Does the Ross Recovery Theorem work Empirically?

Jens Carsten Jackwerth† Marco Menner†

March 15, 2017

Abstract
Starting with the fundamental relationship that state prices are the product of physical probabilities and the pricing kernel, Ross (2015) shows that, given strong assumptions, knowing state prices suffices for backing out physical probabilities and the pricing kernel at the same time. We find that such recovered physical distributions based on the S&P 500 index are incompatible with future realized returns. This negative result remains, even when we add economically reasonable constraints. Reasons for the rejection seem to be numerical instabilities of the recovery algorithm and the inability of the constrained versions to generate pricing kernels sufficiently away from risk-neutrality.

Keywords: Ross recovery, Berkowitz, pricing kernel, risk-neutral density, transition state prices, physical probabilities

*Jens Jackwerth is from the University of Konstanz, PO Box 134, 78457 Konstanz, Germany, Tel.: +49-(0)7531-88-2196, Fax: +49-(0)7531-88-3120, jens.jackwerth@uni-konstanz.de
†Marco Menner is from the University of Konstanz, PO Box 134, 78457 Konstanz, Germany, Tel.: +49-(0)7531-88-3346, Fax: +49-(0)7531-88-3120, marco.menner@uni-konstanz.de

We received helpful comments and suggestions from Guenther Franke, Anisha Ghosh, Eric Renault, and Christian Schlag. We thank seminar participants at the University of Konstanz, the University of Strasbourg, the University of Zurich, the University of St. Gallen, the University of Sydney, the University of Queensland, and participants at the ESSFM Workshop at Gerzensee and at the DGF Conference in Bonn.
I. Introduction

Much of financial economics revolves around the triangular relation between physical return probabilities \( p \), which are state prices \( \pi \) divided by the pricing kernel \( m \):

\[
\text{physical probability } p = \frac{\text{state price } \pi}{\text{pricing kernel } m}
\]  

(1)

Researchers typically pick any two variables to find the third. In option pricing, for example, physical probabilities are changed into risk-neutral ones, which are nothing but normalized state prices, via the pricing kernel. Differently, the pricing kernel puzzle literature, e.g. as surveyed in Cuesdeanu and Jackwerth (2016), starts out with risk-neutral and physical probabilities in order to find empirical pricing kernels. Yet Ross (2015) presents a recovery theorem which allows to back out both the pricing kernel and physical probabilities by only using state prices. To achieve this amazing feat, he needs to make strong assumptions concerning the economy. We investigate his claim and test if the recovered physical probabilities are compatible with future realized S&P 500 returns. We further analyze if the shape of the recovered pricing kernel is in line with utility theory. To understand our sobering results, we discuss in detail why the recovery theorem does not perform well empirically.

The Ross (2015) recovery theorem is based on three underlying assumptions. First, it requires time-homogeneous transition state prices \( \pi_{i,j} \) that represent state prices of moving from any given state \( i \) today to any other state \( j \) in the future. Such transition state prices include the usual spot state prices \( \pi_{0,j} \) with 0 representing the current state of the economy. Spot state prices can be readily found from option prices, see e.g. Jackwerth (2004). Yet Ross recovery also requires as inputs the transition state prices emanating from alternative, hypothetical states of the world.  

---

1See e.g. Cochrane (2000), pp. 50. A state (Arrow-Debreu) price represents the dollar amount an investor is willing to pay for a security that pays out one dollar if the particular state occurs and nothing if any other state occurs. The pricing kernel is closely related to the marginal utility of such investor.

2Imagine that the current state of the world is characterized by the S&P 500 being at 1000. Let there be two future states, 900 and 1000, to which the spot state prices (emanating from 1000) relate. The required other transition state prices are the ones emanating (hypothetically) from 900 and ending at 900 or 1000.
There is, however, no market with actively traded financial products, which allow backing out the
(non-spot) transition state prices \( \pi_{i,j} \) directly. We will therefore use information on spot state prices
with different maturities to obtain transition state prices linking \( N \) states. This requires us to first
estimate an interpolated spot state price surface from sparse option data. There is no unique way
for estimating transition state prices \( \pi_{i,j} \) from spot state prices, and so we suggest several different
methods. We also introduce a version that allows us to apply the recovery theorem directly to spot
state prices \( \pi_{0,j} \).

Second, all transition state prices need to be positive, which turns out to be a fairly benign
assumption as we can always force some small positive state price.\(^3\)

Finally, the pricing kernel is restricted to be a constant times the ratio of values in state \( j \)
over values in state \( i \). A convenient economic interpretation of these values is to associate them
with marginal utilities in those states. The constant can then be associated with a utility discount
factor.

Taken together, the three assumptions allow Ross (2015) to formulate a unique eigenvalue
problem. Its solution yields the physical transition probabilities \( p_{i,j} \), which represent physical
probabilities of moving from state \( i \) to state \( j \), and the pricing kernel.

After we achieve recovery, we empirically test the hypothesis that future realized S&P 500
returns are drawn from the recovered physical spot distribution \( p_{0,j} \). In each month \( t \), we work out
the percentile of next month’s S&P 500 return based on the recovered physical spot cumulative
distribution function. This leaves us with one value \( x_t \) for each month \( t \) that lies in between 0 and 1.
If our hypothesis holds, the set \( \{x_t\} \) is uniformly distributed. We can strongly reject our hypothesis
for three different statistical tests: The Berkowitz (2001) test, the uniformity test introduced by
Knüppel (2015), and the Kolmogorov-Smirnov test.\(^4\) Thus, Ross (2015) does not recover physical
spot distributions of returns, which are consistent with future realized S&P 500 returns. We further
one period later.

\(^3\)Technically, some zero values could be allowed as long as any state can still be reached from any other

\(^4\)Note that Berkowitz (2001) does not directly test for uniformity of the values but for standard normality
of a standard normal transformation of the uniform values.
find that the recovery theorem does not produce downward sloping pricing kernels (as one would expect based on risk averse preferences) but they are riddled with local minima and maxima. In contrast, we cannot reject the hypotheses that future S&P 500 returns are drawn from simple physical distributions based on a power pricing kernel or the five-year historical return distribution.

The empirical problems of Ross recovery stem from three sources. For one, it is hard to obtain transition state prices from option prices. If we use a basic implementation of Ross recovery, which requires no additional assumptions beside positivity of transition state prices, we obtain unstable transition state prices that, in addition, exhibit unrealistic properties such as multi-modality and imply extremely high or low values for the state-dependent risk-free rate. Yet, if we introduce economically reasonable constraints to calm down the transition state prices, then the recovery theorem generates almost flat pricing kernels. Second, we argue that the strong assumption concerning the functional form of the pricing kernel is rather limiting. We further highlight that the recovered pricing kernel is highly dependent on the structure of the transition state price matrix, which in turn is not well identified from the option prices. Third, the assumption of time-homogeneous transition state prices might not hold and different periods may require different transition state prices.

Only a few papers already investigate the recovery theorem from an empirical perspective. Closest to our work are Audrino et al. (2015), who also implement Ross recovery on S&P 500 index options. Their recovered pricing kernels tend to be rather smooth and u-shaped, as opposed to our wavy pricing kernels. This surprising behaviour seems to be due to a particular modelling choice for a penalty term. Their further empirical focus is on developing profitable trading strategies based on the recovered physical probabilities, but, unlike our work, they do not statistically test if future realized returns are drawn from the recovered distribution.

Note that standard CARA or CRRA pricing kernels are incompatible with Ross recovery as his multi-period pricing kernels are identical but for a constant, while, under CARA or CRRA, they are multiplications of single-period pricing kernels.

The subtle reason is a quadratic penalty term, which they use to force all transition state prices to zero. This penalty is stronger for states further away from the current states as option prices are more sensitive to state prices around the current state. As a result, the implied risk-free rates (1 - sum of transition state prices) increase in the distance to the current state, which in turn mechanically leads to u-shaped pricing kernels.
Also informing our empirical work is a simulation study by Tran and Xia (2015), who show that the recovered probabilities vary substantially for different state space dimensions. We check for this prediction in our robustness tests but it empirically does not matter much for our paper.

As an alternative to the numerically difficult recovery of transition state prices from spot state prices, we suggest an additional, implicit method, which obviates that recovery and works directly with the spot state prices. In independent work, Jensen et al. (2016) suggest the same method and add further restrictions of the pricing kernel. They then focus on analyzing the theoretical properties of their generalized recovery. In a short empirical study, they use the mean of the recovered physical probabilities to predict S&P 500 return with a small yet significant $R^2$ of 1.28%. They also apply a Berkowitz test and reject that future realized S&P 500 returns are drawn from the recovered physical distribution based on their particular model.

Further work on the implementation of Ross recovery is Massacci et al. (2016), who propose a fast non-linear programming approach to apply Ross Recovery such that demanded economical constraints like positive state dependent risk-free rates and the unimodality of transition state prices are satisfied.

To our knowledge, we are the first to show that several plausible implementations of the Ross recovery theorem are not compatible with future realized returns of the S&P 500 index and to analyze why this is the case.

On the theoretical side and predating Ross (2015), Hansen and Scheinkman (2009) already showed that, given the underlying Markovian environment, the Perron-Frobenius theorem can be used to recover probabilities $p$. In a second step, they relate the recovered probabilities to physical probabilities. Borovicka et al. (2016) show that Ross (2015) recovers the same probabilities $p$ and, by implicitly assuming that the second step adjustment is simply one, interprets these as physical probabilities. The probabilities $p$ from Ross recovery may thus differ from the true physical probabilities and our paper shows that they indeed do.\(^7\)

\(^7\)Hansen and Scheinkman (2009) do not interpret their constructed measure as the physical probability measure, but simply as a probability measure that may provide useful long-term insights into risk pricing. Related, Martin and Ross (2013) show that Ross (2015) recovers a pricing kernel that depends on the time series behavior of the long bond and on the long end of the yield curve.
Related to that literature, Bakshi et al. (2016) use options on 30-year Treasury bond futures and solve convex minimization problems to extract the martingale component that is missing in Ross recovery. They find that this martingale component features considerable dispersion and thus conflicts with the unity constraint in the recovery theorem.

While Ross recovery works on a discrete state space using the Perron-Frobenius theorem, Carr and Yu (2012) show that recovery can be achieved in continuous time for univariate time-homogeneous bounded diffusion process by using Sturm-Liouville theory. Walden (2016) further investigates an extension of the recovery theorem to continuous time if the diffusion process is unbounded. He derives necessary and sufficient conditions that enable recovery and finds that recovery is still possible for many of these unbounded processes. Additional works on Ross recovery in continuous time are Qin and Linetsky (2016), Qin et al. (2016), and Dubynskiy and Goldstein (2013).

Related to the recovery literature, Schneider and Trojani (2016) extract physical moments from spot state prices in a unique minimum variance pricing kernel framework with mild economic assumptions on specific risk premia.

The remainder of the paper proceeds as follows. Section II explains the Ross recovery theorem. In Section III we introduce our methods to obtain spot state prices, to back out transition state prices, and to apply the theorem without using transition state prices. We further explain the Berkowitz test, the Knüppel test, and the Kolmogorov-Smirnov test, which we use to test our hypothesis. Section IV describes our data set. In Section V we present the empirical results of our study. Reasons for why the Ross recovery theorem empirically fails are given in Section VI. Section VII provides several robustness checks, while Section VIII concludes.

II. The Ross Recovery Theorem

An application of the recovery theorem requires the following assumptions: (i) The transition state prices \( \pi_{i,j} \) follow a time homogeneous process, which means that they are independent of calender
time, (ii) the transition state prices need to be strictly positive, and (iii) the corresponding pricing kernel $m_{i,j}$ is transition independent, which means that it can be written as:

$$m_{i,j} = \frac{u'_j}{u'_i}$$

(2)

for a positive constant $\delta$ and positive state dependent (marginal utility) values $u'_j$ and $u'_i$.

With this structure for the pricing kernel, the physical transition probabilities $p_{i,j}$ have the form:

$$p_{i,j} = \frac{\pi_{i,j}}{m_{i,j}} = \frac{1}{\delta} \cdot \frac{\pi_{i,j} \cdot u'_i}{u'_j}$$

(3)

The Ross recovery theorem then allows to uniquely determine $\delta$, all the $u'_i$, and the physical transition probabilities $p_{i,j}$ from the transition state prices $\pi_{i,j}$.

We illustrate the recovery theorem in a simple example with two states, state 0 and state 1. For any of the two possible initial states, the physical transition probabilities have to sum up to one:

$$p_{0,0} + p_{0,1} = 1 \iff \frac{1}{\delta} \cdot \pi_{0,0} \cdot \frac{u'_0}{u'_0} + \frac{1}{\delta} \cdot \pi_{0,1} \cdot \frac{u'_0}{u'_1} = 1$$

(4)

$$p_{1,0} + p_{1,1} = 1 \iff \frac{1}{\delta} \cdot \pi_{1,0} \cdot \frac{u'_1}{u'_0} + \frac{1}{\delta} \cdot \pi_{1,1} \cdot \frac{u'_1}{u'_1} = 1$$

We can rewrite this system of equations in matrix form and obtain the following eigenvalue problem:

$$\begin{pmatrix} \pi_{0,0} & \pi_{0,1} \\ \pi_{1,0} & \pi_{1,1} \end{pmatrix} \cdot \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} = \delta \cdot \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} \quad \text{where} \quad z_0 = \frac{1}{u'_0} \quad \text{and} \quad z_1 = \frac{1}{u'_1}$$

(5)

Given assumptions (i), (ii), and (iii), an application of the Perron-Frobenius theorem leads to the result that there is only one eigenvector $z$ with strictly positive entries $z_0$ and $z_1$. That eigenvector corresponds to the largest (and positive) eigenvalue $\delta$ of the eigenvalue problem. This property
implies a unique positive pricing kernel $m_{i,j}$ as in Equation 2 and unique physical transition probabilities $p_{i,j}$ for $i, j = 0, 1$.

In the general setting with $N$ different states, Ross defines the state transition matrix $\Pi$ with entries $\pi_{i,j}$. Each row $i$ in $\Pi$ represents state prices of moving from a particular state $i$ to any other state $j$. We always label the current state with $i = 0$ out of a set $I = \{-N_{\text{low}}, ..., 0, ..., N_{\text{high}}\}$ where $N = N_{\text{low}} + N_{\text{high}} + 1$. The ending transition state $j$ is drawn from the same set $I$. The 0-th row of $\Pi$ contains the one period transition state prices, starting from the current state, which coincide with the one period spot state prices $\pi_{0,j}$. Analogous to the two state example, we solve the following $N$-dimensional eigenvalue problem:

$$\Pi z = \delta z, \quad \text{where} \quad z_i = \frac{1}{u_i}$$

Knowing $z_i$ and $\delta$, we are then able to recover the physical transition probabilities $p_{i,j}$ from Equation 3.

### III. Methodology

The basic ingredient missing at this point is the matrix $\Pi$ of transition state prices, which are not readily observable in the market. Yet, transition state prices link spot state prices at different maturities with each other. The spot state prices $\pi_{0,j}^t$ (i.e., the subset of transition state prices that start at the current state 0 and have a time to maturity $t$) can be readily obtained from observed option quotes. We next detail our method for finding spot state prices before returning to the task of finding transition state prices from spot state prices. We finally introduce our statistical tests.

#### A. Obtaining spot state prices from observed option quotes

We collect European put and call options quotes on the S&P 500. We average bid and ask quotes to obtain midpoint quotes, which we then transform into implied volatilities. Quotes are only available for specific moneyness levels and maturities, yet we would like to obtain state prices on a
grid compatible with the recovery theorem, which often requires different levels of moneyness and different maturities. Thus, we generate a smooth implied volatility surface on a fine auxiliary grid, from which we later interpolate to the grid required for the recovery theorem.

We start with the fast and stable method of [Jackwerth (2004)], which finds smooth implied volatilities $\sigma_i$ on a fine grid of states $i$ for a fixed maturity. The fast and stable method minimizes the sum of squared second derivatives of implied volatilities (insuring smoothness of the volatility smile) plus the sum of squared differences between model and observed implied volatilities (insuring the fit to the options data) using the trade-off parameter $\lambda$.

We extend the method to volatility surfaces by adding a maturity dimension to the S&P 500 level dimension. Again, we minimize the sum of squared local total second implied volatility derivatives $\sigma_{i,t}''$ (insuring smoothness of the volatility surface and not only of the volatility smile) plus the sum of squared deviations of the model from the observed implied volatilities (insuring the fit of the surface) by using the trade-off parameter $\lambda$. We weight the squared second derivatives $(\sigma_{i,t}'')^2$ with maturity $t$ to compensate for the stronger curvature of the short-term volatility smile. The optimization problem is:

$$\min_{\sigma_{i,t}} \frac{1}{TN} \sum_{t=1}^{T} \sum_{i \in I} (\sigma_{i,t}'')^2 \cdot t + \lambda \cdot \frac{1}{L} \sum_{l=1}^{L} \left( \sigma_{i(l),t(l)} - \sigma_{i(l),t(l)}^{obs} \right)^2$$

s.t.

$$\sigma_{i,t} \geq 0$$

where $\sigma_{i(l),t(l)}^{obs}$ is the $l$–th observed implied volatility (out of $L$ observations) and $I$ is a fine set of indexes for states $i$ with a total number of $N$ states. For the state space dimension, we discretize option strike prices with a step size of $5$. We then convert strike prices into moneyness levels by normalizing them with the current level of the S&P 500 index. We set sufficient upper and lower bounds for the moneyness such that all possible states with a non-zero probability of occurrence are covered. We discretize the maturity with ten steps per month with a maximum maturity of twelve months, which gives us a total number of 120 maturity steps. This discretization insures
that all available option quotes lie on our fine grid. We then solve Equation 7 to obtain the implied volatility surface on the fine grid. We provide more details in Appendix A.

To obtain state prices on the coarser grid suitable for the recovery theorem, we linearly interpolate the fine implied volatility surface. From the implied volatilities on the coarser grid, we compute call option prices on the coarser grid and apply the Breeden and Litzenberger (1978) approach to find the spot state prices. Namely, for each maturity $t$, we take the numerical second derivative of the call prices to obtain spot state prices.

B. Finding transition state prices from spot state prices

Now that we have the spot state prices in place, we return to the task of finding transition state prices from spot state prices. Following (Ross, 2015), we can identify the transition state prices $\pi_{i,j}$ since they link spot state prices at different maturities with one another. Using monthly transition state prices and spot state prices with monthly maturities of up to one year, we have the following relations:

$$\pi_{0,j}^{t+1} = \sum_{i \in I} \pi_{0,i}^t \cdot \pi_{i,j} \quad \forall j \in I, t = 0, \ldots, 11,$$

where today’s spot state prices with a maturity of zero ($\pi_{0,i}^0$) are 0 for all states but the current state, for which the spot state price is one.

Equation 8 states that one can find the spot state price $\pi_{0,j}^{t+1}$ of reaching state $j$ at maturity $t + 1$ by adding up all the state price contributions of visiting state $i$ one month earlier at maturity $t$ ($\pi_{0,i}^t$) times the transition state price from $i$ to $j$ ($\pi_{i,j}$). We can exactly solve for all the transition state prices $\pi_{i,j}$, if the number of states $N$ is equal to the number of (non-overlapping) transitions. With twelve transitions of one month each, we cover a whole year of maturities and can solve for

---

8 As we lose the lowest and the highest index level due to the numerical second derivative, we choose the initial coarse grid to be one index level too wide on either end. The spot state prices will then be exactly on the desired coarser grid.
only twelve states. Ross (2015) already mentions that such coarse non-overlapping grid leads to poorly discretized transition state prices and to coarse discrete physical probabilities based on the recovery theorem.

**Ross Basic**

Instead of using the coarse non-overlapping grid of dimension twelve-by-twelve, we follow Audrino et al. (2015) and apply an overlapping approach to determine transition state prices. Based on steps of one-tenth of a month (and a state price transition lasting one month, i.e., ten steps), our new relation is:

\[
\pi_{t+10}^{0,j} = \sum_{i \in I} \pi_{t}^{0,i} \cdot \pi_{i,j} \quad \forall j \in I, t = 0, \ldots, 110.
\]  

This results in a total number of 111 overlapping transitions and thus allows \( N = 111 \) states, which we choose equidistant and where we include the current state \( i = 0 \). Directly solving Equations 9 is not advisable as the problem is ill-conditioned. Rather, we impose an additional non-negativity constraint on the transition state prices \( \pi_{i,j} \) and we back them out from the the following least squares problem, which penalizes violations of Equations 9:

\[
\min_{\pi_{i,j}} \sum_{j \in I} \sum_{t=0}^{110} \left( \pi_{t+10}^{0,j} - \sum_{i \in I} \pi_{t}^{0,i} \cdot \pi_{i,j} \right)^2 \quad s.t. \quad \pi_{i,j} > 0 \]  

We collect the transition state prices \( \pi_{i,j} \) in the transition state price matrix \( \Pi \) and recover the matrix \( P = [p_{i,j}]_{i,j \in I} \) of physical transition probabilities by applying the recovery theorem of Equation 6 to \( \Pi \). We label this version of recovery **Ross Basic**.

We depict the results of our implementation for a typical day in our sample, February 17, 2010. Figure 1, Panel A, shows the interpolated smoothed implied volatility surface on the 111 by 111

---

9We use the non-overlapping approach on the twelve-by-twelve state space in the robustness section VII.A. As expected, the approach does not work and the realized future returns do not seem to be drawn from the recovered physical probability distribution.
state space. We note that the volatility surface is smooth yet passes through the observed implied volatilities (black squares). The volatility smile is clearly visible for short maturities; flattening out at longer maturities. Panel B shows the related spot state price surface, which also turns out to be smooth. Spot state prices tend to be high around the current state (moneyness of one) and are more spread out at larger maturities.

Figure 2 Panel A, illustrates the transition state prices, which best relates spot state prices at one maturity to those at a maturity one month later. We would expect large transition state prices on the main diagonal, as it is more likely to end up at states \( j \) which are close to the initial state \( i \). However, the optimization quite often generates large state prices for states that are far away from the current state. This pattern of spurious transition state prices away from the main diagonal carries over to some extent to the recovered physical transition probabilities in Panel B of Figure 2.

The large transition state prices away from the main diagonal can occur because short maturity option quotes are hardly affected by such irrelevant transition state prices, which link states that are not important for the short maturity spot state prices and thus the value of the short maturity options. The optimization can thus allocate mass to these irrelevant states to minimize the objective function, while not changing the short maturity spot state prices much in the process. As a result, some rowsums in \( \Pi \) have values much higher than one, which would imply high negative risk-free rates for some initial states. On a typical sample day, February 17, 2010, we observe that one fourth of all one-month state-dependent risk-free rates are lower than -50% with a lowest value of -92% (-100% annualized), and that one fourth of the rates are higher than 30% with a highest value of 165% (11,993,526% annualized). As the problem is very ill-conditioned, we impose further economic restrictions.

**Ross Bounded**

Namely, we first demand all rowsums of \( \Pi \) to lie in the interval \([0.9, 1]\). As the inverse of the rowsum is equal to one plus the risk-free rate for this state, we limit the monthly risk-free rates to between 0% and 11.11% (0% and 254.07% annualized). We again solve Equation 10 but additionally restrict...
Figure 1. **Implied volatility and spot state price surfaces.** We show the interpolated implied volatility surface and the observed implied volatilities (as black dots) in Panel A. We show the related spot state price surface in Panel B. Based on data from February 17, 2010, we depict these surfaces for S&P 500 index options across maturities and moneyness levels.
Figure 2. Transition state prices and recovered physical transition probabilities, Ross Basic. Transition state prices (Panel A) and corresponding recovered transition probabilities (Panel B) on the February 17, 2010 identified by using Ross Basic.
the rowsums, before continuing with the recovery theorem. We label this version **Ross Bounded**. Figure 3 illustrates the transition state prices in Panel A and the corresponding recovered physical transition probabilities in Panel B. Both surfaces are now highly concentrated around the current state. On the positive side, this eliminates high values in irrelevant states (i.e., far away from the main diagonal). Yet worryingly, even the values on the main diagonal fall off as we move away from the current state. The optimization allocates the restricted state prices in an almost uniform way for very low and very high states. As a result, we do not obtain the economically reasonable diagonal structure for the transition state prices with Ross Bounded.

**Ross Unimodal**

Next, we also outright force the rows in matrix $\Pi$ to be unimodal with maximal values on the main diagonal. Once more, we solve Equation 10 but add the requirement of unimodality and that all rowsums of $\Pi$ lie in the interval $[0.9, 1]$. We then proceed with the recovery theorem and label this version **Ross Unimodal**.

Related to our Ross Unimodal approach, Massacci et al. (2016) extract a 11 by 11 transition state price matrix from intraday S&P 500 option data and from Apple stock option data, respectively, and also force transition state price matrices to have unimodal rows with maximum values on the main diagonal.

Figure 4 shows the transition state prices in Panel A and the corresponding recovered physical transition probabilities in Panel B for Ross Unimodal. By construction, the highest values line the main diagonal, steeply falling off further away from the main diagonal. However, higher values are again concentrated around the current state. This behavior is again driven by the optimization.
Figure 3. Transition state prices and recovered physical transition probabilities, Ross Bounded. Transition state prices (Panel A) and corresponding recovered transition probabilities (Panel B) on February 17, 2010 identified by using Ross Bounded.
Figure 4. Transition state prices and recovered physical transition probabilities, Ross Unimodal. Transition state prices (Panel A) and corresponding recovered transition probabilities (Panel B) on February 17, 2010 identified by using Ross Unimodal.
C. Recovery without using transition state prices, Ross Stable

The computation of transition state prices is a key challenge in the application of the recovery theorem. Yet, the one row in the transition state price matrix $\Pi$ associated with the current state $i = 0$ offers a novel way out. For this current state, the transition state prices ought to coincide with the one period spot state prices, which we readily obtain from option quotes. We use this insight to suggest an alternative recovery method that does not require explicitly solving for the transition state prices.\(^{10}\) The trick is using the eigenvalue problem in the recovery theorem as in Equation 6 and multiplying both sides from the left with the transition state price matrix $\Pi$:

$$\Pi \cdot \Pi z = \Pi \cdot \delta z = \delta (\Pi z) = \delta^2 z. \quad (11)$$

Again, the row of $\Pi^2$ associated with the current state $i = 0$ contains the spot state prices, but now with a maturity of two transition periods. The discount factor $\delta$ now appears in the second power to account for the two periods. Iterating, we obtain the following relation:

$$\Pi^t z = \delta^t z \text{ with } t = 1, ..., T, \quad (12)$$

where $t$ determines how often we apply the transition. For each $t$, we focus on the row in $\Pi^t$ associated with the current state $i = 0$, where the $t$-period transition state prices coincide with the $t$-period spot state prices. We collect all those current rows with different maturities $t$. Full identification requires at least as many equations for different maturities $t$ as there are number of states $N$, which results in the following system of equations (see Appendix B for details):

\(^{10}\)See the independent derivation in [Jensen et al. (2016)].
We are worried that the system of equations is ill-conditioned and, thus, might violate sensible economic constraints. Namely, we want to insure that the utility discount factor $\delta$ and the resulting pricing kernel are non-negative. Thus, we penalize deviations from Equation 13 and include the two new constraints:

$$
\begin{align*}
\min_{\delta} \sum_{t=1}^{T} \left( \sum_{j \in I} \pi_{t,0,j} \cdot \left( \frac{z_{j}}{z_{0}} - \delta^{t} \right) \right)^{2} \quad &s.t.\quad \left( \frac{z_{j}}{z_{0}} \right) > 0, \quad 1 > \delta > 0.
\end{align*}
$$

We still need to decide on the length of the transition period. Using twelve non-overlapping periods of each month gives us only twelve transitions and allows for at most twelve states. This results in too coarse a grid.\(^{11}\) Instead, we use 120 periods of one-tenth of a month each. Here, we make use of the property that the structure of the pricing kernel in the setting of Ross (2015) remains the same for different maturities and only varies by a factor:\(^{12}\)

$$
m_{0,j}^{*} = \frac{1}{\delta^{t-1}} \cdot m_{0,j}, \quad (15)
$$

\(^{11}\)We still use the coarse grid in the robustness Section VII.A but, as expected, future realized returns are incompatible with the recovered physical probability distribution.

\(^{12}\)See Appendix B for details.
where \( m_{t,j} \) is the spot pricing kernel with a maturity of \( t \) transition periods and \( m_{0,j} \) is the spot pricing kernel with a maturity of one transition period. This allows us to use a fine moneyness grid defined on 120 points, corresponding to 120 maturities that are spaced one-tenth of a month apart. Once we solve Equation 14 for the one period pricing kernel, we use Equation 15 to find the ten period (i.e., one-month) pricing kernel. We use this one-month pricing kernel to transform one-month spot state prices into one-month physical probabilities. We label this approach Ross Stable. Note that we cannot provide the corresponding figures for the transition state prices and the transition physical probabilities as we do no longer compute them explicitly.

D. Testing the recovered physical probabilities

We finally have the recovered physical probabilities in hand and want to investigate how good these forecasts are, which are solely based on the option quotes and do not use the historical time series of index returns. Our hypothesis is that:

\[ H_0: \text{Future realized monthly S&P 500 returns are drawn from the recovered physical distribution.} \]

We test our hypothesis as follows. Each month, we find the date \( \tau \) which is 30 calendar days before the option expiration date. Record the realized future return (i.e., the return from date \( \tau \) to the expiration date) on the S&P 500 and label it \( R_\tau \). That return is one realization from the true physical distribution \( p_\tau \). Next, we turn to the recovery theorem. We recover the physical spot distribution \( \hat{p}_\tau \) and the corresponding cumulative distribution \( \hat{P}_\tau \) for date \( \tau \). We can find the percentile of the recovered cumulative distribution \( \hat{P}_\tau \) that corresponds to the realized return \( R_\tau \) and we collect those percentiles \( u_\tau \) for all dates. Under the assumption that the recovered distribution is indeed the one from which the return was drawn (i.e., \( \hat{p}_\tau = p_\tau \)), the percentiles

\begin{equation}
\text{13 As we solve for the pricing kernel with the least squares approach of Equation 14 the system of Equations 13 does not hold exactly. As a result, the recovered physical spot probabilities do not necessarily sum up to one, and so we normalize them.}
\end{equation}

\begin{equation}
\text{14 The values of the recovered cumulative distributions } \hat{P}_\tau \text{ lie on a discrete grid, whereas the future realized returns typically do not lie on the corresponding grid points. We therefore interpolate the recovered cumulative distribution linearly to obtain the transformed points } u_\tau.
\end{equation}
should be i.i.d. uniformly distributed.\footnote{See Bliss and Panigirtzoglou (2004) and Cuesdeanu and Jackwerth (2016).}

To check on the uniformity of the percentiles $u_\tau$, we use three different test: The Berkowitz test, the Knüppel test, and the Kolmogorov-Smirnov test, see Appendix C for details. The Berkowitz test has been used in Bliss and Panigirtzoglou (2004). In that study, the authors argue that it is superior to the Kolmogorov-Smirnov test in small samples with autocorrelated data. While being more powerful than the Kolmogorov-Smirnov test, the Berkowitz test bases its conclusion about standard normality on just the first and the second moments, while ignoring higher moments. The Knüppel (2015) test has the advantage of testing for higher moments, can deal with autocorrelated data, and still has enough power to test smaller samples.\footnote{We do not consider the Cramer-van-Mises test as it is an integrated version of the Kolmogorov-Smirnov test and yields very similar results.}

We apply our three tests of uniformity in combination with all recovery versions (Ross Basic, Ross Bounded, Ross Unimodal, and Ross Stable) to see if the recovered physical distributions are compatible with future realized returns. For comparison, we use two additional benchmark models for the physical distribution. For one, we use the empirical cumulative distribution of the past five years of monthly S&P 500 returns, labeled Historical Return Distribution. For the other, we assume a representative investor having a power utility with a risk aversion coefficient of four. Based on the associated pricing kernel, we transform the spot state prices into a physical distribution. For comparability, we use the same one-month spot state prices as in Ross Stable, which lie on a moneyness grid defined on 120 points. We label this approach Power Utility with $\gamma = 4$.\footnote{The results in Bliss and Panigirtzoglou (2004) suggest that a risk aversion factor of four in a power utility framework is a reasonable choice.}

IV. Data

We use OptionMetrics to obtain end-of-day option data on the S&P 500 index. Consistent with the literature, we only use out-of-the-money put- and call options with a positive trading volume

\begin{footnotesize}
\footnote{See Bliss and Panigirtzoglou (2004) and Cuesdeanu and Jackwerth (2016).}
\footnote{We do not consider the Cramer-van-Mises test as it is an integrated version of the Kolmogorov-Smirnov test and yields very similar results.}
\footnote{The results in Bliss and Panigirtzoglou (2004) suggest that a risk aversion factor of four in a power utility framework is a reasonable choice.}
\end{footnotesize}
and eliminate all options that violate no arbitrage constraints. We fit the implied dividend yield according to put-call parity where the risk free rate is given by the interpolated zero-curve from OptionMetrics. We consider monthly dates $\tau$, which we find by going 30 calendar days back in time from the expiration date. We end up with 223 recovered distributions and realized returns. Our sample period, as well as our option data, ranges from January 1996 to August 2014. We further obtain end-of-day S&P 500 index levels from Datastream. We compute S&P 500 returns from January 1991 to August 2014, which includes a five year period prior to January 1996. That additional data is needed for the historical return distribution.

V. Empirical Results

We are now ready to investigate our central question. Are future realized S&P 500 returns drawn from recovered physical probabilities? The sobering answer is in Table I. All four versions of the recovery theorem (Ross Basic, Ross Bounded, Ross Unimodal, and Ross Stable) strongly reject our null hypothesis (p-values less that 0.028) for all three tests, which we employ (Berkowitz, Knüppel, and Kolmogorov-Smirnov).

In contrast, our simple benchmark models (Power Utility with $\gamma = 4$ and Historical Return Distribution) are not rejected by any of our three tests (p-values of more than 0.294). We are thus facing a complete failure of the recovery theorem, while a power utility setting or even a five year histogram of historical returns cannot be rejected by the data.

To better understand what drives our results, we take a closer look at the recovered probabilities in Figure 5. We start our discussion with Power Utility in Panel E, as it ”works” in explaining the data and is particularly simple. The black line shows the spot state prices, which are derived from the option quotes and are the same in all six panels.

\footnote{There is a tiny difference in how they plot as we have 111 states in Panels A-C and we have 120 states in Panels D and E.} Note that the risk-neutral distribution looks just the same as the the one-month risk-free rate adjustment is invisible in the Figure. The method changes those spot state prices into physical probabilities (dashed gray), for which we cannot
Table I. Tests of the recovered physical probabilities. We present our results if future realized returns are drawn from physical probabilities generated by one of our six approaches: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each approach, we show the $p$-values from the Berkowitz, Knüppel, and Kolmogorov-Smirnov tests for uniformity of the percentiles of future realized returns under the model physical cumulative distribution.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>Berkowitz $p$-value</th>
<th>Knüppel $p$-value</th>
<th>Kolmogorov-Smirnov $p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross Basic</td>
<td>0.018</td>
<td>0.027</td>
<td>0.000</td>
</tr>
<tr>
<td>$\pi_{i,j} &gt; 0$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Bounded</td>
<td>0.005</td>
<td>0.002</td>
<td>0.008</td>
</tr>
<tr>
<td>$\pi_{i,j} &gt; 0$, rowsums $\in [0, 1]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Unimodal</td>
<td>0.001</td>
<td>0.000</td>
<td>0.028</td>
</tr>
<tr>
<td>$\pi_{i,j} &gt; 0$ and unimodal, rowsums $\in [0, 1]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Stable</td>
<td>0.010</td>
<td>0.015</td>
<td>0.004</td>
</tr>
<tr>
<td>Do not use transition state prices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power Utility</td>
<td>0.697</td>
<td>0.320</td>
<td>0.547</td>
</tr>
<tr>
<td>$\gamma = 4$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical Return Distribution</td>
<td>0.294</td>
<td>0.480</td>
<td>0.347</td>
</tr>
</tbody>
</table>
Figure 5. State prices and recovered physical probabilities. We depict spot state prices (black lines), transition state prices (light gray lines), and recovered physical probabilities (gray dashed lines) on February 17, 2010. Our methods are Ross Basic in Panel A, Ross Bounded in Panel B, and Ross Unimodal in Panel C, Ross Stable in Panel D, Power Utility with $\gamma = 4$ in Panel E, and the kernel-smoothed Historical Return Distribution in Panel F.
reject our hypothesis that future realized returns are drawn from it. A successful method thus needs to right-shift the physical probabilities to be compatible with the data (implying a positive market-risk premium). The Historical Return Distribution in Panel F also works in that our main hypothesis cannot be rejected. The physical distribution is the kernel-smoothed histogram of the one-month non-overlapping S&P 500 returns from the past five years.\textsuperscript{19} The physical distribution is less smooth than in Panel E, as some of the jaggedness of the histogram still remains present in the kernel distribution.

We now turn to the recovery methods (Panels A-C) that have an additional light gray line for the transition state prices. This is because the models allow for the transition state prices to differ from the spot state prices, thus incorporating a pricing error for the option quotes. Yet Ross Basic in Panel A does not use the possibility and the spot and transition state prices plot on top of each other. The reason can be found in the lack of economic constraints on the transition state prices (other than positivity), which then allows the optimization to closely follow the spot state prices. As a result, Ross Basic exhibits implausible fluctuations for the transition state prices and rowsums, which imply extremely large negative or positive monthly risk-free rates, ranging from -92\% to 165\% on a typical sample day (February 17, 2010). The physical distribution is somewhat right-shifted but insufficiently so, as we reject our main hypothesis for Ross Basic as for all other recovery methods.

Adding economic constraints on the rowsums in Ross Bounded (Panel B) and unimodality in Ross Unimodal (Panel C) leads to a separation of the transition state prices from the spot state prices. Also, the physical probabilities are almost identical to the transition state prices, caused by an almost flat (model) pricing kernel. As a result, the physical distribution is too close to the spot state prices and we reject our main hypothesis. Finally, in Ross Stable, we do not explicitly compute transition state prices. The physical probabilities are again very close to the spot state prices, caused by an almost flat (implied) pricing kernel and we reject our main hypothesis again.

\textsuperscript{19}The kernel smoothing uses Matlab’s \texttt{ksdensity} with the default bandwidth, see \textit{Bowman and Azzalini} (1997). New York: Oxford University Press Inc., 1997. For plotting, we interpolate the kernel smoothed density onto 120 states that we use for Power Utility and Ross Stable and normalize the result such that it sums to one.
for Ross Stable.

We arrive at two main findings. First, as opposed to our simple benchmark models, all recovery approaches are incompatible with future realized S&P 500 returns. It seems that the recovery theorem cannot generate a sufficiently high risk premium and non-flat pricing kernels. Second, imposing economic structure on the transition state prices reduces the fit to observed option quotes and leads to physical distributions, which are rather similar to the spot state prices.

VI. Reasons for Failure

We now investigate in more detail why the recovery theorem empirically fails. We first analyze the recovered pricing kernels, uncover how those pricing kernels depend on transition state prices, and show that the recovered pricing kernels are often too flat. Second, we look into the empirical implications of time-homogeneous transition state prices, which leads to poorly fitted longer-dated options. Third, we simulate an economy in which Ross recovery holds, complete with option prices and future realized returns. We then perturb the option prices and test if future returns are compatible with the recovered physical distribution based on the perturbed option prices.

A. Recovered Pricing Kernels

To better understand our empirical findings, we first analyze our recovered pricing kernels. Yet how should the pricing kernel look across values of S&P 500 returns? From basic theory, we expect the pricing kernel to be positive and monotonically decreasing, and to reflect the behavior of a risk averse representative investor. Jackwerth (2004), Ait-Sahalia and Lo (2000), and Rosenberg and Engle (2002), however, find that the empirical pricing kernel is locally increasing, a behavior that is referred to as the pricing kernel puzzle. In these papers, physical probabilities are backed out from past S&P 500 index returns, while Ross recovery implies pricing kernels based on forward-looking information. However, Cuesdeanu and Jackwerth (2016) confirm the existence of the pricing kernel puzzle in forward-looking data. Thus, we might expect the recovered pricing kernels to be
monotonically decreasing (standard case) or to be locally increasing (pricing kernel puzzle).

Figure 6 shows the pricing kernels obtained from different recovery approaches on February 17, 2010. The black line shows the \textit{implied} pricing kernel for each approach, measured as \textit{spot} state prices divided by the recovered physical probabilities. We start with Power Utility in Panel E, as the pricing kernel is, by construction, monotonically decreasing and well-grounded theoretically. From our main result, we also know that this pricing kernel translates the spot state prices into physical probabilities that are compatible with future realized returns. In a way, the power pricing kernel ”works”, while the pricing kernels based on the recovery theorem do not. We now try to understand why the latter do not work.

Most similar to the power pricing kernel is the pricing kernel for Ross Basic in Panel A. It is not very smooth but somewhat decreasing, yet less so than the power pricing kernel. As a result, the shift from state prices to physical probabilities is insufficient in that we reject our hypothesis that future realized returns are drawn from the recovered physical distribution. Here, we also depict as a gray line the \textit{model} pricing kernel, measured as \textit{transition} state prices for the current state divided by the recovered physical probabilities. Any difference in the two pricing kernels would be due to the optimization not being able to exactly match the spot state prices (and thus the observed option quotes). For Ross Basic, this is not an issue as the optimization is free to fit option quotes as long as the transition state prices are positive. We learned that this freedom comes at the cost of extreme rowsums, which in turn lead to extreme monthly risk-free rates, ranging from -92\% to 165\%.

Once we implement reasonable economic constraints in Ross Bounded and Ross Unimodal (Panels B and C), the implied pricing kernels become even more wavy overall and flatter for center moneyness levels of about 0.9 to 1.1. The recovered physical probabilities remain closer to the spot state prices and move further away from the distribution of future realized returns\footnote{See e.g. Bliss and Panigirtzoglou\textsuperscript{[2004]}, who confirm established results in the literature, that risk-neutral probabilities perform worse in forecasting the distribution of future returns.}. Interestingly now, the implied and the model pricing kernels fall apart, indicating that the optimization struggles to match the spot state prices as the transition state prices now need to satisfy our economic
Figure 6. Pricing kernels. We present pricing kernels for different recovery approaches on February 17, 2010. Black lines depict *implied* pricing kernels, measured as spot state prices divided by recovered probabilities, while gray lines depict *model* pricing kernels measured as transition state prices for the current state divided by recovered probabilities. Panel A shows the pricing kernels for Ross Basic, Panel B for Ross Bounded, Panel C for Unimodal, Panel D for Ross Stable (only the implied kernel exists), Panel E for Power Utility with $\gamma = 4$ (only the implied kernel exists), and Panel F for Historical Return Distribution (only the implied kernel exists).
constraints. The model pricing kernel (the part driven by the recovery theorem and not due to the fit of option quotes) is now virtually flat. The recovery theorem thus cannot generate a decreasing pricing kernel any more, when it is even gently constrained. Note that the requirement of monthly risk-free rates to lie between 0% and 11.11% is not very onerous.\footnote{The requirement of unimodality is also not very strong and only applies to Ross Unimodal in Panel C, but not to Ross Bounded in Panel B.}

We shed further light on the flat pricing kernels once we realize that the rowsums of the transition state price matrix Π are strongly negatively related to the model pricing kernel.\footnote{Audrino et al. (2015) also noticed this relation.} Thus, for a decreasing pricing kernel (such as the power pricing kernel, which works well empirically), we would need increasing rowsums (i.e., decreasing interest rates) across states. Yet even our modest economic constraints on the rowsums render the model pricing kernels virtually flat. Ross recovery thus does not seem to be capable of generating non-flat model pricing kernels without unreasonably extreme risk-free rates.

We also find a flat implied pricing kernel for Ross Stable in Panel D.\footnote{The model pricing kernel does not exist as we never explicitly compute the transition state prices.} Yet that flatness stems partially from the numerical implementation. We first recover a pricing kernel with maturity of 0.1 months to allow a larger number of 120 states (instead of only twelve), and then convert the 0.1-month pricing kernel into a one-month pricing kernel by the multiplicative adjustment of Equation 15. The negligible curvature of the 0.1-month pricing kernel then directly translates into an almost flat one-month pricing kernel, which again implies that we recover physical probabilities that are close to the spot state prices.

Last, the implied pricing kernel for Historical Return Distribution in Panel F is rather irregular on this particular day, even after we smoothed the historical distribution through a kernel density. Yet in general and across our whole sample, we cannot reject our main hypothesis and the implied pricing kernels manages to translate the spot state prices into generally right-shifted physical distributions.

We conclude that one major problem with Ross recovery seems to be its inability to generate sloped pricing kernels. This is even the case in the unrestricted Ross Basic setting and even more
so in the constrained versions, which put mild economic constraints on the rowsums (i.e., risk-free interest rates).

B. Time-Homogeneity of Transition State Prices and the Fitting of Option Quotes

From Figure 5 we learned that Ross stable and Power Utility use the spot state prices directly, while Ross Basic, Ross Bounded, and Ross Unimodel fit the current transition state prices to those state prices, allowing for some discrepancy. How large is that discrepancy? As a measure, we compare one-month observed implied volatilities $\sigma^{obs}$ with one-month model implied volatilities $\hat{\sigma}$, where we use the one-month state prices of each approach to find model option prices and their associated model implied volatilities; see Appendix D for details. For each date $\tau$, we compute the root-mean-squared error between the two implied volatilities. Our final measure $MRMSE$ is the time-series average of all the $RMSE_{\tau}$. Table II shows the one-month $MRMSE$ for each recovery approach in columns 2 and 3.

We find the lowest one-month $MRMSE$ of 0.008 for the methods Ross Stable and Power Utility, which directly use the spot state prices and which do not use transition state prices. This is not surprising, as the error in that cases only stems from smoothing and interpolating the implied volatility surface. This low error also demonstrates the ability of our smoothing and interpolation methodologies to approximate observed implied volatilities reasonably well.

Among the transition state price versions, Ross Basic has only a slightly higher $MRMSE$ of 0.009. As we only require positivity of the transition state prices, the optimization can freely choose transition state prices to match the spot state price surface. This results in a good fit of one-month spot state prices but comes at the price of having rowsums in the transition state price matrix that (for a typical sample day, February 17, 2010) lead to unreasonably low (-92%) or high (165%) monthly state-dependent risk-free rates.

For the other two transition state price versions (Ross Bounded and Ross Unimodal), we find more than 15 times higher $MRMSE$s, 0.141 and 0.160. The transition state prices for those
Table II. Accuracy of Transition State Prices. We present a measure for how closely the observed option quotes are fitted by each of the five approaches: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, and Power Utility with $\gamma=4$. Our measure is the time-series average of the root-mean-squared errors of model versus observed implied volatilities, $MRMSE$. Columns 2 and 3 present the $MRMSE$ for a one-month maturity, where model implied volatilities are either based on transition state prices or spot state prices. Columns 4 and 5 present the $MRMSE$ for a twelve-month maturity, where model implied volatilities are either based on conflated transition state prices or spot state prices.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>One-month transition state prices MRMSE</th>
<th>One-month spot state prices MRMSE</th>
<th>12-month transition state prices MRMSE</th>
<th>12-month spot state prices MRMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross Basic</td>
<td>$\pi_{i,j} &gt; 0$</td>
<td>0.009</td>
<td>0.050</td>
<td></td>
</tr>
<tr>
<td>Ross Bounded</td>
<td>$\pi_{i,j} &gt; 0$, rowsums $\in [0.9, 1]$</td>
<td>0.141</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td>Ross Unimodal</td>
<td>$\pi_{i,j} &gt; 0$ and unimodal, rowsums $\in [0.9, 1]$</td>
<td>0.160</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td>Ross Stable</td>
<td>Do not use transition state prices</td>
<td>0.008</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Power Utility</td>
<td>with $\gamma = 4$</td>
<td>0.008</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Historical Return Distribution</td>
<td>Not defined</td>
<td>Not defined</td>
<td>Not defined</td>
<td>Not defined</td>
</tr>
</tbody>
</table>
methods are almost useless for pricing, as the MRMSE is of the same magnitude as the implied volatility itself. Adding even mild economic constraints removes many degrees of freedom in the allocation of transition state prices. Therefore, choices for the transition state prices for the current state are more limited, increasing the $MRMSE$.

We further want to investigate how the assumption of time-homogeneous transition state prices influences our empirical results. Given time-homogeneity, we may multiply the one-month-maturity transition state price matrix $\Pi$ $t$-times with itself to again generate transition state prices (and thus also spot state prices) with a maturity of $t$ months. Empirically, we are concerned that the longer-dated transition state prices will not fit the option quotes well, as there is little evidence so far that longer-dated options are priced on the conflated state prices of shorter dated options.

For the recovery methods using transition prices (Ross Basic, Ross Bounded, and Ross Unimodal) and on every date, we conflate the transition state price matrix $\Pi$ twelve times with itself to extract longer-dated spot state prices. Similarly to the one-month horizon, we compute the root-mean-squared differences of model and observed option implied volatilities. Last, we take the time series mean of the $RMSEs$ over all dates with available twelve-month maturity option data to arrive at the twelve-month $MRMSE$. For the methods using the spot state prices directly (Ross Stable and Power Utility), we proceed analogously but use the twelve-months spot state prices instead of the transitions state prices. The results we report in columns 4 and 5 of Table II.

We find the lowest twelve-month $MRSME$ with $0.004$ for Ross Stable and Power Utility, being even lower that the one-month $MRMSE$ of those methods ($0.008$). The reason is that the longer-dated distributions cover a greater range of states and thus approximate the observed spot prices better. The twelve-month $MRMSE$ ($0.062$ and $0.065$) for Ross Bounded and for Ross Unimodal are again some 15 times the values for Ross Stable and Power Utility. Most interestingly, while the transition state prices for Ross Basic were able to approximate one-month option prices well, the conflated 12-month transition state prices imply a large $MRMSE$ of $0.050$.

We conclude that Ross Basic, Ross Bounded, and Ross Unimodal fail to produce transition

\footnote{Again, to keep our discussion simple, we do not distinguish between risk-neutral probabilities and transition state prices because numerically, the small risk-free rate adjustment hardly makes any difference.}
spot state prices that approximate option prices well, with only Ross Basic being able to price the one-month options reasonably well at the cost of implausibly shaped transition state prices and extreme risk-free interest rates.

C. Insights from Simulated Economies

We are concerned that small data errors in the option quotes might cause the recovery theorem to fail. To check, we simulate economies, where a particular recovery method holds true, and generate future realized returns by drawing from the recovered physical distribution. In that case and with a 5% significance level for our statistical tests, we should have a 95% non-rejection rate (i.e., future returns are compatible with the simulated recovery method) and a 5% rejection rate.

Next, we perturb the option quotes. This will decrease the non-rejection rates as the perturbed recovery methods generate physical distributions, which will deviate from the true physical distribution, from which we drew the future realized returns.

Our methodology is as follows. For each date $\tau$ in our sample, we assume that the true risk-neutral economy is represented by the observed implied volatilities $\sigma_{i(l), t(l)}$ with moneyness $i(l)$ and maturity $t(l)$ for observation $l = 1, ..., L$. At date $\tau$, we further draw a one-month future return $R_{\tau}^{true}$ from the physical distributions $p_{\tau}^{true}$ that we recovered with a particular recovery method. That gives us a time series of 223 drawn one-month future realized returns from an economy where the particular recovery method holds.

On each date, we then add $v$ times a standard normally distributed error $\epsilon$ to the observed implied volatilities $\sigma_{i(l), t(l)}$ and obtain perturbed implied volatilities $\hat{\sigma}_{i(l), t(l)}$ as:

$$\hat{\sigma}_{i(l), t(l)} = \sigma_{i(l), t(l)} + v \cdot \epsilon, \quad \epsilon \sim N(0, 1).$$

(16)

We apply our standard algorithm to recover the physical distributions $\hat{p}_{\tau}$, which are now based on the perturbed implied volatilities $\hat{\sigma}_{i(l), t(l)}$. As in our empirical implementation, we then use a
Knüppel test to test the hypothesis that the simulated future returns $R_{true}^\tau$ are compatible with the recovered distributions $\hat{p}_\tau$. We draw 1000 realizations of future return time series, each consisting of 223 returns, and report the non-rejection rates for our hypothesis at the 5% significance level.

We use four different levels of perturbation, based on the multiplier $v$ of the error term $\epsilon$ in Equation 16. We express $v$ on each date $\tau$ as a multiple of the mean implied volatility bid/ask spread on that date. Our four levels of perturbation are none, one-half, one, and twice the mean spread of a particular date.

**Table III. Stability of a Simulated Recovery Methods.** We generate 1000 realizations of a economy, where a particular recovery method holds. We generate 1000 times series of 223 future realized returns. Next, we perturb option implied volatilities by adding $v$ times a standard normally distributed error $\epsilon$. Given $v$, we report the non-rejection rates for the hypothesis that the future realized returns were drawn from the perturbed physical distributions. We use a Knüppel test with a significance level of 5% level.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>$v = 0$-spread</th>
<th>$v = \frac{1}{2}$-spread</th>
<th>$v = 1$-spread</th>
<th>$v = 2$-spread</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>non-rej. rate</td>
<td>non-rej. rate</td>
<td>non-rej. rate</td>
<td>non-rej. rate</td>
</tr>
<tr>
<td>Ross Basic</td>
<td>95.0 %</td>
<td>15.1 %</td>
<td>8.6 %</td>
<td>5.6 %</td>
</tr>
<tr>
<td>Ross Bounded</td>
<td>95.7 %</td>
<td>92.8%</td>
<td>88.8 %</td>
<td>61.4 %</td>
</tr>
<tr>
<td>Ross Unimodal</td>
<td>95.3 %</td>
<td>94.1 %</td>
<td>88.6 %</td>
<td>67.6 %</td>
</tr>
<tr>
<td>Ross Stable</td>
<td>95.0 %</td>
<td>95.6 %</td>
<td>93.4%</td>
<td>74.6%</td>
</tr>
<tr>
<td>Power Utility</td>
<td>96.4 %</td>
<td>91.7 %</td>
<td>88.4%</td>
<td>73.5 %</td>
</tr>
<tr>
<td>Hist. Return Distr.</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table III presents how often our hypothesis is accepted at the 5% significance level for different multipliers $v$ of the error-term $\epsilon$. Without any perturbation, the non-rejection rates are all very close to the theoretical value of 95%. Beyond that, we find that only Ross Basic turns out to be very sensitive to perturbations. Even though we simulate under the assumption that Ross Basic holds, we only have non-rejection rates of around 5% to 15%.

---

25 Using the Berkowitz test at the 5% level instead does not change our results.

26 On each date, we compute the mean across all options of (ask implied volatility - bid implied volatility). The mean spread ranges from 0.007 to 0.061 through our 223 sample date with an average value of 0.016.
with perturbations instead of 95% without. Thus, the failure of Ross Basic could also be
driven by errors in the option implied volatilities.

Yet for all other methods, the decay of the non-rejection rates is much slower as we
increase our perturbations. Using the mean spread of that date as the multiplier \( v \) of our per-
turbation term \( \epsilon \), non-rejection rates are above 88.4%. Thus, Ross Bounded, Ross Unimodal,
Ross Stable, and Power Utility should all be similarly sensitive to perturbation. However,
only Power Utility cannot be rejected as a model of future realized returns, while we reject
all recovery methods. We conclude that the failure of Ross Bounded, Ross Unimodal, and
Ross Stable is not likely to be driven by perturbations of the option quotes.

We conclude that while Ross Bounded, Ross Unimodal, and Ross Stable suffer from an
almost flat model pricing kernel, the methods are less sensitive to perturbations of the option
quotes than Ross Basic. It seems that adding economic constraints results in models, which
are less sensitive to perturbations. This finding also holds for Power Utility.

VII. Robustness

To demonstrate the stability of our results, we implement two types of robustness checks.
First, we investigate if variations in the state space significantly influence our empirical
results. Second, we repeat our study but now exclude the periods related to (i) the Dot-Com
bubble and (ii) the Financial Crisis period from our original sample.

A. Variations in the State Space

The recovery theorem works on a finite state space and we are concerned that our method-
ological choices for the states might drive the results. We proceed to show that this is not the
case when we use (i) log returns instead of straight returns or (ii) a 20% coarser state space

\footnote{The non-rejection rate cannot be computed for Historical Return Distribution.}
than our usual fine state spaces with $N=111$ (120 for Ross Stable and Power Utility). If we (iii) use the non-overlapping state space with $N=12$ from the original [Ross, 2015] paper, then all recovery methods and also Power Utility are being rejected. The fit to the option quotes deteriorates markedly in the process. Note that the Historical Return Distribution is always unaffected by the choice of the state space and we thus always maintain our result that we cannot reject the Historical Return Distribution.

**Log returns**

We repeat our main empirical study but now define our state-space in log-returns. We construct our volatility surface by linearly interpolating the log-moneyness of the fine implied volatility surface. As before, we recover the physical distributions but based on the log state space. We then test if the one-month future realized log-returns are drawn from the recovered distributions. For the Historical Return Distribution we generate the empirical cumulative distribution on each sample date based on a five-year historical sample of one-month non-overlapping log-returns.

Table IV shows the resulting $p$-values for our three test statistics and can be readily compared to Table I. For Ross recovery, all $p$-values are even lower than in the main results except for Kolmogorov-Smirnov in case of Ross Basic, which increases from 0.000 to 0.025. We still reject our hypothesis that future realized returns are compatible with the recovered physical distributions for all recovery approaches but not for our benchmark models. For Power Utility, we find lower $p$-values compared to the main results, but we still cannot reject our hypothesis at the 5% level with any of our three tests. Note that the results for Historical Return Distribution remain virtually unchanged as the distribution is almost independent of the state space, but for some minute interpolation effect in case of the Berkowitz test. We do not report $MRMSE$ values for the log state space, as they are very close to the values of our main run. We conclude that our results are robust when changing to log returns.
Table IV. Tests of the recovered physical probabilities for a log return scaling. We present our results if future realized log returns are drawn from physical probabilities generated by one of the log-scaled six approaches: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each approach, we show the \( p \)-values from the Berkowitz, Knüppel, and Kolmogorov-Smirnov tests for uniformity of the percentiles of future realized log returns under the model physical cumulative distribution.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>Berkowitz ( p )-value</th>
<th>Knüppel ( p )-value</th>
<th>Kolmogorov-Smirnov ( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross Basic</td>
<td>0.001</td>
<td>0.013</td>
<td>0.025</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Bounded</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 ), rowsums ( \in [0.9, 1] )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Unimodal</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 ) and unimodal, rowsums ( \in [0.9, 1] )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Stable</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>Do not use transition state prices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power Utility</td>
<td>0.285</td>
<td>0.159</td>
<td>0.082</td>
</tr>
<tr>
<td>with ( \gamma = 4 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical Return Distribution</td>
<td>0.290*</td>
<td>0.480</td>
<td>0.347</td>
</tr>
</tbody>
</table>

*Small discrepancies with the \( p \)-values of Table I are due to the linear interpolation of the cumulative density function of returns versus log returns.
Changing the dimension of the state-space

In a simulation study, Tran and Xia (2015) show that the recovered market beliefs depend on the dimension of the state-space. In order to see if their concerns apply to our setting, we first reduce our overlapping fine state space by 20% and, second, we use a very coarse non-overlapping state space with just twelve states.

Our main study uses ten maturity steps per month which gives us 111 states for Ross Basic, Ross Bounded, and Ross Unimodal and 120 states for Ross Stable and Power Utility. We now reduce the state space to eight maturity steps per month and end up with 89 states for Ross Basic, Ross Bounded, and Ross Unimodal and with 96 states for Ross Stable and Power Utility.

The results are shown in Table V. Our results hardly change from the main run with the recovery methods being strongly rejected at the 5% level (only the p-value for the Kolmogorov-Smirnov test for Ross Basic is 0.088) and the benchmark models not being rejected. We conclude that a reduction of our state dimension by 20% does not lead to significant changes of our test results.

Next, we use a non-overlapping state space with only twelve states (one maturity step per month), just as in the original paper of Ross (2015). For Ross Basic, Ross Bounded, and Ross Unimodal thus use a twelve by twelve transition state price matrix. For Ross Stable, we solve (14) to recover a one-month pricing kernel on a twelve state moneyness grid (one maturity step being one month).

Now we reject all four recovery methods and even Power Utility with \( p \)-values close to zero. Only the results for Historical Return Distribution are not affected by the state space and we cannot reject our main hypothesis. As the one-month spot state price approximation one twelve states is very coarse, the \( MRMSE \) errors are also much higher with values beyond 0.100 for all five approaches.
Table V. Tests of the recovered physical probabilities for a reduced state-space size. We present our results if future realized returns are drawn from physical probabilities generated by one of the six approaches: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We reduce the original state-space size by approximately 20%. For each approach, we show the \( p \)-values from the Berkowitz, Knüppel, and Kolmogorov-Smirnov tests for uniformity of the percentiles of future realized returns under the model physical cumulative distribution.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>Berkowitz ( p )-value</th>
<th>Knüppel ( p )-value</th>
<th>Kolmogorov-Smirnov ( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross Basic ( \pi_{i,j} &gt; 0 )</td>
<td>0.026</td>
<td>0.035</td>
<td>0.088</td>
</tr>
<tr>
<td>Ross Bounded ( \pi_{i,j} &gt; 0 ), rowsums ( \in [0.9, 1] )</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>Ross Unimodal ( \pi_{i,j} &gt; 0 ) and unimodal, rowsums ( \in [0.9, 1] )</td>
<td>0.001</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>Ross Stable Do not use transition state prices</td>
<td>0.000</td>
<td>0.000</td>
<td>0.006</td>
</tr>
<tr>
<td>Power Utility with ( \gamma = 4 )</td>
<td>0.643</td>
<td>0.256</td>
<td>0.340</td>
</tr>
<tr>
<td>Historical Return Distribution</td>
<td>0.294</td>
<td>0.480</td>
<td>0.347</td>
</tr>
</tbody>
</table>
Choosing twelve equidistant states, which cover the range of the state prices even at the yearly horizon means that the one-month state prices are only non-zero at a few states around the current state. Thus, we implement a second version based on a non-equidistant discretization with more of the twelve states being closer to the current state and fewer of the twelve states being further away (See Appendix E for details). This version leads to a more accurate approximation of the one-month spot state prices at the cost of a worse approximation of higher maturity spot state prices. However, results are as disappointing as for the equidistant twelve state space. We conclude that while small variations to the state space (20% smaller) do not affect the results, going to a twelve state non-overlapping state space leads to numerical problems as the spot state prices can no longer be accurately fitted on such a coarse space.

B. Sub-Samples

For our main study we use monthly data from January 1996 to August 2014 with a sample size of 223 dates. We repeat our main study but now consider two different sub-samples. In our first variation, we exclude the period from February 1998 until February 2000 that we associate with the Dot-Com bubble, which is the two years before the burst of the Dot-Com bubble in March 2000. Results are given in Table VI. For our second variation, we follow Anand et al. (2013) to associate the period July 2007 through March 2009 with the financial crisis and exclude that period from the full sample. Results are given in Table VII.

Results hardly change from our main results and we conclude that our study is robust if we exclude crucial periods such as the Dot-Com bubble or the Financial Crisis.
Tests of the recovered physical probabilities excluding the period associated with the Dot-Com bubble. We present our results if future realized returns are drawn from physical probabilities generated by one of our six approaches: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We use our sample period from January 1996 until August 2014 but exclude months that are associated with the Dot-Com bubble. For each approach, we show the p-values from the Berkowitz, Knüppel, and Kolmogorov-Smirnov tests for uniformity of the percentiles of future realized returns under the model physical cumulative distribution.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>Berkowitz p-value</th>
<th>Knüppel p-value</th>
<th>Kolmogorov-Smirnov p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross Basic</td>
<td>0.016</td>
<td>0.039</td>
<td>0.000</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Bounded</td>
<td>0.007</td>
<td>0.003</td>
<td>0.014</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 ), rowsums ( \in [0.9, 1] )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Unimodal</td>
<td>0.001</td>
<td>0.000</td>
<td>0.027</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 ) and unimodal,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rowsums ( \in [0.9, 1] )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Stable</td>
<td>0.015</td>
<td>0.024</td>
<td>0.007</td>
</tr>
<tr>
<td>Do not use transition state prices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power Utility</td>
<td>0.552</td>
<td>0.378</td>
<td>0.561</td>
</tr>
<tr>
<td>( \gamma = 4 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical Return Distribution</td>
<td>0.745</td>
<td>0.807</td>
<td>0.575</td>
</tr>
</tbody>
</table>
Table VII. Tests of the recovered physical probabilities excluding the period associated with the Financial Crisis. We present our results if future realized returns are drawn from physical probabilities generated by one of our six approaches: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We use our sample period from January 1996 until August 2014 but exclude months that are associated with the Financial Crisis. For each approach, we show the p-values from the Berkowitz, Knüppel, and Kolmogorov-Smirnov tests for uniformity of the percentiles of future realized returns under the model physical cumulative distribution.

<table>
<thead>
<tr>
<th>Recovery Method</th>
<th>Berkowitz p-value</th>
<th>Knüppel p-value</th>
<th>Kolmogorov-Smirnov p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross Basic</td>
<td>0.047</td>
<td>0.021</td>
<td>0.005</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Bounded</td>
<td>0.001</td>
<td>0.003</td>
<td>0.004</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 ), rowsums ( \in [0.9, 1] )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Unimodal</td>
<td>0.000</td>
<td>0.000</td>
<td>0.012</td>
</tr>
<tr>
<td>( \pi_{i,j} &gt; 0 ) and unimodal, rowsums ( \in [0.9, 1] )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross Stable</td>
<td>0.002</td>
<td>0.010</td>
<td>0.002</td>
</tr>
<tr>
<td>Do not use transition state prices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power Utility</td>
<td>0.615</td>
<td>0.359</td>
<td>0.198</td>
</tr>
<tr>
<td>with ( \gamma = 4 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical Return Distribution</td>
<td>0.752</td>
<td>0.700</td>
<td>0.394</td>
</tr>
</tbody>
</table>
VIII. Conclusion

We implement and test several recovery methods in the framework of Ross (2015). We further present a variant of Ross recovery without explicitly estimating the transition state prices, which are numerically hard to determine. We find that future realized S&P 500 returns are incompatible with the recovered physical probabilities. On the other hand, two simple benchmark approaches work much better. For one, we employ the empirical distribution of non-overlapping monthly S&P 500 returns during the past five years. Alternatively, we use the pricing kernel of a power utility function with risk aversion parameter $\gamma = 4$. Both simple benchmark models cannot be rejected in the data.

We further analyze why the recovery theorem fails. We find that the most basic approach, requiring minimal assumptions, delivers unstable transition state price matrices. Backed by a simulation study, we argue that it is very sensitive to small variations in the options data to which the model is fitted. Further, the associated required risk-free rates of this basic approach are economically implausible.

Further implementations of Ross recovery, in which we add economically reasonable constraints to transition state prices, are much more stable than the basic approach but are not able to generate pricing kernels that are sufficiently away from risk-neutrality. Moreover, the assumption of time-homogeneity of the transition state prices leads to poorly fitted option values.

Robustness tests confirm that all our versions of the recovery theorem are incompatible with future realized S&P 500 index returns. Surprisingly simple benchmark models based on historical samples and a power utility cannot be rejected.
REFERENCES


A. Smoothing the Implied Volatility Surface

Analogues to Jackwerth (2004), we minimize the sum of squared local total second derivatives of implied volatilities (insuring smoothness of the volatility surface) plus the sum of squared deviations of the observed from model implied volatilities by using a trade-off parameter $\lambda$. The optimization problem is then defined as:

$$
\min_{\sigma_{i,t}} \frac{1}{TN} \cdot \sum_{t=1}^{T} \sum_{i \in I} (\sigma''_{i,t})^2 \cdot t + \lambda \cdot \frac{1}{L} \cdot \sum_{l=1}^{L} \left( \sigma_{i(l),t(l)} - \sigma_{i(l),t(l)}^{\text{obs}} \right)^2
$$

s.t.

$$
\sigma_{i,t} \geq 0
$$

where $\sigma''_{i,t}$ is the local second derivative (to be defined momentarily) and where $t$ compensates for the lower curvature at higher maturities. $\sigma_{i(l),t(l)}^{\text{obs}}$ is the $l$-th observed implied volatility with a total number of $L$ observations. We start with a high trade-off parameter $\lambda$ and iteratively increase the smoothness of the volatility surface by reducing $\lambda$, thus reducing the fit of observed implied volatility, until we obtain a smooth and positive state price surface.

We define the local second derivatives $\sigma''_{i,t}$ of implied volatilities as:

$$
\sigma''_{i,t} = \frac{\sigma_{i+1,t,t} - 2\sigma_{i,t,t} + \sigma_{i-1,t,t}}{(\Delta i)^2} + \frac{\sigma_{i,t+1,t} - 2\sigma_{i,t,t} + \sigma_{i,t-1,t}}{(\Delta t)^2}
$$

$$
+ \frac{\sigma_{i+1,t+1,t} - \sigma_{i+1,t-1,t} - \sigma_{i-1,t+1,t} + \sigma_{i-1,t-1,t}}{4(\Delta i, \Delta t)}
$$

where the first and second terms are approximations of the partial second derivative with respect to moneyness and maturity. The third term approximates the cross derivative. We evaluate the local second derivatives $\sigma''_{i,t}$ on a fine equidistant grid at time $t$ and moneyness indexed by $i \in I$, where $I = \{-N_{\text{low}}^{\text{fine}}, \ldots, N_{\text{high}}^{\text{fine}}\}$ with $N$ states. Step sizes are $\Delta t$ for time and $\Delta i$ for moneyness.
We impose the following boundary conditions for all $i$ and $t$:

$$\sigma_{i,0} = \sigma_{i,1}, \quad \sigma_{i,T+1} = \sigma_{i,T}$$

and $\sigma_{i-1,t} = \sigma_{i,t}$, for $i = -Nf_{\text{fine}}^{\text{low}}$, $\sigma_{i+1,t} = \sigma_{i,t}$, for $i = Nf_{\text{high}}^{\text{fine}}$

and $\sigma_{i-1,T+1} = \sigma_{i,T}$, for $i = -Nf_{\text{low}}^{\text{fine}}$, $\sigma_{i+1,0} = \sigma_{i,1}$, for $i = Nf_{\text{high}}^{\text{fine}}$

and $\sigma_{i-1,0} = \sigma_{i,1}$, for $i = -Nf_{\text{low}}^{\text{fine}}$, $\sigma_{i+1,T+1} = \sigma_{i,T}$ for $i = Nf_{\text{high}}^{\text{fine}}$

(19)

B. Recovery without Using Transition State Prices

We start with equation (12) and show how it can be used to achieve recovery without explicitly deriving transition state prices. We restate the equation in greater detail:

$$
\begin{pmatrix}
\pi_{N_{\text{low}},-N_{\text{low}}}^{t} & \pi_{N_{\text{low}},-N_{\text{low}}+1}^{t} & \cdots & \pi_{N_{\text{low}},N_{\text{high}}-1}^{t} & \pi_{N_{\text{low}},N_{\text{high}}}^{t} \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
\pi_{N_{\text{low}},-N_{\text{low}}}^{t} & \pi_{N_{\text{low}},-N_{\text{low}}+1}^{t} & \cdots & \pi_{N_{\text{low}},N_{\text{high}}-1}^{t} & \pi_{N_{\text{low}},N_{\text{high}}}^{t} \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
\pi_{N_{\text{high}},-N_{\text{low}}}^{t} & \pi_{N_{\text{high}},-N_{\text{low}}+1}^{t} & \cdots & \pi_{N_{\text{high}},N_{\text{high}}-1}^{t} & \pi_{N_{\text{high}},N_{\text{high}}}^{t}
\end{pmatrix}
\begin{pmatrix}
z_{-N_{\text{low}}} \\
\vdots \\
z_{N_{\text{high}}}
\end{pmatrix}
= \delta^{t}
\begin{pmatrix}
z_{-N_{\text{low}}} \\
\vdots \\
z_{N_{\text{high}}}
\end{pmatrix}
$$

(20)

where $\pi_{i,j}^{t}$ is the transition state price of moving from state $i$ to state $j$ in a time equal to $t$ transition periods. Note that the current row in $\Pi^{t}$ (indexed by $i = 0$) represents the transition state prices with maturity equal to $t$ transition periods:

$$
\begin{pmatrix}
\pi_{0,-N_{\text{low}}}^{t} & \pi_{0,-N_{\text{low}}+1}^{t} & \cdots & \pi_{0,N_{\text{high}}-1}^{t} & \pi_{0,N_{\text{high}}}^{t} \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
\pi_{N_{\text{high}},-N_{\text{low}}}^{t} & \pi_{N_{\text{high}},-N_{\text{low}}+1}^{t} & \cdots & \pi_{N_{\text{high}},N_{\text{high}}-1}^{t} & \pi_{N_{\text{high}},N_{\text{high}}}^{t}
\end{pmatrix}
\begin{pmatrix}
z_{-N_{\text{low}}} \\
\vdots \\
z_{N_{\text{high}}}
\end{pmatrix}
= \delta^{t}
\begin{pmatrix}
z_{-N_{\text{low}}} \\
\vdots \\
z_{N_{\text{high}}}
\end{pmatrix}
$$

(21)
After dividing both sides by $z_0$, we use Equation [21] for every $t = 1, \ldots, T$ and stack the resulting equations to obtain the following system:

$$
\begin{pmatrix}
\pi_{0, N_{low}}^1 & \pi_{0, N_{low}+1}^1 & \cdots & \pi_{0, N_{high}-1}^1 & \pi_{0, N_{high}}^1 \\
\pi_{0, N_{low}}^2 & \pi_{0, N_{low}+1}^2 & \cdots & \pi_{0, N_{high}-1}^2 & \pi_{0, N_{high}}^2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\pi_{0, N_{low}}^{T-1} & \pi_{0, N_{low}+1}^{T-1} & \cdots & \pi_{0, N_{high}-1}^{T-1} & \pi_{0, N_{high}}^{T-1} \\
\pi_{0, N_{low}}^T & \pi_{0, N_{low}+1}^T & \cdots & \pi_{0, N_{high}-1}^T & \pi_{0, N_{high}}^T \\
\end{pmatrix}
\begin{pmatrix}
\frac{z_{N_{low}} - 1}{z_0} \\
\frac{z_{N_{low}} - 1}{z_0} \\
\vdots \\
\frac{z_{N_{low}} - 1}{z_0} \\
\frac{z_{N_{high}} - 1}{z_0} \\
\end{pmatrix}
= 
\begin{pmatrix}
\delta \\
\delta^2 \\
\vdots \\
\delta^{T-1} \\
\delta^T \\
\end{pmatrix}
$$

(22)

To solve this system of equations, we note that the transition state prices starting at the current state but with different maturities can be replaced by the spot state prices of the same maturity, which can be directly obtained from the spot state price surface. The new system of equations has $N$ unknowns ($\frac{z_j}{z_0}$ with $j \in I$ and $\delta$) and as many equations. As $\delta$ is the utility discount factor, we require it to be larger than zero and smaller than one. We force the ratios $\frac{z_j}{z_0}$ to be non-negative, as these ratios are directly linked to the pricing kernel. Now we solve the equation system by means of least squares.

Equation (15) states that the pricing kernel for different maturities of $t$ transition periods will only differ from the fraction $\frac{z_0}{z_j}$ by factor involving $t$ and the discount factor $\delta$:

$$
m_{t,j} = \frac{1}{\delta^t} \cdot \frac{z_0}{z_j}, \quad j \in I
$$

(23)

where $m_{t,j}$ represents the pricing kernel value for state $j$ with a maturity of $t$ transition
periods. We can use this property and link pricing kernels with different maturities in the following way:

\[ m_{t,0,j} = \frac{1}{\delta^t} \cdot \frac{z_0}{z_j} = \frac{1}{\delta^{t-1}} \cdot \frac{1}{\delta} \cdot \frac{z_0}{z_j} = \frac{1}{\delta^{t-1}} \cdot m_{0,j}, \quad j \in I \] (24)

where \( m_{0,j} \) represents the pricing kernel value for state \( j \) with a maturity of one transition period.

C. Test Statistics

Berkowitz test

Formally, we write the transformation of returns \( R_\tau \) into \( u_\tau \) as:

\[ u_\tau = \hat{P}_\tau(R_\tau) = \int_{-\infty}^{R_\tau} \hat{p}_\tau(x)dx \] (25)

where \( u_\tau \sim i.i.d. U(0,1) \). The Berkowitz (2001) test jointly tests uniformity and the i.i.d. property of \( u_\tau \). For this test, the series \( u_\tau \) is transformed by applying the inverse standard normal cumulative density function \( \Phi \) to \( u_\tau \):

\[ z_\tau = \Phi^{-1}(u_\tau) = \Phi^{-1}\left(\int_{-\infty}^{R_\tau} \hat{p}_\tau(x)dx\right) \]

Under the hypothesis \( \hat{p}_\tau = p_\tau \), \( z_\tau \sim i.i.d. N(0,1) \), which suggests the following AR(1) model:

\[ z_\tau - \mu = \rho(z_{\tau-1} - \mu) + \epsilon_\tau \] (26)
where the null requires $\mu = 0$, $\text{Var}(\epsilon_t) = 1$, and $\rho = 0$. Berkowitz then applies the following likelihood ratio test:

$$LR_3 = -2(LL(0, 1, 0) - LL(\hat{\mu}, \hat{\sigma}, \hat{\rho}))$$

where $LL$ characterizes the log likelihood of Equation (26).

**Knüppel test**

The test introduced by Knüppel (2015) first scales the series $u_t$ to $y_t = \sqrt{12} (u_t - 0.5)$. In order to test $u_t$ for standard uniformity, the series $y_t$ is tested for scaled uniformity with mean=0 and variance=1. The test then compares the first $S$ moments of the series $y_t$ to the respective moment of its theoretical counterpart in the following GMM-type procedure with the test statistic $\alpha_S$:

$$\alpha_S = T \cdot D_S^\top \Omega_S^{-1} D_S$$  \hspace{1cm} (27)

where $D_S$ is a vector that consists of the differences between the sample moments $\frac{1}{T} \sum_{\tau=1}^T y_{\tau s}$ and the theoretical moments $m_s$ for $s = 1, ..., S$. $\Omega_S$ is a consistent covariance matrix estimator of all $S$ respective moment differences between $y_t$ and its theoretical counterpart. We follow Knüppel (2015) and set all elements of $\Omega_S$ which represent covariances between odd and even moment differences to zero and apply the test by considering the first four moments ($S=4$). We account for serial correlation of $u_t$ by estimating a Newey-West covariance matrix with automated lag length as proposed by Andrews (1991).
**Kolmogorov-Smirnov test**

The Kolmogorov-Smirnov test looks at the maximum distance between the empirical \( \hat{P} \) and theoretical \( P \) cumulative density function and uses the following test statistic:

\[
KS = \sup_x |P(x) - \hat{P}(x)|
\]  

(28)

**D. In-Sample Pricing Error for Recovery**

To obtain the mean-root-mean-squared-error (MRMSE) for implied volatilities for a specific recovery approach, we use information on one-month spot state prices \( \hat{p} \) that are implied by the recovery method. On each sample date \( \tau \), we compute one-month implied option quotes \( \hat{C} \) that are based on state prices \( \hat{p} \) by numerical integration:

\[
\hat{C}(K_l) = \int_0^\infty \hat{\pi}(S) \cdot \max(S - K_l, 0) dS
\]  

(29)

for every moneyness level \( K_l \) at which we observe a one-month option quote. We then transform the option prices \( \hat{C}(K_l) \) into implied volatilities \( \hat{\sigma}(K_l) \). For each date \( \tau \), we compute the root-mean-squared error \( RMSE_{\tau} \) between one-month observed implied volatilities and model implied volatilities \( \hat{\sigma}(K_l) \). Our final measure MRMSE is the time-series average of all the \( RMSE_{\tau} \).

**E. Non-Equidistant Interpolated Twelve-by-Twelve Spot State Prices**

Our usual twelve-by-twelve state space is equidistant in maturity and moneyness. For a non-equidistant state space in the moneyness dimension, we initially need 14 moneyness level, which will be reduced by the Breeden-Litzenberger approach to the final twelve moneyness
levels. We pick the first six moneyness levels at: the beginning $N_1$ and the end $N_{12}$ of the moneyness range needed to cover the tails of *one-year* spot state prices; the beginning $N_3$ and the end $N_{10}$ of the moneyness range needed to cover the tails of *one-month* spot state prices, a moneyness of zero $N_0$ where we assign the same implied volatility value as at $N_1$, and a moneyness of three $N_{13}$ where we assign the same implied volatility value as at $N_{12}$.

We optimize for the location of the eight remaining moneyness levels in a way that the least squares distance between the interpolated volatility surface and the volatility surface on the fine grid is minimized, while we require $N_2$ to lie in between $N_1$ and $N_3$, $N_4$ to $N_9$ (with one of them being the current state with moneyness of one) to lie in between $N_3$ and $N_{10}$, and $N_{11}$ to lie in between $N_{10}$ and $N_{12}$. 
