Recovery Theorem with a Multivariate Markov Chain*

Anthony Sanford†

September 5, 2017

For the latest version, **click here**

Draft. Please do not circulate or distribute without permission from the author.

Abstract

In this paper, I redefine the prices derived in Ross' Recovery Theorem (Ross, 2015) using a multivariate Markov chain rather than a univariate one. I employ a mixture transition distribution where the proposed states depend on the level of the S&P 500 index and its options' implied volatilities. I include volatility because the transition path between states depends on the propensity of an underlying asset to vary. An asset that is highly volatile is more likely to transition to a far-away state. These higher transition probabilities should lead to higher state prices. The multivariate method improves upon the univariate RT because the latter does not include the volatility inherent in the state transition, which makes its derived prices less precise. The multivariate RT produces forecast results far superior to the univariate RT. Using quarterly forecasts for the 1996-2015 period, the out-of-sample R-square of the RT increases from around 12% to 30%. Moreover, using simulated data, I show that including the implied volatility in the multivariate Markov chain more closely captures the inherent risk in business cycles.

*Financial support for this research was provided by the Fonds de recherche du Québec – Société et culture (FRQSC). Computing resources were provided by the University of Washington's Center for Studies in Demography and Ecology (CSDE). All errors are my own.

†Department of Economics, University of Washington. Email: sanfoan@uw.edu

Contents

1 Introduction

Ross's (2015) Recovery Theorem (RT) is a breakthrough in asset price forecasting. Using the RT, we can obtain the market's best estimate of future expected returns and risk aversions by separating the components of state prices (the discount rate, pricing kernel, and natural probability distribution). Not only does it allow us to use option prices to obtain an out-of-sample non-parameterized expected future distribution of an option's underlying asset, but it is one of the best asset forecasting models available today. However, it has certain shortcomings that this paper aims to address.

This paper's theoretical contribution is that it changes the original univariate model to a multivariate one. The original RT derived contingent state prices using a simple constrained linear regression, which assumed that the probability of transitioning to a new state was dependent on the previous state. But for option prices to truly reflect the conditional variance of the underlying asset (Engle and Mustafa, 1992), the transition path should control for volatility (Page et al., 2006). Controlling for the volatility in the transition path becomes even more important because of the nature of contingent state prices. These prices are not observed in the market. They are a function of observed state prices, which are used to infer prices for states that have not occurred. If contingent state prices were actually observed, they would already contain all available market information, including volatility. However, since we only observe state prices for the current state, it is crucial to derive the contingent state prices contingent on the observed underlying volatilities. Thus, including volatility in the derivation of the contingent state prices is critical to the proper specification of the Recovery Theorem.

One of the key assumptions of the RT is that markets are complete. In reality, markets are not complete. To construct state prices that are complete and behave normally, it is necessary for the data to be as detailed as possible. The original RT was tested empirically using over-the-counter (OTC) data, which is richer[1] than publicly traded

[1]The notional amount for outstanding OTC equity-linked options is estimated to be $4.244 trillion while it is estimated to be $1.972 trillion for exchange traded options BIS (2012).

options data. However, it is unlikely that Ross's OTC dataset includes, for example, options with strike prices at every $1 interval. Moreover, contingent state prices require that we assume time homogeneity. To make this assumption, we must extrapolate option data based on time-to-expiration. I developed a methodology (see companion paper (Sanford, 2016b)) where I extrapolate readily available exchange traded option data on both the strike price and time-to-maturity dimensions by expanding on methods proposed by Figlewski (2008) and Chen (2011). This methodology makes the RT usable in any circumstance where we have sufficient data to estimate smooth splines.

I test the RT both in univariate and multivariate Markov chain settings. The forecast results indicate that the multivariate Markov chain produces results far superior to the univariate RT. Using quarterly forecasts (updated monthly) for the 1996-2015 period, the out-of-sample R-square of the RT increases from around 12% to 30%. Empirically, this paper constitutes one of the first exhaustive analyses of Ross's Recovery Theorem. This paper also provides an intuitive framework by which to understand both the univariate and the multivariate RT.

The paper is divided into four main sections. Section 2 explains the univariate and multivariate RTs, and discusses the steps required to implement the theorem. It also walks through a simple numerical example for both the original univariate and the proposed multivariate RT. Section 4 introduces the data and presents the results. Finally, section 5 explores possible extensions and concludes.

2 Model

The RT's ultimate goal is to obtain the natural probability distribution for asset returns (in this case, equity returns). It accomplishes this goal by deriving state prices using equity options. Using these state prices, we can then disentangle the discount rate, the risk-aversion parameter, and, ultimately, the natural probability distribution without making any parametric or utility function assumptions. I break down the RT into four steps:

1. construct the state prices,

2. construct the contingent state price matrix,

3. use the Perron-Frobenius (Meyer, 2000) theorem to extract the natural probability matrix, and

4. produce the natural marginal distributions, which can be used to obtain the recovered statistics (of which the recovered expected return and expected volatility are of particular interest).

To facilitate comparison, I adopt the same terminology and notation as Ross wherever possible. I do not present all of the proofs from the original RT since those can be found in Ross's paper. I limit the proofs in this paper to those that are new or crucial to the understanding of the model.

2.1 The Recovery Theorem

Financial markets price assets as the present value of all future cash flows (Cochrane, 2009). However, if we are referring to risky assets, as is the case in this paper, prices are subject to adjustments since future payoffs are not guaranteed and, by extension, are considered risky. We call this adjustment for the riskiness of the asset price the risk premium. The risk premium is defined as a function of the risk aversion and the overall level of risk of the asset being priced. We can refer to the price of an asset using the following equation (Cochrane, 2009):

$$p_t = E_t(m_{t+1}x_{t+1}) \tag{1}$$

where p_t is the price of an asset at some time t, E_t is the expectation operator, m_{t+1} is the stochastic discount factor, and x_{t+1} is the future cash flow of the asset. The variable m_{t+1} in equation 1 is what gives us the risk premium because it is the adjustment to the price of an asset that makes it worthwhile for investors to purchase that asset

given its level of risk. Part of the problem in pricing equities, however, is in defining the stochastic discount factor. In markets like the bond market, we can derive the forward rates. We obtain forward rates by comparing the yields of bonds with different expirations, which allows us to obtain the market's estimate of the stochastic discount factor. The same cannot be done with the equity market. So how can we estimate the risk premium? As Ross (2015) notes, we currently estimate the risk premium for equity markets by relying on historical returns or by using opinion polls. Historical returns assume that the past estimate of the risk premium is a good indicator of the future risk premium while opinion polls assume that the opinions of the analysts being polled reflect the entire market's overall sentiment. Both of these methodologies are flawed.

In an effort to address these issues, Ross (2015) proposes to use options. Options, like forward rates, are forward-looking instruments with varying maturities. Hence, there is hope that we may use these securities to estimate the risk premium. That being said, option prices themselves do not explicitly depend on, or allow us to solve for, the risk premium. This is the question that motivates the original Recovery Theorem: how can we use option prices to obtain the risk premium? The RT provides a framework through which we can use options to estimate state prices, which then allow us to estimate the underlying asset's risk premium.

2.2 Estimating state prices (S)

Ross proposes that the starting point in deriving the equity risk premium is to obtain state prices from option prices. Why do we need state prices? We want a security that can be defined as a function of a pricing kernel and the true (or, as Ross calls them, "natural") probabilities. This is in essence a forward rate: a function of a pricing kernel and a probability. However, forward rates are not naturally found in equity markets, so we use option prices instead. Recall the definition for forward rates: today's rate for an asset that has a guaranteed payoff at some future point. Can these types of

securities be obtained using equity options? An option can be defined as a function of the discount rate, the risk aversion parameter, and the probability of downside risk. However, we are not looking for an asset that is only a function of the left side of the returns distribution. Instead, we can construct a portfolio of options. We are going to call these portfolios "state prices." Formally, state prices correspond to the price of a security at some initial time, t_0, such that, at some future time T, the security pays a pre-specified amount (normalized to \$1) if the market is at a pre-specified state of the world and pays nothing otherwise. For example, assuming that the level of the S&P 500 today is 1,000, a state price would be the price of an asset that pays you 1\$ in, say, three months if the level of the S&P 500 is 1,500 at that time. The problem is that this type of security is not readily traded. Breeden and Litzenberger (1978) produce a method to derive state prices, beginning with the continuous time Black-Scholes-Merton equation (Black and Scholes, 1973; Merton, 1973) as follows:

$$Call(K,T) = \int_0^\infty [S_{t,p} - K]^+ p(S_{t,p}, T) dS_{t,p} = \int_K^\infty p(S_{t,p}, T) dS_{t,p}, \tag{2}$$

where $Call(K,T)$ is today's price for a call option with a strike price K and time-to-maturity T. Taking the second derivative with respect to strike price K gives the following result in continuous time:

$$s(K,T) = Call''(K,T) \tag{3}$$

which is Breeden and Litzenberger's (1978) result. In discrete time, we can estimate equation 3 using a butterfly spread. A butterfly spread is a portfolio of three call options: buy a call option at strike price K_1, sell two call options at strike price K_2, and buy a call option at strike price K_3. Mathematically, this corresponds to the following equation:

$$s(K,T) \approx -Call_{K_1} + 2Call_{K_2} - Call_{K_3} \tag{4}$$

which, once standardized, gives a guaranteed payoff of \$1 at expiration T if the market ends at K_2. Hence, we have defined and derived state prices. These state prices are the foundation of the Recovery Theorem.

Knowing the state price of a single state is not enough to solve for the natural probability distribution. We need m equations but only have one set of equations, which implies that we cannot solve the system. However, if we knew the state prices for a complete set of states (m states in this example), we would have m equations and could solve the system of equations (see appendix A for more details). These m equations will be obtained from the estimation of the contingent state prices.

Numerical example Before moving on to the derivation of the contingent state prices, let me introduce a simple numerical example that will be used throughout this paper. The goal of this example is twofold. First, it will provide the intuition behind the RT and its mechanics. Second, the example will show the importance of incorporating volatility in the derivation of contingent state prices (see section 2.3). The example will illustrate that a distribution that has a larger standard deviation will have a probability distribution function (pdf) that is wider (i.e., more probabilities in the tails) than one with a smaller standard deviation. As a result, the probability of a given path is estimated more accurately when we consider volatility as a state variable in the model. This is especially true when we consider the probability of transitioning between states that are far away (e.g., the S&P 500 transitioning between a level of, say, 1,000 to a level of, say, 2,000 in a three-month period). These large movements are more likely to occur (higher probabilities) if the volatility is higher than if it is lower.

To begin, let us assume that we have a set of observable state prices in the economy.

In particular, let us assume that we observe the following state prices:

$$\mathbf{S} = \begin{matrix} \text{m} \\ \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \end{matrix} \begin{matrix} \\ +25\% \\ \text{-}10\% \end{matrix}$$

where S represents an observable vector of state prices, m represents the time-step which coincides with the expiration of the observed state prices (e.g., option time-to-maturity), and $+25\%/-10\%$ represent the possible future states.

The example is purposefully kept very simple: there are only two possible future states (25% gain or 10% loss). Let us assume that the observed state prices are set at \$0.5 for both the bearish (-10%) and bullish ($+25\%$) states. The entire system, up to this point, can be characterized using figure 1:

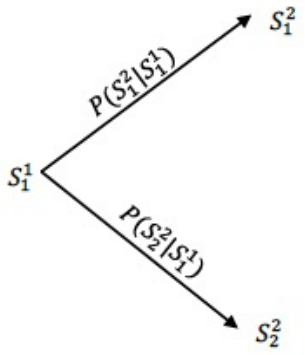

Figure 1: Generalized Setup

where S_1^1 represents the initial price or level of an underlying asset, S_1^2 and S_2^2 represent the two possible future states in our simplified world, and $P()$ represents the contingent state prices. From the state price vector, m is what controls the difference in time between S^1 and S^2. This m will also be the horizon for our forecast. For this example, I assume that the time-step is set to three months. This simple world is one where we have a current level for the S&P 500, say $S_1^1 = \$1,000$, and where the possible future outcomes could be either $S_2^2 = \$900$ or $S_1^2 = \$1,250$. The next step involves the

estimation of the contingent state prices $P()$.

2.3 Estimating contingent state prices (P)

Contingent state prices are nothing more than state prices for initial states that are not currently observed in the market. This paper distinguishes between two derivations for these prices: the univariate (or "naïve") and multivariate contingent state prices.

2.3.1 The univariate model

In equation 1, I defined state prices as a function of a pricing kernel, m, and some future payout, x. Formally, contingent state prices are defined in the exact same way as state prices with the exception that these are now for states that are not observed in the market. We can think of these as state prices for some future state, i, to some other future state, j. More intuitively, we can define the contingent state price matrix as an intermediate-step forward rate. In other words, it is the price of an asset in the future that guarantees a payoff of \$1 if the state of the world transitions from state i to state j at an intermediate time-step $t+\tau$, where $\tau > 0$. This is analogous to obtaining the forward rate at some future time-step. An intermediate time-step forward rate is the expected rate at time t_0 for rolling over a bond at some future time $t+\tau$ for a desired investment horizon that is at time T. This bond price is not known at the initial time, t_0. For example, if we assume an investment horizon of one year, we can decompose the forward rate into two six-month periods. We have the choice between investing in a one-year bond or investing in a six-month bond today and investing in another six-month bond in six months (rolling over the investment). The forward rate is thus the price at time zero (or the rate in this case) of the six-month bond that we will purchase six months from now for our total investment horizon of one year. The intuition for the contingent state price is the same. If we think about contingent state prices using the same horizons as the example for the forward rates, we have the price of a security that pays \$1 if the market starts at state i in six months and expires at

state j in 12 months. Compared to the state prices estimated in the previous section, here we are estimating state prices for state levels that are hypothetical, rather than the current state level. This understanding might seem trivial but it will be important later when I derive the multivariate Markov chain.

Before deriving the contingent state price matrix, I need to introduce an assumption that is crucial to its derivation.

Assumption 1 (Time-Homogeneity)**.** Time homogeneity implies that the contingent state price matrix, P, is not dependent on time.

Using assumption 1, Ross (2015) estimates the contingent state price matrix using the following equation:

$$s_{t+1} = s_t P,\ t = 1, ..., m-1 \tag{5}$$

$$1 \geq P \geq 0$$

where m is the number of states and P is the contingent state price matrix. Assumption 1 allows me to obtain the contingent state prices using equation 5. Time homogeneity assumes that the contingent state prices are the same regardless of which time-step we are trying to estimate.

Now that I have derived the contingent state prices, I can rewrite equation 1 as follows:

$$p_{i,j} = \phi(\theta_i, \theta_j) f_{i,j} \tag{6}$$

where $p_{i,j}$ is a contingent state price, $\phi(\theta_i, \theta_j)$ is the kernel factor, and $f_{i,j}$ is the natural probability that we are ultimately trying to derive.

Once the contingent state price matrix has been obtained, the rest of the RT is derived using the Perron-Frobenius theorem along with some matrix algebra. At this point, we have all of the necessary components to solve for the natural probability matrix. However, a question still remains: can we improve on the estimation of the contingent state prices proposed by Ross? The next section extends the derivation of the contingent state prices to a multivariate Markov chain. This Markov chain controls

for the volatility as well as the current level of the underlying asset.

Numerical example (continued) To maintain simplicity, I assume that there are only two possible hypothetical states of the world. Let us assume the following naïve system:

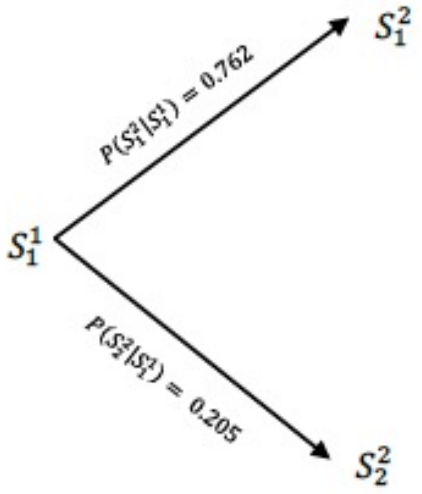

Figure 2: Univariate System One

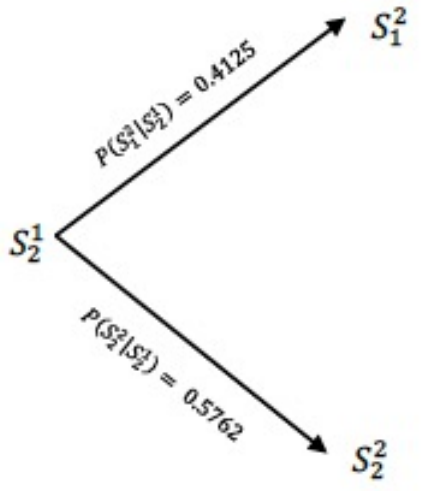

Figure 3: Univariate System Two

In figure 2, the contingent state price of staying in state one is equal to 0.762 and the price of moving to state two is equal to 0.205. In other words, in this system, the price associated with transitioning from S_1^1 to S_2^2 is 0.205. Similarly, the price associated with transitioning from S_1^1 to S_1^2 is 0.762. In the first hypothetical state, S_1^1, investors believe that the market is more likely to stay in the original state (state 1) over the next three months.

Notice that in figures 2 and 3, the contingent state prices are not dependent on anything other than the initial state for that hypothetical world (S_1^1 or S_2^1). This is the major distinction between the naïve setup of Ross and the setup proposed in this paper, and it will lead to a significant difference in the resulting expected natural distribution of returns.

2.3.2 The multivariate model

Including the volatility in the derivation of the contingent state prices removes the assumption that volatility between periods is constant. This is the major contribution of this paper. Volatilities are different depending on the state path probability that we are trying to estimate. Hence, it becomes critical to control for these different changes in volatility in the contingent state price estimation.

Let us derive the multivariate Markov chain. The general specification for the multivariate Markov chain used in this paper was first introduced by Raftery (1985) and is as follows:

$$\min_{\lambda_{i,j}} \min_{P} [[\sum \lambda_{i,j} s_t P - s_{t+1}]_P] \tag{7}$$

where it must, by definition, be the case that:

$$1 \geq P \geq 0 \text{ and } \beta \geq 0$$
$$\sum \lambda_{i,j} = 1$$

More specifically, for the purposes of this paper, I can rewrite the general specification in equation 7 to a two-variable Markov chain as follows:

$$\min_{\lambda_{i,j}} \min_{P,\beta} [[\lambda_{i,j} s_t P + (1 - \lambda_{i,j}) \Phi_t \beta - s_{t+1}]_{P,\beta}] \tag{8}$$

$$1 \geq P \geq 0 \text{ and } \beta \geq 0$$

where Φ is an additional variable necessary for a more accurate derivation of the contingent state price matrix. A simple specification of the multivariate model is to assume that the contingent state price is solely defined by the state levels, but that we need to condition on the the volatility in the regression. This implies that we estimate the contingent state prices using a multivariate Markov chain as follows:

$$s_{t+1} = s_t P + \Delta Ivol_t \beta, \quad t = 1, ..., m-1 \tag{9}$$

where $Ivol_t$ is the implied volatility state at time t. In other words, equation 9 assumes that $\lambda = 1$ in equation 8. Implied volatility is used because it is the market's best estimate of the future volatility state. Equation 9 gives us a third dimension in the Markov chain and therefore results in a matrix of size $(m-1)^3$. Theoretically, we could add more variables to the regression equation. Since I estimate the Markov chain based on 11 states, however, it is best not to add too many variables to the regression equation because there will be too few degrees of freedom to consider the resulting contingent state price matrix reliable. Moreover, and this will be discussed in greater detail in section 3, volatility acts as a proxy for the uncertainty in the macroeconomy. Hence, controlling for volatility in contingent state prices gives us a better sense of the uncertainty of future state paths.

Including the volatility into the model, I solve the following equation:

$$\min_{P,\beta} \|s_{t+1} - s_t P - \Delta vol_t \beta\|^2 \tag{10}$$

where it must, by definition, be the case that:

$$1 \geq P \geq 0 \text{ and } \beta \geq 0 \tag{11}$$

Equation 11 includes a non-negativity condition in our regression such that $P \geq 0$. This is a necessary assumption for us to apply the Perron-Frobenius theorem in the next section. The assumption also makes intuitive sense since prices, by definition, are nonnegative. The upper bound on the contingent state price ensures that there are no prices that lead to arbitrage.

Numerical example (continued) The key insight from this paper is that the naïve state space model of the RT is not accurately specified. This idea is akin to one of an omitted variable bias. There may be a multitude of variables that affect the probabilities of transitioning from one state to another, but one of the most important

variables is the volatility of the underlying asset. Volatility plays a critical role in specifying the probabilities of transitioning between states accurately. Extending the naïve example will show the impact of omitting volatility in deriving contingent state prices. Note that the time-step here is still m (three months). In this example, there are only two possible volatility states, high or low. The resulting contingent state prices are now a function of both the initial state, S_1^1, and the volatility state, σ_H or σ_L. I now assume that we have the following multivariate system:

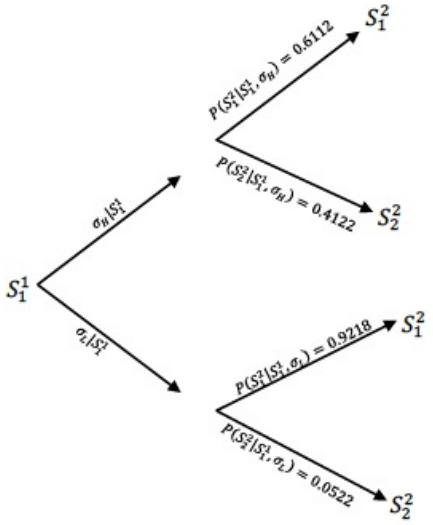

Figure 4: Multivariate System One

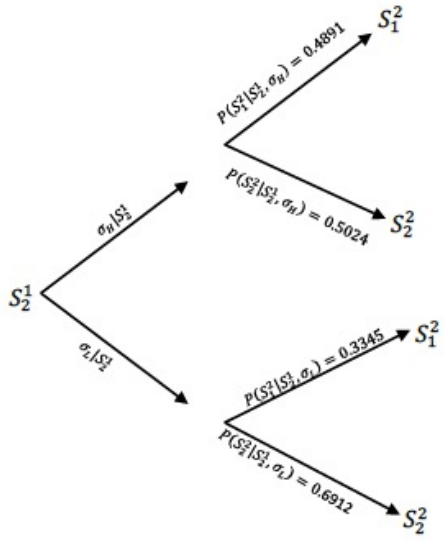

Figure 5: Multivariate System Two

where the probabilities of being in a high-volatility state are simply equal to 0.5 (in both figures). The contingent state price of S_1^2 given S_1^1 and σ_H, $P(S_1^2|S_1^1, \sigma_H)$, is 0.6112. It is best to focus on what the contingent state prices represent and their intuition. For example, $P(S_2^2|S_1^1, \sigma_L)$ is equal to 0.0522 because it is unlikely that the market will transition to a far away state given a low volatility state. By contrast, the contingent state price of moving from an initial state one to the future state two is much more likely given a high volatility state. As such, the contingent state price, $P(S_1^2|S_1^1, \sigma_H)$, is 0.4122.

Figures 2 to 5 can be summarized in matrix form as follows:

$$\mathbf{P_{naive}} = \begin{array}{c} \begin{matrix} \text{+25\%} & \text{-10\%} \end{matrix} \\ \begin{bmatrix} 0.762 & 0.205 \\ 0.4125 & 0.5762 \end{bmatrix} \end{array} \begin{matrix} \text{+25\%} \\ \text{-10\%} \end{matrix}$$

$$\mathbf{P}_{\sigma_\mathbf{H}} = \begin{array}{c} \begin{matrix} \text{+25\%} & \text{-10\%} \end{matrix} \\ \begin{bmatrix} 0.6112 & 0.4122 \\ 0.4891 & 0.5024 \end{bmatrix} \end{array} \begin{matrix} \text{+25\%} \\ \text{-10\%} \end{matrix}$$

$$\mathbf{P}_{\sigma_\mathbf{L}} = \begin{array}{c} \begin{matrix} \text{+25\%} & \text{-10\%} \end{matrix} \\ \begin{bmatrix} 0.9218 & 0.0522 \\ 0.3345 & 0.6912 \end{bmatrix} \end{array} \begin{matrix} \text{+25\%} \\ \text{-10\%} \end{matrix}$$

where P_{naive} represents the contingent state price matrix for the naïve recovery method, P_{σ_H} represents the contingent state price matrix for the high volatility state, and P_{σ_L} represents the contingent state price matrix for the low volatility state. Once we have the contingent state prices, we can apply the RT to recover a natural probability distribution and our estimate of the expected return of an asset, as shown in the next section.

2.4 Estimating the natural probability distribution (F)

At this point in the derivation, we are combining all of the elements from the previous sections to obtain the natural probability matrix. The natural probability matrix represents the market's best estimate of the future distribution of returns for the original option's underlying asset. This section describes the required theorem, assumptions, intuition, and methodologies to obtain the natural probability matrix. The first assumption is time-separable utility, which can be defined as follows:

Assumption 2 (Time-Separable Utility)**.** Time-separable utility implies that we can define the pricing kernel $\phi()$ as:

$$\phi(\theta_i, \theta_j) = \delta \frac{U'(c(\theta_j))}{U'(c(\theta_i))} \tag{12}$$

where δ is a discount rate such that $\delta \in (0, 1]$, and $U' > 0$ is the marginal utility for state j or i.

Intertemporal additive utility is assumed because it generates a transition independent kernel. It follows from the setup of an intertemporal model with a representative agent that has additive time-separable preferences. Once we have obtained the contingent state price matrix from section 2.3, we can apply Ross's RT (for proof, see Ross (2015)). Using a discrete time setup and assumption 2, I can rearrange equation 6 as:

$$U'_i p_{i,j} = \delta U'_j f_{i,j}, \tag{13}$$

where U'_i is the marginal utility such that:

$$U'_i \equiv U'(c(\theta_i)) \tag{14}$$

which can then be written in terms of the normalized kernel:

$$\phi_j \equiv \phi(\theta_1, \theta_j) = \delta(\frac{U'_j}{U'_1}) \tag{15}$$

where θ_1 is the current state. In continuous time, Ross defines the kernel as:

$$\phi(\theta_i, \theta_j) = \delta \frac{h(\theta_j)}{h(\theta_i)} \tag{16}$$

Using equation 16 and assuming transition independence, we have:

$$p(\theta_i, \theta_j) = \phi(\theta_i, \theta_j) f(\theta_i, \theta_j) = \delta \frac{h(\theta_j)}{h(\theta_i)} f(\theta_i, \theta_j) \tag{17}$$

where $h(\theta) = U'(c(\theta))$, and $p(\theta_i, \theta_j)$ is the state price transition function that was derived in section 2.3. From there, the objective is to solve the unknowns: the natural probability transition function $f(\theta_i, \theta_j)$, the kernel $\phi(\theta_i, \theta_j) = \delta \frac{h(\theta_j)}{h(\theta_i)}$, and the discount rate δ. Back to the discrete time specification, we can rewrite equation 17 in matrix form as:

$$DP = \delta FD \tag{18}$$

where P is the m x m state price matrix defined in section 2.2, F is the m x m matrix that we are calling the natural probabilities and is the matrix of interest for this section, and D is the diagonal matrix of undiscounted kernels or a diagonal of marginal rates of substitution as follows:

$$D = \frac{1}{U_1'} \begin{bmatrix} U_1' & 0 & 0 \\ 0 & U_i' & 0 \\ 0 & 0 & U_m' \end{bmatrix} = \begin{bmatrix} \phi_1 & 0 & 0 \\ 0 & \phi_i & 0 \\ 0 & 0 & \phi_m \end{bmatrix} \frac{1}{\delta} \tag{19}$$

Rearranging equation 18, we get:

$$F = \frac{1}{\delta} DPD^{-1} \tag{20}$$

We obtained P in section 2.3, so now D must be estimated. Up to this point, the RT has not provided us with additional insight into disentangling the discount rate, pricing kernel, and natural probability distribution because there were not enough variables and equations to solve our system of equations. The key, however, is to notice that F is a stochastic matrix which, by definition, implies that the rows of F are transition probabilities and so they must sum to 1. Hence, we have the following equation:

$$Fe = e \tag{21}$$

where e is simply a vector of ones. Substituting equation 21 into equation 20, we

obtain:

$$Fe = \frac{1}{\delta} DPD^{-1}e = e \tag{22}$$

and if we define $z \equiv D^{-1}e$, we can rewrite equation 22 as:

$$Pz = \delta z \tag{23}$$

This still does not allow us to solve for D. However, we can make some assumptions about P that will allow us to use the Perron-Frobenius Theorem (Meyer, 2000). Namely, we can assume that the option prices have no arbitrage opportunities (which, by definition, must be the case). No arbitrage implies that the contingent state price matrix will be nonnegative and less than one. Prices are, by definition, nonnegative, which was specified in the derivation of the contingent state price matrix in section 2.3. The second necessary assumption is that the matrix P be irreducible. A matrix is said to be irreducible if we can reach any state in k-steps. As Ross (2015) argues, even if some of the prices in P correspond to a transition probability equal to zero, it should still be possible to reach the desired state via an intermediary state (or states). As such, since P is nonnegative and irreducible, we can apply the Perron-Frobenius Theorem (Meyer, 2000), which states that all nonnegative and irreducible matrices have a unique positive characteristic root (eigenvector) z, and a Perron root δ. This allows us to solve for D, which we can introduce in the natural probability distribution equation:

$$F = \frac{1}{\delta} DPD^{-1} \tag{24}$$

The previous paragraph explains the mechanics of obtaining the true distribution. But what has the application of the Perron-Frobenius theorem allowed us to accomplish? The Perron-Frobenius theorem provides us with two pieces of information critical to the derivation of the natural probability distribution: the discount factor (δ) and the risk aversion (D). We obtain the discount factor and risk aversion using the marginal rate of substitution defined in equation 19. The components of the marginal

rate of substitution are the marginal utilities of consuming today versus consuming tomorrow. The Perron-Frobenius theorem allows us to determine the single unique discount factor and marginal utilities that dictate the transition paths between states. In other words, under the assumptions of the Perron-Frobenius theorem, only one set of marginal utilities and one discount factor will hold. Basically, they are relating the discounted willingness for the representative agent to consume today versus consuming at some other period in the future given certain transition probabilities.

Once we have the true probability matrix, obtaining the market forecast becomes trivial. We divide state prices by the kernel to obtain the natural marginal probabilities. We multiply the natural marginal probabilities by the state levels to obtain an expected return for each time interval.

Numerical example (continued) Continuing from where the numerical example left off in section 2.3.2, recall the contingent state price matrix for the high volatility state:

$$\mathbf{P}_{\sigma_{\mathbf{H}}} = \begin{array}{c} \begin{array}{cc} \text{+25\%} & \text{-10\%} \end{array} \\ \begin{bmatrix} 0.6112 & 0.4122 \\ 0.4891 & 0.5024 \end{bmatrix} \end{array} \begin{array}{l} \text{+25\%} \\ \text{-10\%} \end{array}$$

Applying the Perron-Frobenius theorem as in equation 23, we get the following result for the high volatility state:

$$\delta = 1.0091$$

$$z = \begin{bmatrix} 0.5088 \\ 0.4912 \end{bmatrix}$$

Using the resulting values, we can verify that the equality in equation 23 holds:

$$\begin{bmatrix} 0.6112 & 0.4122 \\ 0.4891 & 0.5024 \end{bmatrix} \begin{bmatrix} 0.5088 \\ 0.4912 \end{bmatrix} = 1.0091 \begin{bmatrix} 0.5088 \\ 0.4912 \end{bmatrix} = \begin{bmatrix} 0.51345 \\ 0.49567 \end{bmatrix}$$

Plugging these numbers into equation 24, we get the following:

$$F_{\sigma_H} = \frac{1}{\delta} DPD^{-1} = \frac{1}{1.0091} \begin{bmatrix} 1.9653 & 0 \\ 0 & 2.0360 \end{bmatrix} \begin{bmatrix} 0.6112 & 0.4122 \\ 0.4891 & 0.5024 \end{bmatrix} \begin{bmatrix} 0.5088 & 0 \\ 0 & 0.4912 \end{bmatrix} = \begin{bmatrix} 0.6057 & 0.3943 \\ 0.5021 & 0.4979 \end{bmatrix}$$

The same set of operations can be applied to the other contingent state price matrices to obtain the rest of the natural probability matrices. Now, we can outline the importance of controlling for volatility in the model. Once we have applied the RT as above, we get the following natural probability distributions:

$$\mathbf{F}_{\sigma_\mathbf{H}} = \begin{array}{c} \begin{array}{cc} \text{+25\%} & \text{-10\%} \end{array} \\ \begin{bmatrix} 0.6057 & 0.3943 \\ 0.5021 & 0.4979 \end{bmatrix} \begin{array}{l} \text{+25\%} \\ \text{-10\%} \end{array} \end{array}$$

$$\mathbf{F}_{\sigma_\mathbf{L}} = \begin{array}{c} \begin{array}{cc} \text{+25\%} & \text{-10\%} \end{array} \\ \begin{bmatrix} 0.9388 & 0.0612 \\ 0.296 & 0.704 \end{bmatrix} \begin{array}{l} \text{+25\%} \\ \text{-10\%} \end{array} \end{array}$$

$$\mathbf{F}_{\mathbf{naive}} = \begin{array}{c} \begin{array}{cc} \text{+25\%} & \text{-10\%} \end{array} \\ \begin{bmatrix} 0.782 & 0.218 \\ 0.4086 & 0.5914 \end{bmatrix} \begin{array}{l} \text{+25\%} \\ \text{-10\%} \end{array} \end{array}$$

The natural probability distribution resulting from the high volatility state has a higher probability of a large positive return compared to the naïvely obtained natural probability distribution when the initial state was bearish (0.5021 compared to 0.4086). In a state of high volatility, the probability of reaching a far away state increases, all else equal. From here, we can obtain our expected return by summing the result of the multiplication of the natural probabilities by the expected outcomes. Continuing our

previous example, we obtain the following expected return:

$$E(r_{\sigma_H}) = \left(\begin{bmatrix} 0.6112 & 0.4122 \\ 0.4891 & 0.5024 \end{bmatrix} / \begin{bmatrix} 0.6057 & 0.3943 \\ 0.5021 & 0.4979 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \right)^{\intercal} \begin{bmatrix} 0.25 \\ -0.10 \end{bmatrix} = 0.15770 = 15.770\%$$

We will call the difference between the naïve expected return and the multivariate expected return an error (error(σ)). I define:

$$error(\sigma) = \hat{E}[S^2|S^1] - \hat{E}[S^2|S^1, \sigma] \tag{25}$$

where $\hat{E}[S^2|S^1]$ is the expected return obtained from the naïve RT and $\hat{E}[S^2|S^1, \sigma]$ is the expected return obtained from the multivariate RT. The naïve RT produces an expected return of approximately 14.02%, while the multivariate RT produces an expected return of 15.77% in the high volatility state and 12.38% in the low volatility state. Expected returns vary widely depending on the volatility state assumptions (or lack thereof). In this simple example, the error can range from -1.65% to 1.75%.

$$\begin{bmatrix} error(\sigma_H) \\ error(\sigma_L) \end{bmatrix} = \begin{bmatrix} 1.75\% \\ -1.65\% \end{bmatrix}$$

Hence, the example demonstrates that not controlling for volatility in the estimation of the transition probabilities used in the RT can have a significant impact on both the resulting expected natural distribution and the resulting expected return.

3 Modelling uncertainty

In section 2, I alluded to the fact that the inclusion of the implied volatility in the derivation of the contingent state prices acted as a proxy for uncertainty in the macroeconomy. In this section, I show that including the implied volatility allows us to capture uncertainty in the business cycle. For example, intuitively, we should expect that, when the probability of a recession is high, the expected return would be low. The -0.5 corre-

lation between the Federal Reserve's estimated U.S. Recession Probabilities (Chauvet and Piger, 2008) and the realized risk-premium illustrates that fact. Now, if we correlate the Federal Reserve's estimated U.S. Recession Probabilities with the univariate RT and the MVRT, we obtain correlations of 0.11 and -0.21 respectively. Hence, the MVRT seems to capture more of the uncertainty than the univariate RT (as it is much closer to -0.5).

To test the idea put forth in this section, I simulate data using Monte Carlo simulations similar to the ones proposed by Heston (1993). In this setup, we obtain the simulated stock price from a Geometric Brownian Motion (GBM) and the stochastic volatility from a stochastic process as in Cox et al. (1985). The parameters used in these simulations can be found in section 4.3. Figures 6 and 7 illustrate ten series of simulated stock prices and volatilities:

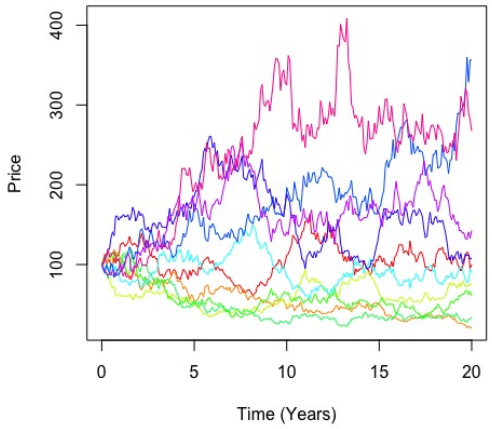

Figure 6: Stock Prices

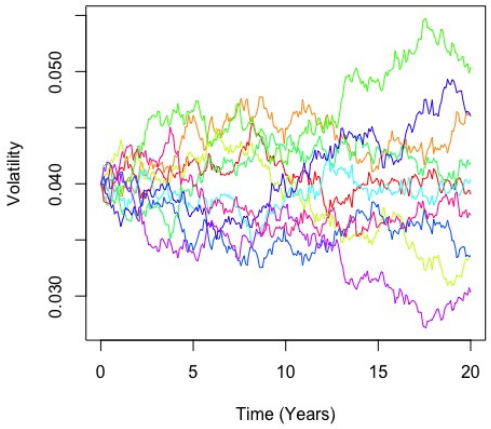

Figure 7: Stochastic Volatility

Once the data has been generated, I derive a binomial model with a representative agent that has heterogeneous habit formation (Campbell and Cochrane, 1999). The habit formation from Campbell and Cochrane (1999) is what generates the time-varying risk-premium. I start by defining the binomial model based off of Cox et al. (1979). We

first define the initial stock prices as the state-dependent value of a stock as follows:

$$S = p_u \cdot S_u + p_d \cdot S_d \tag{26}$$

where p is the risk-neutral probability of an up (u) or down (d) movement in the market and S is defined as:

$$\begin{aligned} S_u &= u \cdot S_0 \\ S_d &= d \cdot S_0 \end{aligned} \tag{27}$$

where S_0 is the current stock price (or initial stock price), u and d represent up or down movements in the market over a specific horizon and S_u (S_d) represents the stock price after an a hypothetical up (down) movement. The up or down movements depend on whether we are trying to model the univariate or the multivariate RT. For the univariate RT, the movements are defined as:

$$\begin{aligned} u &= 1 + \sigma\sqrt{T} \\ d &= 1 - \sigma\sqrt{T} \end{aligned} \tag{28}$$

where σ is the actual volatility observed in the market. For the multivariate RT, the movements are defined as:

$$\begin{aligned} u &= 1 + \sigma_{IVOL}\sqrt{T} \\ d &= 1 - \sigma_{IVOL}\sqrt{T} \end{aligned} \tag{29}$$

where the implied volatility is defined as the next period's volatility, σ_{t+1}, plus or minus an error term. The error term is a value taken from a standard normal distribution: $\epsilon_t \sim N(0, 1)$. The implied volatility is defined as the market's best estimate of the future volatility. By taking the next period's volatility and adjusting it by some error term, I am suggesting that the market has some sense of future volatility, but that its estimation is imperfect.

Recall from equation 6 that we defined the price of an asset as:

$$p_{t+1} = \phi_{t+1} f_{t,t+1} \tag{30}$$

where ϕ_{t+1} is the intertemporal marginal rate of substitution and $f_{t,t+1}$ is the natural probability measure. In order to obtain a forecast, we must first derive the intertemporal marginal rate of substitution. This is done using defining preferences as a function of external habit formations. These habit formations are a function of aggregate consumption, C_t^a, and an individual's habit, X_t, as follows:

$$S_t^a = \frac{C_t^a - X_t}{C_t^a} \tag{31}$$

which can be specified as the log surplus consumption ratio $s_t^a = ln\ S_t^a$ which evolves as a heteroskedastic AR(1) process:

$$s_t^a = (1-\omega)\bar{s} + \omega s_{t-1}^a + \lambda(s_{t-1}^a)(c_t^a - c_{t-1}^a - g) \tag{32}$$

where ω and g are parameters from Campbell and Cochrane (1999) (summarized in section 4.3). Parameter $\bar{s}$ represents the log of the steady state surplus consumption ratio and is defined as:

$$\bar{S} = \sigma\sqrt{\frac{\gamma}{1-\omega}} \tag{33}$$

where γ is the risk-aversion parameter. The sensitivity function, $\lambda(s_t^a)$, is defined as:

$$\lambda(s_t^a) = \begin{cases} \frac{1}{\bar{S}}\sqrt{1-2(s_t-\bar{s})} - 1, & s_t \leq s_{max} \\ 0, & s_t \geq s_{max} \end{cases} \tag{34}$$

where s_{max} is defined as:

$$s_{max} = \bar{s} + \frac{1}{2}(1-\bar{S}^2) \tag{35}$$

Consumption growth is modeled as an i.i.d. lognormal process:

$$\Delta c_{t+1} = g + v_{t+1} \tag{36}$$

where $v_{t+1} \sim i.i.d.\ N(0, \sigma^2)$. The intertemporal marginal rate of substitution, in this case, is as follows:

$$\phi_{t+1} = \delta\Big(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\Big)^{-\gamma} \tag{37}$$

which can then be used in equation 30. Once we have obtained the intertemporal marginal rate of substitution, we can apply the RT derived in earlier sections to obtain the natural probability distribution $f_{t,t+1}$.

4 Data and results

4.1 Overview of data

I collected the data for this paper from the Wharton Research Data Services (WRDS) database. I use daily option prices on the S&P 500, the S&P 500's closing price, and the risk-free rate. The risk-free rate is the one-month Treasury Bill rate, which can be found in the Fama & French factors data. S&P 500[2] prices are from the CRSP dataset. The S&P 500 is generally thought to be the best proxy for the market portfolio. All of the option data are from OptionMetrics. The data are used to obtain forecasts at intervals that range from one day to one quarter. This paper covers the time period from January 1996 to July 2015, the entire timeframe included in the OptionMetrics database. I use this sample for two major reasons. First, one of the forecast horizons in this paper is quarterly. A quarterly forecast requires a large enough sample size to test the efficacy of the RT and this twenty-year sample provides me with approximately 80 data points. Second, it allows me to divide the sample into subsamples and test my model in periods that experience various shocks (such as the tech bubble and the

[2]SECID 108105

recent financial crisis).

Strike prices on the options obtained from OptionMetrics are quoted for lots of 1,000 securities. The Black-Scholes-Merton equation requires strike prices that are on a per-stock basis, so I divided the strike price by 1,000. Time-to-maturity is converted from a date to a fraction of years to expiration, also a required input for the Black-Scholes-Merton equation. Option price is replaced with the midpoint of the bid-ask spread. This is consistent with Figlewski (2008), who argues that bid and ask prices are continuously quoted for almost all strikes regardless of whether a trade takes place. The alternative, transaction prices, occurs irregularly (Figlewski, 2008) and would make it more difficult to extract a proper implied volatility curve (see Appendix A). I compare my estimated implied volatilities to those provided by OptionMetrics. Since the difference between the two is negligible, I use my more complete set of estimates instead of the OptionMetrics data.

One of the difficulties of applying/replicating the RT is in constructing state prices. Ross (2015) uses over-the-counter data rather than the more limited publicly available data because it offers a significantly larger number of traded strikes and maturities. This paper uses readily available data from WRDS instead. Despite this difference in data source, I produce results that are very close to Ross's (see section 4). Another difficulty is that Ross (2015) does not explain how he derives state prices. Theoretically, state prices are easy to understand, but in practice, there is a lot of debate on how to construct them. Appendix A proposes a way to derive the extrapolated data required to construct state prices for this paper.

4.2 Empirical results

This section presents the empirical results for the univariate and the multivariate recovery theorems. I divide the samples into three subsamples to show the impact of different volatility states on the results. The first set of results is for the entire sample (April 1996 to August 2015). The high volatility subsample is from April 1996 to April

2002. The low volatility subsample is from January 2004 to January 2007. I selected the subsamples by examining time series plots to determine which periods had high volatility and which had low volatility.

4.2.1 Full sample results

Table 1 compares the results of Ross (Ross UVRT – first column) with the results of the multivariate RT (MVRT – second column) proposed in this paper to illustrate the superiority of the MVRT. Please note that the univariate results are the closest possible proxy for the results of Ross (since I did not have access to the data to replicate Ross's results exactly). All results presented in this section are out-of-sample. The very nature of the RT is such that in-sample results are not possible. Comparing the out-of-sample adjusted R^2, the MVRT method produces results superior to Ross's methodology.

	Ross UVRT (Apr 09–Apr 13)	MVRT (Apr 09–Apr 13)
	(1)	(2)
Intercept	-0.06054^{*}	0.027675^{**}
	(0.035068)	(0.009153)
Coefficient	5.710293^{**}	0.338864^{***}
	(1.95258)	(0.070478)
Observations	46	49
R^2	0.2162744	0.329701
Adjusted R^2	0.143715	0.315439
F statistic	0.005436	$1.6e^{-05}$

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 1: Ross Subsample - Summary Results

The tables below have four columns, each representing the result for a specific forecasting methodology. The first column is the univariate RT (UVRT), the proxy for Ross's original RT. The second column is the multivariate RT (MVRT), the new method proposed in this paper. The third column is the dividend-price ratio (D/P).

The fourth column is the consumption-wealth ratio (CAY). The forecast regression equation is as follows:

$$R_t = \alpha + \beta E_{t-1}[R_t] + \epsilon_t \tag{38}$$

where α is the intercept, β is the forecast coefficient, and $E_{t-1}[R_1]$ is the previous period's RT forecast. The forecast horizon is held to a quarter (three months) so t corresponds to 0.25 years. One of the criteria for forecast efficiency is the forecast error. This error is defined as the residual, ϵ_t, found in equation 38 and graphed in section 4.2.4. The errors are used as a way to ensure that the model is accurately specified. In general, the smaller the errors, the better the forecast.

Table 2 presents the results for the entire sample (April 1996 to August 2015).

	UVRT (Apr 96–Aug 15)	MVRT (Apr 96–Aug 15)	D/P (Apr 96–Aug 15)	CAY (Apr 96–Aug 15)
	(1)	(2)	(3)	(4)
Intercept	0.01040	0.00482	−0.00378	0.01936
	(0.00930)	(0.00465)	(0.01557)	(0.00836)
Coefficient	1.66110***	0.42471***	13.96761	0.65015
	(0.29290)	(0.04259)	(9.45251)	(0.48717)
Observations	235	235	235	78
R^2	0.12267	0.30187	0.00928	0.02290
Adjusted R^2	0.11885	0.29884	0.00503	0.01004
F statistic	$4.244e^{-08}$	$1.069e^{-19}$	0.14085	0.18601

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 2: Results for the four methods, full sample

The MVRT clearly outperforms all other benchmark results presented in table 2. The out-of-sample adjusted R^2 is 0.29884 compared to the UVRT's adjusted R^2 of 0.11885. This significant increase is consistent across samples, indicating that the MVRT provides significantly better results than previous methods. The MVRT results are also significantly better than the results for other benchmark forecasting methodologies such as the dividend-price ratio and the CAY ratio.

The ideal coefficients in a forecast are for the intercept to be zero and the slope coefficient to be one. In table 2, the slope coefficient in the MVRT is closer to one while maintaining the same level of significance as the UVRT. Both the UVRT and the MVRT seem to indicate that the intercept coefficient is equal to zero. Overall, the results look promising.

To test for robustness, the next set of results break down the original sample into smaller periods with either high or low volatilities. High-volatility subsamples represents periods where the volatility was constant at around 10% while low-volatility samples were periods where the volatility was around 5%. I also add periods (i.e., several months of data) of large changes in volatility to examine the effect on the forecast regression results. Based on the theory, the model should perform best when volatility remains relatively unchanged over time.

4.2.2 High-volatility subsample results

The first subsample is from April 1996 to April 2002. This subsample is the first period of time in the data where the volatility remains relatively high (and unchanged) throughout the sample ($\approx 8\%$).

	UVRT (Apr 96–Apr 02)	MVRT (Apr 96–Apr 02)	D/P (Apr 96–Apr 02)	CAY (Apr 96–Apr 02)
	(1)	(2)	(3)	(4)
Intercept	0.05871	0.00352	−0.04873	0.02190
	(0.01405)	(0.00675)	(0.02976)	(0.01884)
Coefficient	3.15148***	0.58939***	58.51909**	0.61102
	(0.86772)	(0.06343)	(21.94959)	(1.26921)
Observations	73	73	73	24
R^2	0.15668	0.54871	0.09100	0.01042
Adjusted R^2	0.14480	0.54236	0.07820	-0.03456
F statistic	0.00053	$6.814e^{-14}$	0.00950	0.63497

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 3: Results for the four methods, April 1996 to April 2002

In table 3, the results for the MVRT are quite impressive. The out-of-sample adjusted R^2 is almost 55% compared to about 16% for the UVRT. This is quite large for a forecast, likely because there are very little changes both in the mean and the volatility of returns during this time period. We can see this by looking at the D/P ratio, which also shows a significant forecasting ability. Normally, we would expect the dividend-price ratio to forecast long-term changes in asset prices. However, it seems to perform quite well during this period. Much like in the entire sample, the slope coefficient for the MVRT is getting closer to the desired coefficient of one. Moreover, the intercept does seem to be zero as we would hope.

4.2.3 Low-volatility subsample results

This next subsample, shown in table 3, is from April 2004 to January 2007. This period has a relatively low and constant volatility of around 4.7%.

	UVRT (Jan 04–Jan 07)	MVRT (Jan 04–Jan 07)	D/P (Jan 04–Jan 07)	CAY (Jan 04–Jan 07)
	(1)	(2)	(3)	(4)
Intercept	0.09510	0.01896*	0.02177	0.02896
	(0.14100)	(0.00770)	(0.02461)	(0.02706)
Coefficient	5.64410	0.23165*	2.46158	0.30591
	(4.74010)	(0.08897)	(15.82837)	(1.97238)
Observations	37	37	37	13
R^2	0.03895	0.16225	0.00069	0.00218
Adjusted R^2	0.01149	0.13832	-0.02786	-0.08853
F statistic	0.24170	0.01344	0.87731	0.87955

Note: $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 4: Results for the four methods, January 2004 to January 2007

During this time period, all forecasting methodologies perform miserably with the exception of the MVRT. The best performance was from the UVRT which had an out-of-sample adjusted R^2 of about 1.1% while the MVRT's adjusted R^2 is about 14%. The

statistical significance of the slope coefficient has decreased when compared with other sample periods. That being said, it is the only result during this period to achieve any level of statistical significance.

The following table examines what happens when I add months in the sample that have large changes in volatility. Using the sample from table 4 above as a starting point, I added eight months of data before and two years of data after. In total, the sample size went from 37 months to 73 months. Again, the purpose here is to study the impact of adding periods where the volatility changes on the results. These months changed the volatility for the period from about 4.7% to about 9%.

	UVRT (Apr 03–Apr 09)	MVRT (Apr 03–Apr 09)	D/P (Apr 03–Apr 09)	CAY (Apr 03–Apr 09)
	(1)	(2)	(3)	(4)
Intercept	−0.04171	−0.00404	0.08148*	−0.00540
	(0.01940)	(0.00993)	(0.03112)	(0.01768)
Coefficient	1.80831**	0.33532**	−50.92840**	−1.36018
	(0.63681)	(0.10376)	(18.57465)	(1.20377)
Observations	73	73	73	25
R^2	0.10200	0.12823	0.09574	0.05259
Adjusted R^2	0.08935	0.11595	0.08301	0.01140
F statistic	0.00588	0.00187	0.00773	0.27016

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 5: Results for the four methods, April 2003 to April 2009

From table 5, it is clear that the change in the volatilities has led to a decrease in the MVRT's forecasting ability. That being said, the difference is not substantial. The adjusted R^2 has decreased from around 14% to around 11.5%. The most dramatic change in this table appears in the other forecasting models. Specifically, the UVRT and the D/P results have substantially improved. Intuitively, these results should not be surprising. The UVRT is not as affected by changes in the volatility levels as the MVRT. It takes time for the MVRT to improve after a substantial change in the volatilities. This is not necessarily the case for the UVRT. That being said, the MVRT

still outperforms all of the benchmark forecasts presented in this table. So it is still performing quite well, just not as well as we might have hoped.

The next subsample is from April 2010 to the end of the sample period: August 2015. Much like the previous period, this subsample shows a relatively small volatility of about 5%.

	UVRT (Apr 10–Aug 15)	MVRT (Apr 10–Aug 15)	D/P (Apr 10–Aug 15)	CAY (Apr 10–Aug 15)
	(1)	(2)	(3)	(4)
Intercept	0.00130	0.01967**	0.01480	0.03070
	(0.00981)	(0.00604)	(0.02297)	(0.01746)
Coefficient	2.00011***	0.24430***	9.80293	0.20893
	(0.49340)	(0.04754)	(12.62916)	(0.75418)
Observations	65	65	65	23
R^2	0.20697	0.29538	0.00947	0.00364
Adjusted R^2	0.19439	0.28420	-0.00625	-0.04380
F statistic	0.00014	$2.892e^{-06}$	0.44053	0.78446

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 6: Results for the four methods, April 2010 to August 2015

In table 6, both the UVRT and the MVRT perform quite well (although the MVRT does outperform the UVRT again). The out-of-sample adjusted R^2s were about 20% and 28% for the UVRT and MVRT respectively.

For this next subsample, I added 24 months to the subsample. The additional 24 months displayed higher volatility (from the financial crisis), which added a shift in the volatility to the sample. The volatility increased from about 5% to almost 9%.

	UVRT (Apr 08–Aug 15)	MVRT (Apr 08–Aug 15)	D/P (Apr 08–Aug 15)	CAY (Apr 08–Aug 15)
	(1)	(2)	(3)	(4)
Intercept	−0.01510	0.00342	0.00232	−0.00596
	(0.01521)	(0.00896)	(0.02938)	(0.01744)
Coefficient	1.59471**	0.34140*	8.65916	−1.47744
	(0.52310)	(0.06878)	(15.38079)	(0.84335)
Observations	89	89	89	31
R^2	0.09651	0.22067	0.00363	0.09570
Adjusted R^2	0.08613	0.21172	-0.00782	0.06452
F statistic	0.00305	$3.412e^{-06}$	0.57489	0.09037

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 7: Results for the four methods, April 2008 to August 2015

This last sample includes part of the financial crisis. As such, there was a major change in the volatility levels. This is reflected in the relatively worse results of the MVRT when comparing the results from table 7 to those from table 6. Moreover, the statistical significance of the slope coefficient substantially decreases despite the larger sample size.

4.2.4 Varying the forecast horizon

In the previous subsection, I showed the results for various time periods while keeping the forecast horizon the same. Here I show the results for a monthly, quarterly, and yearly forecast. In this section, however, the quarterly forecast is updated every quarter instead of every month as in the previous section. The overlap causes a slight upward bias on the adjusted R^2 results. This serves the purpose of showing that although there is bias, it is quite small. The results for the various forecast horizons are summarized in figures 8 and 9. Figure 8 shows the coefficients for the UVRT and the MVRT only. Both models perform quite well (small errors) in the medium-term forecasts (monthly to quarterly) but the results start to deteriorate at the yearly forecast level. This is to be expected since options are not liquid at the annual time-to-maturity. This results

in a forecast that is unreliable. Although the daily forecast result is not shown here, the forecast performs as poorly as the yearly forecast for the same reason.

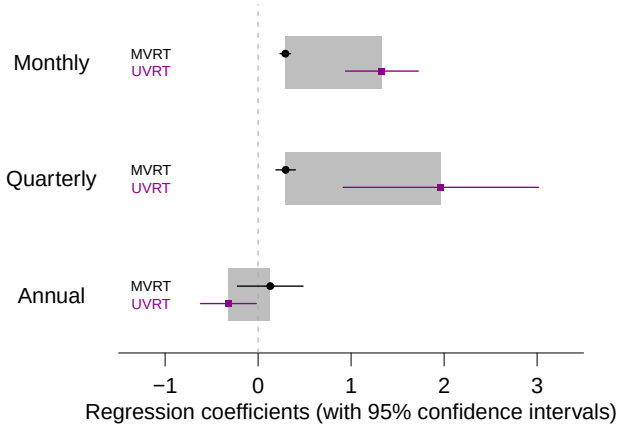

Figure 8: Regression Coefficients

Figure 9 shows the adjusted R^2 results at the various forecast horizons and compares those results to those of the dividend-price ratio. As was the case for the coefficients, the UVRT and the MVRT both perform well in the monthly and the quarterly forecast but are outperformed by the dividend-price ratio at the yearly forecast.

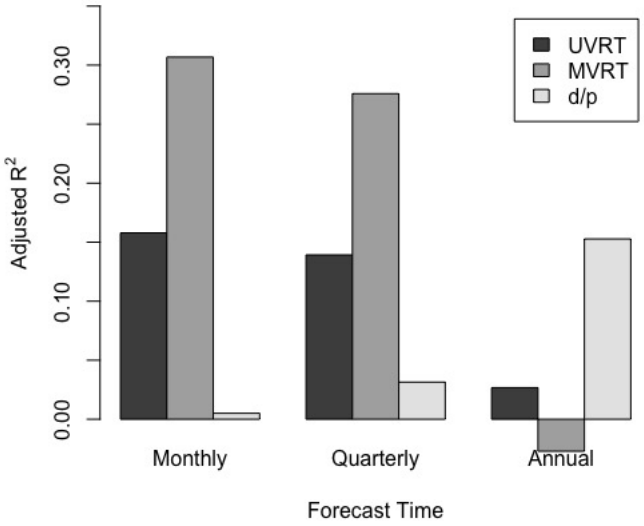

Figure 9: Adjusted R^2

4.3 Simulated results

This section presents the results using simulated data (see section 3). The goal is twofold: to show 1) that the results are not merely a construct of the empirical data, and 2) that the MVRT captures some of the uncertainty in the business cycle. The uncertainty in the business cycle comes from the time-varying risk-premium. A model that successfully captures the uncertainty in the business cycle would be the model that has the highest predictive power. Table 8 below shows the values used for the parameters required in the simulations.

Parameter	Variable	Value
Assumed:		
Mean consumption growth (%)*	g	1.89
Standard deviation of consumption growth (%)*	σ	1.50
Log risk-free rate (%)*	r^f	0.94
Persistence coefficient*	ω	0.87
Initial stock price	S_0	100
Number of simulations	n	10000
Volatility mean-reversion speed	κ	0.003
Volatility of volatility	$\sigma(\sigma)$	0.009
Correlation between stochastic volatility and spot prices	ρ	-0.5
Initial variance	σ_0^2	0.04
Long-term variance	θ	0.04
Reproducibility seed	NA	123

* Annualized values

Table 8: Parameters for simulations

Figures 10 and 11 below show the simulation results for the UVRT. Figure 10 shows the regression coefficient and figure 11 shows the adjusted R^2 for various risk-aversion parameters. The horizontal line represents the coefficient from the regression using empirical data. The goal is to determine which risk-aversion coefficient matches the empirical results. For the coefficient, the risk-aversion parameter that gives us the same results for the simulated data as the empirical data is between 7.5 and 15. The adjusted R^2 is presented for completeness. For some reason, it takes a very large risk-aversion parameter in order to be able to replicate the empirical forecastability results. Nevertheless, the model does seem to have forecasting power whenever a "realistic" risk-aversion parameter is considered.

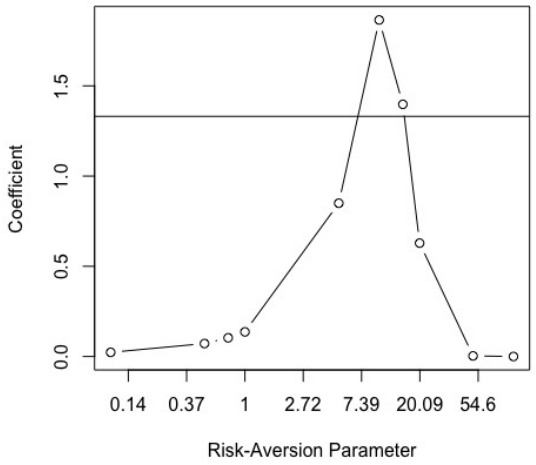

Figure 10: UVRT Simulations - Coefficient

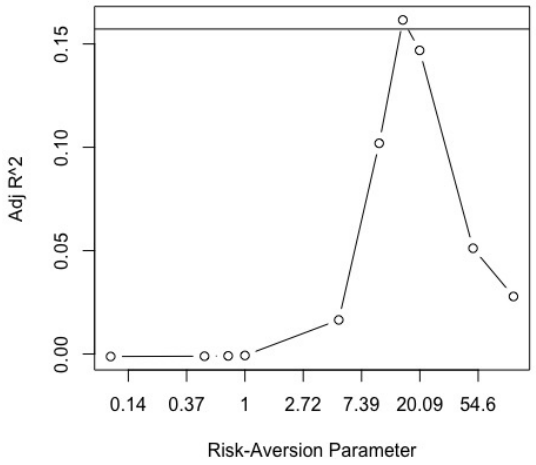

Figure 11: UVRT Simulations - Adj R^2

Figures 12 and 13 show the simulation results for the MVRT. The risk-aversion parameter where the simulated data and the empirical data converge is between 4.5 and 7.5. These values are much closer to what we would expect in reality than the UVRT values.

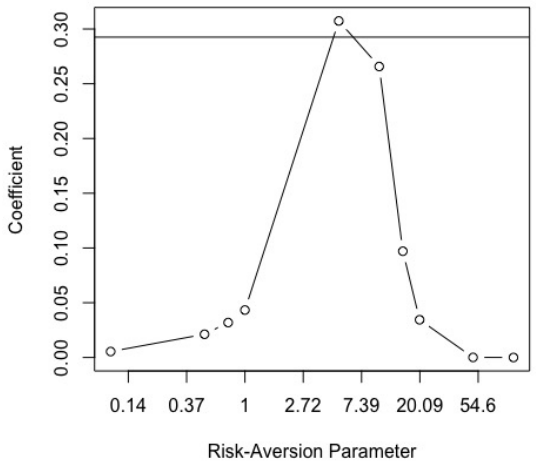

Figure 12: MVRT Simulations - Coefficient

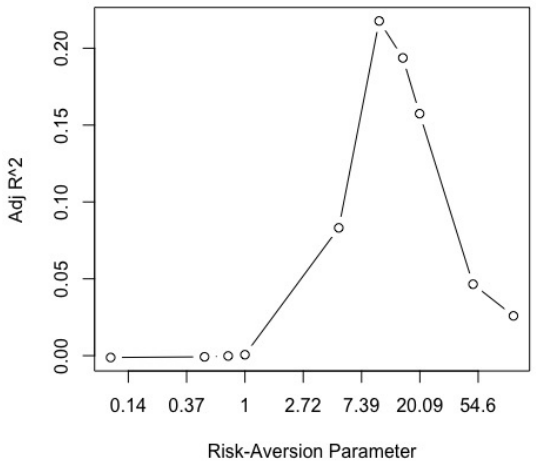

Figure 13: MVRT Simulations - Adj R^2

4.4 Market timing

The true test of whether a forecasting model is valuable boils down to its applicability. In other words, can investors use the model to make money? This section illustrates how the multivariate RT performs when a simple trading strategy is implemented. I outline how the trading strategy was implemented and I present the results in the form of a cumulative returns plot as well as a time-series plot showing the profits generated by each trade for the strategy. This strategy is compared to the cumulative returns plot for a buy and hold strategy on the S&P 500.

The MVRT strategy has an initial investment of $1. Each month, the MVRT gives the investor a signal to either buy (positive signal) or sell (negative signal) the S&P 500. If the signal is negative and the investor currently holds the asset, the asset is sold and shorted. Similarly, if the signal from the MVRT is positive and the investor is short, then the investor closes the current position and buys the asset. This exercise is repeated each time a new signal is obtained (every month in this example). The MVRT occasionally outputs an error. If the signal is an error, then the signal on the following day will be used. In the interest of simplicity, trading costs are not considered. However, since the signals are only obtained once a month, there are a limited number of rebalances, which implies that there are also a limited number of trades. Hence, trading costs for this type of strategy would be negligible. The results can be seen in figures 14 and 15.

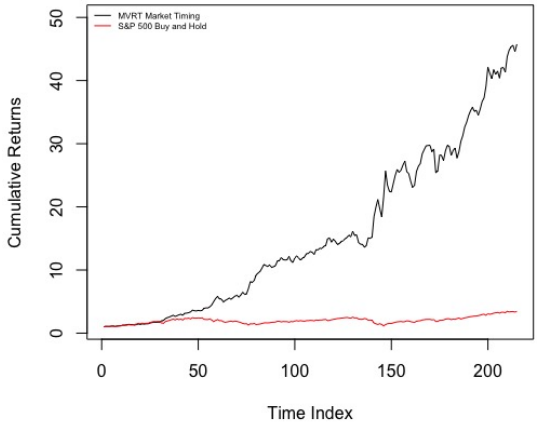

Figure 14: Cumulative Returns Plot

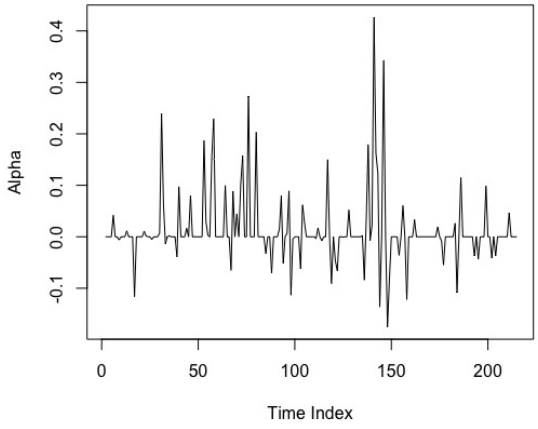

Figure 15: Profit and Loss Plot

Notice that, in figure 14, the cumulative returns from the MVRT (black line) outperform the S&P 500 buy and hold strategy (red line). This is accentuated by the fact that the cumulative returns consider compounding from reinvestment. A better depiction of the superiority of the MVRT can be seen in figure 15. Here we can see that, on average, the positive profits outnumber the negative profits. In fact, almost 57% of the trades are positive. Furthermore, the magnitude of the profits is substantially larger than that of the losses. The average profit is about 5% per trade compared to the average loss which is about 2.6% per trade.

5 Conclusion

This paper aimed to improve the estimation of the natural probabilities derived from the Recovery Theorem (RT). Its major contribution is that it extends the RT by changing the univariate derivation of the contingent state price matrix to a multivariate one. By changing the derivation of the contingent state price matrix to a multivariate Markov chain, the inherent transition probabilities are more accurately defined. In the multivariate chain, I added the volatility, which results in significant improvements in the RT results. The out-of-sample forecast regression's adjusted R^2 increases from

about 0.12 using Ross's specification to about 0.30 using the MVRT method. I show, using a simple numeric and intuitive example, that although the multivariate model performs better than the univariate model, it does much better whenever the changes in volatility are minimal. When changes in the underlying volatility occurs, it takes time for this new information to be fed into the model. As such, the multivariate model's performance does seem to suffer in instances when there are significant changes in volatility.

The Recovery Theorem was a giant leap forward in the forecasting of asset returns. This paper improves on the original specification and will make it possible to use this methodology for other asset pricing endeavors. A number of extensions are possible. For example, since the multivariate RT extracts the market's true distribution of returns, we can extend this research to the question of hedging. A future research direction would be to explore whether firms change their hedging behavior in response to certain future expectations, where the expectations are derived from the RT's natural distribution (Fillebeen and Sanford (2016)).

The multivariate RT could also be used in portfolio construction applications. For instance, we could use the true distribution obtained from the multivariate RT as an actual returns distribution for a portfolio optimization problem. The portfolio weights can then be selected such that a measure that uses the distribution of returns (e.g. expected tail loss) is minimized (see for example Sanford (2016a)). We may also want to use the exponential GARCH model (Bollerslev, 1986) to model the behavior of volatility. We can expect to obtain a better forecast if we incorporate a forward-looking volatility model rather than looking only at current volatility, as I do in this paper.

Finally, research should focus on whether the Recovery Theorem might apply in a setting where markets are incomplete. The RT assumes that the market is complete and, by extension, that it is possible to construct state prices. A natural question therefore arises: what assumptions would be necessary to apply the Recovery Theorem to an incomplete market? This would be a valuable extension to the current literature.

References

BIS (2012). Bis quarterly review, june 2012. *Bank for International Settlements*.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, pages 637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, pages 621–651.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, 107(2):205–251.

Chauvet, M. and Piger, J. (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics*, 26(1):42–49.

Chen, T. (2011). Improve OVDV long-term volatilities. *Bloomberg Research*.

Cochrane, J. H. (2009). *Asset Pricing:(Revised Edition)*. Princeton university press.

Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407.

Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263.

Engle, R. F. and Mustafa, C. (1992). Implied ARCH models from options prices. *Journal of Econometrics*, 52(1):289–311.

Figlewski, S. (2008). Estimating the implied risk neutral density. In Bollerslev, T., Russell, J. R., and Watson, M., editors, *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*. Oxford University Press, Oxford.

Fillebeen, T. and Sanford, A. (2016). Do small firms hedge: Forward looking beliefs using the recovery theorem. *Work in Process*.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

Jackwerth, J. C. and Rubinstein, M. (1996). Recovering probability distributions from option prices. *The Journal of Finance*, 51(5):1611–1631.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, pages 141–183.

Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*, volume 2. SIAM: Society for Industrial and Applied Mathematics, Philadelphia, PA.

Page, S. E. et al. (2006). Path dependence. *Quarterly Journal of Political Science*, 1(1):87–115.

Raftery, A. E. (1985). A model for high-order markov chains. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 528–539.

Ross, S. (2015). The recovery theorem. *The Journal of Finance*, 70(2):615–648.

Rubinstein, M. (1994). Implied binomial trees. *The Journal of Finance*, 49(3):771–818.

Sanford, A. (2016a). Forward-looking expected tail loss: An application of the recovery theorem. *Working Paper*.

Sanford, A. (2016b). State price density estimation with an application to the recovery theorem. *Working Paper*.

Stoll, H. R. (1969). The relationship between put and call option prices. *The Journal of Finance*, 24(5):801–824.

A Appendix – Implied volatility extrapolation

In this section, I introduce my proposed implied volatility extrapolation method and show how extrapolated prices lead to a dense set of option prices. I then briefly define and derive the benchmark extrapolation method used in this paper: the Aït-Sahalia and Lo model. For more information on the extapolation methodology defined in this section, see Sanford (2016b).

A.1 Strike price extrapolation

The first step for the MVRT involves extrapolating the volatility surface with respect to two dimensions: strike prices and time-to-maturity. We extrapolate in terms of strike prices because there are only a certain number of strikes that are traded on any given day. For example, table 9 shows the (unique) strike prices for call options on the S&P 500 for 1 April 1996. However, for this specific day, we would need a set of strike prices ranging from about 350 to 1,200 in order to produce a complete volatility surface. Thus, extrapolation is necessary.[3]

400.00	425.00	450.00	475.00	500.00	510.00	520.00	525.00	530.00	540.00	545.00
550.00	560.00	565.00	570.00	575.00	580.00	585.00	590.00	595.00	600.00	605.00
610.00	615.00	620.00	625.00	630.00	635.00	640.00	645.00	650.00	655.00	660.00
665.00	670.00	675.00	680.00	685.00	690.00	695.00	700.00	725.00	750.00	

Table 9: Strike Prices on S&P 500 call options for 1 April 1996

The strike price extrapolation is based on a slightly modified risk-neutral density estimation methodology proposed by Figlewski (2008). Figlewski (2008) shows that one of the more precise ways to extrapolate a volatility surface is to use a smoothed quartic spline regression with a single at-the-money (ATM) knot. That being said, I have found that using smoothed B-splines rather than quartic splines provides a better overall fit. This is what I used in this paper.

[3]Extrapolation based on strike price is common practice in the volatility surface literature (Jackwerth and Rubinstein, 1996; Rubinstein, 1994; Figlewski, 2008).

We can derive the coefficient estimate for the smoothed spline by first defining the criterion function to be minimized as follows:

$$\min_{\beta} ||C - G\beta||^2 + \lambda\beta'\Omega\beta \tag{39}$$

where

$$G_{i,j} = g_j(\sigma_{IV,i}), \quad i,j = 1,...,n \tag{40}$$

$$\Omega_{i,j} = \int g_i''(t)g_j''(t)dt, \quad i,j = 1,...,n \tag{41}$$

where n is the number of knots, x is the actual knot, $g()$ are the B-spline basis functions, Ω is the penalty matrix, and λ is the smoothing parameter. Next, we need to define what we mean by a B-Spline basis function.[4] We can define the B-Spline function as follows:

$$G_{i,j} = \sum_{i=1}^{n+1} B_j(\sigma_{IV,i})G_i, \qquad \sigma_{IV,min} \leq \sigma_{IV,i} < \sigma_{IV,max} \tag{42}$$

where G_i corresponds to the control points, $B()$ is the basis function of order j, and x corresponds to the knots. Then, we can define the basis function from the B-spline as follows:

$$B_{i,1}(\sigma_{IV}) = \begin{cases} 1, & \text{if } \sigma_{IV,i} \leq \sigma_{IV} < \sigma_{IV,(i+1)} \\ 0, & \text{otherwise} \end{cases} \tag{43}$$

$$B_{i,j}(\sigma_{IV}) = \frac{\sigma_{IV} - \sigma_{IV,i}}{\sigma_{IV,(i+j-1)} - \sigma_{IV,i}} B_{i,j-1}(\sigma_{IV}) + \frac{\sigma_{IV,(i+j)} - \sigma_{IV}}{\sigma_{IV,(i+j)} - \sigma_{IV,(i+1)}} B_{i+1,j-1}(\sigma_{IV}) \tag{44}$$

Finally, we obtain the smoothing spline estimate at the knot C:

$$\hat{r}(C) = \sum_{j=1}^{n} \hat{\beta}_j g_j(\sigma_{IV}) \tag{45}$$

[4]Note that the notation here is slightly different from traditional notation in order to be consistent with the notation in the rest of the paper.

A.2 Time-to-maturity extrapolation

Table 10 shows the TTM on S&P 500 call options for 1 April 1996 in number of years. The time interval between each of the TTMs is not constant. Therefore, I need to extrapolate the data such that TTM follows a constant interval (for now, this interval is set to a constant three-months).[5]

0.05	0.13	0.23	0.47	0.72	0.97	1.22	1.72

Table 10: Time-to-maturity on S&P 500 call options for 1 April 1996

For the TTM extrapolation, I use a method devised by Bloomberg (Chen, 2011) as an extension of Heston (1993). First, let us define the extrapolated call price as follows[6]:

$$C(T,K) = \sum_{l=1}^{N} p_l(T) \cdot BSP(\xi_l(T)S_{0,p}, K, r_f, \Sigma_l(T)/\sqrt{T}) \tag{46}$$

where BSP corresponds to the traditional Black-Scholes equation (Black and Scholes, 1973) where each variable is a regular Black-Scholes input with certain parameters adjusted for extrapolation. The extrapolation details and the parameters in equation 46 are discussed in greater detail later in this section.

I start by defining two functions, $\alpha(t)$ and $\eta_l(t)$, for notational simplicity:

$$\varphi(t) = \frac{T_{i+1} - t}{T_{i+1} - T_i} \tag{47}$$

$$\eta_l(t) = log(\frac{\xi_{l+1}(t)}{\xi_l(t)}) \tag{48}$$

where $\eta_l(t)$ uniquely determines $\xi_l(t)$ under the assumption that $\sum_l p_l(t)\xi_l(t) = 1$, $\xi_l(T) \geq 0$ is the time-dependent multiplicative means of the l-th lognormal, $0 \leq p_l(T) \leq 1$ is the time-dependent weight of the l-th lognormal, t is the market maturity at which we want to extrapolate, and i is the index for each of the observed time-to-maturities.

[5]Later in the paper, I test various interval lengths.

[6]Note that it is trivial to show that extrapolating the option price is the same as extrapolating the option price as long as the inputs for the equation are the same but where the volatility is, in fact, the implied volatility.

If we assume a Poisson default process and a survival probability $D(t) = 1 - Q(t)$, we obtain the hazard rate $\Lambda(t)$ that is consistent with the survival probability:

$$D(t) = 1 - Q(t) = \sum_l p_l(t) = e^{-\Lambda(t)t} \tag{49}$$

where the initial $\Lambda(t)$ is obtained from the Bloomberg survival probability data. Once we have the benchmark hazard rate and survival probability, we need to estimate four equations (the new $\Lambda()$, $p_l()$, $\eta_l()$, and $\Sigma_l()$) and use the values as inputs for equation 46. The specific equations are dependent on whether we are extrapolating between TTMs, we are doing a shorter-term TTM extrapolation (less than three months), or a longer-term TTM extrapolation (greater than six months).[7] Each of these is derived and discussed in its own section below.

Shorter-term extrapolation A shorter-term extrapolation is an extrapolation that occurs either within three months of an available datapoint, or an extrapolation at a TTM below the lowest available TTM (but still less than six months from the lowest available TTM). First, we need the hazard rate $\lambda(t)$ in order to obtain $p_l(t)$. This is obtained as follows:

$$\Lambda_{new} = \Lambda e^{\frac{x_m^2 - x^2}{2T_t}} \tag{50}$$

$$\hat{\Lambda}_{new} = \Lambda_{new} e^{\frac{x^2}{2}(\frac{1}{T_0} - \frac{1}{t})} \tag{51}$$

where $x_m = K_{min}/F(T_i)$, $x = K/F(T_i)$, T_i is the closest TTM, $F()$ is obtained from the Put-Call Parity: $C() - P() = \frac{1}{r_f}(F - K)$ (Stoll, 1969), T_0 is the smallest TTM, and t is the TTM of interest. Here, we are effectively dampening the hazard rate estimate. Once we have adjusted this hazard rate, we can easily obtain $p_l(t)$ by ensuring that

[7] The longer-term extrapolation is used only occasionally since we usually have data within six months of extrapolations of interest.

its weights have the same ratio as what we would have at the lowest TTM.[8] Then, we can obtain the time-dependent standard deviation of the l-th lognormal, $\Sigma_l(t)$, and the means of each lognormal as:

$$\Sigma_l(t) = \frac{\Sigma_l(T_1)t}{T_1} \tag{52}$$

$$\eta_l(t) = \eta_l(T_1)\sqrt{\frac{t}{T_1}} \tag{53}$$

Now, we have all of the necessary components to solve equation 46 (Black and Scholes, 1973).

Extrapolation between time-to-maturities Here, we need to extrapolate between available TTMs. First, we derive the dampened hazard rate using equation 50. The only difference is that we adjust K_{min} by defining it as follows:

$$K_{min} = \varphi(t)K^i_{min} + (1 - \varphi(t))K^{i+1}_{min} \tag{54}$$

Once we have estimated the dampened hazard rate, we can proceed to estimate the multiplicative means, $\xi_l(T)$, the time-dependent weight, $p_l(T)$, and the time-dependent standard deviation, $\Sigma_l(T)$ using the following equations:

$$p_l(t) = (\frac{p_l(T_i + 1)}{D(T_{i+1})}\frac{\sqrt{t} - \sqrt{T_i}}{\sqrt{T_{i+1}} - \sqrt{T_i}} + \frac{p_l(T_i)}{D(T_i)}\frac{\sqrt{T_{i+1}} - \sqrt{t}}{\sqrt{T_{i+1}} - \sqrt{T_i}})D(t) \tag{55}$$

$$\Sigma^2_l(t) = (1 - \varphi(t))\Sigma^2_l(T_{i+1}) + \varphi(t)\Sigma^2_l(T_i) \tag{56}$$

$$\eta^2_l(t) = (1 - \varphi(t))\eta^2_l(T_{i+1}) + \varphi(t)\eta^2_l(T_i) \tag{57}$$

Longer-term extrapolation At longer time horizons, we do not dampen the hazard function. We want the full effects of the potential for default. We obtain the time-

[8] In other words, we are making sure that the weights at $p_l(t)$ are the same as the ratio of weights $\frac{p_{l+1}}{p_l}$ that we would have at T_1.

dependent weights as:

$$p_l(t) = p_l(T_n)\frac{D(t)}{D(T_n)} \tag{58}$$

where T_n is the largest available datapoint with respect to TTM and recalling that we define the survival probability, $D(t)$, using equation 49. We then obtain the time-dependent volatility as:

$$\Sigma_l^2(t) = \Sigma_l^2(T_n)\frac{t}{T_n} \tag{59}$$

Finally, we need to derive the means as follows:

$$\eta_l(t) = \eta_l(T_n)\sqrt{\frac{t}{T_n}} \tag{60}$$

A.3 Implied volatility surface and option prices

Implied volatility surface Figure **??** illustrates the skew of the extrapolated implied volatilities on 1 April 1996. The implied volatility increases at low strike prices, decreases as the strike price becomes higher, and finally increases again at higher strike prices, displaying a volatility skew (although in this case it is almost a volatility smirk). The figure confirms that the extrapolation produced the desired characteristics.

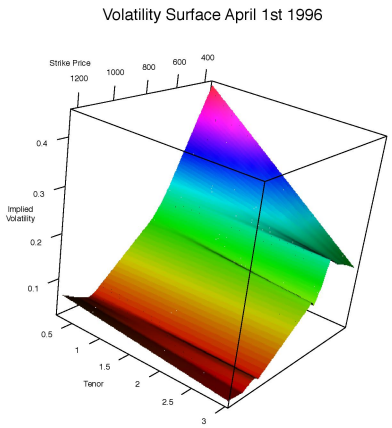

Figure 16: Implied Volatility Surface, 1 April 1996

Option prices Once we have obtained a matrix with implied volatilities at the required strike prices (outlined in section A.1)[9] and TTMs (outlined in section A.2), we can proceed to obtain option prices by inputting the data in the Black-Scholes-Merton equation (Black and Scholes, 1973):

$$C(S_{0,p}, t) = N(d_1)S_{0,p} - N(d_2)Ke^{-r_f(T-t)} \tag{61}$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}[ln(\frac{S_{0,p}}{K} + (r_f + \frac{\sigma^2}{2})(T-t)]$$
$$d_2 = \frac{1}{\sigma\sqrt{T-t}}[ln(\frac{S_{0,p}}{K} - (r_f + \frac{\sigma^2}{2})(T-t)]$$

where $N()$ is a value from the normal distribution. The above produces a matrix of call prices at our required strike prices and TTMs.

[9]In this paper, I use \$1 increments for strike prices.