

*Online Appendix: Does the Ross recovery theorem
work empirically?*

Contents

A	The Ross recovery theorem with two states	2
B	Data	3
C	Obtaining the implied volatility surface	4
D	Recovery without using transition state prices	7
E	A stepwise illustration of the testing procedure	9
F	Power analysis of the statistical tests	13
G	Robustness	18
	G.1. Variations of the state space	18
	G.2. Removing recession periods	26
	G.3. Weekly options	30
H	The relation between transition state prices and the stochastic discount factor	34
I	Insights from simulated economies	36

A. The Ross recovery theorem with two states

In the text of our paper, we highlight that in the Ross recovery setting, the physical transition probabilities $p_{i,j}$ of moving from state i to state j have the form:

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

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

$$p_{0,0} + p_{0,1} = 1 \quad \Leftrightarrow \quad \frac{1}{\delta} \cdot \pi_{0,0} \cdot \frac{u'_0}{u'_0} + \frac{1}{\delta} \cdot \pi_{0,1} \cdot \frac{u'_0}{u'_1} = 1, \quad (2)$$

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

We can rewrite this system of equations in matrix form and obtain an eigenvalue problem as follows:

$$\begin{pmatrix} \pi_{0,0} & \pi_{0,1} \\ \pi_{1,0} & \pi_{1,1} \end{pmatrix} \cdot \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} = \delta \cdot \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} \quad \text{where} \quad z_0 = \frac{1}{u'_0} \quad \text{and} \quad z_1 = \frac{1}{u'_1}. \quad (3)$$

Solving this eigenvalue problem for the greatest positive eigenvalue δ gives us the values for u'_0 , u'_1 , and δ , which we can insert in Eq. (1) to obtain the physical probabilities.

B. Data

We obtain end-of-day option data on the S&P 500 index from the Berkeley Options Database (prior to 1996) and OptionMetrics (after 1995). We use the midpoints of bid and ask option quotes as our option prices. Consistent with the literature, we use only out-of-the-money put and call options with positive trading volume and eliminate all options that violate no-arbitrage constraints. We find the average implied dividend yield from put-call parity pairs for a given maturity where the risk-free rate is given by the interpolated zero-curve from OptionMetrics or the Berkeley Options Database. We consider monthly dates τ , which we find by going 30 calendar days back in time from the expiration date.

On October 21, 1987, there are no option prices available for options expiring 30 calendar days later on November 20, 1997, and so we use October 22, 1997 as the corresponding sample date. In all of September 1992, there are no option prices available for options expiring in October, 1992. We end up with 380 dates. Our sample period and our option data ranges from April 1986 through December 2017.

For the historical return distribution, we further obtain end-of-day S&P 500 index levels from Datastream. We compute monthly S&P 500 returns from April 1981 through December 2017, which includes a five-year period prior to April 1986.

C. Obtaining the implied volatility surface

To generate a smooth implied volatility surface, we apply an extension of the fast and stable method of Jackwerth (2004), which finds smooth implied volatilities σ_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 fit to the option data) using a trade-off parameter λ .

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 fit of the surface) by using the trade-off parameter λ . The optimization problem is:

$$\min_{\sigma_{i,t}} \quad \frac{1}{TN} \cdot \sum_{t=1}^T \sum_{i \in I} (\sigma''_{i,t})^2 \cdot t \quad + \quad \lambda \cdot \frac{1}{L} \cdot \sum_{l=1}^L (\sigma_{i(l),t(l)} - \sigma_{i(l),t(l)}^{\text{obs}})^2 \quad (4)$$

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

where $\sigma''_{i,t}$ is the local second derivative (defined below), and where the maturity t compensates for the higher curvature at shorter maturities. $\sigma_{i(l),t(l)}^{\text{obs}}$ is the l^{th} observed implied volatility with a total number of L observations. We define the local second derivatives $\sigma''_{i,t}$ of implied volatilities as:

$$\sigma''_{i,t} = \frac{\sigma_{i+1,t} - 2\sigma_{i,t} + \sigma_{i-1,t}}{(\Delta_i)^2} + \frac{\sigma_{i,t+1} - 2\sigma_{i,t} + \sigma_{i,t-1}}{(\Delta_t)^2} + \frac{\sigma_{i+1,t+1} - \sigma_{i+1,t-1} - \sigma_{i-1,t+1} + \sigma_{i-1,t-1}}{(4\Delta_i\Delta_t)}, \quad (5)$$

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 $I = \{-N_{low}^{fine}, \dots, 0, \dots, N_{high}^{fine}\}$ with N states. We impose boundary

conditions for all i and t as follows:

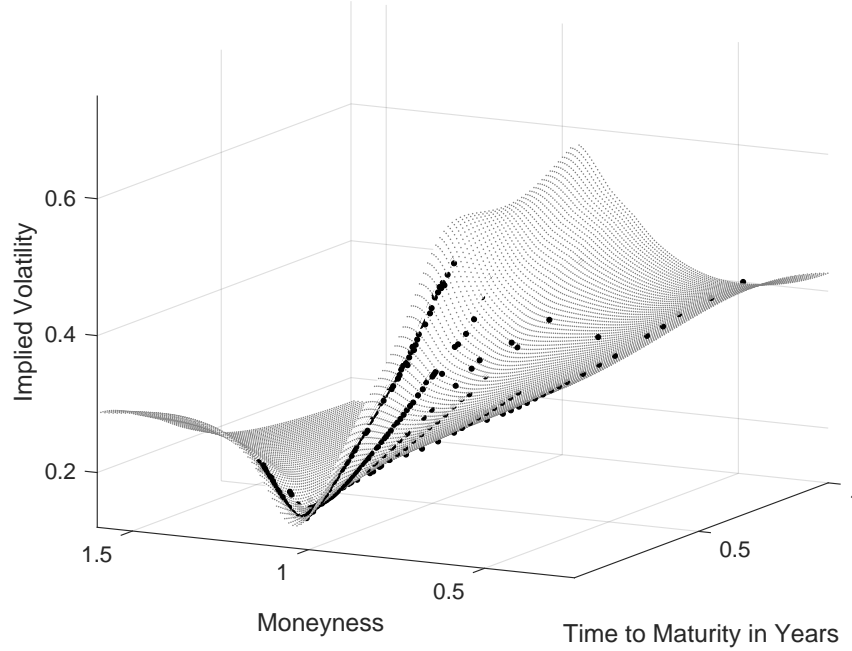
$$\begin{aligned}
& \sigma_{i,0} = \sigma_{i,1}, \quad \sigma_{i,T+1} = \sigma_{i,T} \\
& \text{and } \sigma_{i-1,t} = \sigma_{i,t}, \text{ for } i = -N_{low}^{fine}, \quad \sigma_{i+1,t} = \sigma_{i,t}, \text{ for } i = N_{high}^{fine} \\
& \text{and } \sigma_{i-1,T+1} = \sigma_{i,T}, \text{ for } i = -N_{low}^{fine}, \quad \sigma_{i+1,0} = \sigma_{i,1}, \text{ for } i = N_{high}^{fine} \\
& \text{and } \sigma_{i-1,0} = \sigma_{i,1}, \text{ for } i = -N_{low}^{fine}, \quad \sigma_{i+1,T+1} = \sigma_{i,T} \text{ for } i = N_{high}^{fine}.
\end{aligned} \tag{6}$$

We then solve Eq. (4) to obtain the implied volatility surface on the fine grid. We start with a high trade-off parameter λ and iteratively increase the smoothness of the volatility surface by reducing λ , thus reducing the fit of observed implied volatility, until we obtain a smooth and positive state price surface.

To obtain state prices on the coarser grid suitable for the recovery theorem, we linearly interpolate the fine implied volatility surface. The number of fine states N varies between 227 for earlier sample days and 2579 for later sample days. The number of coarse states is either 111 (Ross Basic, Ross Bounded, Ross Unimodal) or 120 (Ross Stable, Power Utility) with two additional states of padding for taking derivatives in the next step. From the implied volatilities on that coarser grid, we compute call option prices and, at each maturity t , apply the Breeden and Litzenberger (1978) approach to find the spot state prices on the coarser grid. Namely, for each maturity t , we take the numerical second derivative of the call prices to obtain spot state prices, where we lose the lowest and highest index levels due to the numerical second derivative. The spot state prices will then be exactly on the desired coarser grid.

We depict the results of our implementation for a typical day in our sample, February 17, 2010. Fig. 1, Panel A, shows the interpolated smoothed implied volatility surface on the coarser state space suitable for the recovery theorem. 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 and flattens 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 longer maturities.

Panel A: Implied Volatility Surface, 17-Feb-2010



Panel B: Spot State Price Surface, 17-Feb-2010

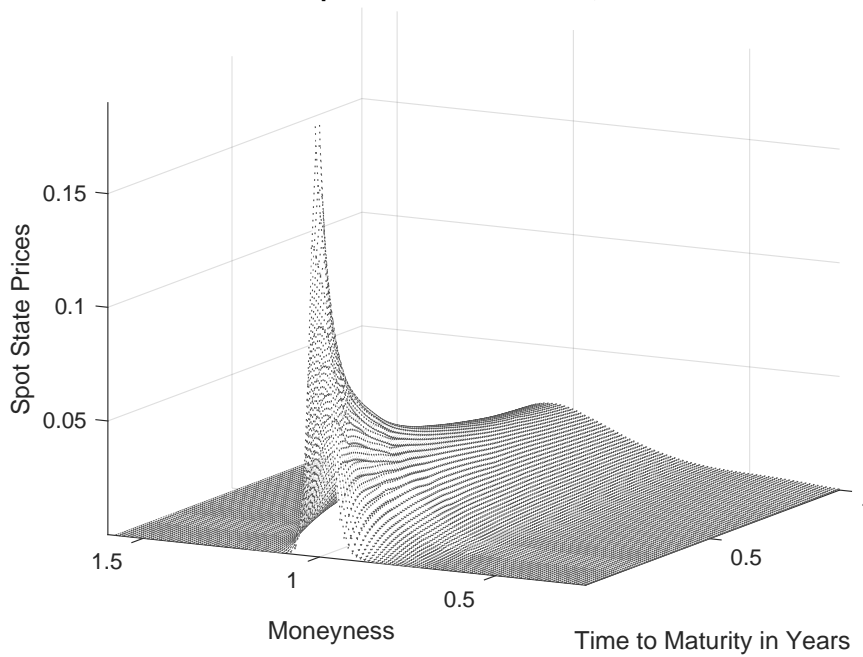


Fig. 1. Implied volatility and spot state price surfaces. We show the interpolated implied volatility surface and the observed implied volatilities (as black dots) in Panel A. We show the related spot state price surface in Panel B. Using data for February 17, 2010, we depict these surfaces for S&P 500 index options across maturities and moneyness levels.

D. Recovery without using transition state prices

Starting with the eigenvalue problem in Ross (2015), we can multiply both sides from the left with the transition state price matrix Π :

$$\Pi \cdot \Pi z = \Pi \cdot \delta z = \delta(\Pi z) = \delta^2 z. \quad (7)$$

Iterating, we obtain the following relation:

$$\Pi^t z = \delta^t z \text{ with } t = 1, \dots, T. \quad (8)$$

We show how Eq. (8) can be used to achieve recovery without explicitly deriving transition state prices. We restate the equation in greater detail:

$$\begin{pmatrix} \pi_{-N_{low}, -N_{low}}^t & \pi_{-N_{low}, -N_{low}+1}^t & \cdots & \pi_{-N_{low}, N_{high}-1}^t & \pi_{-N_{low}, N_{high}}^t \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \pi_{0, -N_{low}}^t & \pi_{0, -N_{low}+1}^t & \cdots & \pi_{0, N_{high}-1}^t & \pi_{0, N_{high}}^t \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \pi_{N_{high}, -N_{low}}^t & \pi_{N_{high}, -N_{low}+1}^t & \cdots & \pi_{N_{high}, N_{high}-1}^t & \pi_{N_{high}, N_{high}}^t \end{pmatrix} \cdot \begin{pmatrix} z_{-N_{low}} \\ \vdots \\ z_0 \\ \vdots \\ z_{N_{high}} \end{pmatrix} = \delta^t \cdot \begin{pmatrix} z_{-N_{low}} \\ \vdots \\ z_0 \\ \vdots \\ z_{N_{high}} \end{pmatrix}, \quad (9)$$

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

$$\begin{pmatrix} \pi_{0, -N_{low}}^t & \pi_{0, -N_{low}+1}^t & \cdots & \pi_{0, N_{high}-1}^t & \pi_{0, N_{high}}^t \end{pmatrix} \cdot \begin{pmatrix} z_{-N_{low}} \\ \vdots \\ z_0 \\ \vdots \\ z_{N_{high}} \end{pmatrix} = \delta^t \cdot z_0. \quad (10)$$

After dividing both sides by z_0 , we use Eq. (10) for every $t = 1, \dots, T$ and stack the

resulting equations to obtain a system of equations:

$$\begin{pmatrix} \pi_{0,-N_{low}}^1 & \pi_{0,-N_{low}+1}^1 & \cdots & \pi_{0,N_{high}-1}^1 & \pi_{0,N_{high}}^1 \\ \pi_{0,-N_{low}}^2 & \pi_{0,-N_{low}+1}^2 & \cdots & \pi_{0,N_{high}-1}^2 & \pi_{0,N_{high}}^2 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \pi_{0,-N_{low}}^{T-1} & \pi_{0,-N_{low}+1}^{T-1} & \cdots & \pi_{0,N_{high}-1}^{T-1} & \pi_{0,N_{high}}^{T-1} \\ \pi_{0,-N_{low}}^T & \pi_{0,-N_{low}+1}^T & \cdots & \pi_{0,N_{high}-1}^T & \pi_{0,N_{high}}^T \end{pmatrix} \cdot \begin{pmatrix} \frac{z_{-N_{low}}}{z_0} \\ \vdots \\ \frac{z_{-1}}{z_0} \\ 1 \\ \frac{z_1}{z_0} \\ \vdots \\ \frac{z_{N_{high}}}{z_0} \end{pmatrix} = \begin{pmatrix} \delta \\ \delta^2 \\ \vdots \\ \delta^{T-1} \\ \delta^T \end{pmatrix}. \quad (11)$$

To solve this system of equations, note that all 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 $\left(\frac{z_j}{z_0}$ with $j \in I$ and $\delta\right)$ and as many equations.¹ As δ is the utility discount factor, we require it to be greater than zero and less than one. We force the ratios $\frac{z_j}{z_0}$ to be non-negative, as these ratios are directly linked to the SDF. We solve the system of equations by means of least squares. Then, the t -period SDF for state j can be found as:

$$m_{0,j}^t = \delta^t \frac{z_0}{z_j}, \quad j \in I. \quad (12)$$

The SDFs for different maturities of t transition periods differ from the fraction $\frac{z_0}{z_j}$ only by a factor involving the discount factor δ and t . We can use this property and link SDFs with different maturities thus:

$$m_{0,j}^t = \delta^t \frac{z_0}{z_j} = \delta^{t-1} \delta \frac{z_0}{z_j} = \delta^{t-1} m_{0,j}, \quad j \in I, \quad (13)$$

where $m_{0,j}$ represents the SDF value for state j with a maturity of one transition period.

¹Note that I has N elements, but there are only $N - 1$ fractions as $\frac{z_0}{z_0} = 1$.

E. A stepwise illustration of the testing procedure

On each sample date τ , we proceed as follows:

Step 1: Generate an implied volatility surface from option prices

- Exclude options that violate no-arbitrage conditions
- Exclude options with maturity t of less than one month or more than one year
- Find the dividend yield for each maturity as the average option-implied dividend yields from put-call parity pairs
- Remove in-the-money options and transform remaining option prices into (observed) option-implied volatilities $\sigma_{i(l),t(l)}$ with states $i(l)$ and maturities $t(l)$ for observations $l = 1, \dots, L$ using the Black-Scholes formula. For the transformation, use the option-implied dividend yields above and the interpolated zero curve from OptionMetrics as the risk-free rate
- Discretize option strike prices with a step size of \$1.25 and set low and high state bounds so that all positive spot state prices occur in the interior of the state set. The number of auxiliary states N varies between 227 for earlier sample days and 2579 for later sample days
- Apply Eq. (4) (with the conditions of Eq. (5) and Eq. (6) in place) to obtain a fine auxiliary volatility surface $\sigma_{i,t}$. That auxiliary surface is defined on N states (ranging from 227 to 2579 depending on the sample date) and on $T = 120$ maturity steps (10 maturity steps each month)
- Interpolate the fine volatility surface to obtain a volatility surface on a coarse grid suitable for a given recovery method:

For Ross Basic, Ross Bounded, and Ross Unimodal:

Equidistantly select 113 states where the current state (moneyness = 1) is included. The coarse volatility surface is then defined on $N = 113$ states and on $T = 120$ maturities (10 maturity steps each month). The maturity discretization is the same as for the fine auxiliary grid

For Ross Stable and Power Utility:

Equidistantly select 122 states where the current state (moneyness = 1) is included. That coarse volatility surface is then defined on $N = 122$ states and on $T = 120$ maturities (10 maturity steps each month). The maturity discretization is the same as for the fine auxiliary grid

Step 2: Generate spot state price surfaces

- Transform the coarse implied volatility surfaces into call option price surfaces using the Black-Scholes formula
- Transform call option price surfaces into spot state price surfaces by applying the approach of Breeden and Litzenberger (1978) for each maturity. Taking numerical second derivatives in the moneyness dimension leads to the first and last state being dropped in that dimension
- The spot state price surfaces are defined on $N = 111$ states and on $T = 120$ maturities for Ross Basic, Ross Bounded, and Ross Unimodal and on $N = 120$ states and on $T = 120$ maturities for Ross Stable and Power Utility

Step 3: Extract transition state prices from spot state prices

Only for Ross Basic, Ross Bounded, and Ross Unimodal:

- Use the spot state price surface to generate a transition state price matrix Π for methods Ross Basic, Ross Bounded, and Ross Unimodal as explained in detail in the Methodology section of the main paper. For those methods, we have 111 states, which results in 111×111 - dimensional transition state price matrices

Step 4: Recover one-month cumulative physical distributions \hat{P}_τ

For Ross Basic, Ross Bounded, and Ross Unimodal:

- Apply the Ross recovery theorem to the transition state prices $\pi_{i,j}$ to obtain the physical transition probabilities $p_{i,j}$
- Obtain one-month *spot* physical probabilities $\hat{p}_\tau = p_{0,j}$ from the recovered physical *transition* probabilities $p_{i,j}$ with current initial state $i = 0$

For Ross Stable, Power Utility, and Historical Return Distribution:

- Ross Stable: Make direct use of the spot state price surface to recover one-month (spot) physical probabilities p_τ as described in Online Appendix D
- Power Utility: Read out one-month spot state prices and transform them into one-month physical probabilities p_τ using a power utility SDF with risk-aversion coefficient $\gamma = 3$
- Historical Return Distribution: Collect five years of monthly S&P 500 returns prior to date τ and construct the one-month empirical pdf p_τ with those returns. The resulting discrete distribution is based on 60 irregularly spaced states (returns) and probabilities of $1/60$ in each state

Having recovered one-month physical distributions p_τ for all dates $\tau = 1, \dots, 380$ at hand, we then:

Step 5: Transform recovered one-month probabilities \hat{p}_τ into a one-month physical cumulative distribution function \hat{P}_τ

- Summing the physical probabilities \hat{p}_τ results in a discrete cumulative distribution that is made up of piecewise constant parts with jumps at the defined states. To construct a continuous cumulative distribution function \hat{P}_τ , we fit a piecewise linear function with breakpoints at the midpoints of those jumps. We use a left limit at a return of -100% with a value of zero and a right limit at 200% with a value of one

Step 6: Test the recovered physical distributions

- For each date τ , collect the one-month *future* return r_τ
- For each date τ , plug the one-month future return r_τ into the recovered cumulative distributions function \hat{P}_τ to obtain one percentile value x_τ for each recovery method, where $x_\tau = \hat{P}_\tau(1 + r_\tau)$
- Apply the Knüppel tests and the Kolmogorov-Smirnoff test to test the percentile set x_τ , $\tau = 1, \dots, 380$, for uniformity and obtain the p -values, which we report in the main text

- For the Berkowitz test, first convert a given percentile set x_τ into a standard normal percentile set z_τ with the inverse standard normal cumulative distributions function Φ^{-1} , where $z_\tau = \Phi^{-1}(x_\tau)$ for $\tau = 1, \dots, 380$. Then apply the Berkowitz test to test the set z_τ for standard normality and obtain the p -values, which we report in the main text

Step 7: Run mean predictions

- For each recovery method, compute a time-series of means μ_τ of the recovered physical distributions \hat{p}_τ with $\tau = 1, \dots, 380$
- Regress the recovered means μ_τ on the time-series of future returns r_τ :

$$r_\tau = a + b\mu_\tau + \epsilon_\tau, \quad \tau = 1, \dots, 380. \quad (14)$$

- Apply Eq. (14) using three different regression models: the Intercept model by setting $b = 1$ and testing $a = 0$, the Slope model by setting $a = 0$ and testing $b = 1$, and the Joint model by testing both $a = 0$ and $b = 1$

Step 8: Run variance predictions

- For each recovery method, compute a time-series of variances σ_τ^2 of the recovered physical distributions \hat{p}_τ with $\tau = 1, \dots, 380$
- For each date τ , compute the realized variance RV_τ as the sample variance of all future daily returns in the month following date τ multiplied by the number of days in that month
- Regress the recovered variances σ_τ^2 on the time-series of realized variances RV_τ :

$$RV_\tau = a + b\sigma_\tau^2 + \epsilon_\tau, \quad \tau = 1, \dots, 380. \quad (15)$$

- Apply Eq. (15) using three different regression models: the Intercept model by setting $b = 1$ and testing $a = 0$, the Slope model by setting $a = 0$ and testing $b = 1$, and the Joint model by testing both $a = 0$ and $b = 1$

F. Power analysis of the statistical tests

We provide a power analysis for each of our four statistical tests by analyzing how well they perform in rejecting the hypothesis that returns drawn from a particular distribution come from an alternative distribution. For any given recovery method, we simulate 10,000 economies, where each economy is generated by drawing 380 future returns from that method’s recovered distributions (one for each sample date). For each economy, we then apply our main testing procedure to test if the generated returns, which are based on a particular recovery method, are drawn from the recovered distributions of any other method. We report the number of rejections (p -value ≤ 0.05) within the 10,000 economies for each combination of our methods. Table 1 reports results for the Berkowitz test, Table 2 for the Kolmogorov-Smirnov test, Table 3 for the Knüppel test with three moments, and Table 4 for the Knüppel test with four moments.

Our results show that tests do not summarily reject alternative distributions. Rather, the tests correctly reject the true distribution in 5% of all cases (the Knüppel test with three moments rejects only 2% of all cases). Alternative distributions are rejected more frequently in general and less similar distributions are rejected more frequently than more similar distributions. As we expected, some distributions cluster: Ross Bounded, Ross Unimodal, and Ross Stable can be grouped into a “risk-neutral” cluster. Power Utility and the Historical Return Distribution can be clustered into a “similar to the future returns” cluster. Ross Basic seems to be closer to the second than the first cluster but is not obviously part of either.

Concerning our four tests, we note that the Knüppel test with three moments has a rejection rate which is a bit too low (2% when it should be 5%). This however is conservative in that it makes rejections harder. Thus, given that we can still reject our null for the four Ross recovery methods, we can rest assured that our main empirical findings are correct.

The Kolmogorov-Smirnov test tends to have the least power to reject the null when a similar distribution is true (see the low percentages comparing Ross Bounded against Ross Unimodal, Ross Bounded against Ross Stable, and Ross Unimodal against Ross Stable). More powerful are in turn the Knüppel test with three moments, and then the Knüppel with four moments, while the Berkowitz test is the most powerful. Still, as all our results are strongly supported by all four tests, we simply report them all.

Table 1

Power Analysis of the Berkowitz Test.

We generate 10,000 realizations of an economy, where a particular recovery method holds. For each realization, we draw 380 future returns. For any simulated economy, we report the rejection rates for the hypothesis that the future returns are drawn from another particular recovery method's physical distributions. We use a Berkowitz test with a significance level of 5%.

<i>Rejection Rates</i>	<i>Ross Basic cdf</i>	<i>Ross Bounded cdf</i>	<i>Ross Unimodal cdf</i>	<i>Ross Stable cdf</i>	<i>Power Utility with $\gamma = 3$ cdf</i>	<i>Hist. Return Distribution cdf</i>
<i>Ross Basic</i> return draws	5%	94%	79%	100%	100%	100%
<i>Ross Bounded</i> return draws	98%	5%	10%	97%	100%	100%
<i>Ross Unimodal</i> return draws	100%	13%	5%	100%	100%	100%
<i>Ross Stable</i> return draws	94%	33%	87%	5%	66%	97%
<i>Power Utility with $\gamma = 3$</i> return draws	26%	84%	98%	69%	5%	85%
<i>Hist. Return Distribution</i> return draws	47%	56%	55%	90%	77%	5%

Table 2

Power Analysis of the Kolmogorov-Smirnov Test.

We generate 10,000 realizations of an economy, where a particular recovery method holds. For each realization, we draw 380 future returns. For any simulated economy, we report the rejection rates for the hypothesis that the future returns are drawn from another particular recovery method's physical distributions. We use a Kolmogorov-Smirnov test with a significance level of 5%.

<i>Rejection Rates</i>	<i>Ross Basic cdf</i>	<i>Ross Bounded cdf</i>	<i>Ross Unimodal cdf</i>	<i>Ross Stable cdf</i>	<i>Power Utility with $\gamma = 3$ cdf</i>	<i>Hist. Return Distribution cdf</i>
<i>Ross Basic</i> return draws	5%	93%	88%	97%	49%	94%
<i>Ross Bounded</i> return draws	85%	5%	5%	6%	56%	68%
<i>Ross Unimodal</i> return draws	78%	7%	5%	10%	55%	80%
<i>Ross Stable</i> return draws	91%	5%	7%	5%	62%	68%
<i>Power Utility with $\gamma = 3$</i> return draws	20%	60%	58%	68%	5%	22%
<i>Hist. Return Distribution</i> return draws	53%	52%	54%	57%	16%	5%

Table 3

Power Analysis of the Knüppel (3 Moments) Test.

We generate 10,000 realizations of an economy, where a particular recovery method holds. For each realization, we draw 380 future returns. For any simulated economy, we report the rejection rates for the hypothesis that the future returns are drawn from another particular recovery method's physical distributions. We use a Knüppel (3 moments) test with a significance level of 5%.

<i>Rejection Rates</i>	<i>Ross Basic cdf</i>	<i>Ross Bounded cdf</i>	<i>Ross Unimodal cdf</i>	<i>Ross Stable cdf</i>	<i>Power Utility with $\gamma = 3$ cdf</i>	<i>Hist. Return Distribution cdf</i>
<i>Ross Basic</i> return draws	2%	88%	70%	97%	72%	96%
<i>Ross Bounded</i> return draws	70%	2%	6%	4%	43%	61%
<i>Ross Unimodal</i> return draws	62%	5%	2%	17%	55%	78%
<i>Ross Stable</i> return draws	88%	5%	24%	2%	48%	54%
<i>Power Utility with $\gamma = 3$</i> return draws	19%	48%	63%	53%	2%	16%
<i>Hist. Return Distribution</i> return draws	25%	48%	56%	49%	21%	2%

Table 4

Power Analysis of the Knüppel (4 Moments) Test.

We generate 10,000 realizations of an economy, where a particular recovery method holds. For each realization, we draw 380 future returns. For any simulated economy, we report the rejection rates for the hypothesis that the future returns are drawn from another particular recovery method's physical distributions. We use a Knüppel (4 moments) test with a significance level of 5%.

<i>Rejection Rates</i>	<i>Ross Basic cdf</i>	<i>Ross Bounded cdf</i>	<i>Ross Unimodal cdf</i>	<i>Ross Stable cdf</i>	<i>Power Utility with $\gamma = 3$ cdf</i>	<i>Hist. Return Distribution cdf</i>
<i>Ross Basic</i> return draws	5%	91%	82%	99%	82%	100%
<i>Ross Bounded</i> return draws	87%	5%	19%	17%	63%	98%
<i>Ross Unimodal</i> return draws	92%	14%	5%	61%	86%	100%
<i>Ross Stable</i> return draws	91%	29%	93%	5%	57%	94%
<i>Power Utility with $\gamma = 3$</i> return draws	26%	75%	99%	60%	5%	76%
<i>Hist. Return Distribution</i> return draws	50%	59%	67%	77%	41%	5%

G. Robustness

In robustness checks for our study, we first investigate whether variations of the state space change our empirical results. Next, we repeat our study but exclude the Early 1990s Recession, the Early 2000s Recession, and the Great Recession from the original sample.

G.1. Variations of the state space

The recovery theorem works on a finite state space and we are concerned that our choices for the states might drive the results; see Tran and Xia (2015). This is not the case when we use (i) log returns instead of straight returns or (ii) a 20% coarser state space than our usual fine state spaces with $N = 111$ (120 for Ross Stable and Power Utility). If we use (iii) the non-overlapping state space with $N = 12$ from Ross Original, then we reject all Ross recovery methods and also Power Utility. The fit to the option prices deteriorates markedly in the process. Note that the Historical Return Distribution is always unaffected by choice of the state space, and we thus always maintain our result that we cannot reject the Historical Return Distribution.

Log returns.

We now define our state-space in log-returns and construct our volatility surface by linearly interpolating the log-moneyness of the fine implied volatility surface. In our density tests, we still reject our hypothesis that future returns are compatible with the recovered physical distributions for all Ross recovery methods at the 5% level. For Power Utility, we find lower p -values than for the main results, but we still cannot reject our hypothesis at the 5% level for three out of our four tests. Table 5 presents the density test results. Results for our moment prediction studies barely change as well, except that Ross Basic now predicts means more poorly than before, while Power Utility predicts means slightly better and variances slightly worse; see Table 6 for the mean prediction results and Table 7 for the variance prediction results.

Table 5

Density tests of the recovered physical probabilities for log returns.

We present our results when future log returns are drawn from physical probabilities generated by one of the six log-scaled methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each method, we show the p -values from the Berkowitz, Kolmogorov-Smirnov, and Knüppel (using both three and four moments) tests for uniformity of the percentiles of future returns under the method's physical cumulative distribution.

$H0: p_\tau = \hat{p}_\tau$	<i>Berkowitz</i>	<i>Kolmogorov- Smirnov</i>	<i>Knüppel 3 moments</i>	<i>Knüppel 4 moments</i>
<i>Recovery Method</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	0.000	0.000	0.000	0.000
<i>Ross Bounded</i> $\pi_{i,j} > 0, \text{rowsums} \in [0.9, 1]$	0.000	0.001	0.000	0.000
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, $\text{rowsums} \in [0.9, 1]$	0.000	0.000	0.000	0.000
<i>Ross Stable</i> No transition state prices	0.000	0.006	0.002	0.000
<i>Power Utility</i> with $\gamma = 3$	0.100	0.620	0.415	0.049
<i>Historical Return Distribution</i>	0.762	0.663	0.939	0.833

Table 6

Mean predictions using log returns.

We test whether future log returns r_τ can be predicted by recovered means μ_τ generated by one of our six log-scaled methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regression models.

$r_\tau = a + b\mu_\tau + \epsilon_\tau$	<i>Intercept</i> <i>Model</i> <i>Set $b = 1$</i> <i>Test $a = 0$</i>	<i>Slope</i> <i>Model</i> <i>Set $a = 0$</i> <i>Test $b = 1$</i>	<i>Joint</i> <i>Model</i> <i>Test $a = 0$</i> <i>and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	0.098 [± 0.029] <i>0.000</i>	-0.009 [± 0.015] <i>0.000</i>	0.007 [± 0.005] -0.001 [± 0.016] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	0.024 [± 0.005] <i>0.000</i>	-0.235 [± 0.204] <i>0.000</i>	0.008 [± 0.007] 0.022 [± 0.299] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.028 [± 0.005] <i>0.000</i>	-0.154 [± 0.167] <i>0.000</i>	0.010 [± 0.007] 0.108 [± 0.247] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	0.011 [± 0.005] <i>0.000</i>	-0.351 [± 0.518] <i>0.000</i>	0.008 [± 0.005] 0.029 [± 0.572] <i>0.000</i>
<i>Power Utility</i> with $\gamma = 3$	0.001 [± 0.005] <i>0.600</i>	0.487 [± 0.529] <i>0.057</i>	0.009 [± 0.007] -0.284 [± 0.758] <i>0.004</i>
<i>Historical Return</i> <i>Distribution</i>	0.001 [± 0.005] <i>0.776</i>	0.583 [± 0.499] <i>0.101</i>	0.008 [± 0.007] -0.057 [± 0.744] <i>0.020</i>

Table 7

Variance predictions using log returns.

We test if future realized variances RV_τ based on log-returns can be predicted by recovered variances σ_τ^2 generated by one of our six log-scaled methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regression models.

$RV_\tau = a + b\sigma_\tau^2 + \epsilon_\tau$	<i>Intercept</i> <i>Model</i> <i>Set $b = 1$</i> <i>Test $a = 0$</i>	<i>Slope</i> <i>Model</i> <i>Set $a = 0$</i> <i>Test $b = 1$</i>	<i>Joint</i> <i>Model</i> <i>Test $a = 0$</i> <i>and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	-0.097 [± 0.047] <i>0.000</i>	0.001 [± 0.001] <i>0.000</i>	0.003 [± 0.000] 0.000 [± 0.001] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0, \text{rowsums} \in [0.9, 1]$	-0.014 [± 0.002] <i>0.000</i>	0.127 [± 0.015] <i>0.000</i>	0.001 [± 0.001] 0.110 [± 0.018] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, $\text{rowsums} \in [0.9, 1]$	-0.016 [± 0.002] <i>0.000</i>	0.118 [± 0.013] <i>0.000</i>	0.000 [± 0.001] 0.104 [± 0.017] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	-0.002 [± 0.000] <i>0.000</i>	0.540 [± 0.042] <i>0.000</i>	0.000 [± 0.000] 0.509 [± 0.050] <i>0.000</i>
<i>Power Utility</i> with $\gamma = 3$	0.000 [± 0.000] <i>0.376</i>	1.088 [± 0.095] <i>0.063</i>	-0.001 [± 0.001] 1.275 [± 0.141] <i>0.000</i>
<i>Historical Return</i> <i>Distribution</i>	0.001 [± 0.000] <i>0.026</i>	1.063 [± 0.205] <i>0.547</i>	0.002 [± 0.001] 0.272 [± 0.452] <i>0.001</i>

Reducing the state-space by 20%.

Next, we reduce the overlapping fine state space by 20%. We thus reduce the state space to eight (instead of ten) maturity steps per month and end up with 89 (instead of 111) states for Ross Basic, Ross Bounded, and Ross Unimodal, and 96 (instead of 120) states for Ross Stable and Power Utility. Table 8 presents the density test results. Again, the Ross recovery methods are strongly rejected at the 5% level, while our benchmark methods are not. Our moment predictions barely change at all; see Table 9 and Table 10.

Table 8

Density tests of the recovered physical probabilities using a reduced state-space size.

We present our results when future returns are drawn from physical probabilities generated by one of the six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We reduce the original state-space size by 20%. For each method, we show the p -values from the Berkowitz, Kolmogorov-Smirnov, and Knüppel (using both three and four moments) tests for uniformity of the percentiles of future returns under the method's physical cumulative distribution.

$H0: p_\tau = \hat{p}_\tau$	<i>Berkowitz</i>	<i>Kolmogorov-</i>	<i>Knüppel</i>	<i>Knüppel</i>
<i>Recovery Method</i>	<i>p-value</i>	<i>Smirnov</i>	<i>3 moments</i>	<i>4 moments</i>
	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	0.000	0.010	0.000	0.000
<i>Ross Bounded</i> $\pi_{i,j} > 0, \text{ rowsums} \in [0.9, 1]$	0.000	0.039	0.010	0.000
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, $\text{rowsums} \in [0.9, 1]$	0.000	0.027	0.002	0.000
<i>Ross Stable</i> No transition state prices	0.000	0.003	0.000	0.000
<i>Power Utility</i> with $\gamma = 3$	0.117	0.447	0.459	0.062
<i>Historical Return</i> <i>Distribution</i>	0.763	0.663	0.939	0.832

Table 9

Mean predictions using a reduced state-space size.

We test if future returns r_τ can be predicted by recovered means μ_τ generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We reduce the original state-space size by 20%. For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regression models.

$r_\tau = a + b\mu_\tau + \epsilon_\tau$	<i>Intercept Model Set $b = 1$ Test $a = 0$</i>	<i>Slope Model Set $a = 0$ Test $b = 1$</i>	<i>Joint Model Test $a = 0$ and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	-0.002 [± 0.006] <i>0.616</i>	-0.011 [± 0.102] <i>0.000</i>	0.009 [± 0.005] -0.058 [± 0.103] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	0.007 [± 0.004] <i>0.001</i>	0.280 [± 1.105] <i>0.201</i>	0.009 [± 0.005] -0.216 [± 1.118] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.008 [± 0.004] <i>0.000</i>	0.164 [± 0.990] <i>0.098</i>	0.008 [± 0.004] -0.009 [± 0.978] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	0.007 [± 0.004] <i>0.004</i>	0.592 [± 1.127] <i>0.477</i>	0.010 [± 0.005] -0.559 [± 1.267] <i>0.001</i>
<i>Power Utility</i> with $\gamma = 3$	-0.002 [± 0.005] <i>0.427</i>	0.394 [± 0.314] <i>0.000</i>	0.009 [± 0.006] -0.065 [± 0.448] <i>0.000</i>
<i>Historical Return Distribution</i>	0.001 [± 0.004] <i>0.788</i>	0.688 [± 0.459] <i>0.156</i>	0.009 [± 0.007] -0.077 [± 0.765] <i>0.022</i>

Table 10

Variance predictions using a reduced state-space size.

We test if future realized variances RV_τ can be predicted by recovered variances σ_τ^2 generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We reduce the original state-space size by 20%. For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regression models.

$RV_\tau = a + b\sigma_\tau^2 + \epsilon_\tau$	<i>Intercept Model Set $b = 1$ Test $a = 0$</i>	<i>Slope Model Set $a = 0$ Test $b = 1$</i>	<i>Joint Model Test $a = 0$ and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	-0.007 [± 0.002] <i>0.000</i>	0.131 [± 0.027] <i>0.000</i>	0.002 [± 0.001] 0.078 [± 0.029] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	-0.005 [± 0.001] <i>0.000</i>	0.398 [± 0.037] <i>0.000</i>	-0.001 [± 0.001] 0.493 [± 0.058] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	-0.008 [± 0.001] <i>0.000</i>	0.304 [± 0.036] <i>0.000</i>	-0.002 [± 0.001] 0.401 [± 0.060] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	-0.002 [± 0.000] <i>0.000</i>	0.683 [± 0.054] <i>0.000</i>	0.000 [± 0.000] 0.730 [± 0.074] <i>0.000</i>
<i>Power Utility</i> with $\gamma = 3$	0.000 [± 0.000] <i>0.518</i>	1.058 [± 0.085] <i>0.185</i>	-0.001 [± 0.000] 1.149 [± 0.118] <i>0.039</i>
<i>Historical Return Distribution</i>	0.001 [± 0.000] <i>0.010</i>	1.124 [± 0.217] <i>0.261</i>	0.002 [± 0.001] 0.252 [± 0.507] <i>0.001</i>

Using a non-overlapping state space with twelve states: Ross Original.

Finally, we use a very coarse non-overlapping state space with just 12 states per month, as in the original work of Ross (2015). For Ross Basic, Ross Bounded, and Ross Unimodal, we now use a 12 by 12 transition state price matrix. For Ross Stable, we recover a one-month stochastic discount factor on a 12-state moneyness grid. For both the density tests and for the moment predictions, unreported results show that we can reject all four Ross recovery methods and also Power Utility with p -values close to zero. Only the results for the Historical Return Distribution are unaffected, as it does not rely on the changed state space.

So far, we chose our 12 states equidistant, which leads to many zeros for the short-maturity spot state prices. Using 12 non-equidistant states leads to a better approximation of the shorter-maturity spot state prices and a poorer approximation of the longer-maturity spot state prices. Yet, results are as disappointing as for the equidistant state space. We construct the non-equidistant state space in the moneyness dimension, by picking 14 moneyness levels, which is reduced by the Breeden-Litzenberger approach to the final 12 moneyness levels. We pick the first 6 moneyness levels at: the beginning $N1$ and the end $N12$ of the moneyness range needed to cover the tails of *one-year* spot state prices; the beginning $N3$ and the end $N10$ of the moneyness range needed to cover the tails of *one-month* spot state prices; a moneyness of zero $N0$ where we assign the same implied volatility value as at $N1$; and a moneyness of three $N13$ where we assign the same implied volatility value as at $N12$.

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 $N2$ to lie in between $N1$ and $N3$; $N4$ to $N9$ (one of them being the current state with moneyness of one) to lie in between $N3$ and $N10$; and $N11$ to lie in between $N10$ and $N12$. For the time dimension we use 12 monthly steps.

Fig. 2 shows one-month spot state prices (black) and 12-month spot state prices (gray dotted) for a non-equidistant state space with $N = 12$ on February 17, 2010.

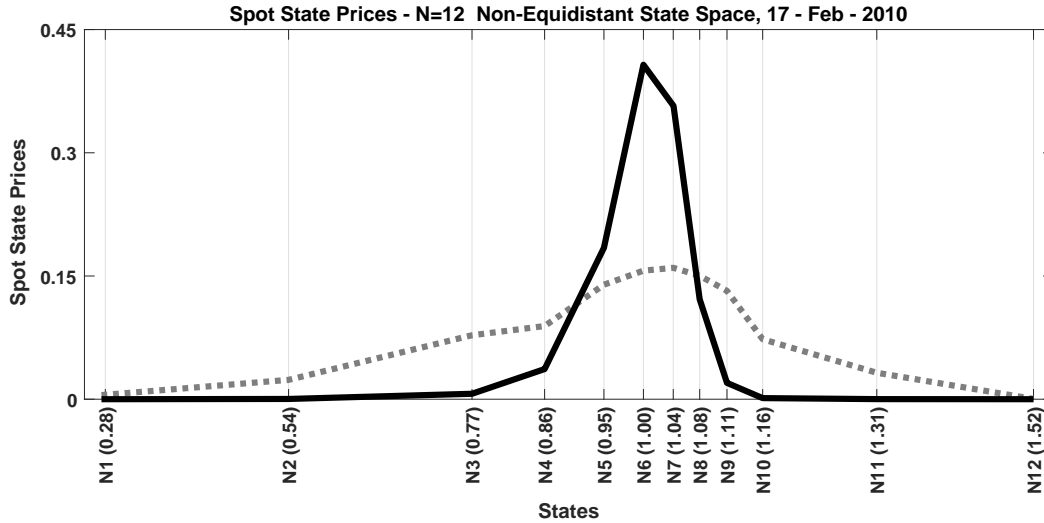


Fig. 2. Spot state prices for a non-equidistant $N = 12$ state-space. We show one-month spot state prices (black) and twelve-month spot state prices (gray dotted) on February 17, 2010 when using a non-equidistant state space with $N = 12$ states.

G.2. Removing recession periods

Our main sample consists of 380 monthly dates from April 1987 through December 2017. This period includes three recession periods: the Early 1990s Recession (July 1990 until March 1991), the Early 2000s Recession (March 2001 until November 2001), and the Great Recession (December 2007 until June 2009).² We want to insure that our results are not driven by these recession periods and consider the reduced sample with only 343 dates.

Table 11 shows results for our density tests under our original setting but excluding recession periods. Table 12 shows the corresponding mean prediction results, and Table 13 the corresponding variance prediction results.

Our density tests do not change much. Also, the moment predictions barely change for the Ross recovery methods. Yet for Power Utility and the Historical Return Distribution, removing recession periods leads to much better mean and poorer variance predictions.

²See, e.g.: <https://fred.stlouisfed.org/series/VIXCLS>

Table 11

Density tests of the recovered physical probabilities when excluding recession periods.

We present our results when future returns are drawn from physical probabilities generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We use our sample period from April 1986 until December 2017 but exclude months that are associated with the Early 1990s Recession (July 1990 until March 1991), the Early 2000s Recession (March 2001 until November 2001), and the Great Recession (December 2007 until June 2009). For each method, we show the p -values from the Berkowitz, Kolmogorov-Smirnov, and Knüppel (using both 3 and 4 moments) tests for uniformity of the percentiles of future returns under the method physical cumulative distribution.

$H0: p_\tau = \hat{p}_\tau$	<i>Berkowitz</i>	<i>Kolmogorov-Smirnov</i>	<i>Knüppel</i> <i>3 moments</i>	<i>Knüppel</i> <i>4 moments</i>
<i>Recovery Method</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	0.001	0.038	0.000	0.000
<i>Ross Bounded</i> $\pi_{i,j} > 0, \text{rowsums} \in [0.9, 1]$	0.000	0.033	0.008	0.000
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.000	0.032	0.000	0.000
<i>Ross Stable</i> No transition state prices	0.000	0.016	0.013	0.001
<i>Power Utility</i> with $\gamma = 3$	0.093	0.689	0.532	0.057
<i>Historical Return</i> <i>Distribution</i>	0.805	0.522	0.946	0.871

Table 12

Mean predictions when excluding recession periods.

We test if future returns r_τ can be predicted by recovered means μ_τ generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We use our sample period from April 1986 until December 2017 but exclude months that are associated with the Early 1990s Recession (July 1990 until March 1991), the Early 2000s Recession (March 2001 until November 2001), and the Great Recession (December 2007 until June 2009). For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regressions.

$r_\tau = a + b\mu_\tau + \epsilon_\tau$	<i>Intercept Model Set $b = 1$ Test $a = 0$</i>	<i>Slope Model Set $a = 0$ Test $b = 1$</i>	<i>Joint Model Test $a = 0$ and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	0.001 [± 0.006] <i>0.659</i>	0.062 [± 0.095] <i>0.000</i>	0.009 [± 0.004] 0.035 [± 0.094] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	0.009 [± 0.004] <i>0.000</i>	0.117 [± 0.888] <i>0.051</i>	0.010 [± 0.004] -0.124 [± 0.870] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.009 [± 0.004] <i>0.000</i>	0.073 [± 0.716] <i>0.011</i>	0.010 [± 0.004] -0.097 [± 0.701] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	0.008 [± 0.004] <i>0.000</i>	1.176 [± 1.248] <i>0.781</i>	0.009 [± 0.005] 0.075 [± 1.331] <i>0.000</i>
<i>Power Utility</i> with $\gamma = 3$	0.001 [± 0.004] <i>0.785</i>	0.689 [± 0.370] <i>0.099</i>	0.009 [± 0.007] 0.079 [± 0.594] <i>0.010</i>
<i>Historical Return Distribution</i>	0.002 [± 0.004] <i>0.486</i>	0.732 [± 0.425] <i>0.216</i>	0.011 [± 0.007] -0.138 [± 0.718] <i>0.007</i>

Table 13

Variance predictions when excluding recession periods.

We test if future realized variances RV_τ can be predicted by recovered variances σ_τ^2 generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. We use our sample period from April 1986 until December 2017 but exclude months that are associated with the Early 1990s Recession (July 1990 until March 1991), the Early 2000s Recession (March 2001 until November 2001), and the Great Recession (December 2007 until June 2009). For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regressions.

$RV_\tau = a + b\sigma_\tau^2 + \epsilon_\tau$	<i>Intercept Model Set $b = 1$ Test $a = 0$</i>	<i>Slope Model Set $a = 0$ Test $b = 1$</i>	<i>Joint Model Test $a = 0$ and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	-0.006 [± 0.002] <i>0.000</i>	0.077 [± 0.017] <i>0.000</i>	0.002 [± 0.000] 0.037 [± 0.015] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	-0.007 [± 0.000] <i>0.000</i>	0.241 [± 0.021] <i>0.000</i>	0.000 [± 0.000] 0.286 [± 0.039] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	-0.007 [± 0.001] <i>0.000</i>	0.221 [± 0.019] <i>0.000</i>	0.000 [± 0.000] 0.230 [± 0.034] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	-0.001 [± 0.000] <i>0.000</i>	0.698 [± 0.053] <i>0.000</i>	0.000 [± 0.000] 0.640 [± 0.078] <i>0.000</i>
<i>Power Utility</i> with $\gamma = 3$	0.000 [± 0.000] <i>0.003</i>	0.830 [± 0.064] <i>0.000</i>	0.000 [± 0.000] 0.780 [± 0.098] <i>0.000</i>
<i>Historical Return Distribution</i>	0.000 [± 0.000] <i>0.724</i>	0.893 [± 0.120] <i>0.080</i>	0.001 [± 0.001] 0.318 [± 0.271] <i>0.000</i>

G.3. Weekly options

We check whether Ross recovery with weekly instead of monthly options leads to the same results. The weekly S&P 500 options were introduced in 2005 and expire each Friday. We process weekly option data analogous to how we process monthly option data; see Online Appendix B and Section 3.1 in the main paper. On each expiry date, we go seven calendar days back to find our sample dates. The weekly options are concentrated in the short maturities (one to 13 weeks) whereas the monthly options are spread out across longer maturities (four to 52 weeks). In total, and up with 597 weekly sample dates from October 14, 2005 through the December 29, 2017.

Next, we compute spot state prices analogous to our approach for monthly options. We discretize the maturity dimension with 10 steps each week resulting in a total of 130 maturity steps. Following our approach in the main paper, we then recover physical distributions for our four Ross recovery methods and our two benchmark models in the weekly setting. The weekly physical distributions recovered from Ross Basic, Ross Bounded, and Ross Unimodal are defined on $N = 121$ states. For Ross Stable and Power Utility with $\gamma = 3$, the distributions are defined on $N = 130$ states. We construct the Historical Return Distribution based on the last 60 weekly observed returns.

We then repeat our main study. Density test results are shown in Table 14, mean prediction results are shown in Table 15, and variance prediction results are shown in Table 16. Again, we strongly reject all Ross recovery methods using density tests (only Ross Bounded cannot be rejected by the KS test, and Ross Stable cannot be rejected by the Knüppel 3 moments test). Power Utility and Historical Return Distribution are only rejected by the Berkowitz test.

In the weekly mean predictions, there are three changes compared to the main results. Ross Bounded, Ross Stable, and the Historical Return Distribution are now rejected in the Slope model. In the weekly variance predictions, results are similar to the main results except that Ross Stable is no longer rejected in the Intercept model, and Power Utility is no longer rejected in the Joint model.

In untabulated results, we repeated the density tests in the machine learning setting. Ross Basic ML is now rejected by all four density tests. Results for the other methods remain unchanged. Altogether, results do not change much one way or the other and are robust with respect to using weekly options.

Table 14

Density tests of the recovered physical probabilities using weekly options.

We present our results when future returns are drawn from physical probabilities generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each method, we show the p -values from the Berkowitz, Kolmogorov-Smirnov, and Knüppel (using both 3 and 4 moments) tests for uniformity of the percentiles of future returns under the method physical cumulative distribution.

$H0: p_\tau = \hat{p}_\tau$	<i>Berkowitz</i>	<i>Kolmogorov- Smirnov</i>	<i>Knüppel 3 moments</i>	<i>Knüppel 4 moments</i>
<i>Recovery Method</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	0.000	0.007	0.000	0.000
<i>Ross Bounded</i> $\pi_{i,j} > 0, \text{rowsums} \in [0.9, 1]$	0.000	0.093	0.042	0.000
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.000	0.000	0.001	0.000
<i>Ross Stable</i> No transition state prices	0.000	0.034	0.054	0.002
<i>Power Utility</i> with $\gamma = 3$	0.039	0.472	0.500	0.067
<i>Historical Return Distribution</i>	0.006	0.680	0.848	0.457

Table 15

Mean predictions when using weekly options.

We test if future returns r_τ can be predicted by recovered means μ_τ generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regressions.

$r_\tau = a + b\mu_\tau + \epsilon_\tau$	<i>Intercept Model Set $b = 1$ Test $a = 0$</i>	<i>Slope Model Set $a = 0$ Test $b = 1$</i>	<i>Joint Model Test $a = 0$ and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	-0.001 [± 0.003] <i>0.471</i>	0.041 [± 0.078] <i>0.000</i>	0.002 [± 0.002] 0.035 [± 0.078] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	0.005 [± 0.002] <i>0.000</i>	-0.415 [± 0.540] <i>0.000</i>	0.001 [± 0.003] -0.160 [± 0.936] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	-0.123 [± 0.005] <i>0.000</i>	0.012 [± 0.014] <i>0.000</i>	0.001 [± 0.005] 0.004 [± 0.038] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	0.002 [± 0.002] <i>0.037</i>	-0.389 [± 1.372] <i>0.047</i>	0.002 [± 0.002] -0.014 [± 1.455] <i>0.045</i>
<i>Power Utility</i> with $\gamma = 3$	0.000 [± 0.002] <i>0.617</i>	0.384 [± 0.399] <i>0.003</i>	0.001 [± 0.002] 0.298 [± 0.442] <i>0.007</i>
<i>Historical Return Distribution</i>	0.000 [± 0.002] <i>0.874</i>	0.078 [± 0.570] <i>0.002</i>	0.002 [± 0.002] -0.143 [± 0.627] <i>0.002</i>

Table 16

Variance predictions when using weekly options.

We test if future realized variances RV_τ can be predicted by recovered variances σ_τ^2 generated by one of our six methods: Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility, and Historical Return Distribution. For each method, we report the p -values and estimated coefficients with the 95% confidence intervals for three different regressions.

$RV_\tau = a + b\sigma_\tau^2 + \epsilon_\tau$	<i>Intercept Model Set $b = 1$ Test $a = 0$</i>	<i>Slope Model Set $a = 0$ Test $b = 1$</i>	<i>Joint Model Test $a = 0$ and $b = 1$</i>
<i>Recovery Method</i>	Intercept [95% CI] <i>p-value</i>	Slope [95% CI] <i>p-value</i>	Intercept [95% CI] Slope [95% CI] <i>p-value</i>
<i>Ross Basic</i> $\pi_{i,j} > 0$	-0.003 [± 0.001] <i>0.000</i>	0.075 [± 0.013] <i>0.000</i>	0.001 [± 0.000] 0.059 [± 0.013] <i>0.000</i>
<i>Ross Bounded</i> $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	-0.002 [± 0.000] <i>0.000</i>	0.343 [± 0.024] <i>0.000</i>	0.000 [± 0.000] 0.415 [± 0.031] <i>0.000</i>
<i>Ross Unimodal</i> $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	-0.008 [± 0.001] <i>0.000</i>	0.112 [± 0.009] <i>0.000</i>	-0.001 [± 0.000] 0.144 [± 0.012] <i>0.000</i>
<i>Ross Stable</i> No transition state prices	0.000 [± 0.000] <i>0.056</i>	0.894 [± 0.052] <i>0.000</i>	0.000 [± 0.000] 0.895 [± 0.060] <i>0.000</i>
<i>Power Utility</i> with $\gamma = 3$	0.000 [± 0.000] <i>0.717</i>	0.966 [± 0.064] <i>0.300</i>	0.000 [± 0.000] 0.950 [± 0.073] <i>0.372</i>
<i>Historical Return Distribution</i>	0.000 [± 0.000] <i>0.025</i>	0.899 [± 0.153] <i>0.195</i>	0.000 [± 0.000] 0.657 [± 0.193] <i>0.000</i>

H. The relation between transition state prices and the stochastic discount factor

The rowsums of the transition state price matrix are inversely related to the Ross recovery SDF. Importantly, any modeling choice that affects rowsums will thus determine the Ross recovery SDF. As the machine learning implementation forces away-from-the-current-state transition state prices (and thus rowsums) downward, a U-shaped SDF obtains. Even though we cannot reject Ross Basic ML, we are nevertheless concerned that the non-rejection is largely due to the mechanics of the machine learning approach. We now document the inverse relation between rowsums and SDF empirically, mathematically in a two-state setting, and in a simulation using our usual state space.

Empirically for our 380 dates, we observe only negative correlations between our recovered SDFs (using Ross Basic ML) and the rowsums of the corresponding transition state price matrices ranging from -0.09 to -0.82 with an average correlation of -0.49. These negative correlations are significantly lower than zero in 372 out of 380 cases (98%). We choose Ross Basic ML here as this is the method that forces the recovered SDF to be U-shaped. For the other methods, we also observe significantly negative correlations.

Mathematically, we prove the negative relation in the following $N = 2$ state example. Let the transition state price matrix be given by Π where, without loss of generality, we assume that the current state is $i = 0$ and the rowsum with initial state $i = 0$ is larger than the rowsum with initial state $i = 1$:³

$$\pi_{0,0} + \pi_{0,1} > \pi_{1,0} + \pi_{1,1}. \quad (16)$$

Now consider the eigenvalue problem from Eq. (3) in the $N = 2$ case:

$$\begin{aligned} \pi_{0,0} \cdot z_0 + \pi_{0,1} \cdot z_1 &= \delta \cdot z_0 \\ \pi_{1,0} \cdot z_0 + \pi_{1,1} \cdot z_1 &= \delta \cdot z_1. \end{aligned} \quad (17)$$

Let $x = \frac{z_0}{z_1}$. We already know that the Perron-Frobenius theorem implies that $x > 0$. Solving each of the two equations in Eq. (17) for δ gives:

$$\pi_{0,0} + \frac{\pi_{0,1}}{x} = \delta = \pi_{1,0} \cdot x + \pi_{1,1} \quad (18)$$

³Note that we do not cover equal rowsums in this proof. Equal rowsums would lead to a flat SDF, as already shown in Ross (2015).

$$\Leftrightarrow \pi_{1,0} \cdot x^2 + (\pi_{1,1} - \pi_{0,0})x - \pi_{0,1} = 0 \quad (19)$$

Rewriting Eq. (16) as $\pi_{0,1} - \pi_{1,0} > \pi_{1,1} - \pi_{0,0}$ and inserting it into Eq. (19) gives:

$$\begin{aligned} \pi_{1,0} \cdot x^2 + (\pi_{0,1} - \pi_{1,0})x - \pi_{0,1} &> 0 \\ \Leftrightarrow \pi_{1,0} \cdot x(x-1) + \pi_{0,1} \cdot (x-1) &> 0. \end{aligned} \quad (20)$$

Given that $x > 0$, this inequality holds only if $x > 1$. We further know that the recovered SDF with current initial state $i = 0$ is given by:

$$SDF = \begin{pmatrix} \delta, \delta \cdot \frac{z_0}{z_1} \end{pmatrix} = (\delta, \delta \cdot x), \quad (21)$$

which is increasing from state 0 to state 1 for $x > 1$. This proves that if we observe *decreasing* rowsums of Π in the $N = 2$ case, we always end up with a recovered SDF that is *increasing* in states. The relation between increasing rowsums and decreasing SDF follows analogously.

Finally, we use a simulation to show the negative relation for larger state spaces, where we cannot show the result mathematically. In line with our empirical study, we use a state space 0 to 1.6 with dimension of $N = 111$, and simulate $M = 10,000$ different transition state price matrices on that state space.

Each simulated matrix Π is constructed as follows. First, we fill each row of Π with a truncated normal distribution function with mean equal to the value of the initial state of that row and standard deviation uniformly drawn between values of 0.15 and 0.3. After each row is filled, we add noise to each value with a probability of 10% and leave the value unchanged otherwise. The noise we add is lognormally distributed and ranges from 0 to 8.60 with an average value of 0.29. For each row of Π , we then determine a target rowsum by uniformly drawing a value between a lower bound $lb = 0$ and an upper bound $ub = 10$. Next, we normalize each row so that it sums to the specified target rowsum.

We generate 10,000 different transition state price matrices, recover the SDF, and compute its correlation with the rowsums of Π . All correlations are negative and lie between -0.15 and -0.82 with an average correlation of -0.42. 97% of the correlations are statistically significant. These findings are in line with the empirical values above. When we restrict rowsums to lie between $lb = 0.9$ and $ub = 1$, as in Ross Bounded, the average correlation is -0.52. We conclude that there is a strong negative relation between rowsums and the SDF.

I. Insights from simulated economies

It is conceivable that small data errors in option prices might cause the recovery theorem to fail. To check, we simulate economies where a particular recovery method holds true and generate future returns by drawing from the recovered physical distributions. In that case and with a 5% significance level for our statistical tests, we should have a 5% rejection rate (i.e., future returns are incompatible with the simulated recovery method).

Next, we perturb the option prices. This will increase the rejection rates, as the perturbed recovery methods generate physical distributions that deviate from the true physical distribution, from which we drew the future returns.

Our methodology is as follows. For each date τ 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 $l = 1, \dots, L$. At date τ , we further draw a one-month future return r_τ^{true} from the *physical* distributions p_τ^{true} that we recovered with a particular recovery method. This gives us a time series of 380 one-month future returns drawn from an economy where the particular recovery method holds.

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

$$\hat{\sigma}_{i(l),t(l)} = \sigma_{i(l),t(l)} + v \cdot \epsilon, \quad \epsilon \sim N(0, 1), \quad (22)$$

where v on each date τ is a multiple of the mean implied volatility bid/ask spread on that date.⁴ Our four levels of perturbation are none, one-half, one, and twice the mean spread of a particular date.

We apply our standard algorithm to recover the physical distributions \hat{p}_τ , which are now based on the *perturbed* implied volatilities $\hat{\sigma}_{i(l),t(l)}$. We then use a Berkowitz test to test the hypothesis that the simulated future returns r_τ^{true} are compatible with the recovered distributions \hat{p}_τ .⁵ We draw 10,000 realizations of future return time series, each consisting of 380 returns, and report the rejection rates for our hypothesis at the 5% significance level.

Table 17 presents how often our hypothesis is rejected at the 5% significance level for different multipliers v of the error-term ϵ . Without any perturbation, the rejection rates all equal the theoretical value of 5%. Beyond that, we find that only Ross Basic turns out to be

⁴On each date, we compute the mean across all options of (ask implied volatility - bid implied volatility). The mean spread ranges from 0.001 to 0.085 for our 380 sample dates, with an average value of 0.013.

⁵Using the Knüppel test with four moments at the 5% level does not change our results qualitatively.

very sensitive to perturbations. Even though the simulations assume that Ross Basic holds, we have rejection rates of around 41% to 99% with perturbations instead of 5% without. Thus, the failure of Ross Basic could be driven by errors in the option implied volatilities.

Yet for all other methods, the increased the rejection rates are barely noticeable up to perturbations by the mean spread and much slower than for Ross Basic as we increase our perturbations to two times the mean spread.⁶ We thus reason that the empirical failure of Ross Bounded, Ross Unimodal, and Ross Stable is not likely to be driven by perturbations of the option prices.

We conclude that, while Ross Bounded, Ross Unimodal, and Ross Stable suffer from rather flat SDFs, the methods are less sensitive to perturbations of the option prices than Ross Basic. It seems that adding economic constraints results in methods that are less sensitive to perturbations. This finding also holds for Power Utility.

Table 17

Stability of simulated recovery methods.

We generate 10,000 realizations of an economy where a particular recovery method holds. For each realization, we draw 380 future returns. Next, we perturb option implied volatilities by adding v times a standard normally distributed error ϵ . The v of each date τ is based on the mean implied volatility bid/ask spread of that date times 0, 0.5, 1, or 2. Given v , we report the rejection rates for the hypothesis that the future returns are drawn from the perturbed physical distributions. We use a Berkowitz test with a significance level of 5%.

<i>Recovery Method</i>	$v = 0 \cdot \text{spread}$ rej. rate	$v = \frac{1}{2} \cdot \text{spread}$ rej. rate	$v = 1 \cdot \text{spread}$ rej. rate	$v = 2 \cdot \text{spread}$ rej. rate
<i>Ross Basic</i>	5%	41%	59%	99%
<i>Ross Bounded</i>	5%	5%	6%	12%
<i>Ross Unimodal</i>	5%	5%	6%	10%
<i>Ross Stable</i>	5%	5%	6%	20%
<i>Power Utility</i>	5%	5%	7%	41%
<i>Hist. Return Distr.</i>	n/a	n/a	n/a	n/a

⁶The rejection rate cannot be computed for the Historical Return Distribution.

REFERENCES

- Breeden, D. T., Litzenberger, R. H., 1978. Prices of state-contingent claims implicit in option prices. *J. Bus.* 51, 621–651.
- Jackwerth, J. C., 2004. Option-implied risk-neutral distributions and risk aversion. Research Foundation of AIMR, CFA Institute.
- Ross, S., 2015. The recovery theorem. *J. Financ.* 70, 615–648.
- Tran, N.-K., Xia, S., 2015. Specified recovery. Working Paper, Washington University in St. Louis.