy-tails', and to be leptokurtic. However, we can steer clear of this point for this paper, as for the most part, we will make only minimal distributional assumptions in applying Ross Recovery.

where

$$d_1 = \frac{\log(S_t/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma \cdot \sqrt{T-t}} \qquad \text{and}$$
$$d_2 = d_1 - \sigma \cdot \sqrt{T-t}.$$

We have evaluated V_t in closed form under the distributional assumption that $S_t/S_0 \sim \log N((\mu - \frac{1}{2}\sigma^2)t, (\sigma t)^2)$, consistent with [13], where μ is a constant denoting the expected return of the stock, and σ is a constant parameter denoting the volatility of the stock[3].

Implied volatility

Looking at markets for European calls today, we can unambiguously determine the parameters for pricing a particular call (by the Black-Scholes-Merton formula) except for the volatility of the stock, σ. However, the market still quotes prices today at which these calls trade.

Implied volatility is the value that when substituted into the Black-Scholes-Merton pricing formula yields today's option value, given knowledge of the rest of the parameters. Implied volatilities are quoted by traders as option 'prices' because of convention and as a matter of convenience. One can use put-call parity to show that the implied volatility for a European call is equal to that of a European put writen on the same underlying with the same parameter values.

There is a one-to-one correspondence between implied volatility and market call option prices, though no closed form solution exists to solve for the implied volatility directly. Hence, numerical methods are used. For instance, Bloomberg calculates implied volatilities for options on US equities and indices on a discrete strike/time to maturity grid by first calculating implied forward prices, and then calculating implied volatilites on the grid consistent with this forward price (that is, substituting such a price for S_t). To yield values off the grid, non-parametric interpolation in variance space is used across strikes, while a Hermite cubic spline interpolation in total implied variance space is used across time to maturity.

The plot of implied volatilities at different strikes and tenors is called the volatility surface. Fixing tenor and looking at implied volatility by strike, we notice that implied volatility is greatest for deep in- and out-of-the-money options, relative to strikes about the at-the-money region. This is called the volatility smile. There exist many views and beliefs that explain the persistence of the smile, for instance, an excess demand for deep out-of-the-money puts to insure against extreme market down swings exists, pushing up prices relative to at-the-money

[3]It is somewhat remarkable that this pricing formula does not depend on μ.

puts, and hence also the left tail of the smile.

Arrow-Debreu securities

A fundamental idea in the seminal work of Black and Scholes [2] is that in an 'ideal' market setting the value of a European style security coincides precisely with the value of a (continuously hedged) replicating portfolio of stock and bond holdings that shares the exact terminal payoff of that security. In this sense, the security is redundant (Indeed, much of mathematical finance is about the pricing of redundant constructs.). From replication as a pricing method came the idea of risk-neutral martingale pricing, as we have seen, which highlights the importance of the risk-neutral distribution $\mathbb{Q}$. Much of the theory was developed in the continuous (bounded) time / continuous (unbounded) state space setting. Now we turn to considering the equivalent machinery in the discrete (bounded) time / discrete (bounded) state space setting, and from the perspective of the financial economist. For a deeper introduction, the reader is referred to [8].

Consider a simple model with discrete time set $T = \{0, 1\}$ and a discrete bounded state space $S = \{S_1, \ldots, S_n : n < \infty\}$. A state price security AD^i (also known as an Arrow-Debreu security) is a derivative that we can purchase at $T = 0$ that pays one unit of numeraire if a particular state of the world S_i occurs at $T = 1$, and nothing otherwise. In our case, the numeraire will be taken to be the money-market account.

Define the collection of state price securities $AD = \{AD^i : 1 \leq i \leq n\}$. A holding of AD guarantees a payoff of \$1 at maturity, regardless of the realised state of the world. Supposing interest rates are 0, and by the principle of no arbitrage, the value of holding AD at time 0 is then obviously also \$1. In fact, by the first fundamental theorem of asset pricing no arbitrage implies the existence of non-negative state price security values, so that $V^i_{AD} \geq 0 \; \forall i$ [6]. Analogous to the definition of a probability measure we call $V_{AD} = \{V^i_{AD} : 1 \leq i \leq n\}$ the state price (probability) density[4] for a given terminal payoff time.

Now, to be clear, each V^i_{AD} does not necessarily reflect the real-world probability that the world transitions to state i – or rather, from S_0 to S_i – at time 1. Indeed, different securities could and often do trade at a discount or premium based on the risk preferences of market participants and their concept of the value of a dollar given the alternative states of the world that could transpire. Further, the prices in V_{AD} depend on the state of the world today. But regardless, V_{AD} does define an equivalent probability measure to the real-world measure, in

[4] Or more correctly, the state price probability mass function.

that it assigns the same null outcomes, by the principle of no arbitage. Assuming market completeness, we have then that V_{AD} is unique[5] by the second fundamental theorem of asset pricing. Indeed, V_{AD} defines the unique (discrete) risk-neutral measure, $\mathbb{Q}$.

If we consider now that interest rates are not necessarily 0, then state prices will convey information about the time value of money. Indeed,

$$\sum_{i=1}^{n} V_{AD_0}^i = \frac{1}{1+r} \sum_{i=1}^{n} V_{AD_1} = \frac{1}{1+r},$$

so that given the value of AD today we can infer the market risk free rate over the next period. Further, we can normalise this price to move to $\mathbb{Q}$, or conversely multiply $\mathbb{Q}$ by $(1+r)$ to move back to the state price density, as required.

Finally, given our assumption of market completeness any terminal European option payoff V_1 can be decomposed into a linear combination of state price securities so that

$$V_1 = \sum_{j=1}^{n} \alpha_j V_{AD_1}^j$$

for a sequence of real constants $\{\alpha_i : 1 \leq i \leq n\}$. Hence by the linearity of option pricing and the principle of no arbitrage, there holds

$$V_0 = \frac{1}{1+r} \sum_{j=1}^{n} \alpha_j V_{AD_0}^j.$$

The ideas of this section can be extended in the obvious way to a multi-period setting.

Breeden-Litzenberger

A chief idea of Breeden and Litzenberger is that risk-neutral state price densities can be constructed from observed market option prices [3]. This is despite state price securities typically not being traded in real-world markets. Hence we have that the volatility surface, the option value surface, the risk-neutral state price density and more generally the risk-netural asset price return distribution are equivalent.

To show why this is the case, we consider an example in the spirit of [3]. We maintain the model similar to the previous section: our time set is $T = \{0, 1\}$ and the world starts in state S_0 before moving into any one of n states $S_i \in S = \{S_1 = 1, S_2 = 2, \ldots, S_n = n : n < \infty\}$. Next, we define a set of n call option prices V written on S, and such that $V_t^i \in V$ is the time

[5] Given today's state of the world.

t value of the i^{th} call with strike i satisfying

$$V_1^i = \max(S-i,0) =: \begin{pmatrix} \max(1-i,0) \\ \max(2-i,,0) \\ \vdots \\ \max(n-i,0) \end{pmatrix}$$

for $i = 0, 1, \ldots, n-1$, with one value realised depending on the the outcome of S.

By constructing portfolios consisting of a linear combination of members of V we create our state price securities. Firstly, we can make all payoffs of the form $(0, \ldots, 0, 1, \ldots, 1)$ for a $1 \times n$ vector consisting of n_0 0's followed by n_1 1's, for $n_0 + n_1 = n$, by taking $V_1^{n_0} - V_1^{n_0+1}$. Next, we can construct the required $1 \times n$ state price securities paying out 1 in state $n_0 + 1$ and zero otherwise by holding

$$V_1^{n_0} - V_1^{n_0+1} - (V_1^{n_0+1} - V_1^{n_0+2}) = V_1^{n_0} - 2V_1^{n_0+1} + V_1^{n_0+2}.$$

Of course, this holding of two long calls with strikes n_0 and $n_0 + 2$ and two short calls both with strike $n_0 + 1$, is a so called butterfly spread. Hence we conclude, by replication, that given the time 0 values of our call options, we also have the time 0 values for our set of state price securities.

More generally, if the increment between consecutive members of S is the constant ΔK, there holds

$$V_1^{n_0} - V_1^{n_0+1} = (\overbrace{0, \ldots, 0}^{n_0 \text{ terms}}, \overbrace{\Delta K, \ldots, \Delta K}^{n-n_0 \text{ terms}}).$$

Hence

$$V_{AD_0}^{n_0+1} = \frac{1}{\Delta K}\Big(V_0^{n_0} - 2V_0^{n_0+1} + V_0^{n_0+2}\Big),$$

and we have constructed the required state price density.

The reader will be unsurprised to learn that the continuous analogue of this result comes from differentiating the Black-Scholes call value function twice with respect to strike:

$$\begin{aligned} \frac{d^2V_0}{dK^2}\Big|_{s=K} &= \frac{d^2}{dK^2}\Big(e^{-rT}\int_0^\infty (s-K)^+ f^{\mathbb{Q}}_{S_T|S_0}(s)ds\Big) \\ &= e^{-rT} f^{\mathbb{Q}}_{S_T|S_0}(K) \\ &\approx e^{-rT}\frac{\mathbb{Q}(S_T = K|S_0)}{\Delta K} \\ &= e^{-rT}\frac{V^K_{AD_0}}{\Delta K}, \end{aligned}$$

where the risk-neutral probability density at K, $f^{\mathbb{Q}}_{S_T|S_0}(K)$, is not a probability, but rather thought of as a factor of an approximation to an element of probability:

$$\mathbb{Q}(S_T \in (K, K+dK]|S_0) = \mathbb{Q}(S_T \leq K+dK|S_0) - \mathbb{Q}(S_T < K|S_0) \approx f^{\mathbb{Q}}_{S_T|S_0}(K)dK.$$

We have seen that the state price security values can not be negative, and that state price securities that pay out on events of measure 0 must have 0 value, but these are our only requirements of them. State price security values greater than 1 can exist in an arbitrage free setting if the economy has a negative interest rate.

The state price by tenor matrix, S

Call the $m \times n$ matrix S, for m states and n tenors, the state price by tenor matrix. It is defined by

$$S = \{S^0_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\},$$

where S^0_{ij} is the time 0 value of the Arrow-Debreu security paying one dollar at time j given that the state of the world i has occured, and zero otherwise. We can construct the entries of S given market option price data via the process of [3], outlined in the preceding section. Hence, the changes in the state of the world that S describes refer to the evolution of the underlying asset on which out options are written. The estimation of S is shown in more detail in part 3.

The state price transition matrix, P

Call the $m \times m$ square matrix P, for m states, the state price transition matrix for a fixed arbitrarily chosen tenor T. It is defined by

$$P = \{p_{ij} : 1 \leq i, j \leq m\},$$

where p_{ij} is the value of an Arrow-Debreu security that pays out if state j of the world occurs at time T, given at time 0 we started in state i. It is important to note that P is *not* a (probability) transition matrix as described earlier, but on normalising the rows of P to sum to unity, it is. Indeed, we recover the risk-neutral transition matrix $\mathbb{Q}$, the $m \times m$ square matrix given by

$$\mathbb{Q} = \left\{ q_{ij} = \frac{p_{ij}}{\sum_k p_{ik}} : 1 \leq i, j \leq m \right\},$$

where q_{ij} is the risk-neutral probability of a transition to state j of the world from time 0 to time T, given at time 0 we started in state i.

We assume that $\mathbb{Q}$ is a time-homogenous Markovian transition matrix. P shares the property of time-homogeneity in the sense that its definition holds over any arbitrarily chosen time interval $[t, t+T]$, $t \geq 0$.

A primary contribution of this dissertation is to discuss robust non-parametric methods of constructing P from S. We refer the interested reader to part 4.

The natural probability transition matrix, F

Call the $m \times n$ matrix F, for m states, the natural probability transition matrix. It is defined by

$$F = \{f_{ij} : 1 \leq i, j \leq m\},$$

where f_{ij} is the real-world probability of a transition to state j of the world from time 0 to time T, given that we started in state i. It is somewhat remarkable that contrary to the long term view suggesting otherwise we can construct the entries of F from P, in our case under the mild assumptions of [15]; the process is explained in more detail below.

We have that F is a time-homogenous Markovian transition matrix. Also, given that the entries of F define the real-world transition probabilities of going from one state of the world to another as calibrated from market prices, which themselves are the cumulation of investor choices and beliefs, it is fair to think of its entries as the market beliefs about the real-world probability mass function describing the evolution of the respective reference underlying asset.

From P to F: a Ross Recovery theorem

Here we give a very brief overview of the key ideas of Ross, which are further explained by Carr and Yu, and which allow the construction of F from P [5, 15]. For more details we refer keen readers to those papers. We stay close to their notation below.

Define the transition kernel ψ by

$$\psi_{ij} = \frac{p_{ij}}{f_{ij}}.$$

Students of mathematical finance will recongnise this as the Radon-Nikodym derivative. An equivalent statement of no arbitrage is that a positive kernel exists [15]. The kernel is called transition independent if we can find a positive function of the states h and a positive constant δ such that for each pair (i, j)

$$\psi_{ij} = \delta \frac{h(S_i)}{h(S_j)},$$

from which it is easy to see that

$$p_{ij} = \delta \frac{h(S_i)}{h(S_j)} f_{ij}. \tag{2.7}$$

Given p_{ij} the idea of Ross is to find three unknowns: each term in the above product. As Ross points out, without restrictions on either the form of the kernel or of f_{ij}, it would not be possible to uniquely determine our unknowns from the information of p_{ij} alone. There exists much research on approaches for solving for our unknowns [10, 11], however the approaches are strong in the sense that they use historical return distributions or parametric assumptions to reach their conclusions. Ross makes the assumption of a transition independent kernel as an alternative that, despite being essentially empirically unverifiable, is arguably somewhat weaker. Refer to [5] to reach the conclusions of Ross making an alternative set of assumptions, that are better verifiable.

Proceeding to solve for equation (2.7), define the diagonal matrix

$$D = \text{diag}(h(S_1), h(S_2), \dots h(S_m))$$

so that we can write (2.7) in the following matrix form:

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

Solving for F,

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

Now F is a probability transition matrix so its row sums are equal to 1. Letting $\mathbf{1}$ be the $m \times 1$ matrix of 1's, there holds

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

and hence

$$Pz = \delta z$$

where $z = D^{-1}\mathbf{1}$. With one more condition on P, we have a chance at solving for D uniquely so that we can recover F. This is summarised in the following important theorem.

Recovery theorem [15]. Assuming no arbitrage, irreducibility of the pricing matrix P, and that the pricing matrix is generated by a transition independent kernel, then given any set of state prices there exists a unique positive solution pair: the pricing kernel and natural measure.

Proof. By the Perron-Frobenius theorem we solve $Pz = \delta z$ where δ is the Perron root of P, and z is the corresponding unique (up to positive scaling) strictly positive eigenvector. Then

the i^{th} diagonal member of D, d_{ii}, is given by z_i^{-1}. Hence we can solve for F by substitution into our above expression, with the solution being unique since the scaling term on z clearly cancels out. □

Hence by the Recovery theorem we can find F uniquely from P. Knowledge of F is very useful in the market place. For instance, it gives us a means to infer the market beliefs of real-world asset returns, a means to quantify the markets risk aversion in that we find the pricing kernel, and even a means to begin to analyse rare-event probabilities. Applications of the Recovery theorem are shown in part 3.

Finally, we make a brief note on the interpretation of δ. In Ross's paper, the representative agent theory of financial economics first explored by Samuelson [16] is referenced. Roughly, the idea is to model intertemporal choices, that is, choices that depend on cost-benefit tradeoffs through time. Samuelson first proposed the discounted utility model, where current expected utility as a function of present and future consumption is a weighted sum of future expected utilities at discrete times as a function of consumption at that point in time:

$$\mathbb{E}U(c_0, c_1, \ldots, c_T) = \sum_{i=1}^{T} \delta^i \mathbb{E}U(c_i).$$

The parameter δ is the discount factor for a representative agent, and represents the rate of pure preference for consumption in the present. It is assumed to lie in $(0, 1)$, so more weight is assigned to current consumption relative to consumption that occurs later in time. The validity of this assumption has been empirically tested and debated in the economic literature – we won't weigh in on this here. We will also assume that for our purposes δ is a fixed constant, though the validity of this is also arguable. Regardless, the δ in the above model is, by the assumptions of Ross, the same δ as we consider here. In our estimation procedure the estimated value of δ cleary depends on P. We have by the Perron-Frobenius theorem that δ is a strictly positive constant, but what else can we say? The answer is little really, except in the following circumstances when P has extra structure. From [14], there holds:

Proposition 2.2.1. (Ex. 8.2.7) If $P > 0$ then

$$\min_i \sum_{j=1}^{m} p_{ij} \leq \delta \leq \max_i \sum_{j=1}^{m} p_{ij}.$$

Proposition 2.2.2. (Ex. 8.3.7) Call a non-negative square matrix with row sums less than or equal to 1, and at least one row sum less than 1, substochastic. Then $\delta \leq 1$ for every substochastic matrix, and $\delta < 1$ for every irreducible substochastic matrix.

There is no reason that any extra conditions will necessarily hold for P so that $\delta \in (0, 1)$. All that we can be assured of is that $\delta > 0$, so that any estimated values of $\delta > 1$ are a reminder of the limitations of our model.

Part 3

The Ross Recovery theorem

A goal of this paper is to explain the econometric method employed in [15] in clear detail. We do this in three steps. The first step is to estimate the state price by tenor matrix S from market data. The second step is to estimate the state price transition matrix P given S. In the final step, we make use of the Recovery theorem to estimate the real-world transition matrix F from P. Step 1 and step 3 are reasonably straightfoward; their exists an extensive literature relating to the the former, and the latter is a straight application of the Recovery theorem. These two steps are the focus of this section. How step 2 is approached by [15] is not so straightforward, owing to two main reasons. This discussion is the focus of part 4, and the chief contribution of this dissertation.

Throughout, we apply the Ross Recovery calculus to S&P500 index options data for April 27, 2011. We make use of the Ross data for steps 2 and 3, as it is implied by his matrix estimations in steps 1 and 2. We use our own market data set for step 1. Our data is the volatility surface for S&P500 index options data for April 27, 2011 compiled from Bloomberg, and we use it as a proxy for the actual data set used by Ross. It is a fair approximation, however in that Ross used a proprietary data set inaccessible to the author, we can only replicate the Ross results up to a margin of error. Errors and inconsistencies that arise will be thoroughly outlined when considering each step.

3.1 Step 1. From option price data to S

In constructing S, we are estimating a 11×12 matrix. The columns are the 12 tenors corresponding to the 12 quarters over the next 3 years. The rows correspond to 11 specific return

levels chosen by Ross, which expressed as percentages are

$$(-35.1, -29.3, -22.9, -15.9, -8.3, 0.0, 9.0, 18.9, 29.7, 41.4, 54.2).$$

We will see how these values are chosen in part 4, as they are specific to the construction of P.

It is curious that Ross takes $n = 12$ rather than 11 in constructing his S. We will also see in part 4 that the purpose of constructing S is so that we might estimate P, however it requires only an 11×11 square matrix S to arrive at Ross's desired P, so that the additional column of information is redundant.

Our volatility surface as accessed on Bloomberg has dimensions of tenor and option moneyness. Option moneyness maps directly to percentage return as considered by Ross. This is simple to see:

$$\text{Option moneyness} := \frac{K}{S_0} \cdot 100\%$$
$$\text{Return} := \log\left(\frac{K}{S_0}\right) \cdot 100\%$$

Now, the tenor structure of our Bloomberg data is the same as Ross's: we have quarterly data extending 3 years forward. However, due to restrictions accessing a wide range of data via Bloomberg, it is impossible for us to match the return range of Ross. Our moneyness range is 80% to 122.5% corresponding to a return range of -22.3% to 20.3%, so we fall widely short in both extremes. Hence, we will only estimate a few rows of S and show how it compares to the corresponding rows of Ross's S.

Another difference our data set has with Ross's is that the Ross moneyness increment is 0.5%, while ours is 2.5%. This is because there are limitations to accessing very granular data in Bloomberg in a straightforward way. Hence, we extend our data by simple linear interpolation, so that our moneyness increment matches Ross's. While this is extremely crude, this part of the estimation process is only intended to be illustrative, and we will see that it still leads to an estimate for S which does not widely vary from Ross's.

Finally, given that Ross has this very granular increment for option moneyness, but only 11 unique return levels of the underlying to which he assigns values, it is assumed that his data is binned on intervals about the returns, though this is not declared in [15]. For the i^{th} state price at tenor T, S_{iT}, $2 \leq i \leq 10$, we sum all state prices corresponding to returns on the interval

$$\left[\frac{R_{i-1} + R_i}{2}, \frac{R_i + R_{i+1}}{2}\right)$$

to construct that entry of S. For $i \in \{1, 11\}$ the situation for calculating S_{iT} is more delicate. In particular, these values must account for the information in the left and right tails of the

state price density. This likely means that Ross had to extrapolate outside of his large data set to approximate these entries. While Ross makes no hint of how he did this, there is a body of literature offering suggestions that make use of a variety of techniques and ideas, both parametric and non-parametric. The interested reader should could, for instance, consult [1, 7, 18, 19].

A consequence of extrapolating in the tails is ultimately increased error in the extremes of our estimates for P and F. This is compounded by the fact that for deep in- and out-of-the-money options we may have relatively less reliable information about option prices since these options are less liquidly traded.

Finally, we construct the middle-most 3 rows of S using the data we have. We interpolate our data in moneyness, convert implied volatilities to call prices, and for each tenor retrieve our state-price densities by the Breeden-Litzenberger result. State prices are then binned and the entries recorded. Refer to table A.1 for the table output in [15], which is included for ease of reference, and table A.2 for the three rows of S we constructed. The estimated rows broadly agree with what is presented in Ross up to a tolerable margin of error. Our row corresponding to negative returns is broadly relatively under estimated, while the row corresponding to positive returns is broadly relatively over estimated.

Finally, the row sums of S in Table A.1 are shown and broadly agree with those given by Ross, with a modicum of rounding error at the third decimal place.

3.2 Step 3. The Recovery theorem: F from P

To illustrate an application of the Recovery theorem we construct the 11×11 square matrix F from the 11×11 square matrix P estimated by Ross in [15]. For convenience, the latter mentioned matrix is shown in Table A.3. The careful reader might notice that the column and row labels denoting the asset price have been labelled differently to how they appear in Ross's paper. This is explained in part 4.

Given the results from the financial theory section of this dissertation it is trivial to apply the Recovery theorem to P to construct F. The result is shown in Table A.4. Entries marked as a '0' denote precisely that, while entries marked as '0.000' denote a non-zero entry truncated to 3 decimal places of accuracy. This is done so there is no confusion or ambiguity regarding the fact the F is equivalent to $\mathbb{Q}$ and hence P, that is, they assign the same events 0 probability. Also, there are small discrepancies in entries of F relative to the corresponding entries in Ross's F, due to rounding error. Regardless, it is the author's opinion that the recovery of a quarterly real-world probability state price transition matrix is an impressive result; this idea is the chief

contribution of Ross [15].

The pricing kernel is also trivially found; the kernel for the middle-most row of F is shown as the bottom row of Table A.4. Again the kernel broadly agrees by the kernel in Ross's paper, though we cannot estimate the kernel for the final 3 entries in the row due to rounding error leading to a divide by 0 error.

Finally, we can check the characteristic root δ for our system, which is found to be 1.0018. Ross notes that the fact that $\delta > 1$ indicates the delicate nature of his estimation procedure, especially in light of the fact that estimating P from an S constructed from monthly observations rather than quarterly[1], we find $\delta = 0.9977$ – a better result [15].

In part 4 we will estimate more real-world probability transition matrices for alternative estimated state price transition matrices, without refering to the underlying process of how.

[1]Ross makes no other reference to how he has done this, or the results he has retrieved regarding S, P and F for monthly observations.

Part 4

Econometric issues in the application of Ross Recovery

In this section, we look at step 2 of our estimation method, and construct the state price transition matrix P from the state price by tenor matrix S. As alluded to earlier, despite the work of Ross [15] this is not a straight-forward task, being made difficult for two main reasons. Firstly, there are typographical errors that make the terse explanations of Ross's method in the Ross paper ambiguous. Secondly, while we highlight some of these errors, it is the author's opinion that reconstructing Ross's result is still impossible without further assumptions about Ross's own process, since subtle but vital details are not described.

In a surprising twist when constructing this dissertation, on forming a very likely hypothesis regarding the method used by Ross to estimate P, the Ross result cannot be recovered to a suitable margin of error. In fact, we will show that employing what we call the Ross method leads to at best the pursuit of a solution to an ill-posed inverse problem, and at worst a very different estimate to P relative to what is presented by Ross.

But all is not lost. This section will also compare nine different robust non-parametric methods for estimating P from S, and in particular show that one method provides a superior estimate for P. Further, we generate real-world transition matrices F for each P, and show that our preferred estimate for P yields an acceptable estimate for F via the Recovery theorem. This comparison, and improved estimation of P from S, is the chief contribution of this dissertation.

4.1 Step 2. Moving from S to P, Ross methodology.

Let $T = 0.25$, a quarter year. Recall from the definition that

$$P = \{p_{ij} : 1 \leq i, j \leq m\}$$

where p_{ij} is the value on an Arrow-Debreu security that pays out if state j of the world occurs at time T, given at time 0 we started in state i. Obviously, it is not possible to solve for P given real-world data since (at least for $m > 1$) such a market for options does not exists. For a given point in time, we can at best find $P_i := \{p_{ij} : 1 \leq j \leq m\}$ for a fixed i denoting the state of the world today. Indeed, for our choice of tenor, P_i is exactly the first column of S. Hence, we have m of the m^2 unknowns required to construct P, from our work in step 1.

Ross has chosen to set $m = 11$, which is likely to be optimal given his data set and to be sufficient for him to illustrate the recovery theorem in reasonable detail. We do the same.

Choosing an odd number of state, Ross chooses his states in a symmetric way about the baseline state that represents the economy 'today', April 27, 2011. It follows that the middle-most row of P corresponds to the information we have about transitioning from today's state, the first column of S. Also, Ross labels his states in three equivalent ways. We will define these states, before explaining their calculation. Firstly, as standard deviations about today's price:

$$S_1 = \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}.$$

Secondly, as a normalised price sequence corresponding to these standard deviations. The values are normalised by the day's asset price level,$S_0 = 1355.6$. We note here that the values as given in Ross's paper are not consistent with the standard deviation Ross has claimed to have used, and we attribute this to a 'copy/paste' error in the paper, where the copying has been done from an earlier table of results. The corrected sequence is:

$$S_2 = \{0.649, 0.706, 0.771, 0.841, 0.917, 1, 1.090, 1.189, 1.296, 1.413, 1.541\}.$$

Finally, we have the states defined as returns corresponding to the normalised price series, keeping the rounding as in the Ross paper:

$$S_3 = \{-35\%, -29\%, -23\%, -16\%, -8\%, 0\%, 9\%, 19\%, 30\%, 41\%, 54\%\}.$$

Ross chooses a standard deviation of 9%, but to be more precise the actual standard deviation implied by S_3 is 8.65% per quarter. This equates to a yearly standard deviation of

$$\begin{aligned}\sqrt{4 \cdot \text{ATM quarterly implied variance}} &= 2 \cdot \text{ATM quarterly implied volatility} \\ &= 2 \cdot 8.65\% \\ &= 17.3\%,\end{aligned}$$

which roughly agrees with the quantity gained from Bloomberg of 20%. From S_2, it is clear that Ross has adopted the convention of using multiplicative standard deviations to define the states; this convention is typical for non-negative and usually log-normal data[1]. Recall that when calculating a level of k multiplicative standard deviations about the mean, the multiplicative factor is $e^{k\sigma}$, where σ is our quarterly implied volatility. Ross chooses ± 5 standard deviations as the range for the state space to provide 'reasonable coverage' out into the tails of the distribution. Hence we have S_1 and S_2, and (quarterly) returns for S_3 are calculated as $(S_T - S_0)/S_0 \times 100\%$.

We can use the assumed time homogeneity of P to solve for the remaining $m(m-1)$ unknowns since

$$S(:, t+1) = S(:, t) \cdot P, \qquad 1 \leq t \leq m-1, \tag{4.1}$$

for $S(:, t)$ the t^{th} column of S. Observe that these equations state that for fixed t, the state price security value for the derivative paying \$1 if state j of the world occurs at time $t+1$ is the sum over all possible states k of the product of the state price value in state k at time t and the transition price of moving from state k to j [15], which can be thought of simply in terms of the law of total probability.

If we construct P making use of the system of equations (4.1), we can yield negative entries in P, which contradicts our definition of P in our no arbitrage setting. A simple example showing why this is the case can be seen in [5]. This necessitates the need to impose additional restrictions on our system. At least, we require that the entries of P are non-negative. Ross also chooses to impose additional restrictions beyond non-negativity on the estimation 'in an effort to minimise the errors in the estimation of P' [15]. He does not elaborate on this statement; in particular it is not clear what errors he is referring to. We address this further in the next section. Regardless, Ross's additional restriction is that the rows of P are unimodal.

At this point, we make an important aside to avoid confusion. The Ross state price by tenor matrix S is presented with 12 columns in Ross's results. The final column is redundant information when it comes to estimating P using the Ross method.

We saw in part 2 how to estimate P by the Ross method. The method corresponds exactly to M3. Rasmus Bro has an influential paper on estimation of this type [4], and deposited useful code the Matlab Central File Exchange for practical implementation. We make use of the Bro code for estimations of this type.

[1] A typical reason for doing this in science, for example, is so that calculated confidence intervals do not contain a negative region of the real line.

As addressed briefly in [4], and discussed earlier, a problem with estimation method M3 is that it produces a solution that is not necessarily unique. In our case, there can be many alternative P's estimated from a given S (This is despite there being exactly one S for any given P: this is clear from (4.1).). Hence moving from S to P and back again is an ill-posed inverse problem. Based on this alone, M3 is hardly a suitable estimation method.

Despite this, we estimate a representative P under M3. It is shown in Table A.9. For comparison, the precise Ross P is given in Table A.3. We observe they are markedly different. Firstly, Ross's P appears both sensible and plausible. The maximum of each row lies on the diagonal, which is intuitively pleasing. The broad trends in prices by row are also intuitive. However *our* M3 estimate has many displeasing features. There is one wildly large price that is hard to intuitively justify, and rows such as that referring to 4σ make little sense. Why, for example, is it implied that from state 4σ of the world, the system can only transition to states $-3\sigma, -4\sigma$ or -5σ? Evaluating the estimation multiple times, including initialising at alternative starting guesses, does not yield us an estimate any closer to Ross's. To account for this, we imagine that Ross has (i) imposed additional restrictions on his system undocumented in his paper, such as the mode of each row must be on the diagonal; or (ii) he has been extremely lucky with the convergence of his algorithm, estimating a very impressive but unlikely P; or (iii) Ross has combined a convenient starting guess for his algorithm with a with low error tolerance for changes in the estimate by an ALS algorithm, to arrive at his P.

The estimation of P from S in our setting is an important part in the Ross Recovery process. We seek to analyse and improve it in the next section.

We make one more important note regarding arbitrage in our model. Consider the final row of the Ross estimate for P. It appears that we have an arbitrage opportunity. Indeed, since the state space is bounded above, and the final row corresponds to the upper-most state of the world, the market can only stay in that state of the world or decline to a lower state. It appears that an arbitrage would be to short the asset. However, it is pointed out in [5] that this is not necessarily an arbitrage. If in the top state of the world, dividend yields exceed the short rate of interest, arbitrage is avoided. In the top state of the economy, this is a realistic assumption. Further, looking at the interest rate implied by our model in the top state, we see it is *negative* (since the row sum of P is greater than 1). Hence, dividend yields must be greater than the short rate here, since they are non-negative.

4.2 A comparison of nine ways to estimate P from S.

In this section we consider nine ways to estimate P from S. Recall the estimation methods M1 – M9 outlined in section 2.1; these are the methods we will consider. Where applicable, a tenth method, the Ross method that we have already seen, and referred to as 'Ross', is also discussed.

We consider M1 – M9 since we seek robust, non-parametric estimation methods for finding P. We will perform all estimations using the Ross estimate for S. We compare the methods based on four main criteria, which we outline below.

S-level error. For a given method, we construct P. It is then easy to retrieve a unique S implied by P: by equation (4.1) it follows that the middle-most row of P^k is the k^{th} column of S. We define the S-level error of the estimation to be the Frobenius 2-norm given by

$$\text{error}_S = ||S - \hat{S}||_2^2 = \sum_{i,j} |s_{ij} - \hat{s}_{ij}|^2.$$

P-level error. This error measures the robustness of an estimation method to error in S. It is important because the S we estimate is subject to error from the estimation method used to construct S and also from the noise in the data we collect from the market. For the given S, we estimate P. We then add a small amount of random noise to S and restimate P. The P-level error is then

$$\text{error}_P = ||P - P^{\text{noise}}||_2^2 = \sum_{i,j} |p_{ij} - p_{ij}^{\text{noise}}|^2.$$

Parsimony. We consider the number of unique parameter estimates required to estimate a particular P.

Estimated δ. Recall that the estimated δ is the Perron root for each estimation method. As we have seen, for consistency in our modelling it is ideal for δ to lie in the interval $(0, 1)$.

We address acceptable S- and P-level errors for our estimation methods in turn. The median price on all S&P500 index options used to construct S is approximately \$US140. It is realistic to assume that each option costs \$US1 to trade [7], implying a bid-ask spread of approximately $2/140 \approx 140$bps. Now, we accept error in the entries of $\hat{S}$ within an interval of half the bid-ask spread about the corresponding entries of S, so that (rounded) prices in $\hat{S}$ are consistent with the midprices implied by S. Hence, an acceptable upper limit for the S-level

error can be given by

$$\begin{aligned}
&\big(\text{tolerable error about mid} \cdot \text{total value of assets in } S\big)^2 \cdot \#\text{ estimated unknowns} \\
&= \Big(\frac{1}{4} \cdot 140\text{bps} \cdot 10.23\Big)^2 \cdot 110 \\
&\approx 0.14.
\end{aligned}$$

Considering the P-level error, we start by noting that the value is subject to the random perturbation in S. Hence, we simulate 1000 perturbations in S to generate 1000 estimates for error_P, and perform our analysis on the statistical measures of centrality and spread gleaned from our estimates. In particular, we choose the robust measures of the $25^{th}, 50^{th}$ and 75^{th} quantile, since some extreme values of error_P that inevitably arise over a large sample of simulations could distort measures sensitive to outliers, such as the mean and standard deviation. We also report the maximum P-level error found over our simulations, to give an idea of how unstable a method can be relative to small errors in S. For a particular method, we would like all of these statistics to be as small as possible. To be more precise, a P-level error of α implies an average absolute error across each entry of P of $\sqrt{\alpha/110}$. The practioner will have to decide what level of α is tolerable based on his application. Finally, we note that our perturbations in S are very small; we add a $N(0, 0.005^2)$ error to every entry in S, then take absolute values. Hence about 95% of members of $\hat{S}$ lie within ± 0.01 of the original value, and on average the members of S are unchanged.

We consider our results in the next subsection.

4.2.1 Results

For ease of reference, P-level error, S-level error, estimated δ and the number of unique parameters estimated for each method are shown in Table 4.1. In Figure 4.1 the S-level error for each of the ten methods is displayed diagrammatically. In Figure 4.2 we have box and whisker plots for the simulated P-level errors for methods M2-M9. The Ross method was not amenable to this latter analysis, and M1 is not displayed so as to not dramatically alter the scale. Finally, the Appendix to this dissertation contains the estimated P and F matrices, along with the pricing kernel corresponding to today's state, for the interested reader's perusal. The corresponding tables from Ross's paper are also included. We make reference to these images and tables throughout to adequately compare and contrast each of the estimation methods, which we do in turn, below.

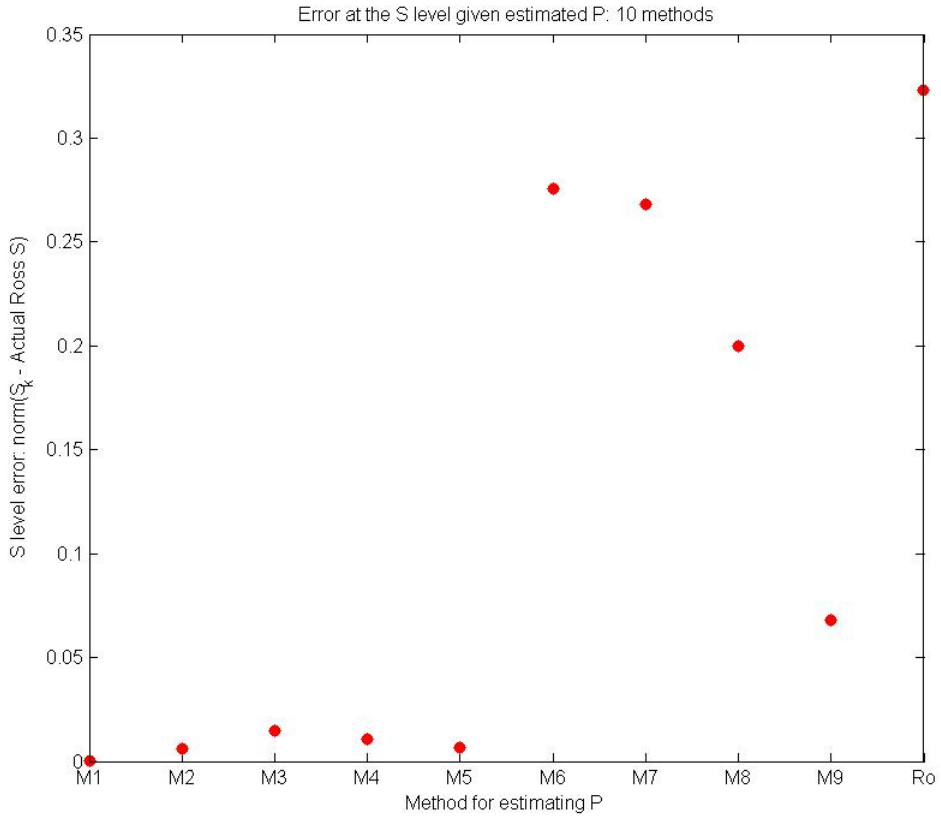


Figure 4.1: This image depicts the S-level error for each of our 10 estimation methods.

Under M1, ordinary least squares with a fixed middle row, we estimate 110 unique parameters. From Table A.5, we see we have an unacceptable estimate for P – many entries are negative. Hence, we cannot use the Perron-Frobenius theorem to generate a corresponding F; Table A.6 is intentionally left blank. Consequently, neither can we recover a Perron root δ. As expected, our S-level error is exactly 0, since this system is fully specified so that P can be estimated exactly from S, given equations (4.1). It is amusing to note that the median P-level errror for this method is 25.3, while the maximum error over 1000 simulations is 17624.3. Clearly, this estimation method is highly sensitive to small errors in S.

We conclude that M1 is an unacceptable estimation method. We proceed by requiring that every other method we trial imposes the non-negativity of estimates.

Under M2, ordinary least squares with a fixed middle row and non-negativity, we estimate 110 unique parameters. From Table A.7, we see that we have a better estimate for P than M1 in that the state price security values sit roughly in a range that we would expect. The P- and

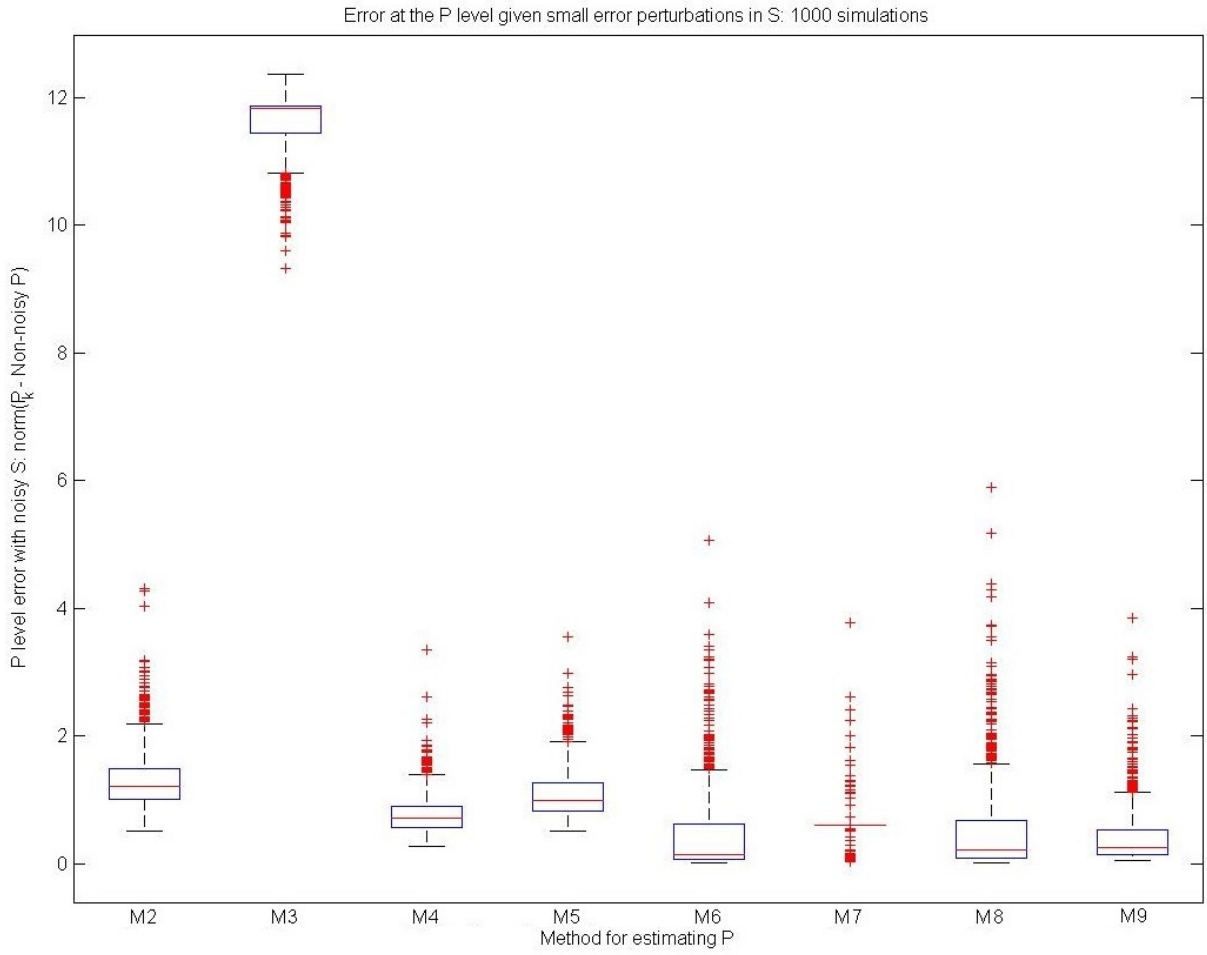


Figure 4.2: This image depicts the P-level error for M2-M9.

S-level error, and δ for this method are broadly acceptable, though the relatively high P-level error might not be acceptable to some practitioners. Also, when we look at Table A.8 we notice many obscure probabilities. Why, for instance, should we move from the state corresponding to -3σ to state 2σ with real-world probability 0.835? Our result is likely of little practical use.

We conclude that despite having some nice properties, we should try and impose extra restrictions on P to yield estimates that are at least slightly more intuitively pleasing. We also hope for a reduced median P-level error.

Under M3, unimodal with fixed row and non-negativity, we estimate 110 unique parameters. We are already somewhat familiar with this method owing to the results of Ross. Despite the method yielding an acceptable δ and S-level error, the P-level error is extremely high, seemingly due to the large entries that persistently occur when estimating P under this method. Also,

the P that we generate it not intuitively pleasing. Finally, given that M3 generates multiple different estimate for P when run multiple times, we report that no other estimated P was sufficient. It is interesting to contrast the results of M3 with those of the Ross method as shown in Table 4.1, given that the methods are supposedly the same. The Ross method yields an unacceptable δ, and the highest S-level error of any estimate. It does, however, yield the most intuitively pleasing estimates for P and F.

We conclude that M3 is unacceptable due to its high P-level error and unintuitive estimates. The Ross method fails based on the S-level error and generated δ, though it does yield pleasing estimates. It is curious that in the Ross estimate for P, the mode of each row falls on the main diagonal. This encourages us to impose some extra structure on our unimodal estimation, leading us to consider M4.

Under M4, unimodal with the row modes fixed to the diagonal, fixed row and non-negativity, we estimate 110 unique parameters. The method generates moderate P-level error which is in the middle of the range compared to other methods, and a very small S-level error. The δ is also accceptable. The estimate for P and F is moderately intutive, though, for instance, in row -2σ there are only two options for the future evolution of the asset. However where this method really fails can be seen from Table A.11. The bottom-most row contains all 0 entries. The estimate satisfies non-negativity, and also trivially unimodality with the mode falling on the diagonal. However, since P is no longer an irreducible matrix, we cannot use the Perron-Frobenius theorem to estimate F. We instead appeal to the extension of the theorem to reducible matrices, and retrieve an estimate of F, albeit with a final row containing no information. On one hand, estimating an informationless final row of P and F doesn't concern us too much, because the data that contributes to its estimation is typically concerning illiquid deep in-the-money call options, which could be regarded as relatively informationless in itself. On the other hand, we would prefer that we did not retrieve a 10×10 matrix for F (Though one could artificially augment an entry $\epsilon > 0$ in the bottom-most corner of P, and have the highest state of the world corresponding to an absorbing state in F.).

It should be noted here that the algorithm used for constructing P from S by M4 was potentially suboptimal. Despite the discussion of estimation under unimodality with fixed mode placement and non-negativity in [4], the Matlab code provided does not implement the non-negativity. Further, the mentioned piece of code (offered as part of a larger package) is deemed temperamental by this author, failing to converge to a solution on occasion due to a

particular non-trivial bug that sometimes attempts to access the 0^{th} entry of a vector[2]. The author fixed the bugs in the code, and imposed non-negativity in what seems a rather ad hoc way: by replacing negative entries with 0's as they arose. Indeed, it is stated (with reference to a proof elsewhere) in [4] that this is *the* optimal method for implemeting non-negativity in unimodal least squares regression[3]. Regardless, this note should clarify for the interested reader how a row of 0's was yielded in the estimate for P under M4.

In conclusion, M4 is a reasonable method for estimating P from S, but has drawbacks in that the P-level error is moderate, and that the estimation output is not as clean as one might like.

Noticing a number of '0' entries in the estimate for P under M4, we thought it might be interesting to restrict P to have '0' entries in certain positions. Under M5, 0's placement with fixed row and non-negativity, we estimate 69 unique parameters setting $u = 3$ and $d = 5$. This means that from a given state, in the next period the underlying asset can move upward up to 3 states or downward up to 5 states. These choices are consistent with the data presented in S.

Though our estimated S-level error and δ are pleasing the P-level error may be unacceptable to some practitioners. Further, the estimates in Table A.13 and A.14 are not pleasing intuitively. In particular, there are many '0' entries on the diagonal and upper and lower off-diagonals that greatly restrict how the underlying can evolve in a time period. We conclude that M5 is not a reasonable estimation method.

Our final estimation method methods M6-M9 are from a class we name 'sliding window'. Methods from this class are restricted to have certain estimated parameters in P to be equal. They require the smallest number of unique parameter estimates of all our methods. Under M6, sliding window with non-negativity, we estimate 21 unique parameters. Recall that the entries of each diagonal of P are equal, and that the middle row of P is not restricted to be the first column of S. Viewing Table's A.15 and A.16, we notice that we have generated the first set of results that are reasonably intuitively pleasing. Also, the P-level error and δ estimate for M6 are within suitable bounds. However a problem of this method, which is enough for us to conclude it is unsuitable, is that its S-level error is remarkably high.

[2]Thirdly, the code available online contains an obvious error: there is a misuse of the Matlab 'break' function within an 'if' statement. Break is only defined in 'for' and 'while' loops. Hence, the code cannot even begin to run without this first being corrected.

[3]While this author is somewhat convinced, it is left as a subject for further research.

Under M7, sliding window with fixed row and non-negativity, we estimate 10 unique parameters, the least of any method. For this method the entries of each diagonal of P are equal, and the middle row of P is fixed to be the first column of S. Hence there are only a few parameters to estimate in the upper right and lower left 'corners' of P. The P-level error of this method shows the least variablity about the middle 50 percentile, which is unsurprising given the rigid structure imposed on P that makes it robust to changes in S. However, again, we can conclude that this method is unsuitable, owing to its large S-level error.

Under M8, sliding window with fixed row and non-negativity, we fix the diagonal entries of P to be the same, though not necessarily equal to the middle-most row. In this model, we estimate 20 unique parameters. Despite seeing a reduction of the S-level error under M8, relative to M6 and M7, we can conclude that this method is unsuitable, given it is still sufficiently large.

Under our final estimation method, M9, sliding window with fixed row, two states and non-negativity, we estimate 30 unique parameters. In this model, the middle row is fixed, and the submatrices lying above and below this row are restricted to have equal entries on their respective diagonals. The motivation for this is to allow for two unique states of the world 'above' and 'below' the current state, which we refer to as the (relatively) 'good' and 'bad' economies. The model allows enough freedom in the parameters so that the S-level error is much less than than the other sliding window methods. Indeed, the error is well below our upper bound for what is tolerable. Further, this method has one of the lowest P-level errors of all methods considered, so that the estimation of P is robust to small errors in S. We have that $\delta = 0.996 \in (0, 1)$, in agreeance with our modelling.

Referring to Table A.21 and A.22, we have what are intuitively plausible estimates for P and F. We see a distinct difference in the parameter estimates for the good and bad states. Given we are in the bad state, there is a large probability (0.363) of staying in that state into the next quarter, and a larger probability (>0.450) of moving up by one state. This is in contrast to a probability of 0.594 of staying in a good state over the next quarter given we started in one, though we have a significantly reduced chance of moving up by one state, given we are in a good state, relative to being in a bad state. In our model, there are no glaring inconsistencies with real-world transition probabilites. It is interesting to note, however, that from state $-\sigma$ there is a 2.5% chance of a crash to the bottom state. There is also a similar chance of 2.9% or falling from the second top state to the bottom, and a 7.3% chance of a crash from the top

state to the -4σ state. In fact, corresponding to these latter two values is state price security value of \$US 0.56, which can be thought of as the value of a tail-risk insurance policy paying out \$US 1 in the respective crash event. Also, the value of these policies is greater than the value of the policy when there is a crash from state -1σ. This indicates either that investors are more risk-averse toward larger crashes, or that they value a dollar more when going from very good times into very bad times, as opposed to bad times into very bad times. Of course, this is all very pleasing intuitively.

We conclude that M9 is our preferred and potentially a very powerful method for estimating P from S.

Finally, the estimation of methods M6–M9 present an interesting trade-off. Loosening the restrictions on P means that we estimate more unique parameters in our model, at the expense of our model being less parsimonius, while on the other hand we reduce the error of our model (in the sense we have described) and benefit from an improved fit. It is pleasing that our preferred method, M9, estimates a P that is much more parsimonius than M1-M4, and that is broadly a better model.

4.3 Remark: testing F

This short section is intended to be illustrative in nature only. We ask the question, given we have an estimate for F, how can we test that it is a useful model for a practical setting?

Firstly, it is well-known that any finite state Markov chain has a stationary distribution: see, for instance, exercises 6.6.1 and 6.6.2 of [9]. Given the assumptions of our model, we might be tempted to analyse the stationary distribution of F. But for practical intents and purposes, this is a useless pursuit: it takes aproximately 100 time steps to find such a stationary distribution – or 25 years – and our model is certainly not designed for forecasting that far into the future. You might be hard pressed to find a practioner who is in the industry long enough to check up on his forecast in the end, anyway!

However, it is instructive to back test our F for predictive power, at least in the short term. On April 26, 2012 the S&P500 index traded at 1399.98, corresponding to a (binned) return of 0%, and on April 26, 2013 the index traded at 1582.24 corresponding to a (binned) return of 19%[4]. Generating real-world return distributions for each of these days, using our model on April 27, 2011 and given our intial state, we respectively recover

$$[0.004,\ 0.004,\ 0.012,\ 0.028,\ 0.077,\ 0.191,\ 0.345,\ 0.251,\ 0.079,\ 0.009,\ 0.000]$$

[4]These trading days are the last for their respective weeks.

Table 4.1: This table records some of the key statistics relating to the estimation of P from S for each of the 10 methods considered.

Method	# parameters	P-level error 25 percentile	P-level error median	P-level error 75 percentile	P-level error maximum	S-level error	δ
M1	110	22.595	25.335	41.277	17624.329	0.000	na
M2	110	1.003	1.213	1.490	4.301	0.006	0.974
M3	110	11.444	11.822	11.862	12.364	0.023	0.982
M4	110	0.572	0.712	0.901	3.352	0.011	0.970
M5	69	0.823	0.997	1.258	3.555	0.007	0.971
M6	21	0.069	0.141	0.627	5.071	0.276	0.989
M7	10	0.609	0.609	0.609	3.776	0.268	0.987
M8	20	0.086	0.218	0.676	5.889	0.200	1.017
M9	30	0.145	0.218	0.535	3.844	0.068	0.996
Ross	110	na	na	na	na	0.323	1.002

and

$$[0.006,\ 0.007,\ 0.012,\ 0.025,\ 0.059,\ 0.127,\ 0.241,\ 0.257,\ 0.174,\ 0.075,\ 0.016].$$

So the chances of seeing our respective results were 19.1% and 25.7%. The author won't read too much into this, but the return distribution also assigns more mass to the upper good states of the world with time. This may be evidence that on April 27, 2011, the market not only forecast positive returns of the S&P500 index with high probability, but also an increased chance of reaching the highest return levels on an even further horizon. The model has also assigned the most probability mass to the 19% (binned) return. Of course, this is precisely the return the index has seen over the past years.

How could we check this latter most result, that seems somewhat remarkable, against the chance it is just a supreme coincidence? In theory, we could construct a very large data sample, and construct estimates F (dynamically updating with time or otherwise), and use the models to make predictions forward. We can then observe what happened in 'the future', and if we gather a sufficient number of prediction/observation pairs, we could perform a statistical test to see if predictions fell in the correct category 'enough' times according to what was implied by the returns distribution. This is a big task; we leave it for future work. However these preliminary results are encouraging.

Part 5

Conclusion

This paper is about inferring market beliefs about the real-world probability distribution describing the future financial returns of an underlying asset from option price data. That this is possible is a triumph of the modern idea of Ross Recovery. This idea guides us to take option price data on a single day, estimate the state price mass functions for different tenors, then estimate the state price transition matrix – which can be normalised to give the risk-neutral transition matrix – before finally applying the Recovery theorem to recover the real-world transition matrix.

The primary goal of this thesis was to understand the work of Ross [15] and to piece together the nuance and subtleties of his econometric process. We did this, while finding some surprising results concerning the limitations of the propositions of Ross. In particular, these occured in the estimation of the state price transition matrix from the state price by tenor matrix. We compared nine alternative methods for performing this estimation, and found one to be superior in terms of error, robustness, parsimony and framework fit. Finally, we found that the resulting real-world transition matrix was useful for making statements about the real-world evolution of the S&P500 between April 2011 and April 2013.

It is the author's view that the relatively new idea of Ross Recovery will have an effect on how modern finance is practised. Ross Recovery is useful, for example, for the risk manager who seeks to hedge risk under the real-world probability measure, and for the investment manager, who seeks to optimally allocate his resources across risky assets.

Also, there is much more scope for interesting theoretical work under Ross Recovery. Firstly, the idea could be extended to an alternative market, such as the interest rates market, where the assumptions of the modelling may better hold. In particular, an interest rate underlying may fit the assumption of existing on a bounded state space more reasonably than a stock index[1].

[1]Indeed, the original intent of this dissertation was to explore this avenue, before other lines of pursuit

Secondly, the least-squares regression methods could be modified so that weighted least squares estimation is used. The most liquidly traded options should contribute most to the estimation of P from S; given real market data we could assign weights to different strike and tenor combinations to reflect this fact. Thirdly, we could dynamically update our estimates through time, and see what story this tells about changing market beliefs and sentiment. Fourthly, as mentioned in Ross [15], our estimation has ignored the potential dependence of state on other variables such as past returns and implied volatility, which could we could include in our modelling. Fifthly, as mentioned in text, we could seek to improve the estimation procedure from S to P – though this will depend on the availability of good data. And finally, we would like to perform statistical tests on F to assess how its forecasts hold up over time.

presented.

Appendix A

A.1 Estimating S

Ross's S as estimated from proprietary data is shown in Table A.1. Our estimate for three rows of S given our limited data set is shown below in Table A.2. We expect the minor discrepencies since we estimated using different data sets (and possibly slightly different techniques).

A.2 Ross P and F

For convenience, Ross's estimates for P and F are displayed in Table A.3 and Table A.4 for ease of reference.

A.3 Estimated P's and corresponding F's

Table A.5 to A.22 give the results referred to in part 4 of the dissertation.

A.4 Selected Matlab code

It is the author's decision to not include the Matlab code used to generate the results in this dissertation as part of the Appendix. Readers interested in the application of Ross Recovery and the work of this dissertation should feel free to contact the author to request elements of his code.

Table A.1: State prices by tenor, April 27, 2011. Ross data.

Tenor:	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3
Return												
-0.351	0.005	0.023	0.038	0.050	0.058	0.064	0.068	0.071	0.073	0.075	0.076	0.076
-0.293	0.007	0.019	0.026	0.030	0.032	0.034	0.034	0.035	0.035	0.035	0.034	0.034
-0.229	0.018	0.041	0.046	0.050	0.051	0.052	0.051	0.050	0.050	0.049	0.048	0.046
-0.159	0.045	0.064	0.073	0.073	0.072	0.070	0.068	0.066	0.064	0.061	0.058	0.056
-0.083	0.164	0.156	0.142	0.128	0.118	0.109	0.102	0.096	0.091	0.085	0.081	0.076
0	0.478	0.302	0.234	0.198	0.173	0.155	0.141	0.129	0.120	0.111	0.103	0.096
0.090	0.276	0.316	0.278	0.245	0.219	0.198	0.180	0.164	0.151	0.140	0.130	0.120
0.189	0.007	0.070	0.129	0.155	0.166	0.167	0.164	0.158	0.152	0.145	0.137	0.130
0.297	0.000	0.002	0.016	0.036	0.055	0.072	0.085	0.094	0.100	0.103	0.105	0.105
0.414	0.000	0.000	0.001	0.004	0.009	0.017	0.026	0.036	0.045	0.053	0.061	0.067
0.542	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.002	0.003	0.003
Discount Factor	1.000	0.993	0.983	0.969	0.953	0.938	0.920	0.900	0.883	0.859	0.836	0.809

Table A.2: State prices by tenor, April 27, 2011. Bloomberg data.

Tenor:	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3
Return												
⋮						⋮						
-0.159						⋮						
-0.083	0.138	0.120	0.106	0.097	0.091	0.084	0.078	0.075	0.070	0.067	0.062	0.062
0	0.473	0.293	0.224	0.189	0.165	0.145	0.131	0.119	0.112	0.099	0.096	0.089
0.09	0.308	0.375	0.342	0.281	0.231	0.202	0.181	0.166	0.152	0.146	0.135	0.128
0.189						⋮						
⋮						⋮						

Table A.3: The state price transition matrix, P, April 27, 2011. Ross Paper (corrected column and row labels).

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.671	0.241	0.053	0.005	0.001	0.001	0.001	0.001	0.001	0.000	0.000
-4	0.706	0.280	0.396	0.245	0.054	0.004	0.000	0.000	0.000	0.000	0.000	0.000
-3	0.771	0.049	0.224	0.394	0.248	0.056	0.004	0.000	0.000	0.000	0.000	0.000
-2	0.841	0.006	0.044	0.218	0.390	0.250	0.057	0.003	0.000	0.000	0.000	0.000
-1	0.917	0.006	0.007	0.041	0.211	0.385	0.249	0.054	0.002	0.000	0.000	0.000
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0.000	0.000	0.000
1	1.090	0.001	0.001	0.001	0.004	0.04	0.204	0.382	0.251	0.058	0.005	0.000
2	1.189	0.001	0.001	0.001	0.002	0.006	0.042	0.204	0.6493	0.243	0.055	0.004
3	1.296	0.002	0.001	0.001	0.002	0.003	0.006	0.041	0.195	0.361	0.232	0.057
4	1.413	0.001	0.000	0.000	0.001	0.001	0.001	0.003	0.035	0.187	0.347	0.313
5	1.541	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.181	0.875

Table A.4: The natural probability transition matrix, F, Ross Recovered from P, Table A.3.

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.670	0.252	0.060	0.006	0.002	0.002	0.002	0.003	0.004	0	0
-4	-29%	0.267	0.395	0.266	0.066	0.006	0	0	0	0	0	0
-3	-23%	0.043	0.206	0.393	0.279	0.073	0.007	0	0	0	0	0
-2	-16%	0.005	0.036	0.193	0.389	0.290	0.082	0.005	0	0	0	0
-1	-8%	0.004	0.005	0.031	0.182	0.384	0.309	0.081	0.004	0	0	0
0	0%	0.003	0.004	0.011	0.031	0.132	0.477	0.332	0.011	0	0	0
1	9%	0.000	0.001	0.001	0.002	0.027	0.169	0.381	0.313	0.095	0.012	0
2	19%	0.000	0.000	0.000	0.001	0.003	0.028	0.163	0.6492	0.317	0.102	0.012
3	30%	0.001	0.000	0.000	0.001	0.001	0.003	0.025	0.149	0.360	0.330	0.130
4	41%	0.000	0	0	0.000	0.000	0.000	0.001	0.019	0.131	0.346	0.501
5	54%	0	0	0	0	0	0	0	0	0.014	0.113	0.873
Kernel	$\Phi =$	1.85	1.77	1.63	1.45	1.25	1.00	0.83	0.841	NaN	NaN	NaN

Table A.5: The state price transition matrix, P, April 27, 2011. M1

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.172	3.734	4.259	3.701	2.628	2.046	0.427	8.256	-0.211	-2.326	-1.482
-4	0.706	-2.424	-1.275	-1.775	3.636	4.888	1.050	0.268	6.394	-3.061	-3.820	0.868
-3	0.771	1.188	0.161	0.553	-1.915	-2.883	-0.770	0.129	-2.346	0.995	0.968	-0.337
-2	0.841	0.565	3.995	4.545	2.566	0.345	2.610	-0.205	6.204	0.502	-1.565	-1.643
-1	0.917	0.180	0.006	0.270	-0.242	-0.229	-0.127	0.848	-0.266	-0.045	-0.144	-0.064
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	-0.148	-0.605	-0.808	-0.1185	0.397	-0.072	0.173	-0.690	-0.040	0.385	0.316
2	1.189	0.173	-1.552	-1.747	-1.491	-0.762	-0.991	0.016	-3.175	0.343	1.239	0.658
3	1.296	0.139	-0.661	-0.877	-0.670	-0.501	-0.445	-0.261	-1.755	0.871	0.792	0.301
4	1.413	0.479	-1.469	-1.504	-1.923	-1.309	-0.986	-0.121	-4.292	0.500	2.324	0.687
5	1.541	-2.296	0.632	-1.103	4.778	5.134	0.455	-0.535	8.757	-3.989	-4.426	-0.572

Table A.6: The natural probability transition matrix, F, estimated from P. M1

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%											
-4	-29%											
-3	-23%											
-2	-16%											
-1	-8%						not					
0	0%											
1	9%						applicable					
2	19%											
3	30%											
4	41%											
5	54%											

Table A.7: The state price transition matrix, P, April 27, 2011. M2

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.034	0.041	0.041	0.049	0.056	0.000	0.000	0.000	0.099	0.011	0.000
-4	0.706	0.062	0.035	0.031	0.093	0.084	0.002	0.000	0.000	0.005	0.002	0.000
-3	0.771	0.068	0.047	0.013	0.151	0.087	0.000	0.000	0.627	0.007	0.005	0.000
-2	0.841	0.046	0.047	0.023	0.081	0.098	0.031	0.000	0.001	0.014	0.000	0.000
-1	0.917	0.007	0.016	0.150	0.008	0.134	0.007	0.482	0.000	0.000	0.000	0.000
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	0.052	0.034	0.019	0.121	0.176	0.256	0.376	0.192	0.001	0.000	0.000
2	1.189	0.162	0.057	0.108	0.076	0.047	0.081	0.054	0.552	0.149	0.000	0.000
3	1.296	0.197	0.079	0.037	0.052	0.024	0.000	0.000	0.000	0.722	0.156	0.000
4	1.413	0.239	0.001	0.092	0.002	0.001	0.000	0.000	0.000	0.000	0.814	0.050
5	1.541	0.006	0.678	0.006	0.144	0.000	0.000	0.004	0.000	0.016	0.202	0.004

Table A.8: The natural probability transition matrix, F, estimated from P. M2

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.035	0.027	0.165	0.041	0.208	0.000	0.000	0.000	0.478	0.046	0.000
-4	-29%	0.010	0.036	0.196	0.121	0.489	0.010	0.000	0.000	0.036	0.013	0.000
-3	-23%	0.018	0.008	0.013	0.032	0.081	0.000	0.000	0.835	0.008	0.006	0.000
-2	-16%	0.059	0.038	0.114	0.083	0.450	0.165	0.000	0.009	0.081	0.000	0.000
-1	-8%	0.002	0.003	0.169	0.002	0.137	0.009	0.678	0.000	0.000	0.000	0.000
0	0%	0.001	0.001	0.017	0.009	0.143	0.491	0.330	0.009	0	0	0
1	9%	0.011	0.005	0.016	0.020	0.132	0.226	0.385	0.205	0.001	0.000	0.000
2	19%	0.032	0.007	0.085	0.012	0.034	0.068	0.053	0.567	0.141	0.000	0.000
3	30%	0.043	0.011	0.032	0.009	0.019	0.000	0.000	0.000	0.741	0.146	0.000
4	41%	0.057	0.000	0.087	0.000	0.001	0.000	0.000	0.000	0.000	0.835	0.019
5	54%	0.004	0.279	0.016	0.075	0.000	0.000	0.013	0.000	0.047	0.560	0.005
Kernel	$\Phi =$	4.137	6.491	1.045	5.126	1.148	0.974	0.837	0.806	NaN	NaN	NaN

Table A.9: The state price transition matrix, P, April 27, 2011. M3

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.718	0.258	0.258	0.117	0.117	0	0	0	0	0	0
-4	0.706	0	0	0.044	0.044	0.146	0	0	0	0	0	0
-3	0.771	0	0	0	0.310	0.014	0.014	0	0	0	0	0
-2	0.841	0.210	0.168	0.168	0.168	0.168	0.108	0	0	0	0	0
-1	0.917	0.041	0.041	0.073	0.073	0.073	0.073	0.479	0	0	0	0
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	0.002	0.002	0.036	0.060	0.200	0.210	0.377	0.234	0.004	0	0
2	1.189	0	0	0.019	0.026	0.026	0.055	0.055	0.692	0.183	0	0
3	1.296	0	0	0	0	0	0	0 0.013	0.736	0.327	0.008	
4	1.413	0.098	0.012	0.012	0	0	0	0	0	0	0	0
5	1.541	0	0	0	0	0	0	0	0.967	0.967	11.770	0.852

Table A.10: The natural probability transition matrix, F, estimated from P. M3

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.731	0.041	0.058	0.075	0.095	0	0	0	0	0	0
-4	-29%	0	0	0.063	0.179	0.758	0	0	0	0	0	0
-3	-23%	0	0	0	0.891	0.049	0.060	0	0	0	0	0
-2	-16%	0.340	0.042	0.061	0.171	0.216	0.170	0	0	0	0	0
-1	-8%	0.053	0.008	0.021	0.059	0.075	0.091	0.694	0	0	0	0
0	0%	0.005	0.001	0.004	0.030	0.137	0.487	0.328	0.009	0	0	0
1	9%	0.001	0.000	0.007	0.034	0.149	0.183	0.383	0.245	0.003	0	0
2	19%	0	0	0.004	0.014	0.018	0.047	0.055	0.704	0.159	0	0
3	30%	0	0	0	0	0	0	0	0.015	0.749	0.035	0.201
4	41%	0.955	0.018	0.026	0	0	0	0	0	0	0	0
5	54%	0	0	0	0	0	0	0	0.045	0.038	0.049	0.867
Kernel	$\Phi =$	0.954	6.126	4.281	1.518	1.198	0.982	0.842	0.820	NaN	NaN	NaN

Table A.11: The state price transition matrix, P, April 27, 2011. M4

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.869	0.285	0.233	0.071	0.056	0	0	0	0	0	0
-4	0.706	0.063	0.063	0.063	0.063	0.063	0	0	0	0	0	0
-3	0.771	0	0	0.216	0.216	0.216	0.116	0	0	0	0	0
-2	0.841	0	0	0	0.212	0.018	0	0	0	0	0	0
-1	0.917	0.061	0.061	0.128	0.128	0.210	0.210	0.210	0	0	0	0
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	0.017	0.017	0.017	0.029	0.137	0.148	0.520	0.236	0.004	0.001	0
2	1.189	0	0.017	0.017	0.053	0.053	0.053	0.053	0.685	0.181	0.003	0
3	1.296	0	0	0	0	0	0	0.018	0.018	0.733	0.157	0
4	1.413	0	0	0	0	0	0	0	0.047	0.047	0.829	0.051
5	1.541	0	0	0	0	0	0	0	0	0	0	0

Table A.12: The natural probability transition matrix, F, estimated from P. M4

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.896	0.033	0.047	0.001	0.024	0	0	0	0	0	0
-4	-29%	0.579	0.065	0.113	0.006	0.238	0	0	0	0	0	0
-3	-23%	0	0	0.223	0.011	0.471	0.295	0	0	0	0	0
-2	-16%	0	0	0	0.219	0.782	0	0	0	0	0	0
-1	-8%	0.154	0.017	0.062	0.003	0.216	0.252	0.296	0	0	0	0
0	0%	0.011	0.002	0.008	0.001	0.145	0.493	0.334	0.007	0	0	0
1	9%	0.032	0.004	0.006	0.001	0.103	0.130	0.537	0.187	0.002	0.000	0
2	19%	0	0.005	0.008	0.001	0.052	0.061	0.071	0.707	0.095	0.002	0
3	30%	0	0	0	0	0	0	0.047	0.036	0.756	0.160	0
4	41%	0	0	0	0	0	0	0	0.096	0.049	0.855	0
5	54%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Kernel	$\Phi =$	0.465	4.171	2.393	47.558	1.130	0.970	0.825	1.076	NaN	NaN	NaN

Table A.13: The state price transition matrix, P, April 27, 2011. M5

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.873	0.277	0.259	0	0	0	0	0	0	0	0
-4	0.706	0	0	0.028	0	0.375	0	0	0	0	0	0
-3	0.771	0.026	0	0	0.196	0	0	0	0	0	0	0
-2	0.841	0	0.148	0.145	0.106	0	0.068	0	0	0	0	0
-1	0.917	0.097	0.026	0.147	0	0.129	0	0.481	0	0	0	0
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	0	0.012	0	0.122	0.194	0.255	0.376	0.235	0.004	0.001	0
2	1.189	0	0	0.028	0.081	0.057	0.072	0.054	0.692	0.180	0.003	0
3	1.296	0	0	0	0.083	0.031	0	0	0	0.740	0.157	0
4	1.413	0	0	0	0	0	0	0	0.044	0.036	0.829	0.051
5	1.541	0	0	0	0	0	0	0	0.634	0	0	0

Table A.14: The natural probability transition matrix, F, estimated from P. M5

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.899	0.087	0.015	0	0	0	0	0	0	0	0
-4	-29%	0	0	0.005	0	0.995	0	0	0	0	0	0
-3	-23%	0.495	0	0	0.5047	0	0	0	0	0	0	0
-2	-16%	0	0.343	0.060	0.110	0	0.488	0	0	0	0	0
-1	-8%	0.127	0.010	0.010	0	0.133	0	0.720	0	0	0	0
0	0%	0.005	0.002	0.001	0.007	0.141	0.492	0.343	0.009	0	0	0
1	9%	0	0.003	0	0.015	0.137	0.217	0.387	0.237	0.003	0.000	0
2	19%	0	0	0.001	0.010	0.041	0.062	0.057	0.713	0.113	0.002	0
3	30%	0	0	0	0.017	0.037	0	0	0	0.762	0.185	0
4	41%	0	0	0	0	0	0	0	0.065	0.033	0.854	0.049
5	54%	0	0	0	0	0	0	0	1.000	0	0	0
Kernel	$\Phi =$	0.919	3.008	16.960	6.771	1.168	0.971	0.804	0.8180	NaN	NaN	NaN

Table A.15: The state price transition matrix, P, April 27, 2011. M6

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.626	0.262	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-4	0.706	0.011	0.626	0.262	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-3	0.771	0.073	0.011	0.626	0.262	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-2	0.841	0.000	0.073	0.011	0.626	0.262	0.000	0.000	0.000	0.000	0.000	0.000
-1	0.917	0.000	0.000	0.0732	0.011	0.626	0.262	0.000	0.000	0.000	0.000	0.000
0	1	0.000	0.000	0.000	0.073	0.011	0.626	0.262	0.000	0.000	0.000	0.000
1	1.090	0.027	0.000	0.000	0.000	0.073	0.011	0.626	0.262	0.000	0.000	0.000
2	1.189	0.000	0.027	0.000	0.000	0.000	0.073	0.011	0.626	0.262	0.000	0.000
3	1.296	0.000	0.000	0.027	0.000	0.000	0.000	0.073	0.011	0.626	0.262	0.000
4	1.413	0.457	0.000	0.000	0.027	0.000	0.000	0.000	0.073	0.011	0.626	0.262
5	1.541	0.006	0.457	0.000	0.000	0.027	0.000	0.000	0.000	0.073	0.011	0.626

Table A.16: The natural probability transition matrix, F, estimated from P. M6

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.633	0.367	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-4	-29%	0.008	0.633	0.359	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-3	-23%	0.040	0.008	0.633	0.319	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-2	-16%	0.000	0.046	0.009	0.633	0.312	0.000	0.000	0.000	0.000	0.000	0.000
-1	-8%	0.000	0.000	0.052	0.009	0.633	0.306	0.000	0.000	0.000	0.000	0.000
0	0%	0.000	0.000	0.000	0.055	0.009	0.633	0.303	0.000	0.000	0.000	0.000
1	9%	0.008	0.000	0.000	0.000	0.056	0.009	0.633	0.293	0.000	0.000	0.000
2	19%	0.000	0.010	0.000	0.000	0.000	0.059	0.010	0.633	0.288	0.000	0.000
3	30%	0.000	0.000	0.012	0.000	0.000	0.000	0.062	0.010	0.633	0.283	0.000
4	41%	0.104	0.000	0.000	0.014	0.000	0.000	0.000	0.064	0.010	0.633	0.175
5	54%	0.002	0.218	0.000	0.000	0.025	0.000	0.000	0.000	0.106	0.016	0.633
Kernel	$\Phi =$	3.010	2.179	1.611	1.339	1.138	0.989	0.866	0.785	0.722	0.678	1.029

Table A.17: The state price transition matrix, P, April 27, 2011. M7

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.478	0.276	0.007	0	0	0	0	0	0.028	0	0
-4	0.706	0.164	0.478	0.276	0.007	0	0	0	0	0	0.028	0
-3	0.771	0.045	0.164	0.478	0.276	0.007	0	0	0	0	0	0.028
-2	0.841	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0	0	0
-1	0.917	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0	0
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	0.005	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0
2	1.189	0	0.005	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0
3	1.296	0.0129	0	0.005	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007
4	1.413	0	0.013	0	0.005	0.005	0.007	0.018	0.045	0.164	0.478	0.276
5	1.541	0.609	0	0.013	0	0.005	0.005	0.007	0.018	0.045	0.164	0.478

Table A.18: The natural probability transition matrix, F, estimated from P. M7

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.484	0.421	0.013	0	0	0	0	0	0.082	0	0
-4	-29%	0.110	0.484	0.344	0.010	0	0	0	0	0	0.052	0
-3	-23%	0.025	0.135	0.484	0.308	0.009	0	0	0	0	0	0.039
-2	-16%	0.009	0.034	0.151	0.484	0.314	0.009	0	0	0	0	0
-1	-8%	0.003	0.012	0.037	0.148	0.484	0.308	0.008	0	0	0	0
0	0%	0.002	0.004	0.013	0.037	0.151	0.484	0.301	0.008	0	0	0
1	9%	0.002	0.003	0.005	0.014	0.039	0.154	0.484	0.292	0.008	0	0
2	19%	0	0.003	0.003	0.005	0.015	0.041	0.159	0.484	0.284	0.007	0
3	30%	0.005	0	0.003	0.004	0.006	0.016	0.043	0.164	0.484	0.270	0.006
4	41%	0	0.007	0	0.004	0.004	0.006	0.018	0.046	0.172	0.484	0.259
5	54%	0.239	0	0.009	0	0.005	0.005	0.008	0.020	0.051	0.180	0.484
Kernel	$\Phi =$	2.497	1.657	1.347	1.222	1.087	0.987	0.918	0.878	NaN	NaN	NaN

Table A.19: The state price transition matrix, P, April 27, 2011. M8

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.567	0.240	0.000	0.000	0.000	0.041	0.000	0.000	0.000	0.000	0.000
-4	0.706	0.149	0.567	0.240	0.000	0.000	0.000	0.041	0.000	0.000	0.000	0.000
-3	0.771	0.000	0.149	0.567	0.240	0.000	0.000	0.000	0.0407	0.000	0.000	0.000
-2	0.841	0.000	0.000	0.149	0.567	0.240	0.000	0.000	0.000	0.041	0.000	0.000
-1	0.917	0.005	0.000	0.000	0.149	0.567	0.240	0.000	0.000	0.000	0.041	0.000
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	0.022	0.000	0.005	0.000	0.000	0.149	0.567	0.240	0.000	0.000	0.000
2	1.189	0.012	0.022	0.000	0.005	0.000	0.000	0.149	0.567	0.240	0.000	0.000
3	1.296	0.000	0.012	0.022	0.000	0.005	0.000	0.000	0.149	0.567	0.240	0.000
4	1.413	0.548	0.000	0.012	0.022	0.000	0.005	0.000	0.000	0.149	0.567	0.240
5	1.541	0.000	0.548	0.000	0.012	0.022	0.000	0.005	0.000	0.000	0.149	0.567

Table A.20: The natural probability transition matrix, F, estimated from P. M8

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.558	0.359	0.000	0.000	0.000	0.083	0.000	0.000	0.000	0.000	0.000
-4	-29%	0.096	0.558	0.289	0.000	0.000	0.000	0.057	0.000	0.000	0.000	0.000
-3	-23%	0.000	0.119	0.558	0.266	0.000	0.000	0.000	0.057	0.000	0.000	0.000
-2	-16%	0.000	0.000	0.130	0.558	0.248	0.000	0.000	0.000	0.064	0.000	0.000
-1	-8%	0.002	0.000	0.000	0.139	0.558	0.222	0.000	0.000	0.000	0.079	0.000
0	0%	0.002	0.005	0.016	0.045	0.172	0.470	0.282	0.009	0	0	0
1	9%	0.010	0.000	0.004	0.000	0.000	0.141	0.558	0.287	0.000	0.000	0.000
2	19%	0.005	0.012	0.000	0.004	0.000	0.000	0.120	0.558	0.301	0.000	0.000
3	30%	0.000	0.005	0.012	0.000	0.003	0.000	0.000	0.115	0.558	0.307	0.000
4	41%	0.124	0.000	0.005	0.010	0.000	0.002	0.000	0.000	0.112	0.558	0.188
5	54%	0.000	0.235	0.000	0.007	0.014	0.000	0.003	0.000	0.000	0.183	0.558
Kernel	$\Phi =$	2.115	1.392	1.136	1.007	0.957	1.017	0.980	0.805	NaN	NaN	NaN

Table A.21: The state price transition matrix, P, April 27, 2011. M9

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	0.649	0.706	0.771	0.841	0.917	1	1.090	1.189	1.296	1.413	1.541
-5	0.649	0.361	0.289	0.014	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-4	0.706	0.057	0.361	0.289	0.014	0.023	0.000	0.000	0.000	0.000	0.000	0.000
-3	0.771	0.028	0.057	0.361	0.289	0.014	0.023	0.000	0.000	0.000	0.000	0.000
-2	0.841	0.000	0.028	0.057	0.361	0.289	0.014	0.023	0.000	0.000	0.000	0.000
-1	0.917	0.180	0.000	0.028	0.057	0.361	0.289	0.014	0.023	0.000	0.000	0.000
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0	0	0
1	1.090	0.000	0.000	0.051	0.027	0.036	0.110	0.592	0.249	0.000	0.000	0.000
2	1.189	0.032	0.000	0.000	0.051	0.027	0.036	0.110	0.592	0.249	0.000	0.000
3	1.296	0.000	0.032	0.000	0.000	0.051	0.027	0.036	0.110	0.592	0.249	0.000
4	1.413	0.556	0.000	0.032	0.000	0.000	0.051	0.027	0.036	0.110	0.592	0.249
5	1.541	0.006	0.556	0.000	0.032	0.000	0.000	0.051	0.027	0.036	0.110	0.592

Table A.22: The natural probability transition matrix, F, estimated from P. M9

	$\#\sigma$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$\#\sigma$	S_T/S_0	-35%	-29%	-23%	-16%	-8%	0 %	9%	19%	30%	41%	54%
-5	-35%	0.363	0.497	0.040	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-4	-29%	0.033	0.363	0.472	0.037	0.095	0.000	0.000	0.000	0.000	0.000	0.000
-3	-23%	0.010	0.035	0.363	0.460	0.037	0.095	0.000	0.000	0.000	0.000	0.000
-2	-16%	0.000	0.011	0.036	0.363	0.470	0.038	0.083	0.000	0.000	0.000	0.000
-1	-8%	0.025	0.000	0.011	0.035	0.363	0.473	0.033	0.061	0.000	0.000	0.000
0	0%	0.000	0.001	0.004	0.017	0.101	0.480	0.385	0.012	0	0	0
1	9%	0.000	0.000	0.009	0.007	0.016	0.079	0.594	0.294	0.000	0.000	0.000
2	19%	0.002	0.000	0.000	0.012	0.010	0.022	0.093	0.594	0.267	0.000	0.000
3	30%	0.000	0.003	0.000	0.000	0.018	0.015	0.029	0.103	0.594	0.238	0.000
4	41%	0.029	0.000	0.005	0.000	0.000	0.031	0.022	0.036	0.115	0.594	0.168
5	54%	0.000	0.073	0.000	0.011	0.000	0.000	0.063	0.039	0.057	0.163	0.594
Kernel	$\Phi =$	11.681	6.816	4.185	2.635	1.626	0.996	0.717	0.608	NaN	NaN	NaN

Bibliography

[1] Y. Ait-Sahalia and A. W. Lo. Nonparametric estimation of state-price densities implicit in financial asset prices. *J. Finance*, 53:499–547, 1998.

[2] F. Black and M. S. Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.*, 81:637–654, 1973.

[3] D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *J. Bus.*, 51:621–651, 1978.

[4] R. Bro and N. D. Sidiropoulos. Least squares algorithms under unimodality and non-negativity constraints. *J. Chemometrics*, 12:223–247, 1998.

[5] P. Carr and J. Yu. Risk, Return and Ross Recovery. *J. Deriv.*, 20:38–59, 2012.

[6] P. H. Dybvig and S. A. Ross. Arbitrage, state prices and portfolio theory. *Handbook of the Economics of Finance*, Elsevier, 2003.

[7] S. Figlewski. Estimating the implied risk neutral density for the U.S. market portfolio. *Volatility and Time Series Economics: Essays in Honor of Robert F. Angle*, Oxford University Press, 2008.

[8] N. Gisiger. Risk-neutral probabilities explained. `http://ssrn.com/abstract=1395390`, 2010. [Online; last accessed 10-May-2013].

[9] G. R. Grimmett and D. R. Stirzaker. *Probability and Random Processes*. Oxford University Press, 3rd edition, 2001.

[10] J. C. Jackwerth. Recovering risk aversion from option prices and realized returns. *Rev. Financ. Stud.*, 13:4333–4451, 2000.

[11] J. C. Jackwerth and M. Rubinstein. Recovering probability distributions from option prices. *J. Finance*, 51:1611–1631, 1996.

[12] E. Kreyszig. *Advanced Engineering Mathematics*. John Wiley & Sons, 9th edition, 2010.

[13] R. C. Merton. Theory of rational option pricing. *Bell J. Econ.*, 4:141–183, 1973.

[14] Carl Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, 2000.

[15] S. Ross. The Recovery theorem. *J. Finance*, (Forthcoming) 2013.

[16] P. A. Samuelson. A note on measurement of utility. *Rev. Econ. Stud.*, 4:155–161, 1937.

[17] N. N. Taleb. *The Black Swan*. Random House Inc., 2008.

[18] A. Yatchew and W. Hardle. Nonparametric state price density estimation using constrained least squares and the bootstrap. *J. Econometrics*, 133:579–599, 2006.

[19] M. Yuan. State price density estimation via nonparametric mixtures. *Ann. Appl. Stat.*, 3:963–984, 2009.