Journal of Financial Economics 137 (2020) 723–739

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Does the Ross recovery theorem work empirically?

Jens Carsten Jackwerth[a,*], Marco Menner[b,1]

[a] University of Konstanz, PO Box 134, Konstanz 78457, Germany
[b] Univ. Ramon Llull, ESADE, Avenida de Torreblanca 59, Sant Cugat 08172, Spain

ARTICLE INFO

Article history:
Received 4 July 2017
Revised 15 August 2019
Accepted 12 September 2019
Available online 19 March 2020

JEL classification:
G13

Keywords:
Ross recovery
Stochastic discount factor
Risk-neutral density
Transition state prices
Physical probabilities

ABSTRACT

Starting with the fundamental relation that state prices are the product of physical probabilities and the stochastic discount factor, Ross (2015) shows that, given strong assumptions, knowing state prices suffices to back out physical probabilities and the stochastic discount factor at the same time. We find that such recovered physical distributions based on the S&P 500 index are incompatible with future returns and fail to predict future returns and realized variances. These negative results are even stronger when we add economically reasonable constraints. Simple benchmark methods based on a power utility agent or the historical return distribution cannot be rejected.

1. Introduction

Much of financial economics revolves around the triangular relation among physical return probabilities p, which are state prices π divided by the stochastic discount factor (SDF) m:

$$\text{physical probability } p = \frac{\text{state price } \pi}{\text{SDF } m}. \tag{1}$$

Researchers typically pick any two variables to find the third. Yet Ross (2015), based on earlier work by Hansen and Scheinkman (2009), presents a recovery theorem that allows a researcher to back out both the SDF and physical probabilities by using state prices only. To achieve this, he needs to make strong assumptions. We investigate this claim and test whether the recovered physical probabilities are compatible with future S&P 500 returns. We further analyze whether the shape of the recovered SDF is consistent with utility theory. To understand our results, we discuss in detail why the recovery theorem does not perform well empirically.

Ross (2015) makes three explicit assumptions. First, he assumes 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. Second, all transition state prices need to be positive. Third, the SDF $m_{i,j}$ is restricted to be a constant times the ratio of values in state j and values in state i.

Taken altogether, the three assumptions allow Ross (2015) to formulate an eigenvalue problem based on transition state prices. Its unique solution yields the physical transition probabilities $p_{i,j}$, which represent physical probabilities of moving from state i to state j, and the SDF.

* Corresponding author.
E-mail addresses: jens.jackwerth@uni-konstanz.de (J.C. Jackwerth), marco.menner@esade.edu (M. Menner).

[1] We received helpful comments and suggestions from Bjorn Eraker, Guenter Franke, Anisha Ghosh, Eric Renault, and Christian Schlag. We thank participants at the 2016 ESSFM Workshop at Gerzensee, 2016 DGF Conference in Bonn, 2018 American Finance Association Meeting in Philadelphia, 2018 AFBC in Sydney, 2018 EFMA conference in Milan, University of Bolzano, University of St. Gallen, University of Konstanz, University of Queensland, University of Strasbourg, University of Sydney, and University of Zurich.

https://doi.org/10.1016/j.jfineco.2020.03.006
0304-405X/

Yet how useful are the spot physical probabilities $p_{0,j}$ emanating from the current state $i = 0$? Are future returns really drawn from those recovered distributions? We reject this hypothesis strongly using four different density tests: the Berkowitz (2001) test, two versions of the uniformity test introduced by Knüppel (2015), and the Kolmogorov-Smirnov test. Further, we show that the means and variances of the recovered physical distributions cannot predict future returns and realized variances. Also, Ross recovery does not produce downward-sloping SDFs (as one would expect given risk-averse preferences) but rather ones that are riddled with local minima and maxima.

Two simple benchmark methods by contrast perform well on all accounts. Our first benchmark uses a power utility SDF to transform option-implied state prices into physical probabilities. Our second benchmark uses simply the five-year historical return distribution. Neither benchmark can be rejected.

The empirical problems of Ross recovery stem from several sources and are related to his strong assumptions. First, it is hard to obtain transition state prices from option prices. Our basic implementation of Ross recovery requires only positivity of transition state prices and leads to unstable transition state prices that exhibit multimodality and imply extreme risk-free rates in different states. Yet if we introduce economically reasonable constraints to restrict the transition state prices (bounding state-dependent risk-free rates and requiring unimodal transition state prices), Ross recovery generates almost flat SDFs. A stable version that uses only spot state prices without the need to estimate transition state prices also does not work.

Second, we argue that the strong assumption concerning the functional form of the SDF is rather limiting. For example, the state-dependent SDF in Ross recovery is incompatible with the state-independent one under power utility. We further show that the recovered SDF is highly dependent on the structure of the transition state price matrix, which in turn is not well-identified from option prices.

Third, the assumption of time-homogeneous transition state prices might not hold, and different periods may require different transition state prices. Indeed, we show empirically a poor simultaneous fit to short- and long-dated options for most Ross recovery methods.

We add to the empirical literature concerning Ross recovery by pinpointing exactly where Ross recovery goes awry. Moreover, we answer the intriguing question as to where Ross recovery, despite its theoretical shortcomings, might still be useful empirically as a rough approximation of reality. Alas, our work confirms the negative theoretical outlook.

As an alternative to the numerically difficult recovery of transition state prices from spot state prices, we suggest an additional implicit method (Ross Stable), which obviates that recovery and works directly with spot state prices. In independent work, Jensen et al. (2019) suggest the same method but add further structure to the SDF. They then focus on analysis of the theoretical properties of their generalized recovery. They also employ a Berkowitz density test (where they reject their model) and predict means and volatilities (where they find little mean predictability in a setting with look-ahead bias and, due to higher persistence, better volatility predictability). We differ by choosing two simple methods close to the original Ross (2015) work, two further economically constrained Ross recovery methods, two benchmark methods, more density tests, more mean and variance predictions, an analysis of the reasons for failure, and the longest S&P 500 option data sample in the literature (April 1986 through December 2017).

Next, we apply machine learning to Ross recovery.[2] We use a cross-validated elastic net regularization, which forces transition state prices to zero. For the basic implementation of Ross recovery using machine learning, we cannot reject Ross recovery. This nonrejection stems from two sources. For one, the elastic net regularization mechanically induces a U-shaped SDF that makes rejection harder. Second, recovered physical probabilities are very noisy, which leads to nonrejection in our statistical tests. Once we add reasonable economic constraints, we are back to a strong rejection of Ross recovery.

This paper relates to several others in a nascent literature on Ross recovery. Borovicka et al. (2016) theoretically analyze the decomposition of the SDF into a permanent and a transitory part. Using long-dated bond option data, Bakshi et al. (2018) empirically find the permanent component to be time varying. This violates an implicit assumption made by Ross (2015), namely that the permanent component is constant at one. Our empirical approach differs significantly, as we directly implement several Ross recovery versions using S&P 500 index option data. Jensen et al. (2019) develop, in parallel with us, a version of Ross recovery without the need to explicitly estimate the transition state prices. They concentrate on the theoretical properties of this particular method yet also test it in a short empirical section and obtain for this method results similar to ours. To our knowledge, we are the first to provide a rigorous empirical investigation of Ross recovery using S&P 500 index option data, thereby linking our empirical findings to the theoretical properties of the recovery theorem and the existing literature.

Closest to our work are Audrino et al. (2020), who also implement Ross recovery on S&P 500 index options while forcing down state prices. Their recovered SDFs tend to be rather smooth and U-shaped, as in our machine learning setting. Their further empirical focus is on development of profitable trading strategies based on recovered physical probabilities, but unlike us, they do not statistically test if future returns are drawn from the recovered distribution.

Dillschneider and Maurer (2018) apply our main approach without the downward-forcing of transition state prices to confirm our density test results. They further generalize Ross recovery to unbounded continuous state spaces. Yao (2018) reproduces our study using both short- and long-dated S&P 500 options and confirms our results. Massacci et al. (2016) use a fast nonlinear programming approach for Ross recovery, which allows for economic constraints such as positive state-dependent risk-free rates and the unimodality of transition state prices.

[2] We thank an anonymous referee for this suggestion.

While we do not add to the theoretical literature, Ross recovery has been theoretically extended by Carr and Yu (2012) and Walden (2017). Additional works on Ross recovery in continuous time are Dubynskiy and Goldstein (2013), Qin and Linetsky (2016), and Qin et al. (2018).

The paper proceeds as follows. Section 2 explains the Ross recovery theorem. In Section 3, 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. Section 4 states our hypothesis, introduces our tests, and describes our results. Section 5 applies machine learning to Ross recovery. In Section 6 we provide reasons why the Ross recovery theorem fails empirically. Section 7 concludes.

2. The Ross recovery theorem

The Ross (2015) recovery theorem is based on three explicit 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. This means they are independent of calendar time. Such transition state prices include the usual spot state prices $\pi_{0,j}$ where zero is 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.[3] We use information on spot state prices with different maturities to obtain these transition state prices and suggest several different methods, including one that allows us to apply the recovery theorem directly to spot state prices $\pi_{0,j}$ without the need to estimate all the other transition state prices.

Second, all transition state prices $\pi_{i,j}$ need to be strictly positive, which turns out to be a fairly benign assumption, as we can easily enforce positive transition state prices.

Third, the SDF $m_{i,j}$ is transition-independent, which means it can be written as

$$m_{i,j} = \delta \frac{u'_j}{u'_i}, \tag{2}$$

for a positive constant δ and positive state-dependent values u'_j and u'_i. Ross (2015) suggests a possible interpretation of those values u' as marginal utilities while viewing δ as a utility discount factor.

Given the three assumptions and using Eq. (1), the physical transition probabilities $p_{i,j}$ have the form

$$p_{i,j} = \frac{\pi_{i,j}}{m_{i,j}} = \frac{1}{\delta} \cdot \frac{\pi_{i,j} \cdot u'_i}{u'_j}. \tag{3}$$

For a fixed initial state i, the physical transition probabilities $p_{i,j}$ across future states j need to sum to one. After rearrangement, this leads to

$$\sum_j \pi_{i,j} \frac{1}{u'_j} = \delta \cdot \frac{1}{u'_i} \quad \forall i. \tag{4}$$

Given N different states with indices i and j both running from 1 to N, Ross collects the transition state prices $\pi_{i,j}$ in an $N \times N$ matrix Π. Ross then collects the N equations from Eq. (4) and formulates an eigenvalue problem, where z is a vector of the inverse marginal utilities:

$$\Pi z = \delta z, \quad \text{where} \quad z_i = \frac{1}{u'_i}. \tag{5}$$

An application of the Perron-Frobenius theorem leads to the result that there is only one eigenvector z with strictly positive entries z_i. That eigenvector corresponds to the highest positive eigenvalue δ of the eigenvalue problem. This property implies a unique positive SDF $m_{i,j}$, as in Eq. (2) and unique physical transition probabilities $p_{i,j}$ for all i and j across N states, as in Eq. (3). We provide a worked example with $N = 2$ states in Online Appendix A.

Relying on the three assumptions, Borovicka et al. (2016) argue that the SDF is the product of a transitory component and a permanent component; Ross (2015) implicitly sets the permanent component to unity. Using long-dated bond option prices instead of index option prices, Bakshi et al. (2018) empirically find the permanent component to be time varying, contradicting the assumptions above and hinting at potential empirical problems. For now, we keep the assumptions in place and interpret, as Ross (2015) did, the recovered transitory SDF to be identical to the total economy-wide SDF, as the permanent SDF component is unity. This allows us to recover SDFs and physical probabilities, where we show the latter to be incompatible with future returns. We later use the recovered quantities to analyze the reasons for failure and link them to violations of the three assumptions.

We always label the current state as $i = 0$ in a set of state indices $I = \{-N_{low}, \ldots, 0, \ldots, N_{high}\}$, where $N = N_{low} + N_{high} + 1$. The ending transition state index j is drawn from the same set I. The $0th$ row of Π contains the one-period transition state prices, starting from the current state, which in theory coincides with the one-period spot state prices.

3. Methodology

The basic ingredient missing for Ross recovery at this point is the matrix Π of transition state prices, which are not easily observable in the market. Yet we can extract transition state prices by exploiting that they link spot state prices at different maturities with each other. We can readily obtain those spot state prices from observed option prices.

3.1. Obtaining spot state prices from observed option prices

We collect European put and call option quotes on the S&P 500 from the Berkeley Options Database (April 1986 through December 1995) and from OptionMetrics (January 1996 through December 2017), see Online Appendix B. We

[3] Imagine that the current state of the world is characterized by the S&P 500 standing 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 one period later.

consider 380 monthly sample dates, which we find by going 30 calendar days back in time from the option expiration date. After applying standard filters, we average bid and ask quotes to obtain midpoint option prices, which we then transform to implied volatilities. For any given sample day, quotes are available only for specific moneyness levels and maturities, yet we would like to obtain state prices on a grid compatible with the recovery theorem. Thus, we first generate a smooth implied volatility surface on a fine auxiliary grid, from which we later interpolate to the grid required for the recovery theorem.

For the auxiliary grid, we discretize option strike prices with a step size of \$1.25 and add additional lower and higher states so that all positive spot state prices occur in the interior of the state set. The number of auxiliary states varies between 227 for earlier sample days and 2,579 for later sample days. We then convert strike prices into moneyness levels by normalizing them by the current level of the S&P 500 index. We discretize the maturity dimension into 120 steps with 10 steps per month. This discretization ensures that all available option prices lie on our fine auxiliary grid.

To generate a smooth implied volatility surface, we apply an extension of the fast and stable method in Jackwerth (2004) that balances minimizing the sum of squared local total second implied volatility derivatives (insuring smoothness of the volatility surface) and minimizing the sum of squared deviations of model implied volatilities from the observed implied volatilities (insuring the fit of the surface).

To obtain state prices on a coarser grid suitable for the recovery theorem, we linearly interpolate the fine implied volatility surface. We typically use more than 100 states, while Ross (2015) originally uses only 12 states. From the implied volatilities on the 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 (see Online Appendix C).

3.2. Finding transition state prices from spot state prices: Ross Original

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 $\pi_{i,j}$ because they link spot state prices $\pi^t_{0,j}$ at different maturities t with one another. Using 12 nonoverlapping, monthly transition state prices and spot state prices with monthly maturities of up to one year, we can state the relations:

$$\pi^{t+1}_{0,j} = \sum_{h\in I} \pi^t_{0,h} \cdot \pi_{h,j} \quad \forall j \in I,\ t = 0, \ldots, 11, \tag{6}$$

where today's spot state prices with a maturity of zero $\left(\pi^0_{0,j}\right)$ are zero for all states except the current state, for which the spot state price is one. Eq. (6) states that one can find the spot state price $\pi^{t+1}_{0,j}$ of reaching state j at maturity $t+1$ by adding up all the state price contributions of visiting state h one month earlier at maturity t $\left(\pi^t_{0,h}\right)$ times the transition state price from h to j $(\pi_{h,j})$.

Directly solving Eq. (6) is not advisable, as the problem is ill-conditioned and does not necessarily deliver positive transition state prices. Rather, we impose an additional non-negativity constraint on the transition state prices $\pi_{i,j}$ and back them out from the least squares problem below, which penalizes violations of Eq. (6):

$$\min_{\pi_{i,j}} \quad \sum_{j\in I} \sum_{t=0}^{11} \left(\pi^{t+1}_{0,j} - \sum_{h\in I} \pi^t_{0,h} \cdot \pi_{h,j} \right)^2 \quad s.t. \quad \pi_{i,j} > 0. \tag{7}$$

We collect the transition state prices $\pi_{i,j}$ in the transition state price matrix Π and recover physical transition probabilities by applying the recovery theorem. We label this recovery method Ross Original.

Ross (2015) observes that such a coarse nonoverlapping grid of dimension 12-by-12 leads to poorly discretized transition state prices and to coarse discrete physical probabilities. We confirm that the method does not work and the future returns do not seem to be drawn from the recovered physical probability distribution (see Section 4.2).

3.3. Finding transition state prices from spot state prices: Ross Basic

Instead of using the coarse nonoverlapping 12-by-12 grid in Ross Original, we follow Audrino et al. (2020) and apply an overlapping approach to determining the transition state prices. Using steps of one-tenth of a month (and a state price transition lasting one month, i.e., ten steps), our new relation is

$$\pi^{t+10}_{0,j} = \sum_{h\in I} \pi^t_{0,h} \cdot \pi_{h,j} \quad \forall j \in I,\ t = 0, \ldots, 110. \tag{8}$$

This results in a total number of 111 overlapping transitions and thus allows $N = 111$ states, which we choose to be equidistant and where we include the current state $i = 0$. Again, we impose an additional nonnegativity constraint on the transition state prices $\pi_{i,j}$, and we recover them by solving a least squares problem that penalizes violations of Eq. (8):

$$\min_{\pi_{i,j}} \quad \sum_{j\in I} \sum_{t=0}^{110} \left(\pi^{t+10}_{0,j} - \sum_{h\in I} \pi^t_{0,h} \cdot \pi_{h,j} \right)^2 \quad s.t. \quad \pi_{i,j} > 0. \tag{9}$$

We label this method Ross Basic and depict the results of our implementation for a typical day in our sample, February 17, 2010. Fig. 1 illustrates the transition state prices, which best relate spot state prices at one maturity to those at a maturity one month later.

We expect high transition state prices on the main diagonal, as it is more likely to end up in states j that are close to the initial state i. In fact, the optimization quite often generates high state prices for states j that are far away from the initial state i.

The high transition state prices away from the main diagonal can occur because short-maturity option prices are little 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 allocates prices to these

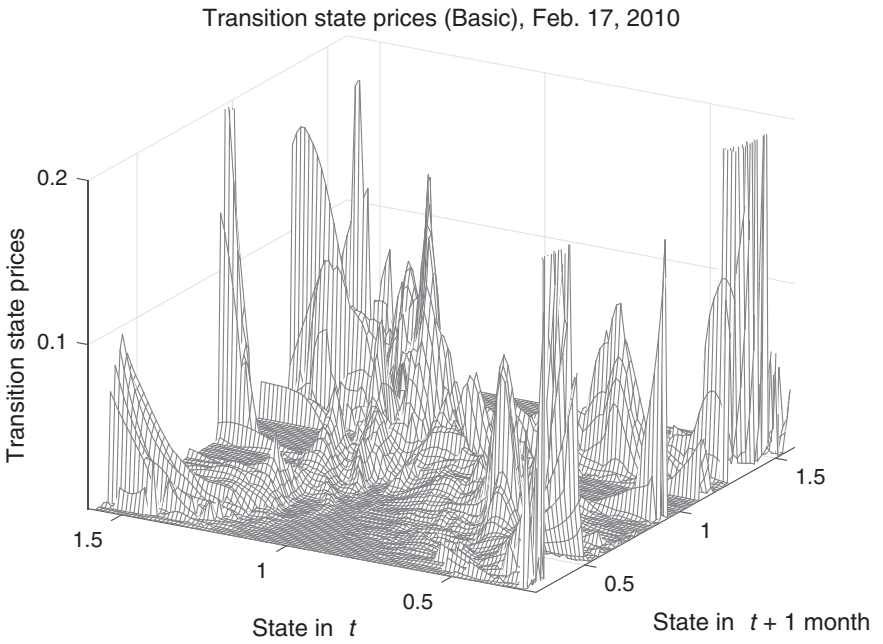

Fig. 1. Transition state prices and recovered physical transition probabilities, Ross Basic. We show the transition state prices as identified by the Ross Basic method. All data are from February 17, 2010. States are expressed in terms of moneyness.

irrelevant states to minimize the objective function while not much changing the short-maturity spot state prices much in the process.

As a result, some rowsums in Π have values much higher than one, which would imply high negative risk-free rates for some initial states. On our sample day, February 17, 2010, we observe that one-third of all one-month state-dependent risk-free rates are lower than -20% (the lowest value being -98%) and one-third of the rates are higher than 80% (the highest value being 576%). We thus impose reasonable economic constraints on the problem in Eq. (9).

3.4. Adding economic constraints: Ross Bounded

We first constrain all rowsums of Π to 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 be between 0% and 11.11% (0% and 254.07% annualized). We again solve Eq. (9) but additionally restrict the rowsums:

$$\min_{\pi_{i,j}} \quad \sum_{j\in I}\sum_{t=0}^{110}\left(\pi_{0,j}^{t+10}-\sum_{h\in I}\pi_{0,h}^{t}\cdot\pi_{h,j}\right)^2 \quad s.t. \quad \pi_{i,j}>0$$
$$\text{and} \quad 0.9\le\sum_{j\in I}\pi_{i,j}\le 1.0 \quad \forall i\in I. \tag{10}$$

We label this method Ross Bounded. Fig. 2 illustrates the transition state prices for Ross Bounded. They are now highly concentrated around the current state (labeled by a moneyness level of one). 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 constrained optimization reduces state prices almost uniformly the farther they are away from the current state. As a result, we do not obtain the economically reasonable diagonal structure for the transition state prices with Ross Bounded.

3.5. Adding economic constraints: Ross Unimodal

Next, in addition to bounding the rowsums, we also constrain the rows in our Π matrix to be unimodal with maximal values on the main diagonal. We solve Eq. (10) but add the requirement of unimodality:

$$\min_{\pi_{i,j}} \quad \sum_{j\in I}\sum_{t=0}^{110}\left(\pi_{0,j}^{t+10}-\sum_{h\in I}\pi_{0,h}^{t}\cdot\pi_{h,j}\right)^2 \quad s.t. \quad \pi_{i,j}>0$$
$$\text{and} \quad 0.9\le\sum_{j\in I}\pi_{i,j}\le 1.0 \quad \forall i\in I$$
$$\text{and} \quad \pi_{i,j}\le\pi_{i,l} \quad \forall j,l,i\in I \quad with \quad j<l\le i$$
$$\text{and} \quad \pi_{i,j}\ge\pi_{i,l} \quad \forall j,l,i\in I \quad with \quad j>l\ge i. \tag{11}$$

We label this method *Ross Unimodal*. Fig. 3 shows the transition state prices for Ross Unimodal. By construction, the modes line the main diagonal, steeply falling off farther away from the main diagonal. As intended, away-from-current states on the main diagonal have higher values than in Ross Bounded. Results are robust to shifting the modes five index levels up or one index level down in parallel to the main diagonal. Here we are guided by the fact that across our sample, the mode of the spot state prices is, at most, five index levels above and one index level below the current state.

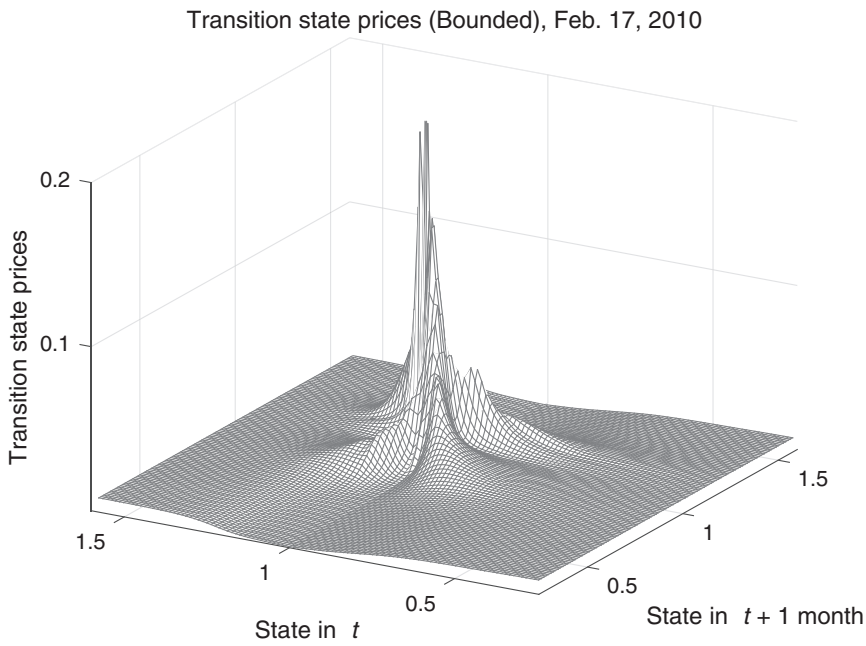

Fig. 2. Transition state prices, Ross Bounded. We show the transition state prices as identified by the Ross Bounded method. All data are from February 17, 2010. States are expressed in terms of moneyness..

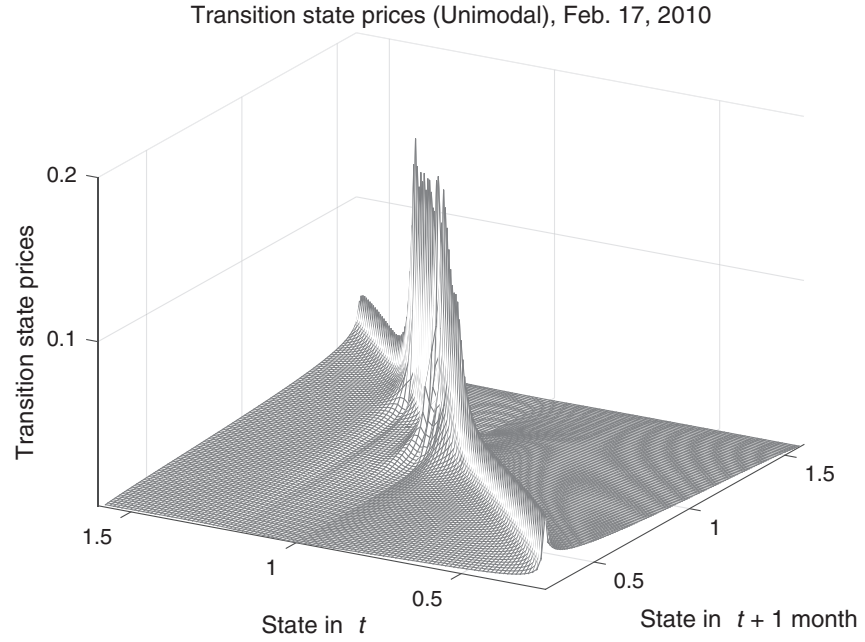

Fig. 3. Transition state prices, Ross Unimodal. We show the transition state prices as identified by the Ross Unimodal method. All data are from February 17, 2010. States are expressed in terms of moneyness..

3.6. Recovery without using transition state prices: Ross Stable

The computation of transition state prices is a key challenge in applying the recovery theorem. Yet the one row in the transition state price matrix Π associated with the current state $i = 0$ offers a novel way out. For the current state, the transition state prices ought to coincide with the spot state prices, which we can readily obtain from option prices. We use this insight to suggest an alternative recovery method that does not require explicitly solving for the transition state prices.[4] The trick is to use the eigenvalue

[4] See the independent derivation in Jensen et al. (2019).

problem in Eq. (5) and multiply both sides from the left with the transition state price matrix Π:

$$\Pi \cdot \Pi z = \Pi \cdot \delta z = \delta(\Pi z) = \delta^2 z. \tag{12}$$

Again, in the row of Π^2 associated with the current state $i = 0$ are the spot state prices but now with a maturity of two transition periods. In this case, the discount factor δ appears in the second power to account for the two periods. Iterating, we obtain the relation

$$\Pi^t z = \delta^t z \text{ with } t = 1, \ldots, T, \tag{13}$$

where t determines how often we apply the transitions. For each t, we focus on the row in Π^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, which results in a system of equations (see Online Appendix D for details):

$$\sum_{j \in I} \pi^t_{0,j} \cdot \left[\frac{z_j}{z_0}\right] = \delta^t, \quad t = 1, \ldots, T. \tag{14}$$

One might worry that the system of equations is ill-conditioned and thus might violate reasonable economic constraints. Namely, we want to insure that the utility discount factor δ and the resulting SDF are nonnegative. So, we penalize deviations from Eq. (14) and include the two new constraints:

$$\min_{\left[\frac{z_j}{z_0}\right], \delta} \sum_{t=1}^{T} \left(\sum_{j \in I} \pi^t_{0,j} \cdot \left[\frac{z_j}{z_0}\right] - \delta^t \right)^2$$
$$s.t. \quad \left[\frac{z_j}{z_0}\right] > 0, \; 1 > \delta > 0. \tag{15}$$

As we solve for the SDF with the least squares approach of Eq. (15), the system of equations in Eq. (14) does not hold exactly. As a result, the recovered physical spot probabilities do not necessarily sum to one, so we normalize them.

Full identification requires at least as many equations with different maturities t as there are number of states N. Using only $N = 12$ as in Ross Original results in a very coarse grid and recovered physical probabilities that are incompatible with future returns. Thus, we use $N = 120$ and as many equations with 1 to 120 periods of one-tenth of a month each and recover the SDF spanning one-tenth of a month. We make use of the property that the structure of the SDF in Ross recovery for different maturities varies by only a factor (see Online Appendix D for details):

$$m^t_{0,j} = \delta^{t-1} \cdot m_{0,j}, \tag{16}$$

where $m^t_{0,j}$ is the spot SDF with a maturity of t transition periods and $m_{0,j}$ is the spot SDF with a maturity of one transition period. We use Eq. (16) to find the ten-period (i.e., one-month) SDF and use it to transform one-month spot state prices into one-month physical probabilities. We label this method Ross Stable. Note that we cannot provide the corresponding figures for the transition state prices and the transition physical probabilities, as we no longer compute them explicitly.

3.7. Competing benchmark methods: Power Utility and Historical Return Distribution

In addition to our methods based on Ross recovery (Ross Basic, Ross Bounded, Ross Unimodal, and Ross Stable), we introduce two competing benchmark methods that are not related to Ross recovery to obtain physical distributions. For the first, we assume a representative investor with a power utility and a risk aversion coefficient of three.[5] Given the power utility SDF, we transform the spot state prices into physical probabilities. 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 method Power Utility with $\gamma = 3$.

In addition, we use the empirical distribution of the past five years of monthly S&P 500 returns. It is irregularly spaced at the historical returns with probability 1/60 at each return. We label this method Historical Return Distribution.

4. Testing the recovered physical probabilities

Each month, we find the date τ that is 30 calendar days before the option expiration date. We obtain the future return r_τ (i.e., the realized return from date τ to the expiration date) on the S&P 500 from Thomson Reuters Datastream. That return is one realization drawn from the true physical distribution p_τ. Our hypothesis is then, for each method in turn:

H0: Future one-month S&P 500 returns (with distribution p_τ) are drawn from the one-month physical distribution $\hat{p}_\tau$ (i.e., $p_\tau = \hat{p}_\tau$).

We test our hypothesis first with density tests and then with mean and variance predictions (see Online Appendix E for details).

4.1. Density tests: methodology

Our six methods give us the physical spot distribution $\hat{p}_\tau$ and the corresponding cumulative distribution $\hat{P}_\tau$ for date τ.[6] We then find the percentile x_τ of the cumulative distribution $\hat{P}_\tau$ that corresponds to the future return r_τ using

$$x_\tau = \hat{P}_\tau(1 + r_\tau) = \int_{-\infty}^{1+r_\tau} \hat{p}_\tau(v) dv, \tag{17}$$

and we collect those percentiles x_τ for all dates τ across moneyness levels $1 + r$. Under our null that future returns are indeed drawn from the physical distribution according to a particular method (i.e., $\hat{p}_\tau = p_\tau$), the percentiles should be independently and identically (i.i.d.) uniformly

[5] Bliss and Panigirtzoglou (2004) find an optimal risk aversion factor of four in a power utility framework for their sample (1983–2001). We repeat their analysis for our much longer sample (1986–2017) and find an optimal risk aversion factor of three.

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

Table 1
Density tests of the recovered physical probabilities.
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 three and four moments) tests for uniformity of the percentiles of future returns under the method's physical cumulative distribution. Our sample runs from April 1986 through December 2017.

H0: $p_\tau = \hat{p}_\tau$ Recovery Method	Berkowitz *p*-value	Kolmogorov-Smirnov *p*-value	Knüppel 3 moments *p*-value	Knüppel 4 moments *p*-value
Ross Basic $\pi_{i,j} > 0$	0.001	0.020	0.000	0.000
Ross Bounded $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	0.000	0.049	0.012	0.000
Ross Unimodal $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.000	0.044	0.000	0.000
Ross Stable No transition state prices	0.002	0.045	0.028	0.002
Power Utility with $\gamma = 3$	0.155	0.530	0.594	0.088
Historical Return Distribution	0.763	0.663	0.939	0.832

distributed. We use four different tests to establish the uniformity of the percentiles x_τ: the Berkowitz test, two versions of the Knüppel test, and the Kolmogorov-Smirnov test.

4.1.1. Berkowitz test

The Berkowitz (2001) test jointly tests uniformity and the i.i.d. property of x_τ. For this test, the series x_τ is transformed by applying the inverse standard normal cumulative density function Φ to x_τ:

$$z_\tau = \Phi^{-1}(x_\tau). \tag{18}$$

As z_τ is distributed standard normally, we test the AR(1) model:

$$z_\tau - \mu = \rho(z_{\tau-1} - \mu) + \epsilon_\tau, \tag{19}$$

where the null requires $\mu = 0$, $\mathrm{Var}(\epsilon_\tau) = 1$, and $\rho = 0$. Berkowitz then applies a likelihood ratio test as follows:

$$LR_3 = -2(LL(0, 1, 0) - LL(\hat{\mu}, \hat{\sigma}, \hat{\rho})), \tag{20}$$

where *LL* characterizes the log likelihood of Eq. (19).

4.1.2. Knüppel test

The Knüppel (2015) test first scales the series x_τ to $y_\tau = \sqrt{12}(x_\tau - 0.5)$. To test x_τ for standard uniformity, the series y_τ is tested for scaled uniformity with zero mean and unit variance. The test then compares the first S moments of the series y_τ to the respective theoretical moments in a generalized method of moments-type procedure with test statistic α_S:

$$\alpha_S = \mathcal{T} \cdot D_S^\intercal \Omega_S^{-1} D_S, \tag{21}$$

where $\mathcal{T}$ is the number of dates in our sample. D_S is a vector that consists of the differences between the sample moments $\frac{1}{\mathcal{T}} \sum_{\tau=1}^{\mathcal{T}} y_\tau^s$ and the theoretical moments μ_s for $s = 1, \ldots, S$. Ω_S is a consistent covariance matrix estimator of all S moment differences. We follow Knüppel (2015) and set all elements of Ω_S that represent covariances between odd and even moment differences to zero and apply the test by considering the first three moments ($S = 3$) and the first four moments ($S = 4$), respectively. We account for serial correlation of x_τ by estimating a Newey-West covariance matrix with automated lag length, as proposed by Andrews (1991).

4.1.3. Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test looks at the maximum distance between the empirical cumulative distribution function $\hat{U}$, which is based on the percentiles x_τ and computed, as in Kaplan and Meier (1958), and the uniform cumulative density function U. It uses the following test statistic:

$$KS = \sup_v |U(v) - \hat{U}(v)|. \tag{22}$$

In comparing the four tests, Bliss and Panigirtzoglou (2004) argue that the Berkowitz test is superior to the Kolmogorov-Smirnov test in small samples with autocorrelated data. While more powerful than the Kolmogorov-Smirnov test, the Berkowitz test uses only the first two moments and ignores higher moments. The Knüppel (2015) test has the advantage of testing for higher moments, can deal with autocorrelated data, and still has power even in small samples.

In Online Appendix F, we provide a power analysis of our four density tests. The tests do not summarily reject too often, and tests differ in their statistical power. Rather than picking among tests, we simply report all results below, as they are similar.

4.2. Density tests: results

Table 1 presents the density test results. Our four original versions of the recovery theorem (Ross Basic, Ross Bounded, Ross Unimodal, and Ross Stable) reject the null hypothesis (*p*-values less than 5%) in all four tests. Future returns are not drawn from physical distributions recovered as per Ross recovery. To the contrary, our simple benchmark methods (Power Utility and Historical Return

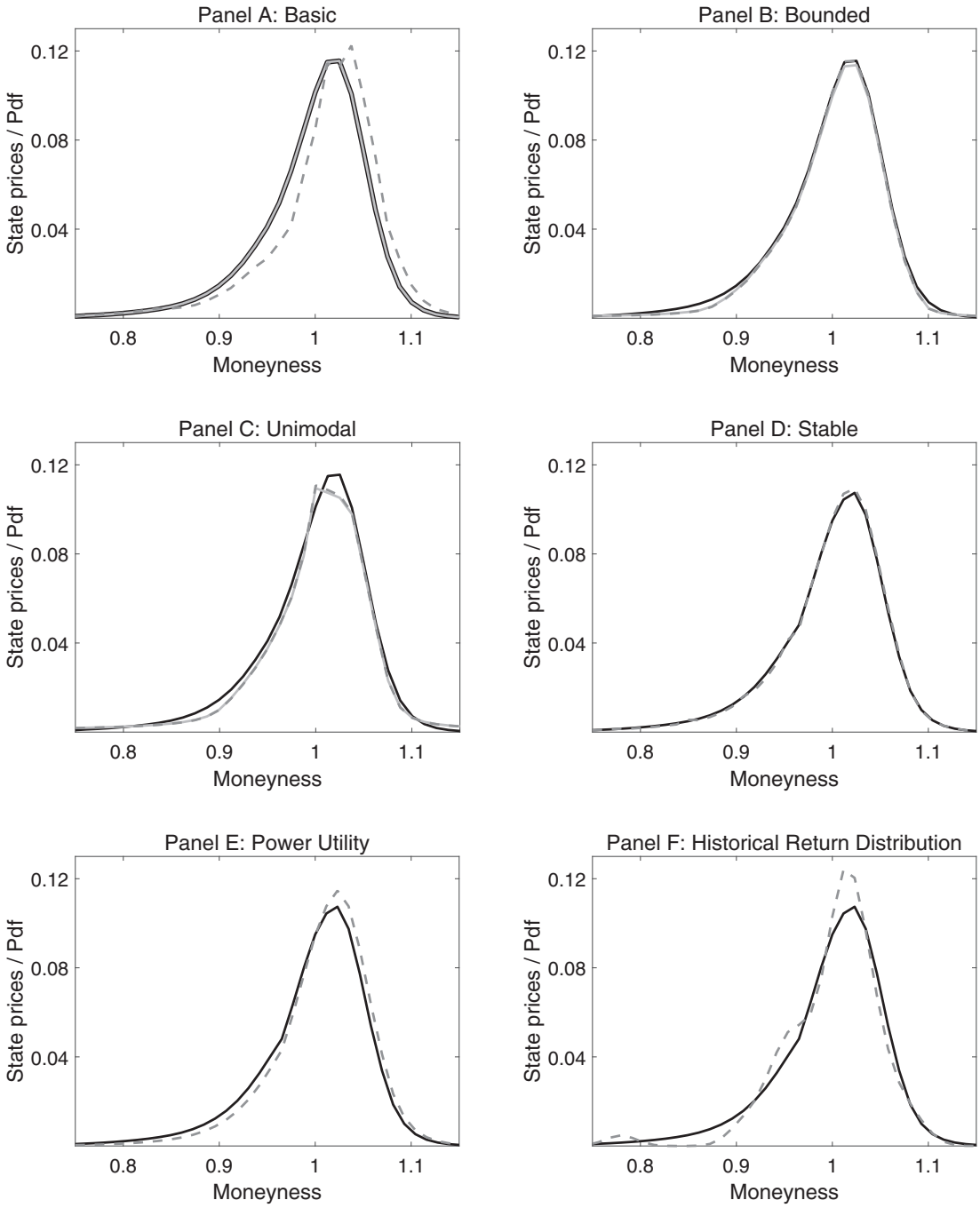

Fig. 4. State prices and recovered physical probabilities. We depict spot state prices (black lines), transition state prices (light gray lines), and physical probabilities (gray dashed lines) on February 17, 2010. Our methods are Ross Basic in Panel A, Ross Bounded in Panel B, Ross Unimodal in Panel C, Ross Stable in Panel D, Power Utility with $\gamma = 3$ in Panel E, and the kernel-smoothed Historical Return Distribution in Panel F.

Distribution) are not rejected by any of our four tests (p-values higher than 8%). We thus see a complete empirical failure of the recovery theorem, while our benchmark methods cannot be rejected.

To understand our results, we take a closer look at the probabilities in Fig. 4. The solid black lines show spot state prices, which are derived from option prices and are virtually the same in all six panels except for small differences in the number of states (Panels A–C use 111 states, Panels D–F 120). The risk-neutral distributions, differing by only a small interest rate adjustment, would also look just the same. The dashed lines show the physical probabilities.

In the Historical Returns Distribution in Panel F, we cannot reject our null that future returns are drawn from the physical distribution of the past 60 monthly returns. The physical distribution has been kernel-smoothed (MAT-

LAB ksdensity with default bandwidth) onto the same state space as Power Utility and Ross Stable ($N = 120$) and has been normalized. The physical distribution is more peaked around at-the-money than the spot state prices and shows some waviness because of the original data.

The Power Utility method in Panel E changes the spot state prices into physical probabilities using a power SDF with $\gamma = 3$, and we cannot reject our hypothesis that future returns are drawn from it. Again, the physical distribution is more peaked around at-the-money than the spot state prices and is somewhat lower than the spot state prices for moneyness levels below one. Such a shift implies a positive market risk premium and thus makes much economic sense. The realized market risk premium during our sample over April 1986–December 2017 is 6.65% per year.

The recovery methods in Panels A–C have an additional light gray line for the transition state prices. This is because the methods allow for transition state prices to differ from the spot state prices, thus incorporating a pricing error for the option prices. Yet, Ross Basic in Panel A does not use this possibility, which is why 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 allows the optimization to closely follow the spot state prices. As a result, Ross Basic produces implausible fluctuations for the transition state prices and rowsums, which imply extreme monthly risk-free rates, ranging from -98% to 576% on a typical sample day (February 17, 2010). The physical distribution is somewhat right-shifted but insufficiently so, as we reject our hypothesis for Ross Basic as well as for all other recovery methods.

Adding economic constraints on the rowsums in Ross Bounded (Panel B) leads to a very slight separation of transition state prices from spot state prices. Further adding the economic constraint of unimodality in Ross Unimodal (Panel C), the separation becomes somewhat stronger (we quantify this mispricing later). Also, the physical probabilities are almost identical to the transition state prices in both cases. As a result, the physical distributions end up being too close to the spot state prices, and we reject our hypothesis that the recovered distributions are compatible with future returns. Finally, in Ross Stable (Panel D), we do not explicitly compute transition state prices. The physical probabilities are again very close to the spot state prices, and we also reject our hypothesis for Ross Stable.

Summing up, all recovery methods, as opposed to our simple benchmark methods, are incompatible with future S&P 500 returns. Ross Basic suggests extreme fluctuations in the transition state prices and risk-free rates in different states. The other recovery methods cannot generate a sufficiently high risk premium, as the recovered physical distributions stay too close to the spot state prices. This closeness implies SDFs that are too flat and are almost risk-neutral.

4.3. Mean predictions: methodology

Our null that the distribution of future returns is the same as the physical distribution based on each of our six methods implies that the one-month mean μ_τ of the physical distribution at date τ should predict the one-month future return r_τ.[7] We collect both time series and run a regression as follows:

$$r_\tau = a + b\mu_\tau + \epsilon_\tau. \tag{23}$$

Our null implies that the intercept $a = 0$ and the slope $b = 1$. So we test three different models. In the Intercept model, we fix the slope at $b = 1$ and test for the intercept at $a = 0$. In the Slope model, we fix the intercept at $a = 0$ and test for the slope at $b = 1$. In the Joint model, we test for the intercept being $a = 0$ and, at the same time, the slope at $b = 1$.

4.4. Mean predictions: results

Table 2 presents the mean prediction results for all recovery methods. For the Intercept model (fixing the slope at $b = 1$), we cannot reject a zero intercept for Ross Basic, Power Utility, and the Historical Return Distribution (p-values above 0.43).

For Ross Bounded, Ross Unimodal, and Ross Stable, intercepts are all positive and are significantly different from zero (p-values below 0.002), which indicates that those means are significantly lower than average returns. This finding supports our insight above that those three methods imply an almost flat SDF and thus do not generate much of a risk premium.

For the Slope model (fixing the intercept at $a = 0$), we cannot reject a unit slope for Ross Bounded, Ross Stable, and the Historical Return Distribution. Yet we note that the point estimates are far away from one for Ross Bounded ($b = 0.069$) and Ross Stable ($b = 0.198$). The slope for the Historical Return Distribution ($b = 0.668$) is much more reasonable in comparison. For Ross Basic, Ross Unimodal, and Power Utility, the slopes are significantly different from one (p-values below 0.022). We reject the Joint model for all methods.

Jensen et al. (2019) also predict means for their version of Ross Stable. They focus on predictive power in a setting with look-ahead bias. Eliminating the bias and using our much longer sample,[8] the adjusted R-squared values for Table 2 are all around 0%.

We conclude that all methods predict mean returns poorly. The Historical Return Distribution performs best, as we cannot reject the Intercept and the Slope models separately, even though we have to reject the Joint model.

4.5. Variance predictions: methodology

Our null further implies that the one-month variance σ_τ^2 of the physical distribution at date τ should predict the one-month future realized variance RV_τ, which we compute as the sample variance of all future daily returns in the month following date τ multiplied by the number of

[7] Bakshi et al. (2018) conduct related tests of predicting the first and second moments of bond returns instead of index returns. Their results are also, in line with ours, negative.

[8] Table 2 does not use excess returns as in Jensen et al. (2019), yet results do not change when doing so.

Table 2
Mean predictions.
We test whether 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 regression models. Our sample runs from April 1986 through December 2017.

$r_\tau = a + b\mu_\tau + \epsilon_\tau$	Intercept model Set $b = 1$ Test $a = 0$	Slope model Set $a = 0$ Test $b = 1$	Joint model Test $a = 0$ and $b = 1$
Recovery method	Intercept [95% CI] p-value	Slope [95% CI] p-value	Intercept [95% CI] Slope [95% CI] p-value
Ross Basic $\pi_{i,j} > 0$	–0.001 [± 0.006] 0.652	0.075 [± 0.101] 0.000	0.008 [± 0.005] 0.035 [± 0.102] 0.000
Ross Bounded $\pi_{i,j} > 0$, rowsums $\in$ [0.9, 1]	0.008 [± 0.004] 0.001	0.069 [± 0.949] 0.055	0.009 [± 0.004] –0.185 [± 0.943] 0.000
Ross Unimodal $\pi_{i,j} > 0$ and unimodal, rowsums $\in$ [0.9, 1]	0.008 [± 0.004] 0.001	0.106 [± 0.764] 0.022	0.009 [± 0.004] –0.080 [± 0.758] 0.000
Ross Stable No transition state prices	0.007 [± 0.004] 0.002	0.198 [± 1.281] 0.219	0.010 [± 0.005] –0.940 [± 1.378] 0.000
Power Utility with $\gamma = 3$	–0.002 [± 0.005] 0.432	0.393 [± 0.314] 0.000	0.009 [± 0.006] –0.067 [± 0.449] 0.000
Historical Return Distribution	0.001 [± 0.004] 0.788	0.668 [± 0.459] 0.156	0.009 [± 0.007] –0.077 [± 0.765] 0.022

days in that month. We collect both time series and run the regression:

$$RV_\tau = a + b\sigma_\tau^2 + \epsilon_\tau . \tag{24}$$

As with the mean predictions, we test the Intercept, Slope, and Joint models.

4.6. Variance predictions: results

Table 3 presents the variance prediction results for all recovery methods. For the Intercept model (fixing the slope at $b = 1$), we cannot reject a zero intercept for Power Utility (p-value 0.524). For all other methods, intercepts are significantly different from zero (p-values below 0.010) because most methods' variances are significantly higher than realized variances (except for the Historical Return Distribution, where they are significantly lower).

For the Slope model (fixing the intercept at $a = 0$), we cannot reject a unit slope for Power Utility and the Historical Return Distribution. For all Ross recovery methods, the slopes are significantly different from one (p-values of 0.000). We again reject the Joint model for all methods.

The adjusted R-squared values for Table 3 range from 0% (Historical Return Distribution) and 5% (Ross Basic) to 40–52%. In line with the high R-squared values for volatility in Jensen et al. (2019), the R-squared values for variance are much higher than for the mean predictions. Exceptions are the Historical Return Distribution, which uses too many stale returns (60) in its realized variance prediction, and the noisy Ross Basic method. The high R-squared values are largely due to the persistence of variance. Simply using the CBOE Volatility Index (VIX) would give similar explanatory power for future realized variance.

We conclude that all methods have a hard time predicting variances. Power Utility performs best, as we cannot reject the Intercept and the Slope models separately, even though we have to reject the Joint model. The Historical Return Distribution cannot be rejected for the Slope model.

4.7. Robustness

Looking at our density tests, as well as our mean and variance predictions, we find that Ross recovery does not work, not in its basic form (Ross Basic), nor in the economically constrained forms (Ross Bounded, Ross Unimodal, or Ross Stable). Our benchmark methods (Power Utility and Historical Return Distribution) perform much better, as they tend not to reject our null in our various tests.

Several robustness checks confirm the stability of our results. We investigate whether variations of the state space influence our empirical findings by (i) using log returns instead of straight returns, by (ii) reducing the dimension of our state space by 20%, and by (iii) using Ross Original, which is based on 12 nonoverlapping monthly periods and 12 states. We also repeat our study but now exclude the early 1990s recession, the early 2000s recession, and the Great Recession from our original sample. Last, we use weekly options instead of monthly options for the sample for which the weeklies are available from 2005–2017. For details on our robustness checks, see Online Appendix G.

Table 3
Variance predictions.
We test whether 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 regression models. Our sample runs from April 1986 through December 2017.

$RV_\tau = a + b\sigma_\tau^2 + \epsilon_\tau$	Intercept model Set $b = 1$ Test $a = 0$	Slope model Set $a = 0$ Test $b = 1$	Joint model Test $a = 0$ and $b = 1$
Recovery method	Intercept [95% CI] p-value	Slope [95% CI] p-value	Intercept [95% CI] Slope [95% CI] p-value
Ross Basic $\pi_{i,j} > 0$	−0.005 [± 0.002] 0.000	0.118 [± 0.027] 0.000	0.002 [± 0.001] 0.067 [± 0.028] 0.000
Ross Bounded $\pi_{i,j} > 0$, rowsums $\in$ [0.9, 1]	−0.007 [± 0.001] 0.000	0.326 [± 0.032] 0.000	−0.002 [± 0.001] 0.436 [± 0.055] 0.000
Ross Unimodal $\pi_{i,j} > 0$ and unimodal, rowsums $\in$ [0.9, 1]	−0.008 [± 0.001] 0.000	0.304 [± 0.030] 0.000	−0.002 [± 0.001] 0.401 [± 0.049] 0.000
Ross Stable No transition state prices	−0.001 [± 0.000] 0.000	0.871 [± 0.068] 0.000	0.000 [± 0.000] 0.915 [± 0.091] 0.000
Power Utility with $\gamma = 3$	0.000 [± 0.000] 0.524	1.058 [± 0.085] 0.180	−0.001 [± 0.000] 1.150 [± 0.118] 0.037
Historical Return Distribution	0.001 [± 0.000] 0.010	1.124 [± 0.217] 0.261	0.002 [± 0.001] 0.252 [± 0.507] 0.001

5. Applying machine learning to Ross recovery

Machine learning takes advantage of large amounts of data and tries to detect patterns that might evade human investigators. We apply machine learning to Ross recovery to optimally regularize the transition state prices and to tease out information in the options data.[9] As Ross Stable fits the SDF and the discount factor (and not transition state prices) to the data, we focus on Ross Basic, Ross Bounded, and Ross Uniform. We address potential overfitting in the process.

In our first approach, we use absolute instead of squared deviations in Eqs. (9)–(11). Unreported results show that our empirical findings barely change for Ross Basic, Ross Bounded, and Ross Unimodal.

In our second approach, we borrow from elastic net regression and augment our equations not only with a term forcing the sum of absolute transition state prices to zero but also with an additional term forcing the sum of squared transition state prices to zero.

In the case of Ross Basic, this changes the optimization procedure for finding transition state prices to

$$\min_{\pi_{i,j}} \quad \sum_{j\in I}\sum_{t=0}^{110}\left(\pi_{0,j}^{t+10} - \sum_{h\in I}\pi_{0,h}^{t}\cdot\pi_{h,j}\right)^2 + \lambda_1\sum_{j,i\in I}\pi_{i,j} + \lambda_2\sum_{j,i\in I}\left(\pi_{i,j}\right)^2 \quad s.t. \quad \pi_{i,j} > 0, \tag{25}$$

where the positive regularization parameters λ_1 and λ_2 are determined by cross-validation on a ten-day training set, which mitigates overfitting. We compute a reduced spot state price surface that incorporates only 50% of the maturities of the training set to determine transition state prices $\pi_{i,j}$. For each combination of λ_1 and λ_2, we then evaluate the full set of available maturities during the training set using the associated transition state prices and record the resulting sum of squared errors between model and observed spot state prices. We search for the optimal combination of λ_1 and λ_2 on a fine and large grid so that we minimize the sum of squared errors in the training set. Finally, we use the optimal parameters on our remaining data set.

For Ross Bounded and Ross Unimodal, the elastic net regularization is implemented accordingly. For each of these new methods (labeled Ross Basic ML, Ross Bounded ML, and Ross Unimodal ML) we recover physical probabilities and test whether future returns are compatible with those probabilities. Table 4 presents the density test results.

For Ross Basic ML, none of the four tests can reject our hypothesis that future returns are drawn from the recovered distribution. To understand this result, we take a closer look at the effect of including an elastic net penalty. The penalty pushes transition state prices toward zero, especially away-from-the-current states that are not as relevant in explaining the spot state price surface. The resulting rowsums are thus lower in away-from-the-current states and are higher in near-the-current states. Empirically, we observe a strong inverse relation between the rowsums of the transition state price matrix and the Ross

[9] We thank an anonymous referee for pointing us toward machine learning and loss functions other than sum of squares. See, e.g., Gu et al. (2018) for details on applying machine learning to asset pricing.

Table 4
Density tests of the recovered physical probabilities (machine learning).
We present our results when future returns are drawn from physical probabilities generated by Ross Basic, Ross Bounded, and Ross Unimodal when an elastic net regularization is included. We show the p-values from the Berkowitz, Kolmogorov-Smirnov, and Knüppel (both three and four moments) tests for uniformity of the percentiles of future returns under the method physical cumulative distribution. Our sample runs from April 1986 through December 2017.

H0: $p_\tau = \hat{p}_\tau$ Recovery method	Berkowitz p-value	Kolmogorov-Smirnov p-value	Knüppel 3 moments p-value	Knüppel 4 moments p-value
Ross Basic ML $\pi_{i,j} > 0$	0.281	0.485	0.486	0.315
Ross Bounded ML $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	0.000	0.065	0.001	0.000
Ross Unimodal ML $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.000	0.056	0.000	0.000

recovery SDF, which have an average correlation of -0.49. Empirically, we observe only negative correlations between our 380 recovered SDFs and the rowsums of the corresponding transition state price matrices with correlations ranging from -0.09 to -0.82 and an average of -0.49. Statistically, these negative correlations are significantly lower than zero in 372 out of 380 cases. The inverse-U-shaped rowsums therefore mechanically translate into a U-shaped SDF.[10] We provide more details and a simulation study to illustrate that property in Online Appendix H.

The U-shaped SDF, even though determined largely by forcing down transition state prices in away-from-the-current states, turns out to be compatible with future returns; see, e.g., Bakshi et al. (2010) and Chaudhuri and Schroder (2015). Alas, Ross Basic ML delivers implausible transition state prices, which imply extreme state-dependent risk-free rates. For our sample day of February 17, 2010, the risk-free rates vary from -62% to 6,207% and are even more extreme on the upside than the values of -98% to 576% for Ross Basic. As a result, the recovered physical probabilities entail much noise, which leads to nonrejection in our statistical tests. Yet, as the nonrejection of Ross Basic ML is driven by the particular choice of penalty terms and by the noise in the physical probabilities, we do not trust the recovered physical probabilities.

After constraining the rowsums of the transition state price matrix (Ross Bounded ML) and additionally forcing unimodal rowsums (Ross Unimodal ML), transition state prices are calmed significantly, which reduces the variability of the SDF. Hence, recovered physical probabilities are again close to risk-neutral ones and are not compatible with future returns anymore (only the low-power Kolmogorov-Smirnov test cannot reject with p-values of 0.065 and 0.056).

We additionally perform our moment prediction study for the elastic net methods. In unreported results, we find that Ross Basic ML compares to the Power Utility benchmark method, while Ross Bounded ML and Ross Unimodal ML fail to predict means and variances. Untabulated results show that our empirical findings for the elastic net versions are robust to varying the state space and eliminating recession periods. For weekly options, we can additionally reject Ross Basic with all four density tests.

[10] Audrino et al. (2020) use a similar regularization technique for their implementation of Ross recovery and also find U-shaped SDFs.

6. Reasons for failure

We now investigate why the recovery theorem empirically fails by looking at SDFs, the time-homogeneity of transition state prices, simulated economies, and the particular issues in machine learning methods.

6.1. Recovered stochastic discount factors

Assuming a theoretical risk-averse representative investor, we expect SDFs to be positive and monotonically decreasing. However, Ait-Sahalia and Lo (2000), Jackwerth (2004), and Rosenberg and Engle (2002) find the empirical SDF to be locally increasing, the so-called pricing kernel puzzle. Yet how do the SDFs of our six methods in Fig. 5 line up with these findings? The solid black line shows the implied SDF for each method, measured as spot state prices divided by the physical probabilities.

For Power Utility in Panel E, the SDF is, by construction, monotonically decreasing and theoretically well-founded. From our main results, we already know that this SDF translates the spot state prices into physical probabilities that are compatible with future returns. The SDF for Ross Basic in Panel A looks somewhat similar. It is not very smooth yet is decreasing for moneyness levels higher than one. As a result, the shift from state prices to physical probabilities is insufficient in that we reject our hypothesis that future returns are drawn from the recovered physical distribution.

Here, we also depict as a gray line the method SDF, measured as transition state prices for the current state divided by the recovered physical probabilities. Any difference between method and implied SDFs is because the optimization fails to exactly match the spot state prices (and thus the observed option prices). For Ross Basic, this is not an issue, as the optimization is free to fit option prices as long as the transition state prices are positive.

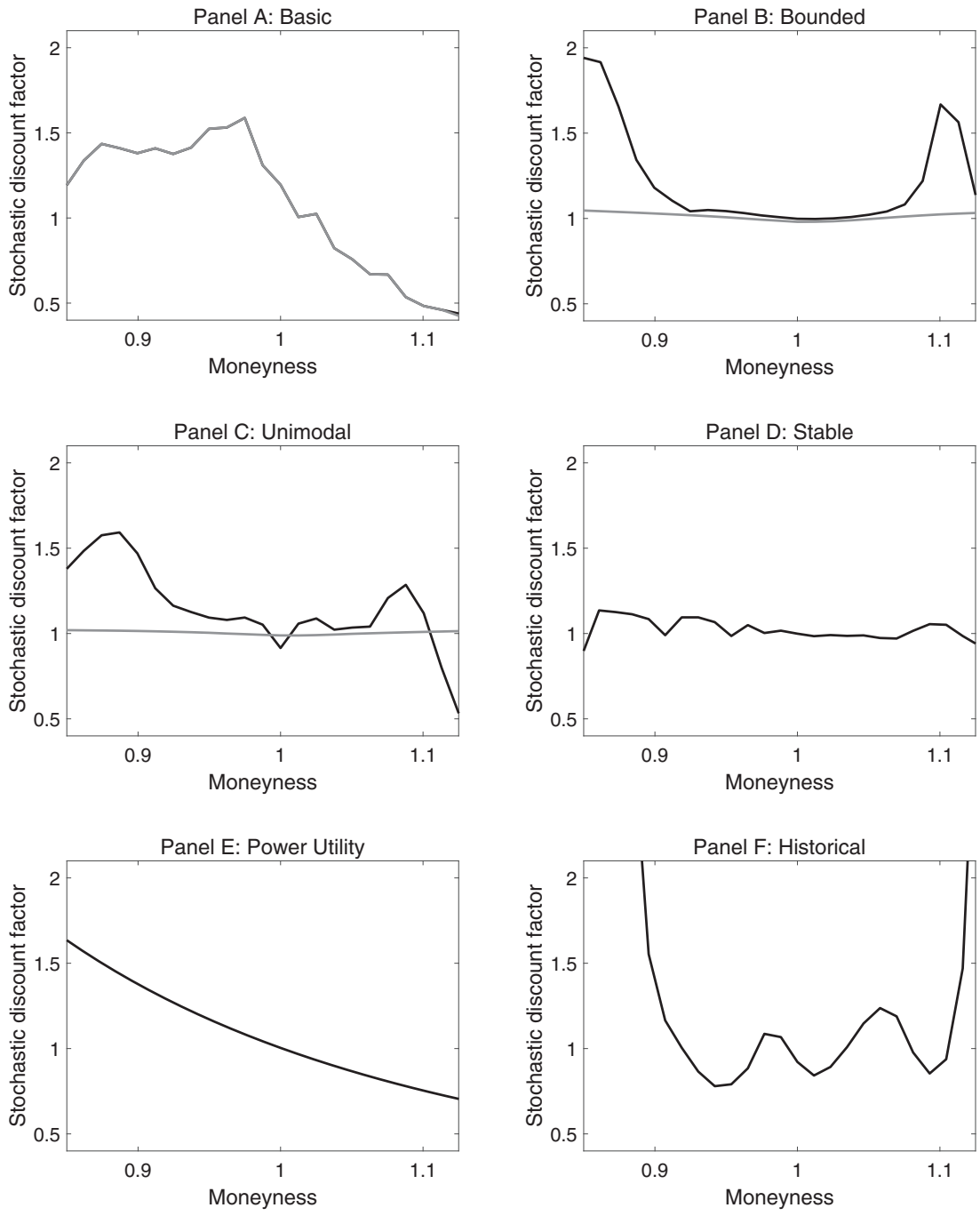

Fig. 5. Stochastic discount factors. We present SDFs for six different methods on February 17, 2010. Black lines depict implied SDFs, measured as spot state prices divided by physical probabilities, while gray lines depict method SDFs measured as transition state prices for the current state divided by physical probabilities. Panel A shows the SDFs for Ross Basic, Panel B for Ross Bounded, Panel C for Ross Unimodal, Panel D for Ross Stable, Panel E for Power Utility with $\gamma = 3$, and Panel F for Historical Return Distribution. Only the implied SDF exists for methods D–F.

We learned that this freedom comes at the cost of extreme rowsums, which in turn lead to extreme monthly risk-free rates, ranging from -98 to 576%.

Once we implement reasonable economic constraints in Ross Bounded and Ross Unimodal (Panels B and C), the implied SDFs become even more wavy overall and flatter for center moneyness levels of about 0.9 to 1.1. The physical probabilities remain closer to the spot state prices and are incompatible with future returns. Interestingly, now the implied and the method SDFs diverge, indicating that the optimization struggles to match the spot state prices, as the transition state prices now need to satisfy our eco-

Table 5
Accuracy of transition state prices.
We report how well our methods (Ross Basic, Ross Bounded, Ross Unimodal, Ross Stable, Power Utility with $\gamma = 3$, and the Historical Return Distribution not being applicable) fit observed option prices. Our measure is the time-series average of the RMSE of method versus observed implied volatilities, MRMSE. Columns 1 and 2 present the MRMSE for a one-month maturity, where method-implied volatilities are based on either transition state prices or on spot state prices. Columns 3 and 4 present the MRMSE for a 12-month maturity, where method-implied volatilities are based on either conflated transition state prices or on spot state prices. Our sample runs from April 1986 through December 2017.

Recovery method	1-month transition state prices MRMSE	1-month spot state prices MRMSE	12-month transition state prices MRMSE	12-month spot state prices MRMSE
Ross Basic $\pi_{i,j} > 0$	0.006		0.050	
Ross Bounded $\pi_{i,j} > 0$, rowsums $\in [0.9, 1]$	0.124		0.052	
Ross Unimodal $\pi_{i,j} > 0$ and unimodal, rowsums $\in [0.9, 1]$	0.134		0.065	
Ross Stable No transition state prices		0.005		0.003
Power Utility with $\gamma = 3$		0.005		0.003
Historical Return Distribution	n/a	n/a	n/a	n/a

nomic constraints. The method SDF (the part driven by the recovery theorem and not due to the fit of option prices) is now U-shaped but is virtually flat in the center. U-shapes could occur because of heterogeneity in beliefs about return outcomes and short-selling (Bakshi et al., 2010) or jumps in the stock return processes (Chaudhuri and Schroder, 2015). In the case of Ross recovery, the flat center part of the SDFs dominates the recovered physical probabilities, and they are incompatible with observed future returns.

We also find a flat implied SDF for Ross Stable in Panel D.[11] Yet that flatness stems partially from the numerical implementation. We first recover the SDF with maturity of 0.1 months to allow a larger number of 120 states (instead of only 12) and then convert the 0.1-month SDF into a 1-month SDF by the multiplicative adjustment of Eq. (16). The negligible curvature of the 0.1-month SDF then directly translates into an almost flat 1-month SDF, which again implies that we recover physical probabilities that are close to the spot state prices.

Last, the implied SDF for the Historical Return Distribution in Panel F is rather irregular on this particular day, even after we kernel-smooth the historical distribution. Yet in general, and across our whole sample, we cannot reject our main hypothesis, and the implied SDFs manage to translate spot state prices into generally right-shifted physical probabilities.

We conclude that the Ross recovery methods (except Ross Basic) cannot generate a sufficiently sloped SDF to make the physical distribution consistent with future returns. In the case of Ross Basic, the SDF is sloped but is too wavy, and we still reject our hypothesis.

[11] The method SDF does not exist, as we never explicitly compute transition state prices.

6.2. Time homogeneity of transition state prices and option pricing errors

How well do the Ross recovery methods fit option prices? Ross Stable (and also Power Utility) uses only spot state prices, and both methods fit the one-month option prices very well in terms of the mean root-mean-squared-errors (MRMSE). This error is the time-series average of the root-mean-squared-errors (RMSE) between one-month method implied and observed implied volatilities at each date τ.[12] Table 5, column 2 shows these one-month MRMSEs to be 0.005, which is very low compared to a typical implied volatility of about 0.170. Note that the Historical Return Distribution does not use state prices at all and cannot be used to price options.

Among the transition state price methods in Table 5, column 1, Ross Basic has only a slightly higher MRMSE of 0.006. As we require only positivity of the transition state prices, the optimization can freely choose transition state prices to match the spot state prices. This good fit comes at the price of implying extreme monthly state-dependent risk-free rates.

For the other two transition state price methods (Ross Bounded and Ross Unimodal), we find more than 20 times higher MRMSEs (0.124 and 0.134), which are large when compared to the typical implied volatility of 0.17. Adding even mild economic constraints removes many degrees of freedom in the allocation of transition state prices and increases the error between method and observed implied volatilities.

We further want to investigate how the assumption of time-homogeneous transition state prices influences our

[12] For the method implied volatilities, we first compute method call option prices C with the corresponding transition state prices π for the current state using numerical integration: $C(K) = \int_0^\infty \pi(S) \cdot \max(S - K, 0) dS$. We then transform method call option prices into method implied volatilities σ.

empirical results. Given time-homogeneity, we may multiply the one-month transition state price matrix Π t-times with itself to generate transition state prices (and thus also spot state prices) with a maturity of t months.

The third column of Table 5 reports the MRMSEs of the transition state price methods for the 12-month horizon, which we can compare to the MRMSEs of 0.003 in the last column for the spot state price methods.[13] Ross Bounded and Ross Unimodal again show errors of up to 20 times the value of 0.003. Interestingly, the error for Ross Basic is 17 times the value of 0.003 and is much worse at the 12-month than at the 1-month horizon.

We conclude that Ross Basic, Ross Bounded, and Ross Unimodal fail to produce transition spot state prices that approximate option prices well. Only Ross Basic prices the one-month options reasonably well but at the cost of implausible transition state prices and extreme risk-free interest rates.

6.3. Insights from simulated economies

It is possible that small data errors in the option prices 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 distribution. Then, at the 5% significance level for our statistical tests, we should have a 95% nonrejection rate (i.e., future returns are compatible with the simulated recovery method) and a 5% rejection rate.

Next, we perturb the option prices; for details, see Online Appendix I. Such perturbation will increase the rejection rates, as the perturbed recovery methods generate physical distributions, which will deviate from the true physical distribution, from which we drew the future returns.

Without any perturbation, the rejection rates are all very close to the theoretical value of 5%. Beyond that, we find that only Ross Basic turns out to be very sensitive to perturbations of up to two times the mean bid-ask spread. For Ross Bounded, Ross Unimodal, Ross Stable, and Power Utility, rejection rates stay below 7% for perturbations of up to the mean bid-ask spread and only increase for even greater perturbations. Thus, the failure of Ross Basic could also be driven by its sensitivity with respect to small data errors in the option-implied volatilities. With added economic constraints, methods are less sensitive to perturbations.

7. Conclusion

We implement and test several recovery methods to examine the Ross (2015) framework. We further present a variant of Ross recovery without explicitly estimating the transition state prices, which are numerically hard to determine. Using density tests, we find that future S&P 500 returns are incompatible with recovered physical probabilities. Two simple benchmark methods (a power utility with $\gamma = 3$ and the empirical distribution of non-overlapping monthly S&P 500 returns during the past five years) are not rejected by the data. We confirm our results when we use moment predictions instead of density tests. Our results are robust to variations in the state space and the sample.

Next, we apply machine learning to Ross recovery, which regularizes transition state prices and forces them to zero. In the unconstrained basic implementation, we cannot reject Ross recovery any longer, but this result is likely mechanically related to the regularization. Once we add reasonable economic constraints, we reject all Ross recovery methods again.

We further analyze why the recovery theorem fails. We find that the most basic method, requiring minimal assumptions, delivers unstable transition state prices. Backed by a simulation study, we argue that it is very sensitive to small variations in the option prices to which the method is fitted. Further, the extreme risk-free rates implied by this basic method are economically implausible.

Alternative implementations of Ross recovery with mild economic constraints are more stable than the basic method but cannot generate SDFs that are far enough away from risk-neutrality to be compatible with future index returns. Further, the assumption of time homogeneity of the transition state prices leads to poorly fitted option prices for all Ross recovery methods that explicitly estimate transition state prices.

[13] The reason for a lower 12-month horizon error of 0.003 than a 1-month error of 0.005 for spot state prices is that the longer dated distributions cover a greater range of states and thus approximate the observed spot prices better.

References

Ait-Sahalia, Y., Lo, A.W., 2000. Nonparametric risk management and implied risk aversion. J. Econometr. 94, 9–51.

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.

Audrino, F., Huitema, R., Ludwig, M., 2020. An empirical implementation of the Ross recovery theorem as a prediction device. J. Financ. Econometr. Forthcoming

Bakshi, G., Chabi-Yo, F., Gao, X., 2018. A recovery that we can trust? Deducing and testing the restrictions of the recovery theorem. Rev. Financ. Stud. 31, 532–555.

Bakshi, G., Madan, D., Panayotov, G., 2010. Returns of claims on the upside and the viability of u-shaped pricing kernels. J. Financ. Econ. 97, 130–154.

Berkowitz, J., 2001. Testing density forecasts, with applications to risk management. J. Bus. Econ. Stat. 19, 465–474.

Bliss, R., Panigirtzoglou, N., 2004. Option-implied risk aversion estimates. J. Finance 59, 407–446.

Borovicka, J., Hansen, L.P., Scheinkman, J.A., 2016. Misspecified recovery. J. Finance 71, 2493–2544.

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

Carr, P., Yu, J., 2012. Risk, return, and Ross recovery. J. Derivat. 20, 38–59.

Chaudhuri, R., Schroder, M., 2015. Monotonicity of the stochastic discount factor and expected option returns. Rev. Financ. Stud. 28, 1462–1505.

Dillschneider, Y., Maurer, R., 2018. Functional Ross recovery: theoretical results and empirical tests.Unpublished working paper.Goethe University Frankfurt.

Dubynskiy, S., Goldstein, R. S., 2013. Recovering drifts and preference parameters from financial derivatives. Unpublished working paper.University of Minnesota.

Gu, S., Kelly, B. T., Xiu, D., 2018. Empirical asset pricing via machine learning. Unpublished working paper. Chicago Booth School of Business.

Hansen, L.P., Scheinkman, J.A., 2009. Long-term risk: an operator approach. Econometrica 77, 173–234.

Jackwerth, J.C., 2004. Option-Implied Risk-Neutral Distributions and Risk Aversion. CFA Institute, Charlottesville.

Jensen, C.S., Lando, D., Pedersen, L.H., 2019. Generalized recovery. J. Financ. Econ. 133, 154–174.
Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. 53, 457–481.
Knüppel, M., 2015. Evaluating the calibration of multi-step-ahead density forecasts using raw moments. J. Bus. Econ. Stat. 33, 270–281.
Massacci, F., Williams, J., Zhang, Y., 2016. Empirical recovery: Hansen-Scheinkman factorization and Ross recovery from high frequency option prices. University of Trento, Durham Business School. Unpublished working paper.
Qin, L., Linetsky, V., 2016. Positive eigenfunctions of Markovian pricing operators: Hansen-Scheinkman factorization, Ross recovery, and long-term pricing. Operat. Res. 64, 99–117.
Qin, L., Linetsky, V., Nie, Y., 2018. Long forward probabilities, recovery, and the term structure of bond risk premiums. Rev. Financ. Stud. 31, 4863–4883.
Rosenberg, J.V., Engle, R.F., 2002. Empirical pricing kernels. J. Financ. Econ. 64, 341–372.
Ross, S., 2015. The recovery theorem. J. Finance 70, 615–648.
Walden, J., 2017. Recovery with unbounded diffusion processes. Rev. Finance 21, 1403–1444.
Yao, X., 2018. What information can Ross' recovery theory recover? Evidence from S&P 500 index option prices. University of Connecticut. Unpublished working paper.