Generalized Recovery*

Christian Skov Jensen, David Lando, and Lasse Heje Pedersen†

First version March, 2015. This version December 21, 2016

Abstract

We characterize when physical probabilities, marginal utilities, and the discount rate can be recovered from observed state prices for several future time periods. We make no assumptions of the probability distribution, thus generalizing the time-homogeneous stationary model of Ross (2015). Recovery is feasible when the number of maturities with observable prices is higher than the number of states of the economy (or the number of parameters characterizing the pricing kernel). When recovery is feasible, our model is easy to implement, allowing a closed-form linearized solution. We implement our model empirically, testing the predictive power of the recovered expected return and other recovered statistics.

*We are grateful for helpful comments from Jaroslav Borovicka, Peter Christoffersen, Horatio Cuesdeanu, Darrell Duffie, Lars Peter Hansen, Jens Jackwerth, Pia Mølgaard, Stephen Ross, Jose Scheinkman, Paul Schneider, Paul Whelan as well as from seminar participants at University of Konstanz, University of Toronto, HEC-McGill winter finance workshop 2015, ATP, SED summer conference 2016, EFA 2016, and Copenhagen Business School, and especially to Dan Petersen for guiding us through Sard's theorem. All three authors gratefully acknowledge support from the FRIC Center for Financial Frictions (grant no. DNRF102) and Jensen and Pedersen from the European Research Council (ERC grant no. 312417).

†Jensen is at Copenhagen Business School; Lando is at Copenhagen Business School and CEPR; Pedersen is at Copenhagen Business School, NYU, AQR Capital Management, and CEPR; www.lhpedersen.com.

1 Introduction

The holy grail in financial economics is to decode probabilities and risk preferences from asset prices. This decoding has been viewed as impossible until Ross (2015) provided sufficient conditions for such a recovery in a time-homogeneous Markov economy (using the Perron-Frobenius Theorem). However, his recovery method has been criticized by Borovicka, Hansen, and Scheinkman (2015) (who also rely on Perron-Frobenius and results of Hansen and Scheinkman (2009)), arguing that Ross's assumptions rule out realistic models.

This paper sheds new light on this debate, both theoretically and empirically. Theoretically, we generalize the recovery theorem to handle a general probability distribution which makes no assumptions of time-homogeneity or Markovian behavior. We show when recovery is possible – and when it isn't – using a simple "counting" argument (based on Sard's Theorem). When recovery is possible, we show that our recovery inversion from prices to probabilities and preferences can be implemented in closed form, making our method simpler and more robust. We implement our method empirically using option data from 1996-2014 and study how the recovered expected returns predict future actual returns.

To understand our method, note first that Ross (2015) assumes that state prices are known not just in each final state, but also starting from each possible current state as illustrated in Figure 1, Panel A. Simply put, he assumes that we know all prices today and all prices in all "parallel universes" with different starting points. Since we clearly cannot observe such parallel universes, Ross (2015) proposes to implement his model based on prices for several future time periods, relying on the assumption that all time periods have identical structures for prices and probabilities (time-homogeneity), illustrated in Figure 1, Panel B. In other words, Ross assumes that, if S&P 500 is 2000, then one-period option prices are the same regardless of the time period.

We show that the recovery problem can be simplified by starting directly with the state prices for all future times given only the *current* state (Figure 1, Panel C). We impose no dynamic structure on the probabilities, allowing the probability distribution to be fully general at each future time, thus relaxing Ross's time-homogeneity assumption which is unlikely to be met empirically.

We first show that when the number of states S is no greater than the number of time periods T, then recovery is possible. To see the intuition, consider simply the number of equations and the number of unknowns: First, we have S equations at each time period, one for each Arrow-Debreu price, for a total of ST equations. Second, we have 1 unknown discount rate, $S-1$ unknown marginal utilities, and – for each future time period – we have $S-1$ unknown probabilities. In conclusion, we have ST equations with $1+S-1+(S-1)T = ST+S-T$ unknowns. These equations are not linear, but we provide a precise sense in which we can essentially just count equations. Hence, recovery is possible when $S \leq T$.

To understand the intuition behind this result, note that, for each time period, we have S equations and only $S-1$ probabilities. Hence, we have one extra equation that can help us recover the marginal utilities and discount rate — and the number of marginal utilities does not grow with the number of time periods.

By focusing on square matrices, Ross's model falls into the category $S = T$ so our counting argument explains why he finds recovery. However, our method applies under much more general conditions. We show that, when Ross's time-homogeneity conditions are met, then our solution is the same as his. The converse is not true: when Ross's conditions are not met, then our model can be solved while Ross's cannot. Further, we illustrate that our solution is far simpler and allows a closed-form solution that is accurate when the discount rate is close to 1.

To understand the condition $S \leq T$, consider what happens if the economy evolves in a standard multinomial tree with no upper or lower bound on the state space: For each extra time period, we get at least two new states since we can go up from highest state and down from the lowest state. Therefore, in this case $S > T$, so we see that recovery is impossible because of the number of states is higher than the number of time periods. Hence, achieving recovery without further assumptions is typically impossible in most standard models of finance where the state space grows in this way. In other words, our model provides a simple alternative way – via our counting argument – to understand the critique of Borovicka, Hansen, and Scheinkman (2015) that recovery is impossible in standard models.

Nevertheless, we show that recovery is possible even when $S > T$ under certain conditions. While maintaining that probabilities can be fully general (and, indeed, allow growth), we assume that the utility function is given via a limited number of parameters. Again, we simply need to make our counting argument work. To do this, we show that, if the pricing kernel can be written as functions of N parameters, then recovery is possible as long as $N + 1 < T$. This large state-space framework is what we use empirically as discussed further below.

We illustrate how our method works in the context of three specific models, namely Mehra and Prescott (1985), Black and Scholes (1973), and a simple non-Markovian economy. For each economy, we generate model-implied prices and seek to recover natural probabilities and preferences using our method. This provides an illustration of how our method works, its robustness, and its shortcomings. For Mehra and Prescott (1985), we show that $S > T$ so general recovery is impossible, but, when we restrict the class of utility functions, then we achieve recovery. For the binomial model in the spirit of Black and Scholes (1973), we show that recovery is impossible even under restrictive utility specifications because consumption growth is iid., which leads to a flat term structure, a pricing matrix of a lower rank, and a continuum of solutions for probabilities and preferences. Finally, we show how recovery is possible in the non-Markovian setting, which falls outside the framework of Borovicka, Hansen, and Scheinkman (2015) and Ross (2015), illustrating the generality of our framework in terms of the allowed probabilities.

Finally, we implement our methodology empirically using the large set of call and put options written on the S&P 500 stock market index over the time period 1996-2014. We estimate state price densities for multiple future horizons and apply our closed-form method to recover probabilities and preferences each month. Based on the recovered probabilities, we derive the risk and expected return over the future month from the physical distribution of returns. Our empirical results suggest that the recovered statistics have predictive power for the distribution of future realized returns, although we caution that these results are based on a relatively short sample of 18 years and we are able to reject that the full distribution of recovered probabilities exactly matches the true distribution using a Berkowitz test.

The literature on recovery theorems is quickly expanding. Hansen and Scheinkman (2009) provide general results of their operator approach to long term risk. Bakshi, Chabi-Yo, and Gao (2015) empirically test the restrictions of the Recovery Theorem. Audrino, Huitema, and Ludwig (2014) and Ross (2015) consider how to extract a full transition state price matrix from current option prices, relying on time-homogeneity and additional restrictions and approximations. Martin and Ross (2013) apply the recovery theorem in a term structure model in which the driving state variable is a stationary Markov chain and they show how recovery can be done using the (infinitely) long end of the yield curve. Several papers focus on generalizing the underlying Markov process to a continuous-time process with a continuum of values: Carr and Yu (2012) use Long's portfolio to show a recovery result using Sturm-Liouville theory as the equivalent to Ross's use of Perron-Frobenious theory. Walden (2013) shows how recovery is possible in an unbounded diffusion setting, and Linetsky and Qin (2015) show a recovery theorem assuming that the driving state process belongs to a general class of continuous-time Markov processes (Borel right processes) which include multidimensional processes in bounded and unbounded state spaces. These papers all impose time-homogeneity of the underlying Markov process. Schneider and Trojani (2015) focus on recovering moments of the physical distribution and choose among potential pricing kernels matching these moments the kernel giving rise to a minimal variance physical measure. What these papers have in common is that they attempt to recover physical probabilities and the pricing kernel using forward looking information. Malamud (2016) shows that knowledge of investor preferences is not necessarily enough to recover physical probabilities when option supply is noisy, but shows how recovery can may be feasible when the volatility of option supply shocks is also known. Prior to Ross (2015), the dynamics of the risk-neutral density and the physical density along with the pricing kernel has been extensively researched using historical option or equity market data. A partial list of prominent papers includes, Jackwerth (2000), Jackwerth and Rubinstein (1996), Bollerslev and Todorov (2011), Ait-Sahalia and Lo (2000), Rosenberg and Engle (2002), Bliss and Panigirtzoglou (2004) and Christoffersen, Heston, and Jacobs (2013).

Our paper contributes to the literature by characterizing recovery for any probability distribution, not just time-inhomogeneous Markov process, by proving a simple solution and its closed-form approximation, and by providing natural empirical tests. Rather than relying on specific probabilistic assumptions (Markov processes and ergodocity) as in Ross (2015) and Borovicka, Hansen, and Scheinkman (2015), we follow the tradition of general equilibrium (GE) theory, where Debreu (1970) pioneered the use of Sard's theorem and differential topology. Bringing Sard's theorem into the recovery debate provides new economic insight on when recovery is possible.[1]

The remainder of the paper is structured as follows. Section 2 briefly reviews Ross's Recovery Theorem. Section 3 develops our Generalized Recovery Theorem, showing how and when marginal utilities, physical probabilities, and the discount rate can be decoded from prices. Section 4 provides a closed-form solution to the recovery problem. Section 5 generalizes our model to capture a large state space in which marginal utilities are given by a lower-dimensional set of (risk aversion) parameters. Section 6 illustrates our method in the context of three specific models. Section 7 describes our data and empirical methodology and Section 8 provides our empirical results.

[1]We thank Steve Ross for pointing out the historical role of Sard's theorem in general equilibrium theory.

2 Ross's Recovery Theorem

This section briefly describes the mechanics of the recovery theorem of Ross (2015) as a background for understanding our generalized result in which we relax the assumption that transition probabilities are time-homogeneous.

The idea of the recovery theorem is most easily understood in a one-period setting. In each time period 0 and 1, the economy can be in a finite number of states which we label $1, \ldots, S$. Starting in any state i, there exists a full set of Arrow-Debreu securities, each of which pay 1 if the economy is in state j at date 1. The price of these securities is given by $\pi^{i,j}$.

The objective of the recovery theorem is to use information about these observed state prices to infer physical probabilities $p^{i,j}$ of transitioning from state i to j. We can express the connection between Arrow-Debreu prices and physical probabilities by introducing a pricing kernel m such that for any $i, j = 1, ..., S$

$$\pi^{i,j} = p^{i,j} m^{i,j} \tag{1}$$

It takes no more than a simple one-period binomial model to convince oneself, that if we know the Arrow-Debreu prices in one and only one state at date 0, then there is in general no hope of recovering physical probabilities. In short, we cannot separate the contribution to the observed Arrow-Debreu prices from the physical probabilities and the pricing kernel.

The key insight of the recovery theorem is that by assuming that we know the Arrow-Debreu prices for *all* the possible starting states, then with additional structure on the pricing kernel, we can recover physical probabilities. We note that knowing the prices in states we are not currently in ("parallel universes") is a strong assumption.

In any event, under this assumption, Ross's result is that there exists a unique set of physical probabilities $p^{i,j}$ for all $i, j = 1, \ldots, S$ such that (1) holds if the matrix of Arrow-Debreu prices is irreducible and if the pricing kernel m has the form known from the standard representative agent models:

$$m^{i,j} = \delta \frac{u^j}{u^i} \tag{2}$$

where $\delta > 0$ is the discount rate and $u = (u^1, \ldots, u^S)$ is a vector with strictly positive elements representing marginal utilities.

The proof can be found in Ross (2015), but here we note that counting equations and unknowns certainly makes it plausible that the theorem is true: There are S^2 observed Arrow-Debreu prices and hence S^2 equations. Because probabilities from a fixed starting state sum to one, there are $S(S-1)$ physical probabilities. It is clear that scaling the vector u by a constant does not change the equations, and thus we can assume that $u^1 = 1$ so that u contributes with an additional $S-1$ unknowns. Adding to this the unknown δ leaves us exactly with a total of S^2 unknowns. The fact that there is a unique strictly positive solution hinges on the Frobenius theorem for positive matrices.

The most troubling assumption in the theorem above is that we must know state prices also from starting states that we are currently not in. It is hard to imagine data that would allow us to know these in practice. Ross's way around this assumption is to leave the one-period setting and assume that we have information about Arrow-Debreu prices from several future periods and then use a time-homogeneity assumption to recover the same information that we would be able to obtain from the equations above.

We therefore consider a discrete-time economy with time indexed by t, states indexed by $s = 1, ..., S$, and $\pi^{i,j}_{t,t+\tau}$ denoting the time-t price in state i of an Arrow-Debreu security that pays 1 in state j at date $t+\tau$. The multi-period analogue of Eqn. (1) becomes

$$\pi^{i,j}_{t,t+\tau} = p^{i,j}_{t,t+\tau}\, m^{i,j}_{t,t+\tau} \tag{3}$$

Similarly, the multi-period analogue to equation (2) is the following assumption, which again follows from the existence of a representative agent with time-separable utility:

Assumption 1 (Time-separable utility) *There exists a* $\delta \in (0, 1]$ *and marginal utilities* $u^j > 0$ *for each state* j *such that, for all times* τ*, the pricing kernel can be written as*

$$m^{i,j}_{t,t+\tau} = \delta^{\tau} \frac{u^j}{u^i} \tag{4}$$

Critically, to move to a multi-period setting, Ross makes the following additional assumption of time-homogeneity in order to implement his approach empirically:

Assumption 2 (Time-homogeneous probabilities) *For all states* i, j *and time horizons* $\tau > 0$*,* $p^{i,j}_{t,t+\tau}$ *does not depend on* t*.*

This assumption is strong and not likely to be satisfied empirically. We note that Assumptions 1 and 2 together imply that risk neutral probabilities are also time-homogeneous, a prediction that can also be rejected in the data.

In this paper, we dispense with the time-homogeneity Assumption 2. We start by maintaining Assumption 1, but later consider a broader assumption that can be used in a large state space.

3 A Generalized Recovery Theorem

The assumption of time-separable utility is consistent with many standard models of asset pricing, but the assumption of time-homogeneity is much more troubling. It restricts us from working with a growing state space (as in standard binomial models) and it makes numerical implementation extremely hard and non-robust, because trying to fit observed state prices to a time-homogeneous model is extremely difficult. Furthermore, the main goal of the recovery exercise is to recover physical transition probabilities from the current states to all future states over different time horizons. Insisting that these transition probabilities arise from constant one-period transition probabilities is a strong restriction. We show in this section that by relaxing the assumption of time-homogeneity of physical transition probabilities, we can obtain a problem which is easier to solve numerically and which allows for a much richer modeling structure. We show that our extension contains the time-homogeneous case as a special case, and therefore ultimately should allow us to test whether the time-homogeneity assumption can be defended empirically.

3.1 A Noah's Arc Example: Two States and Two Dates

To get the intuition of our approach, we start by considering the simplest possible case with two states and two time-periods. Consider the simple case in which the economy has two possible states $(1, 2)$ and two time periods starting at time t and ending on dates $t + 1$ and $t + 2$. If the current state of the world is state 1, then equation (3) consists of four equations:

$$\begin{array}{rcll}
\pi^{1,1}_{t,t+1} & = & p^{1,1}_{t,t+1} & m^{1,1}_{t,t+1} \\
\pi^{1,2}_{t,t+1} & = & (1 - p^{1,1}_{t,t+1}) & m^{1,2}_{t,t+1} \\
\pi^{1,1}_{t,t+2} & = & p^{1,1}_{t,t+2} & m^{1,1}_{t,t+2} \\
\pi^{1,2}_{t,t+2} & = & \underbrace{(1 - p^{1,1}_{t,t+2})}_{\text{2 unknowns}} & \underbrace{m^{1,2}_{t,t+2}}_{\text{4 unknowns}}
\end{array} \tag{5}$$

We see that we have 4 equations with 6 unknowns so this system cannot be solved in full generality. However, the number of unknowns is reduced under the assumption of time-separable utility (Assumption 1). To see that most simply, we introduce the notation h for the normalized vector of of marginal utilities:

$$h = \left(1, \frac{u^2}{u^1}, \ldots, \frac{u^S}{u^1}\right)' \equiv (1, h_2, \ldots, h_S)'. \tag{6}$$

where we normalize by u^1. With this notation and the assumption of time-separable utility, we can rewrite the system (5) as follows:

$$
\begin{aligned}
\pi^{1,1}_{t,t+1} &= p^{1,1}_{t,t+1}\delta \\
\pi^{1,2}_{t,t+1} &= (1 - p^{1,1}_{t,t+1})\delta h_2 \\
\pi^{1,1}_{t,t+2} &= p^{1,1}_{t,t+2}\delta^2 \\
\pi^{1,2}_{t,t+2} &= (1 - p^{1,1}_{t,t+2})\delta^2 h_2
\end{aligned}
\tag{7}
$$

This system now has 4 equations with 4 unknowns, so there is hope to recover the physical probabilities p, the discount rate δ, and the ratio of marginal utilities h.

Before we proceed to the general case, it is useful to see how the problem is solved in this case. Moving h_2 to the left side and adding the first two and the last two equations gives us two new equation

$$
\begin{aligned}
\pi^{1,1}_{t,t+1} + \pi^{1,2}_{t,t+1}\frac{1}{h_2} - \delta = 0 \\
\pi^{1,1}_{t,t+2} + \pi^{1,2}_{t,t+2}\frac{1}{h_2} - \delta^2 = 0
\end{aligned}
\tag{8}
$$

Solving equation (8) for h_2 yields $\frac{1}{h_2} = (\delta - \pi^{1,1}_{t,t+1})/\pi^{1,2}_{t,t+1}$ and we can further arrive at

$$
\pi^{1,1}_{t,t+2} - \frac{\pi^{1,2}_{t,t+2}\pi^{1,1}_{t,t+1}}{\pi^{1,2}_{t,t+1}} + \frac{\pi^{1,2}_{t,t+2}}{\pi^{1,2}_{t,t+1}}\delta - \delta^2 = 0 \tag{9}
$$

Hence, we can solve the 2-state model by (i) finding δ as a root of the 2nd degree polynomial (9); (ii) computing the marginal utility ratio h_2 from (8); and (iii) computing the physical probabilities by rearranging (7).

3.2 General Case: Notation

Turning to the general case, recall that there are S states and T time periods. Without loss of generality, we assume that the economy starts at date 0 in state 1. This allows us to introduce some simplifying notation since we do not need to keep track of the starting time or the starting state — we only need to indicate the final state and the time horizon over which we are considering a specific transition.

Accordingly, let $\pi_{\tau s}$ denote the price of receiving 1 at date τ if the realized state is s and collect the set of observed state prices in a $T \times S$ matrix Π defined as

$$\Pi = \begin{bmatrix} \pi_{11} & \dots & \pi_{1S} \\ \vdots & & \vdots \\ \pi_{T1} & \dots & \pi_{TS} \end{bmatrix} \tag{10}$$

Similarly, letting $p_{\tau s}$ denote the physical transition probabilities of going from the current state 1 to state s in τ periods, we define a $T \times S$ matrix P of physical probabilities. Note that $p_{\tau s}$ is *not* the probability of going from state τ to s (as in the setting of Ross (2015)), but, rather, the first index denotes time for the purpose of the derivation of our theorem.

From the vector of normalized marginal utilities h defined as in (6) we define the $S-$dimensional diagonal matrix $H = \text{diag}(h)$. Further, we construct a $T-$dimensional diagonal matrix of discount factors as $D = \text{diag}(\delta, \delta^2, \dots, \delta^T)$.

3.3 Generalized Recovery

With this notation in place, the fundamental TS equations linking state prices and physical probabilities, assuming utilities depend on current state only, can be expressed in matrix form as

$$\Pi = DPH \tag{11}$$

Note that the (invertible) diagonal matrices H and D depend only on the vector h and the constant δ so, if we can determine these, we can find the matrix of physical transition probabilities from the observed state prices in Π:

$$P = D^{-1}\Pi H^{-1} \tag{12}$$

Since probabilities add up to 1, we can write $Pe = e$, where $e = (1, \ldots, 1)'$ is a vector of ones. Using this identity, we can simplify (12) such that it only depends on δ and h:

$$\Pi H^{-1}e = DPe = De = (\delta, \delta^2, \ldots, \delta^T)' \tag{13}$$

To further manipulate this equation it will be convenient to work with a division of Π into block matrices:

$$\Pi = \begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix} = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} \tag{14}$$

Here, Π_1 is a column vector of dimension T, where the first $S-1$ elements are denoted by Π_{11} and the rest of the vector is denoted Π_{21}. Similarly, Π_2 is a $T \times (S-1)$ matrix, where the first $S-1$ rows are called Π_{12} and the last rows are called Π_{22}. With this notation and the fact that $H(1,1) = h(1) = 1$, we can write (13) as

$$\Pi_1 + \Pi_2 \begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{bmatrix} \delta \\ \vdots \\ \delta^T \end{bmatrix} \tag{15}$$

where of course $h_s^{-1} = \frac{1}{h_s}$. Given that these equations are linear in the inverse marginal utilities h_s^{-1}, it is tempting to solve for these. To solve for these $S-1$ marginal utilities, we consider the first $S-1$ equations

$$\Pi_{11} + \Pi_{12} \begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{bmatrix} \delta \\ \vdots \\ \delta^{S-1} \end{bmatrix} \tag{16}$$

with solution[2]

$$\begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \Pi_{12}^{-1} \left(\begin{bmatrix} \delta \\ \vdots \\ \delta^{S-1} \end{bmatrix} - \begin{bmatrix} \pi_{11} \\ \vdots \\ \pi_{S-1,1} \end{bmatrix} \right) \tag{17}$$

Hence, if δ were known, we would be done. Since δ is a discount rate, it is reasonable to assume that it is close to one over short time periods. We later use this insight to derive a closed-form approximation which is accurate as long as we have a reasonable sense of the size of δ. For now, we proceed for general unknown δ.

We thus have the utility ratios given as a linear function of powers of δ. The remaining $T-S+1$ equations give us

$$\Pi_{21} + \Pi_{22} \begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{bmatrix} \delta^S \\ \vdots \\ \delta^T \end{bmatrix} \tag{18}$$

and from this we see that if we plug in the expression for the utility ratios found above, we end up with $T-S+1$ equations, each of which involves a polynomium in δ of degree a most T. If $T=S$, then δ is a root to a single polynomium so at most a finite number of solutions exist. If $T>S$, then typically no solution exists for general Arrow-Debreu prices Π since δ must simultaneously solve several polynomial equations. However, if the prices are generated by the model, then a solution exists and it will almost surely be unique. To be precise, we say that Π has been "generated

[2] Of course, to invert Π_{12} it must have full rank. As long as Π_2 has full rank, we can re-order the rows to ensure that Π_{12} also has full rank.

by the model" if there exist δ, P, and H such that Π can be found from the right-hand side of (11). The following theorem formalizes these insights (using Sard's Theorem):

Proposition 1 (Generalized Recovery) *Consider an economy satisfying Assumption 1 with Arrow-Debreu prices for each of the T time periods and S states. The recovery problem has*

1. *a continuum of solutions if $S > T$;*

2. *at most S solutions if the submatrix Π_2 has full rank and $S = T$;*

3. *no solution generically in terms of an arbitrary positive matrix Π and $S < T$;*

4. *a unique solution generically if Π has been generated by the model and $S < T$.*

Proof. We have already provided a proof for 1 and 2 in the body of the text. Turning to 3, we note that the set X of all (δ, h, P) is a manifold-with-boundary of dimension $S \cdot T - T + S$. The discount rate, probabilities and marginal utilities map into prices, which we denote by $F(\delta, h, P) = DPH = \Pi$, where, as before, $D = \text{diag}(\delta, ..., \delta^T)$ and $H = \text{diag}(1, h_2, ..., h_S))$, and F is C^∞. If $S < T$, the image $F(X)$ has Lebesgue measure zero in $\mathbb{R}^{T \times S}$ by Sard's theorem, proving 3. Indeed, this means that the prices that are generated by the model $F(X)$ have measure zero relative to all prices Π.

Turning to 4, we first note that P and H can be uniquely recovered from (δ, Π) (given that Π is generically full rank). Indeed, H is recovered from (17) and P is recovered from (12). Therefore, we can focus on (δ, Π).

For two different choices of the discount rate (δ_a, δ_b) and a single set of prices Π, we consider the triplet $(\delta_a, \delta_b, \Pi)$. We are interested in showing that the different discount rates cannot both be consistent with the same prices, generically. To show this, we consider the space M where the reverse is true, hoping to show that M is "small." Specifically, M is the set of triplets where Π is of full rank and both discount rates are consistent with the prices, that is, there exists (unique) P_i and H_i ($i = a, b$) such that $D_a P_a H_a = D_b P_b H_b = \Pi$.

Given that probabilities and marginal utilities can be uniquely recovered from prices and a discount rate (as explained above), we have a smooth map G from M to X by mapping any triplet $(\delta_a, \delta_b, \Pi)$ to (δ_a, h_a, P_a), where (h_a, P_a) are the recovered marginal utility and probabilities. The image of this map consists exactly of those elements of X for which F is not injective. The proof is complete if we can show that this image has Lebesgue measure zero, which follows again by Sard's theorem if we can show that the dimension of M is strictly smaller than $ST - T + S$.

To study the dimension of M, we note that we can think of M as the space of triplets such that the span of Π contains both the points $(\delta_a, \delta_a^2, ..., \delta_a^T)'$ and $(\delta_b, \delta_b^2, ..., \delta_b^T)'$. The span of Π is given by $V_\Pi := \{\Pi \cdot (1, h_2, h_3, ..., h_S)' | h_s > 0\}$, which is an affine $(S-1)$-dimensional subspace of $\mathbb{R}^T$ for Π of full rank. The set of all those $\Pi \in \mathbb{R}^{T \times S}$ such that V_Π passes through two given points of $\mathbb{R}^T$ (in general position with respect to each other) form a subspace of dimension $ST - 2(T - S + 1)$ since each point imposes $T - S + 1$ equations (and saying that the points are in general position means that all these equations are independent). Therefore, M is a manifold of dimension $ST - 2T + 2S$ since the pair (δ_a, δ_b) depends on two parameters, and, for a given pair, there is a $(ST - 2T + 2S - 2)$-dimensional subspace of possible Π (any two distinct points are always in general position). Hence, we see that $\dim(M) = ST - 2T + 2S < ST - T + S = \dim(X)$ since $S < T$, which implies that $G(M)$ has measure zero in X. Further, the prices where recovery is impossible, $F(G(M))$, have measure zero in the space of all prices generated by the model $F(X)$ where we use the Lebesgue measure on X to define a measure[3] on $F(X)$. ■

Proposition 1 provides a simple way to understand when recovery is possible, namely, essentially when the number of time periods T is at least as large as the number of states S.

[3]We can define a measure on $F(X)$ by $\mu^*(A) := \mu(F^{-1}(A))$ for any set A, where μ is the Lebesgue measure on X.

Proposition 1 also sheds an alternative light on the critique of Borovicka, Hansen, and Scheinkman (2015) that recovery is infeasible in standard models. Indeed, we provide a simple counting argument: Suppose that the economy has growth such that, for each extra time period, the economy can increase from the previously highest state and go down from the previously lowest state. Then we get two new states for each new time period, which implies that $S > T$ such that recovery is impossible. Nevertheless, we can still achieve recovery in such a large state space if we consider a class of pricing kernels that is sufficiently low-dimensional as we discuss below in Section 5.

3.4 Further Results

We next show that our problem is indeed a generalized problem in the sense that if a solution exists satisfying the more restrictive assumptions in Ross (2015), then it is also a solution to our problem. The reverse is *not* true: a solution to the generalized recovery problem cannot be achieved in Ross's framework if the world is not time-homogeneous.

Proposition 2 (Strictly More General Method) *Suppose that we observe T periods of state prices given the current state at date 0 and Assumption 1 applies (time-separable utility).*

1. *If Assumption 2 also applies (time-homogeneity) then a solution to Ross's Recovery problem produces a solution to our generalized recovery problem as well. Generically among price matrices for Ross's problem, the corresponding price matrix Π for the generalized recovery problem is full rank.*

2. *A solution to the generalized recovery problem is not in general a solution to Ross's recovery problem without Assumption 2. With $S = T$, there exists set of parameters with positive Lebesgue measure for the generalized recovery problem where no solution exists for Ross's recovery problem. With $S > T$, generically*

among price matrices for the the generalized recovery problem, there exists no solution to Ross's recovery problem.

Proof. For part 1, let $\bar{\Pi}$ denote an $S \times S$ matrix of one-period state prices as considered in Ross (2015), i.e., $\bar{\pi}_{ij}$ is the value in state i at date 0 of receiving 1 in the next period if the state is j. Let F denote the corresponding matrix of one-period physical transition probabilities. A solution to Ross' problem satisfies

$$\bar{\Pi} = \delta H^{-1} F H \tag{19}$$

and therefore also by time-homogeneity for all $k = 1, \ldots, T$

$$\bar{\Pi}^k = \delta^k H^{-1} F^k H \tag{20}$$

If the starting state is 1 (without loss of generality) then the equations of our generalized recovery problem are the subset obtained by considering the first row of each equation obtained by varying k above. The equations above show that by setting the k'th row of our matrix of physical transition probabilities P equal the first row of F^k, we have a solution to the equations for our generalized recovery problem.

To see that Π is full rank, we first diagonalize Ross's price matrix as $\bar{\Pi} = VZV'$, where $Z = \text{diag}(z_1, ..., z_S)$ is the matrix of eigenvalues and V is the matrix of eigenvectors. The k'th row in the generalized-recovery pricing matrix is the first row (still assuming that the starting state is 1) of $\bar{\Pi}^k = VZ^kV'$. Letting v denote the first row in V, we see that the k'th row of Π is $vZ^kV' = (v_1 z_1^k, ..., v_S z_S^k)V'$ so

$$\Pi = \begin{bmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ z_1^{T-1} & \ldots & z_S^{T-1} \end{bmatrix} \begin{bmatrix} v_1 z_1 & & 0 \\ & \ddots & \\ 0 & & v_S z_S \end{bmatrix} V' \tag{21}$$

Therefore, Π is full rank generically because it is the product of three full-rank matrices. Indeed, the first matrix is a Vandermonde matrix, which is full rank when the

z's are non-zero and different, which is true generically. The second matrix is clearly also full-rank since the v's are also non-zero generically, and the third matrix is full rank by construction.

For part 2, consider first the case where $S < T$. The dimension of the parameter set (transition probabilities + utility parameters) generating the generalized-recovery price matrix Π is $ST - T + S$, which is strictly greater than the dimension S^2 of the parameter space generating price matrices in Ross's homogeneous case. Hence, generically no time-homogeneous solution can generate a generalized recovery price Π.

Our framework is also more general in the the case $S = T$. Recalling that $p_{\tau i}$ denotes the probability of going from the current state 1 to state i in τ periods, it is clear that in a time-homogeneous setting we must have $p_{22} \geq p_{11}p_{12}$, i.e., the probability of going from state 1 to state 2 in two periods is (conservatively) bounded below by the probability obtained by considering the particular path that stays in state 1 in the first time period and then jumps to state 2 in the second. However, such a bound need not apply for the true probabilities if the transition probabilities are not time-homogeneous. The set of parameters that can generate Π matrices that are not attainable from homogeneous transition probabilities is clearly of Lebesgue measure greater than zero in the S^2-dimensional parameter space. ■

Part 1 of the proposition shows that, when Ross's assumptions are met, a solution to his problem is also a solution to our generalized problem. Further, our method can also recover the underlying parameters (as per Proposition 1) since the price matrix Π is full rank. Part 2 of the proposition shows that for many "typical" price matrices (e.g., those observed in the data), no solution exists for Ross's recovery problem even though a solution exists for the generalized recovery problem.

We finally note that the very special case of an observed flat term structure of interest rates has some special properties. In particular, with a flat term structure there exists a solution to the problem in which the representative agent is risk neutral, echoing an analogous result by Ross.

To see this result, note that the price of a zero-coupon bond with maturity τ is equal to the sum of the τ'th row of Π, which we write as $(\Pi e)_\tau$. Having a flat term structure means that the yield on the zero-coupon bonds does not depend on maturity, i.e., that there exists a constant r such that

$$\frac{1}{(1+r)^\tau} = (\Pi e)_\tau \tag{22}$$

Let the $T \times S$ matrix Q contain the risk-neutral transition probabilities seen from the starting state, i.e., the k'th row of Q gives us the risk-neutral probabilities of ending in the different states at date k.

Proposition 3 (Flat Term Structure) *Suppose that the term structure of interest rates is flat, i.e., there exists $r > 0$ such that $\frac{1}{(1+r)^\tau} = (\Pi e)_\tau$ for all $\tau = 1, \ldots, T$. Then the recovery problem is solved with equal physical and risk-neutral probabilities, $P = Q$. This means that either the representative agent is risk neutral or the recovery problem has multiple solutions.*

Proof. Let R denote the diagonal matrix whose k'th diagonal element is $\frac{1}{(1+r)^k}$. Having a flat term structure means that the matrix Π of state prices as seen from a particular starting state can be written as

$$\Pi = RQ$$

which defines Q as a stochastic matrix (i.e., with rows that sum to 1). Clearly, by letting $\delta = 1/(1+r)$ and having risk-neutrality, i.e. $H = I_S$ (the identity matrix of dimension S), we obtain a solution to our recovery problem

$$\Pi = RQ = DPH = RPI_S = RP$$

by setting $P = Q$. ■

We note that this result should be interpreted with caution. The knife-edge (i.e., measure zero) case of a flat term structure may well be generated by the knife-edge case of a price matrix Π with low rank, which implies that a continuum of solutions may exists and the representative agent may well be risk averse (as one would expect). Intuitively, a flat term structure may be generated by a Π with so much symmetry that it has a low rank.

4 Closed-Form Recovery

The recovery problem is almost linear, except for the powers of the discount rate δ which enter into the problem as a polynomial. In practical implementations over the time horizons where options are liquid, a linear approximation provides an accurate approximation given that δ is close to one. For instance, we know from the literature that δ is close to 0.97 at an annual horizon.

The linear approximation is straightforward. To linearize the discounting of δ^τ around a point δ_0 (say, $\delta_0 = 0.97$), we write $\delta^\tau \approx a_\tau + b_\tau \delta$ for known constants a_τ and b_τ. Based on the Taylor expansion $\delta^\tau \approx \delta_0^\tau + \tau\delta_0^{\tau-1}(\delta - \delta_0)$, we have $a_\tau = -(\tau - 1)\delta_0^\tau$ and $b_\tau = \tau\delta_0^{\tau-1}$. As seen in Figure 2, the approximation is accurate for $\delta \in [0.94, 1]$ for time horizons less than 2 years.

With the linearization of the polynomials in δ, the equations for the recovery problem (13) become the following:

$$\begin{pmatrix} \pi_{11} \\ \vdots \\ \pi_{T1} \end{pmatrix} + \begin{pmatrix} \pi_{12} & \ldots & \pi_{1S} \\ \vdots & & \vdots \\ \pi_{T2} & \ldots & \pi_{TS} \end{pmatrix} \begin{pmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{pmatrix} = \begin{pmatrix} a_1 + b_1\delta \\ \vdots \\ a_T + b_T\delta \end{pmatrix} \tag{23}$$

which we can rewrite as a system of T equations in S unknowns as

$$\begin{pmatrix} -b_1 & \pi_{12} & \ldots & \pi_{1S} \\ \vdots & \vdots & & \vdots \\ -b_T & \pi_{T2} & \ldots & \pi_{TS} \end{pmatrix} \begin{pmatrix} \delta \\ h_2^{-1} \\ \vdots \\ h_S^{-1} \end{pmatrix} = \begin{pmatrix} a_1 - \pi_{11} \\ \vdots \\ a_T - \pi_{T1} \end{pmatrix} \tag{24}$$

Rewriting this equation in matrix form as

$$Bh_\delta = a - \pi_1 \tag{25}$$

we immediately see the closed-form solution

$$h_\delta = \begin{cases} B^{-1}(a - \pi_1) & \text{for } S = T \\ (B'B)^{-1}B'(a - \pi_1) & \text{for } S < T \end{cases} \tag{26}$$

We see that, when $S = T$, we simply need to solve S linear equations with S unknowns. When $S < T$, we could simply just consider S equations and ignore the remaining $T - S$ equations.

More broadly, if $S < T$ and we start with prices Π that are not exactly generated by the model (e.g., because of noise in the data), then (26) provides the values of δ and the vector h that best approximate a solution in the sense of least squares.

The following theorem shows that the closed-form solution is accurate as long as the value of δ_0 is close to the true discount rate:

Proposition 4 (Closed-Form Solution) *If prices are generated by the model and B has full rank $S \leq T$ then the closed-form solution (26) approximates the true solution in the following sense: The distance between the true solution $(\bar{\delta}, \bar{h}, \bar{P})$ and the approximate solution (δ, h, P) approaches 0 faster than $(\delta_0 - \bar{\delta})$ as δ_0 approaches $\bar{\delta}$.*

Proof. The approximation result follows from Lemma 1 in the appendix. ■

5 Recovery in a Large State Space

A challenge in implementing the Ross Recovery Theorem is that it does not allow for an expanding set of states as we know it, for example, from binomial models and multinomial models of option pricing. Simply stated, the expanding state space in a binomial model adds more unknowns for each time period than equations even under the assumption of utility functions that depend on the current state only. We next show how we handle an expanding state space in our model.

We have in mind a case where the number of states S is larger than the number of time periods T. In a standard binomial model, for example, with two time periods we need five states corresponding to the different values that the stock can take over its path. The key to solving this problem is to reduce the dimensionality of the utility ratios captured in the vector h. To do that, we replace Assumption 1 with the following assumption that the pricing kernels belong to a parametric family with limited dimensionality.

Assumption 1* (General utility with N parameters) *The pricing kernel at time τ in state s (given the initial state 1 at time 0) can be written as*

$$m_{0,\tau}^{1,s} = \delta^{\tau} h_s(\theta) \tag{27}$$

where $\delta \in (0, 1]$ and $h(\cdot) > 0$ is a one-to-one C^{∞} smooth function of the parameter $\theta \in \Theta$, an embedding from $\Theta \subset \mathbb{R}^N$ to $\mathbb{R}^S$, and Θ has a non-empty interior.

With a large number of unknowns compared to the number of equations, we need to restrict the set of unknowns, and this is done by assuming that the utilities are parameterized by a lower-dimensional set Θ.

5.1 A Large Discrete State Space

Let us first consider two simple examples of how we can parameterize marginal utilities with a low-dimensional set of parameters. First, we consider a simple linear expression for the marginal utilities and then we discuss the case of constant relative risk aversion (a non-linear mapping from risk aversion parameters Θ to marginal utilities).

We start with a simple linear example of how the parametrization works. We consider a matrix B of full rank and dimension $(S-1) \times N$ such that

$$\begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_{S-1} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1N} \\ \vdots & & \vdots \\ b_{S-1,1} & \dots & b_{S-1,N} \end{pmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} = A + B\theta \tag{28}$$

Combining this equation with the recovery problem (15) gives

$$(\Pi_1 + \Pi_2 A) + \Pi_2 B \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_N \end{pmatrix} = \begin{pmatrix} \delta \\ \vdots \\ \delta^T \end{pmatrix} \tag{29}$$

This equation has *exactly* the same form as our original recovery problem (15), but now $\Pi_1 + \Pi_2 A$ plays the role of Π_1, similarly $\Pi_2 B$ plays the role of Π_2, and θ plays the role of $(h_2^{-1}, ..., h_S^{-1})'$. The only difference is that the dimension of the unknown parameter has been reduced from $S-1$ to N. Therefore, Proposition 1 holds as stated with S replaced by $N+1$.

Hence, while before we could achieve recovery if $S \leq T$, now we can achieve recovery as long as $N+1 \leq T$. In other words, recovery is possible as long as the representative agent's utility function can be specified by a number of parameters that is small relative to the number of time periods for which we have price data.

Assumption 1* also allows for the marginal utilities to be non-linear function of the risk aversion parameters θ. This generality is useful because standard utility functions may give rise to such a non-linearity. As a simple example, consider an economy with a representative agent with CRRA preferences. In this economy, the pricing kernel in state s at time τ (given the current state 1 at time 0) is

$$m_{0,\tau}^{1,s} = \delta^\tau \left(\frac{c_s}{c_1}\right)^{-\theta} \tag{30}$$

where c_s is the known consumption in state s of the representative agent and θ is the unknown risk aversion parameter. Hence, Assumption 1* is clearly satisfied with

$h_s^{-1}(\theta) = (\frac{c_s}{c_1})^\theta$. Our generalized recovery result extends to the large state space as stated in the following proposition.

Proposition 5 (Generalized Recovery in a Large State Space) *Consider an economy satisfying Assumption 1* with Arrow-Debreu prices for each of the T time periods and S states such that $N + 1 < T$. The recovery problem has*

1. *no solution generically in terms of an arbitrary Π matrix of positive elements;*

2. *a unique solution generically if Π has been generated by the model.*

Proof. Following the same logic as the proof of Proposition 1, we note that the set X of all (δ, θ, P) is a manifold-with-boundary of dimension $S \cdot T - T + N + 1$. The discount rate, marginal utility parameters, and probabilities map into prices, which we denote by $F(\delta, \theta, P) = DPH = \Pi$, where, as before, $D = \mathrm{diag}(\delta, ..., \delta^T)$ and $H = \mathrm{diag}(h_1(\theta), h_2(\theta), ..., h_S(\theta)))$, and F is C^∞. Since $N + 1 < T$, the image $F(X)$ has Lebesgue measure zero in $\mathbb{R}^{T\times S}$ by Sard's theorem, proving part 1.

Turning to part 2, we first note that P can be uniquely recovered from $(\bar{\theta}, \Pi)$ using equation (12), where $\bar{\theta} = (\delta, \theta)$. Therefore, we can focus on $(\bar{\theta}, \Pi)$, studying the solutions to $\Pi(h_1^{-1}(\theta), ..., h_S^{-1}(\theta))' = (\delta, ..., \delta^T)'$.

For two different choices of the parameters $(\bar{\theta}_a, \bar{\theta}_b)$ and a single set of prices Π, we consider the triplet $(\bar{\theta}_a, \bar{\theta}_b, \Pi)$. We are interested in showing that the different parameters cannot both be consistent with the same prices, generically. To show this, we consider the space M where the reverse is true, hoping to show that M is "small." Specifically, M is the set of triplets where Π is of full rank and both discount rates are consistent with the prices, that is, there exists (unique) P_i $(i = a, b)$ such that $D_a P_a H_a = D_b P_b H_b = \Pi$.

Given that probabilities can be uniquely recovered from prices and parameters, we have a smooth map G from M to X by mapping any triplet $(\bar{\theta}_a, \bar{\theta}_b, \Pi)$ to $(\delta_a, \theta_a, P_a)$. The image of this map consists exactly of those elements of X for which F is not injective. The proof is complete if we can show that this image has Lebesgue measure zero, which follows again by Sard's theorem if we can show that the dimension of M is strictly smaller than $S \cdot T - T + N + 1$.

To study the dimension of M, consider first $V_\Pi := \{\Pi(h_1^{-1}(\theta), ..., h_S^{-1}(\theta))' | \theta \in \Theta\}$, which is an N-dimensional submanifold of $\mathbb{R}^T$ for Π of full rank and given that h is a one-to-one embedding. We note that we can think of M as the space of triplets such that V_Π contains both the points $(\delta_a, \delta_a^2, ..., \delta_a^T)'$ and $(\delta_b, \delta_b^2, ..., \delta_b^T)'$, where the corresponding θ's are given uniquely from the definition of V_Π since Π is full rank and h is one-to-one. The set of all those $\Pi \in \mathbb{R}^{T \times S}$ such that V_Π passes through two given points of $\mathbb{R}^T$ form a subspace of dimension $ST - 2(T - N)$ since each point imposes $T - N$ equations. Therefore, M is a manifold of dimension $ST - 2T + 2N + 2$. Hence, we see that $G(X)$ has measure zero in X and $F(G(X))$ has measure zero in $F(X)$. ∎

5.2 Continuous State Space

Finally, we note that our framework also easily extends to a continuous state space under Assumption 1*. We start with a continuous state-space density $\pi_\tau(s)$ at each time point $\tau = 1, \ldots, T$ (given the current state at time 0). As before, $\pi_\tau(s)$ represents Arrow-Debreu prices or, more precisely, $\pi_\tau(s)ds$ represents the current value of receiving 1 at time τ if the state is in a small interval ds around s. Similarly, we let $p_\tau(s)$ denote the physical probability density of transitioning to s in τ periods. The fundamental recovery equations now become

$$\pi_\tau(s) = \delta^\tau h_s(\theta) p_\tau(s) \tag{31}$$

By moving h to the left-hand side and integrating, we can eliminate the natural probabilities as before.

$$\int \pi_\tau(s) h_s^{-1}(\theta) ds = \delta^\tau \tag{32}$$

For each time period τ, this gives an equation to help us recover the $N+1$ unknowns, namely the discount rate δ and the parameters $\theta \in \mathbb{R}^N$. Hence, we are in the same situation as in the discrete-state model of Section 5.1, and we have recovery if there are enough time periods as stated in Proposition 5.

As before, the linear case is particularly simple. Suppose that the marginal utilities can be written as[4]

$$h_s^{-1}(\theta) = A(s) + B(s)\theta \tag{33}$$

where, for each s, $A(s)$ is a known scalar and $B(s)$ is a known row-vector of dimension N. Using this expression, we can rewrite equation (32) as a simple equation of the same form as our original recovery problem (15):

$$\pi_\tau^A + \pi_\tau^B\ \theta = \delta^\tau \tag{34}$$

where $\pi_\tau^A = \int \pi_\tau(s)A(s)ds$ and $\pi_\tau^B = \int \pi_\tau(s)B(s)ds$. Hence, as before, we have T equations that are linear except for the powers of the discount rate.

6 Recovery in Specific Models: Examples

In this section we investigate recovery of specific models of interest. In a controlled environment, we show when, given state prices, our model recovers the true underlying risk-aversion parameter, time-preference parameter along with the true multiperiod physical probabilities.

6.1 Recovery in the Mehra and Prescott (1985) model

The Mehra and Prescott (1985) model works as follows. The aggregate consumption either grows at rate $u = 1.054$ or shrinks at rate $d = 0.982$ over the next period.

[4] Note that $h_s^{-1}(\theta)$ denotes $\frac{1}{h_s(\theta)}$, i.e., it is *not* the inverse function of $h_s(\theta)$.

This consumption growth between time $t-1$ and t is captured by a process X_t. The aggregate consumption process can be written as

$$Y_t = \prod_{s=1}^{t} X_s \tag{35}$$

where the initial consumption is normalized as $Y_0 = 1$.

Consumption growth X_t is a Markov process with two states, up and down. The probability of having an up state after an up state is $\phi_{uu}; = Pr(X_t = u | X_{t-1} = u) = 0.43$ and, equally, the probability of staying in the down state is $\phi_{dd} = 0.43$. Hence, the probability of switching state is $\phi_{ud} = \phi_{du} = 0.57$.

The Arrow-Debreu price of receiving 1 at time t in a state $s_t = (y_t, x_t)$ is computed based on the CRRA preferences for the representative agent with risk aversion $\gamma = 4$ as

$$\pi_{0,t}^{1,s_t} = \delta^t y_t^{-\gamma} Pr(X_t = x_t, Y_t = y_t) \tag{36}$$

where the time-preference parameter is $\delta = 0.98$ and the physical probabilities $Pr(X_t = x_t, Y_t = y_t)$ of each state are computed based on the Markov probabilities above.[5]

Based on this model of Mehra and Prescott (1985), we compute Arrow-Debreu prices in each state over $T = 20$ time periods and examine whether we can recover probabilities and preferences based on knowing only these prices (we have also performed the recovery for other values of T).

Impossibility of general recovery. We first notice from equation (35) that consumption has growth, which immediately implies that $S > T$. This means that recovery is impossible without further assumptions. Hence, we proceed using the method concerning a large state space of Section 5.

[5] We note that prices of long-lived assets, for example the overall stock market, depends on both X_t and Y_t (even if the aggregate consumption Y_t is the aggregate dividend). Therefore, stock index options would provide information on Arrow-Debreu prices on each state $s_t = (y_t, x_t)$. Alternatively, we could consider recovery based only on Arrow-Debreu securities that depend on y_t. This would correspond to observing options on "dividend strips." Either way, we get the same recovery results in the Mehra and Prescott (1985) model.

Recovery under CRRA. The simplest way to proceed is to assume that we know the form of the pricing kernel (36), but we don't know the risk aversion γ, the discount rate δ, or the probabilities. We can then write the Generalized Recovery equation set on the form

$$\Pi h^{-1}(\gamma) = \begin{bmatrix} \delta & \delta^2 & \dots & \delta^T \end{bmatrix}' \tag{37}$$

where h is a one-to-one C^∞ smooth function of the parameter γ based on (36), see Appendix B for details.[6] Therefore, we are in the domain of Assumption 1* and, as long as $T > 2$ (since $N = 1$ is the number of risk aversion parameters and 2 is the total number of variables, δ and γ) then by Proposition 5 we know that the Generalized Recovery equation set generically has a unique solution.

We first seek to recover γ and δ by minimizing the pricing errors (again, see Appendix B for details). Panel A of Figure 3 shows the objective function for this minimization problem. As seen from the figure, there is a unique solution to the problem, which naturally equals the true parameters $\hat{\delta} = 0.98, \hat{\gamma} = 4$.

Finally, we turn to the recovery of natural probabilities. It is worth noticing that we do *not* recover the Markov switching probabilities $\phi_{uu}, \phi_{dd}, \phi_{ud}$ or ϕ_{du}. Rather, what is recovered is the multi-period probabilities $p_{0,t}^{1,s_t}$ of transitioning from the initial state to each future state (consistent with the intuition conveyed in Figure 1).[7] The probabilities $p_{0,t}^{1,s_t}$ are recovered exactly. Fortunately, these multi-period probabilities are all we need for making predictions about such statistics as expected returns, variances, and quantiles across different time horizons.

6.2 Black-Scholes-Merton and iid. consumption growth

We can capture a binomial model in the spirit of Black-Scholes-Merton and Cox, Ross, and Rubinstein (1979) as follows. We consider the same model for aggregate

[6] Matlab code is available from the authors upon request.

[7] Recovery of the underlying path-dependent probabilities is possible if we have access to Arrow-Debreu prices for all paths or if we assume that we know the structure of the underlying tree.

consumption Y_t, but now X_t is iid. (corresponding to $\phi_{uu} = \phi_{du}$ and $\phi_{dd} = \phi_{ud}$). In other words, the standard binomial Black-Scholes-Merton model has iid consumption growth. Specifically, we assume that up and down probabilities are always 50% ($\phi_{uu} = \phi_{du} = \phi_{dd} = \phi_{ud} = 0.5$).

This binomial model implies a flat term structure which puts us in the case of Proposition 3, where recovery is impossible.[8] Concretely, the problem is that the price matrix Π from (37) is not full rank. Hence, as seen in Figure 3 Panel B, the objective of minimizing pricing errors has a continuum of solutions. In other words, recovery is not feasible.

6.3 A non-stationary model without Markov structure

Lastly, we consider a model where the consumption growth X_t is not Markov. Specifically, we still consider the binomial tree described above in Sections 6.1–6.2, but now we let the probability of transitioning up/down from any state s at any time t depend on the path taken from time 0 to time t. At each node at each path, we draw a random uniformly distributed probability for an "up" move, and, of course, assign one minus this probability to the next "down" node.

We now seek to recover δ and γ. As seen in Figure 3 Panel C, the objective function has a unique solution which again equals the true parameters $\hat{\delta} = 0.98$ and $\hat{\gamma} = 4$. Hence, recovery can be possible even when the driving process is non-stationary and non-Markovian, again under parametric assumptions about the utility function (i.e., a model outside the scope of Ross (2015) and Borovicka, Hansen, and Scheinkman (2015)).

[8] Iid. consumption growth and standard utility functions generally lead to a flat term structure because the price of a bond with τ periods to maturity can be written as $E_t(\delta^\tau \frac{u_{t+\tau}}{u_t}) = E_t(\prod_{s=1,\dots,\tau} \delta \frac{u_{t+s}}{u_{t+s-1}}) =: (\frac{1}{1+r})^\tau$, where the expected utility increments are the same for all s because they depend on consumption growth $\frac{c_{t+s}}{c_{t+s-1}}$, which has constant expected value when it is iid.

7 Data and Empirical Methodology

In this section we describe our data and empirical methodology.

7.1 Data and Sample Selection

We use the Ivy DB database from OptionMetrics to extract information on standard call and put options written on the S&P 500 index for every Wednesday from January 1996 to august 2014. We obtain implied volatilities, strikes, and maturities, allowing us to back out market prices. As a proxy for the risk-free rate, we use the zero-coupon yield curve of the Ivy DB database, which is derived from from LIBOR rates and settlement prices of CME Eurodollar futures. We also obtain expected dividend payments, calculated under the assumption of a constant dividend yield over the life time of the option. We consider options with time to maturity between 10 and 360 days and apply standard filters, excluding contracts with zero open interest, zero trading volume, and quotes with best bid below \$0.50, and options with implied volatility higher than 100%.

7.2 Recovery Methodology

The Generalized Recovery Theorem relies on the knowledge of state prices from the current initial state to all possible future states for several future time periods. Unfortunately, there is currently no market trading pure Arrow-Debreu securities. Therefore, we use options to back out Arrow-Debreu prices as described in detail in Appendix C. This method yields state prices for each day t that we consider across 34 future time horizons, namely 30, 40, 50,..., 350, 360 calender days, and across a range of index levels. Given the current index level S_t and the current VIX_t volatility index, we consider index levels from $S_t - 2.5S_t \times \mathrm{VIX}_t$ to $S_t + 4S_t \times \mathrm{VIX}_t$. For example, on March 12, 2014 the S&P 500 index value was at 1868 and VIX was at 0.1447, so on this day we consider state prices for future index levels from 1192 to 2949.

Given these observed state prices, we recover preferences and probabilities as follows. We apply the closed-form approximation method of Proposition 4 in the context of a large state space as in Section 5. Indeed, as discussed above, we have $T = 34$ time periods and more than 1000 index values so there are many more states than time periods, i.e., in the notation of Proposition 1 we have $S >> T$. Following the discussion in Section 5 we impose a linear, lower-dimensional, structure on the inverse pricing kernel, $H^{-1}e$. We do this by letting $H^{-1}e = B\theta$, where θ is an 11-dimensional column vector and B is a known $S \times 11$ "design matrix."

We use the design matrix B shown in Figure 4. Panel A illustrates the columns of our design matrix, which are piecewise linear. The first column is constant, meaning that the first parameter θ_1 determines the initial level of the inverse pricing kernel $H^{-1}e = B\theta$. The next column slopes up and is then flat, so θ_2 is the initial slope of $B\theta$. Similarly, θ_3 is the slope of the next line segment generated by $B\theta$.

The resulting inverse pricing kernel is shown in Figure 5 Panel A for March 12, 2014. The piecewise linear structure is visible. Panel B shows the corresponding pricing kernel, which has piecewise constant curvature. We impose that $\theta_1, ..., \theta_{11} \geq 0$ which means that the inverse pricing kernel is monotonically increasing or, equivalently, that the pricing kernel is monotonically decreasing[9] i.e., that marginal utility decreases at higher levels of wealth.

The design matrix is characterized by its "break points" that separate the state space into 10 regions. These regions are chosen as follows. The lowest region ranges over states from $(1 - 2.5\text{VIX}_t)S_t$ to $(1 - 2\text{VIX}_t)S_t$. The highest region covers states ranging from $(1 + 2\text{VIX}_t)S_t$ to $(1 + 4\text{VIX}_t)S_t$. In between these extremes, we consider 8 regions of equal size in the range $(1 - 2\text{VIX}_t)S_0$ to $(1 + 2\text{VIX}_t)S_t$. When using this specification of B and the estimated Arrow-Debreu prices, we obtain an $S \times N$ matrix ΠB with full rank for every last Wednesday of the month for the period 1/1996 to

[9] There is an ongoing debate in the literature of whether or not the pricing kernel is monotonically decreasing. Jackwerth (2000), Ait-Sahalia and Lo (2000) and Rosenberg and Engle (2002) suggest that the kernel is not monotonically decreasing in wealth whereas Barone-Adesi, Engle, and Mancini (2008), Bliss and Panigirtzoglou (2004) and Linn, Shive, and Shumway (2015) argue that it is. Our framework allows for both monotonically decreasing and non-monotonically decreasing pricing kernels.

7/2014.

At each date t we impose a first order Taylor expansion of δ^τ, for $\tau \in \{1, ..., 34\}$, in D around the initial guess

$$\bar{\delta} = \frac{1}{1 + r^f_{t,t+1}} \tag{38}$$

where $r^f_{t,t+1}$ is the time t 1-month risk-neutral interest rate. This implies that our starting guess for the physical discounting is risk-neutrality discounting. With this in place we set up the following minimization problem

$$\begin{aligned} &\min_{\theta,\delta} \text{norm}\left(\Pi B\theta - (a + b\delta)\right) \\ &\qquad s.t. \quad \theta > 0 \\ &\qquad\qquad\;\; \delta \in (0, 1] \end{aligned} \tag{39}$$

where a and b are known vectors coming from the linearization of $\delta^\tau \approx a_\tau + b_\tau\delta$ around $\bar{\delta}$ given in (38) as discussed in Section 4. Given a state price matrix Π and a design matrix B we estimate the θ and δ that best fits the model in a squared error sense.

Once the pricing kernel and discount rate have been recovered, we back out the multi-period physical probabilities as

$$P = D^{-1}\Pi \text{ diag}(B\theta) \tag{40}$$

where D is a diagonal matrix with elements $D_{ii} = \delta^i$ and $\text{diag}(B\theta)$ is a diagonal matrix with elements $\text{diag}(B\theta)_{jj} = B_j\theta$ where B_j is the j'th row of B. We normalize P to have row sums of one, this is necessary since θ and δ are found from the minimization problem in (39) and not solved perfectly.

7.3 Computing Statistics under the Physical Probability Distribution

Once we have recovered the probabilities of each state for each future time period, it is straightforward to compute any statistic under the physical probability distribution. If the level of the index at time t is S_t, then the state space consists of all integer values of the index between the minimum value $(1 - 2.5\mathrm{VIX}_t)S_t$ and $(1 + 4\mathrm{VIX}_t)S_t$. Let N_t denote the number of states as seen from time t and think of state 1 as the lowest state and N_t as the highest state. We compute time t physical expectation of one month returns by summing over the N_t states as

$$E_t^{\mathbb{P}}[r_{t,t+1}] = \sum_{\nu=1}^{N_t} p_{t+1,\nu} r_{t+1,\nu} \tag{41}$$

where $p_{t+1,\nu}$ is the estimated time t conditional physical probability for the transition to state ν at time $t+1$, i.e., in one month. Similarly, $r_{t+1,\nu} = \frac{S_{t+1}(\nu)}{S_t} - 1$ is the return over the period t to $t+1$ if state ν is realized at time $t+1$. Here, $S_{t+1}(\nu)$ is the integer-value of the index at time $t+1$ if state ν is realized.

If $r_{t,t+1}$ is the 1-month return on the index in period t to $t+1$ and $r^f_{t,t+1}$ is the 1-month risk-free rate, then we compute

$$\mu_t = E_t^{\mathbb{P}}[r_{t,t+1}] - r^f_{t,t+1} \tag{42}$$

as the conditional expected 1-month excess return over period t to $t+1$. Furthermore, we let

$$\sigma_t = \sqrt{\mathrm{VAR}_t^{\mathbb{P}}(r_{t,t+1})} \tag{43}$$

be time t conditional 1-month volatility. We compute the contemporaneous unpredictable innovation in the conditional expected return as

$$\Delta\mu_{t+1} = \mu_{t+1} - E_t[\mu_{t+1}] \tag{44}$$

where we impose an AR(1)-process on the innovation to the risk premium $E_t[\mu_{t+1}] = \alpha_0 + \alpha_1\mu_t$ based on the regression

$$\mu_{t+1} = \alpha_0 + \alpha_1\mu_t + \varepsilon_{t+1} \tag{45}$$

The estimated persistence parameter α_1 is 0.31 at the monthly horizon.

8 Empirical Results

As described in Section 7, we are able to recover physical probabilities for each state and each date that we consider, under our given assumptions. We next investigate the quality of these recovered probabilities. We investigate the recovered probabilities by examining their ability to predict the future market return and the future market volatility and by examining the full distribution via a Berkowitz test, cf. Berkowitz (2001).

We first consider the expected return and risk implied by the recovered probabilities. Figure 6 shows how the conditional expected monthly excess returns varies over time. The average conditional expected monthly excess return is 0.32%, that is, 3.86% on an annualized basis, with significant variation over time.

The estimated physical volatility is plotted in Figure 7. It is not surprising that volatilities can be recovered, so we report these as a simple reality check of our method. The recovered volatility looks reasonable and is 97% correlated with the VIX index as seen in Table 1.

Table 1 also shows that the recovered volatility is highly correlated with the SVIX variable of Martin (2015), which in turn is also highly correlated with VIX. Hence, the following tests we only include one of them at a time. We focus on the former two, but our regressions below are qualitatively the same if we control for SVIX.

Lastly, Table 1 shows that the recovered expected return is positively correlated with the future realized return, and more so than any of the other variables. Further, the innovation in the expected return is negatively correlated with the contemporaneous return, consistent with the idea that a higher required return is associated with lower current prices. We next test this predictability more directly.

Table 2 reports the results of regressing the ex post realized excess return on the ex ante recovered expected excess return, μ_t, the ex post innovation in expected return, $\Delta\mu_{t+1}$, and, as controls, the ex ante recovered volatility, σ_t, and the ex ante VIX volatility index:

$$r_{t,t+1} = \beta_0 + \beta_1\mu_t + \beta_2\Delta\mu_{t+1} + \beta_3\sigma_t + \beta_4\mathrm{VIX}_t + \epsilon_{t,t+1} \tag{46}$$

where ϵ_{t+1} is a noise term. To understand this regression, note that we are interested in testing whether the recovered probabilities give rise to reasonable expected returns, that is, time-varying risk premia. For this, we want to test whether a higher ex ante expected return is associated with a higher ex post realized return ($\beta_1 > 0$) and whether an increase in the risk premium is associated with a contemporaneous drop in the price ($\beta_2 < 0$). More specifically, under the null hypothesis of correctly recovered probabilities, the estimates of regression (46) should satisfy

$$\beta_0 = 0 \quad \text{and} \quad \beta_1 = 1 \quad \text{and} \quad \beta_2 < 0 \tag{47}$$

Table 2 reports evidence consistent with this null hypothesis over the full sample from 1/1997 to 7/2014. First, the intercept β_0 is insignificantly different from zero in all specifications. Second, β_1 is positive and marginally significant from 0 in some specifications and never significantly different from 1. The coefficient β_2 is highly significant and has the desired negative sign. Further, as expected the absolute value of β_2 is greater than zero since a shock to the discount rate leads to a larger shock to the price (cf. Gordon's growth model for the extreme example of a permanent shock).

Table 3 reports the result of regression (46) over two sub-samples that have been considered in the literature (e.g., Martin (2015)). Panel A reports regression results for the pre-crisis period (1/1996–8/2008) and Panel B reports the results outside the financial crisis (full sample excluding 9/2008–7/2009). The results are broadly consistent with those over the full sample. All the key parameters have the expected sign, the estimated coefficient β_1 is positive and marginally significant or insignificant, and β_2 is negative and significant.

Following Goyal and Welch (2008) we also compare the out-of-sample predictive ability of our ex-ante expected returns compared to a 'prevailing-mean' estimate. More precisely, letting PM_{t+1} denote the mean of excess returns estimated over 10 years up to time t, we compute

$$R^2 = 1 - \frac{\sum_t \left(r_{t+1} - E_t(r_{t+1})\right)^2}{\sum_t \left(r_{t+1} - PM_{t+1}\right)^2}$$

and

$$\bar{R}^2 = R^2 - (1 - R^2)\left(\frac{T-k}{T-1}\right)$$

find $\bar{R}^2 = 0.0128$ which is positive and large compared to typical values of $\bar{R}^2$, cf. Goyal and Welch (2008). Using the heuristic calculations of Cochrane (1999) this magnitude of R^2 (cf. Kelly and Pruitt (2013)) can be shown to have significant economic implications.

Table 4 reports the results of regressing ex post realized volatility on the ex ante recovered conditional volatility, σ_t, or the VIX volatility index:

$$\sqrt{\text{VAR}(r_{t,t+1})} = \beta_0 + \beta_1 \sigma_t + \beta_2 \text{VIX}_t + \epsilon_{t,t+1} \tag{48}$$

where the realized volatility $\sqrt{\text{VAR}(r_{t,t+1})}$ is computed using close-to-close daily data over the 4 weeks from t to $t+1$ by OptionMetrics. As seen in Table 4, the estimated slope coefficient β_1 is 1.01, close to the predicted value of 1, for the recovered volatility

σ_t. The slope coefficient for VIX is estimated to 0.90. The slope estimate for the recovered volatility is highly significantly different from zero (t-statistic above 16) and not significantly different from one, but the estimated intercept β_0 is marginally significantly different from zero, a rejection of the model. Also, the VIX index has a slightly higher R^2, which may reflect that the recovery method introduces some noise in the volatility measure.

We finish of by testing the full predicted distribution of the recovered probabilities, we consider a so-called Berkowitz test, cf. Berkowitz (2001). Let $\hat{F}_t$ denote the estimated distribution of the excess return r_{t+1} given the information at time t. If the estimated distribution is equal to the true distribution, then the distribution of $u_{t+1} = \hat{F}_t(r_{t+1})$ is uniform and the distribution of $x_{t+1} = \Phi^{-1}(u_{t+1})$ is normal. In the Berkowitz test, we estimate the coefficients in the model $x_{t+1} = c + \beta x_t + \epsilon_t$ and perform a likelihood ratio test of the joint hypothesis that $c = \beta = 0$ and $Var(\epsilon_t) = 1$. We reject the hypothesis that the recovered distribution fully captures the realized excess return distribution. To further shed light on the discrepancy between the estimated and the empirical distribution of excess returns, we compute for each date t the cut-off points for the five quintiles of the estimated conditional distribution of the next period's excess return, and we then record the quintile into which the realized excess return actually fell. Figure 8 plots the number of times the realized excess return, r_{t+1}, ended up in each quintile estimated at time t. If our estimated conditional distribution at time t were accurate, then we would expect excess realized returns to be uniformly distributed across each quintile. Figure 8 suggests that our estimated distribution may put to little mass on very large excess returns.

In summary, we find positive evidence that the recovered probabilities contain information about future expected returns, but we are able to reject that the recovered probabilities provide a perfect description of the future evolution of the market. Our goal has mainly been to illustrate that our general method is easily applicable and can be useful in testing important issues in asset pricing.

9 Conclusion

We characterize when preferences and natural probabilities can be recovered from observed prices using a simple counting argument. We make no assumptions on the physical probability distribution, thus generalizing Ross (2015) who relies on strong time-homogeneity assumptions.

In economies with growth, recovery is generally not feasible as emphasized by Borovicka, Hansen, and Scheinkman (2015). To address this issue, we consider a framework to handle economies with growth, models with more states than time periods, classical multinomial models, models with an infinite state space, and models with non-Markovian behavior. The fundamental assumption is that the pricing kernel can be parameterized by a sufficiently low-dimensional parameter vector which balances the extra information obtained by adding new time periods with the expanding set of unknown state prices. When recovery is feasible, our model is easy to implement, allowing a closed-form linearized solution.

We implement our model empirically, testing the predictive power of the recovered statistics. Our empirical findings indicate that the ex ante expected returns based on our recovered physical probabilities may help predict future returns, but we reject that the full recovered probability distribution is a perfect description of the future market behavior. Future research may further explore the best ways to empirically implement our theory and test its benefits and limitations.

A Appendix: Proofs

Lemma 1 *Suppose that $x^* \in \mathbb{R}^n$ is defined by $f(x^*) = 0$ for a differentiable function $f : \mathbb{R}^n \to \mathbb{R}^n$ with full rank of the Jacobian df in the neighborhood of x^*, and x is defined as the solution to the equation, $f(\bar{x}) + df(\bar{x})(x - \bar{x}) = 0$, where f has been linearized around $\bar{x} = x^* + \Delta x\,\varepsilon$ for $\Delta x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}$. Then $x = x^* + o(\varepsilon)$ for $\varepsilon \to 0$.*

Proof. Since we have $x = \bar{x} - df^{-1}f(\bar{x})$ we see that, as $\varepsilon \to 0$,

$$\frac{x - x^*}{\varepsilon} = \frac{\bar{x} - x^*}{\varepsilon} - df^{-1}\frac{f(\bar{x}) - f(x^*)}{\varepsilon} \to \Delta x - df^{-1}df\,\Delta x = 0 \tag{A.1}$$

■

B Appendix: Details on Recovery in Mehra-Prescott

Let

$$\Pi = \begin{bmatrix} \pi_{0,1}^{0,d} & \pi_{0,1}^{1,u} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \pi_{0,2}^{0,d} & \pi_{0,2}^{1,d} & \pi_{0,2}^{1,u} & \pi_{0,2}^{2,u} & 0 & \dots & 0 & 0 & 0 & & 0\dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \pi_{0,T}^{0,d} & \pi_{0,T}^{1,d} & \pi_{0,T}^{1,u} & \dots & \pi_{0,T}^{T,u} \end{bmatrix} \tag{B.1}$$

where $\pi_{0,t}^{k,u}$ is the state price of making a total of k "up" moves in t periods where the last move was "up," that is, the Arrow-Debreu price for the state $s_t = (y_t, x_t) = (u^k d^{t-k}, u)$. Similarly, $\pi_{0,t}^{k,d}$ is the state price of making a total of k "up" moves in t periods where the last move was "down".

Π has dimension $T \times (\sum_{t=1}^{T} 2t)$. This implies that the $h^{-1}(\gamma)$ vector of inverse marginal utility ratios must be $(\sum_{t=1}^{T} 2t)$-dimensional. We fix this in the following way. We let

$$h^{-1}(\gamma) = \Big[(y_1^0)^\gamma \quad (y_1^1)^\gamma \quad (y_2^0)^\gamma \quad (y_2^1)^\gamma \quad (y_2^1)^\gamma \quad (y_2^2)^\gamma \quad \dots \quad (y_T^T)^\gamma\Big]' \tag{B.2}$$

where $y_t^k = u^k d^{t-k}$ is the level of aggregate consumption when making a total of k "up" moves in t periods and γ is the risk-aversion parameter that we wish to recover.

There is no closed-form solution to the non-linear case of CRRA preferences. In order to obtain model estimates we sort to a numerical exercise, that is to minimize the objective function g:

$$\min_{\gamma,\delta} g(\gamma,\delta) := \text{norm}\left(\Pi h^{-1}(\gamma) - \begin{bmatrix} \delta \\ \delta^2 \\ \vdots \\ \delta^T \end{bmatrix}\right) \tag{B.3}$$
$$\begin{aligned} s.t. \quad & \gamma \in \mathbb{R}_+ \\ & \delta \in (0,1] \end{aligned}$$

Based on the recovered (γ, δ) that solve this minimizition problem, we can recover the natural probabilities from (36).

C Appendix: Computing State Prices Empirically

Before we can recover probabilities, we need to know that Arrow-Debreu prices or, said differently, characterize the risk-neutral distribution. There exist many ways to do this in practice based on observed option prices, including various interpolation methods. To ensure that we start with an arbitrage-free collection of Arrow-Debreu prices by strike and maturity, we use the model of Bates (2000) to derive state prices from observed option prices. This parametric approach puts structure on the tails of the risk-neutral density, which also allows us to extrapolate outside the range of observable option quotes. While the Bates (2000) model may not be the "true" specification of the economy, we simply use this framework as a standard method in the literature to compute state prices, and, consistent with this pragmatic view, we allow parameters to change over time (which also avoids look-ahead bias).

In this model, the risk-neutral process for the price of the underlying asset, S_t, and the instantaneous variance, V_t, are assumed to be of the form

$$dS_t/S_t = (r^f - d - \lambda\bar{k})dt + \sqrt{V_t}dZ_t + kdq_t \tag{C.1}$$

$$dV_t = (\alpha - \beta V_t)dt + \sigma_v\sqrt{V_t}dZ_{vt} \tag{C.2}$$

where Z_t and Z_{vt} are Brownian motions with correlation ρ, and q_t is a Poisson counting process that captures the risk of jumps in the price. The jumps occur with intensity λ and each jump causes the price to be multiplied by the factor $1 + k$, which is lognormally distributed, i.e., $\ln(1 + k) \sim N(\ln(1 + \bar{k})\frac{1}{2}\delta^2, \delta^2)$. Further, r^f is the risk-free rate and d is the dividend yield.

We calibrate these model parameters every fourth Wednesday as follows:[10] On each day, given the current level of the market S_t and the risk-free term structure $r^f_{t,t+\tau}$, we find the model parameters $(\alpha, \beta, \lambda, \bar{k}, \sigma_v, \delta)$ and state variable V_t that minimize the vega-weighted squared pricing errors for fifty call and put options, following the methodology of Trolle and Schwartz (2009). The fifty chosen call/put options are those with the highest volumes. We allow the model parameters to vary over time since we simply use the model to smooth observed option prices (that may be noisy) such that they are arbitrage-free.

Once we have obtained model estimates, we compute the risk-neutral density $f(\tau, S_\tau)$ for any time τ periods into the future and state S_τ given the current time state S_t as:

$$f(\tau; S_\tau) = \frac{1}{\pi}\int_0^\infty \left(\frac{S_\tau}{S_t}\right)^{-iu}\psi(\tau, u)du \tag{C.3}$$

[10] We use data for every fourth Wednesday as a compromise between (i) the tradition in the asset pricing literature on return predictability of focusing on monthly returns, and (ii) the tradition in the option literature of focusing on Wednesdays, where among other reasons option liquidity is high.

that is, by integrating the characteristic function ψ numerically using the Gauss-Laguerre quadrature method. Knowing the risk-neutral density, the corresponding state price density $\pi(\tau; S_T)$ is the density discounted by the τ-period risk-free rate $r^f_{t,t+\tau}$:

$$\pi(\tau; S_\tau) = e^{-r^f_{t,t+\tau}} f(T; S_\tau) \tag{C.4}$$

This completes the computation of state prices. Indeed, we think of $\pi(\tau; S_\tau)$ as the Arrow-Debreu prices we need as starting point for our recovery for each index level. For example $\pi(1, 2000)$ is the Arrow-Debreu price of receiving \$1 in one year of the S&P500 is between 2000 and 2001. We consider the grid of maturities and index levels described in Section 7.2.

References

Ait-Sahalia, Yacine, and Andrew W. Lo, 2000, Nonparametric risk management and implied risk aversion, *Journal of Econometrics* 94, 9–51.

Audrino, Francesco, Robert Huitema, and Markus Ludwig, 2014, An empirical analysis of the Ross recovery theorem, *Working paper.*

Bakshi, Gurdip, Fousseni Chabi-Yo, and Xiaohui Gao, 2015, A recovery that we can trust? deducing and testing the restrictions of the recovery theorem, *Working paper.*

Barone-Adesi, Giovanni, Robert F. Engle, and Loriano Mancini, 2008, A garch option pricing model with filtered historical simulation, *Review of Financial Studies* 21(3), 1223–1258.

Bates, D. S., 2000, Post 87 crash fears in the S&P 500 futures options market, *Journal of Econometrics* 94, 181–238.

Berkowitz, Jeremy, 2001, Testing density forecats, with applications to risk management, *Journal of Business and Economic Statistics* 19, 465–474.

Black, F., and M. S. Scholes, 1973, The pricing of options and corporate liabilities, *The Journal of Political Economy* 81, 637–654.

Bliss, Robert R., and Nikolaos Panigirtzoglou, 2004, Option implied risk aversion estimates, *Journal of Finance* 1, 407–446.

Bollerslev, Tim, and Viktor Todorov, 2011, Tails, fears, and risk premia, *Journal of Finance* 66, 2165–2211.

Borovicka, Jaroslav, Lars Peter Hansen, and Jose A. Scheinkman, 2015, Misspecified recovery, *Journal of Finance* forthcomming.

Carr, Peter, and Jiming Yu, 2012, Risk, return, and Ross recovery, *Working paper.*

Christoffersen, Peter, Steven L. Heston, and Kris Jacobs, 2013, Capturing option anomalities with a variance-dependent pricing kernel, *Review of Financial Studies* 26, 1963–2006.

Cochrane, John H., 1999, New facts in finance, *Economic Perspectives* 23, 3658.

Cox, John C, Stephen A Ross, and Mark Rubinstein, 1979, Option pricing: A simplified approach, *Journal of Financial Economics* 7, 229–263.

Debreu, Gerard, 1970, Economies with a finite set of equilibria, *Econometrica* 38, 387–392.

Goyal, Amit, and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Hansen, Lars Peter, and Jose A. Scheinkman, 2009, Long term risk: An operator approach, *Econometrica* 77, 177–234.

Jackwerth, J., 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* 13, 433–451.

Jackwerth, Jens C., and Mark Rubinstein, 1996, Recovering probability distributions from option prices, *Journal of Finance* 51, 1611–1631.

Kelly, Bryan, and Seth Pruitt, 2013, Market expectations in the cross-section of expected values, *Journal of Finance* 68, 1721–1756.

Linetsky, Vadim, and Likuan Qin, 2015, Positive eigenfunctions of markovian pricing operators: Hansen-Scheinkman factorization, Ross recovery and long-term pricing, *Working paper.*

Linn, Matthew, Sophie Shive, and Tyler Shumway, 2015, Pricing kernel monotonicity and conditional information, *Working paper.*

Malamud, Semyon, 2016, Noisy recovery, *Working paper.*

Martin, Ian, 2015, What is the expected return on the market?, *Working paper*.

Martin, Ian, and Steve Ross, 2013, The long bond, *Working paper*.

Mehra, Rajnish, and Edward C Prescott, 1985, The equity premium: A puzzle, *Journal of monetary Economics* 15, 145–161.

Rosenberg, Joshua V., and Robert F. Engle, 2002, Empirical pricing kernels, *Journal of Financial Economics* 64, 341–372.

Ross, Steve, 2015, The recovery theorem, *The Journal of Finance* 70, 615–648.

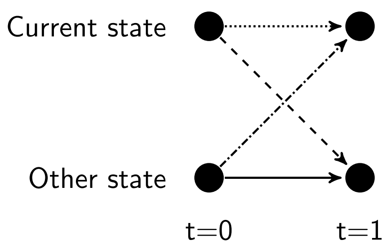
Schneider, Paul, and Fabio Trojani, 2015, (almost) model-free recovery, *Working paper*.

Trolle, Anders B., and Eduardo S. Schwartz, 2009, Unspanned stochastic volatility and the pricing of commodity derivatives, *Review of Financial Studies* 22, 4423–4461.

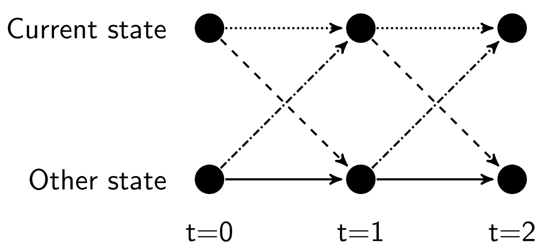
Walden, Johan, 2013, Recovery with unbounded diffusion processes, *Working paper*.

D Tables and Figures

Panel A. Ross's Recovery Theorem: one period, two "parallel universes"



Panel B. Ross's Recovery Theorem: time-homogeneous dynamic setting



Panel C. Our Generalized Recovery: No assumptions about probabilities

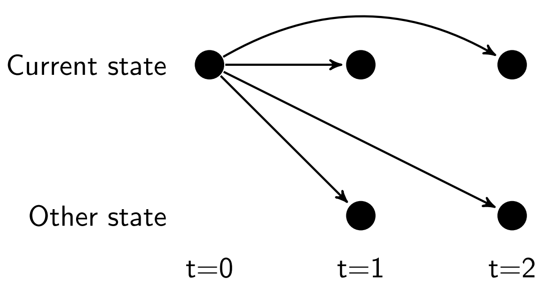
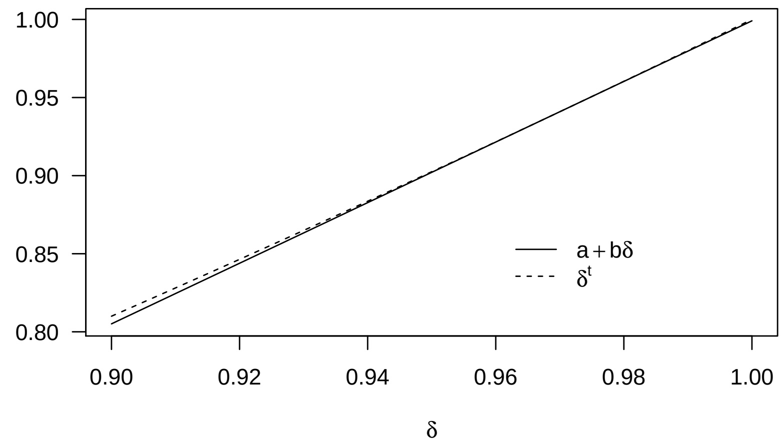
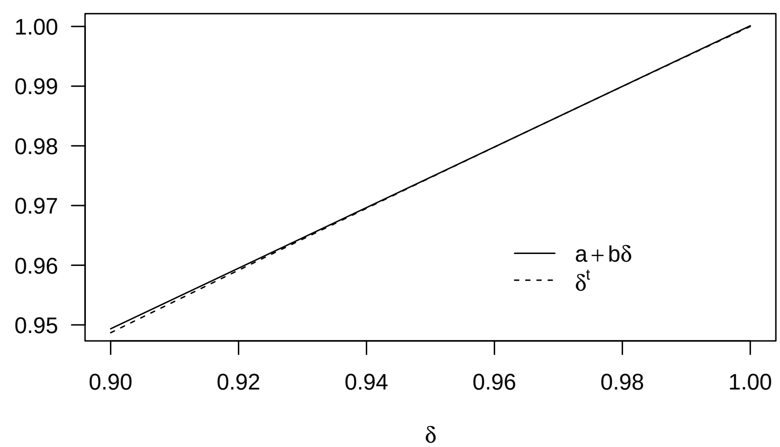


Figure 1: **Generalized Recovery Framework.** Panel A illustrates the idea behind Ross's Recovery Theorem, namely that we start with information about all Arrow-Debreu prices in *all* initial states (not just the state we are currently in, but also prices in "parallel universes" where today's state is different). Panel B shows how Ross moves to a dynamic setting by assuming time-homogeneity, that is, assuming that the prices and probabilities are the same for the two dotted lines, and so on for each of the other pairs of lines. Panel C illustrates our Generalized Recovery method, where we make no assumptions about the probabilities.



Panel A: t = 2 years

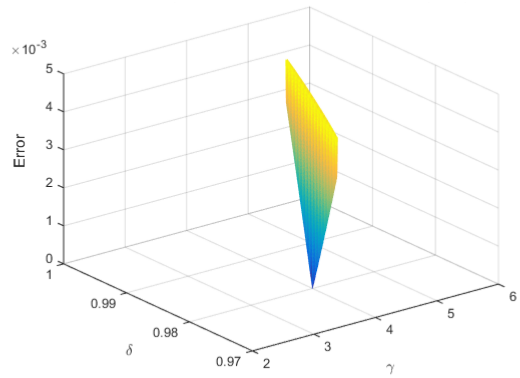


Panel B: t = 0.5 years

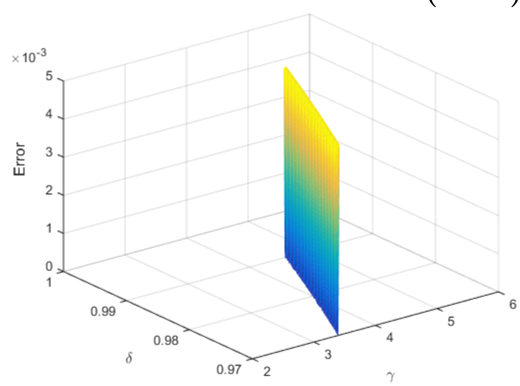
Figure 2: **Closed-Form Solution: Approximation Error.** The figure shows that the generalized recovery problem is very close to being linear. We show that the only non-linearity comes from the discount rate δ due to the powers of time, δ^t. However, the function $\delta \to \delta^t$ is very close to being linear for the relevant range of annual discount rates, say $\delta \in [0.94, 1]$, and the relevant time periods that we study. Panel A plots the discount function and the linear approximation around $\delta_0 = 0.97$ given a horizon of $t = 2$ years. Panel B plots the same for a horizon of a half year.

Table 1: **Correlation Matrix.** This tables shows the pairwise correlations between the recovered conditional expected excess return, μ_t, the recovered conditional volatility, σ_t, the VIX_t index, the lower boundary on the equity premium, SVIX_t, due to Martin (2015), the ex post innovation in the expected return, $\Delta\mu_{t+1}$, and the ex post realized excess return, r_{t+1}.

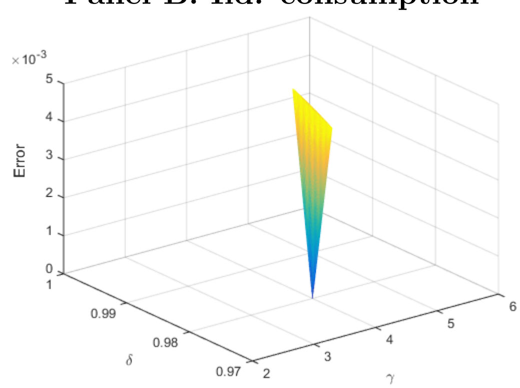
	μ_t	σ_t	VIX_t	SVIX_t	$\Delta\mu_{t+1}$	r_{t+1}
μ_t	1	0.563	0.543	0.512	0	0.118
σ_t		1	0.971	0.937	0.224	0.032
VIX_t			1	0.963	0.249	0.023
SVIX_t				1	0.209	0.003
$\Delta\mu_{t+1}$					1	-0.394
r_{t+1}						1



Panel A: Mehra Prescott (1985)

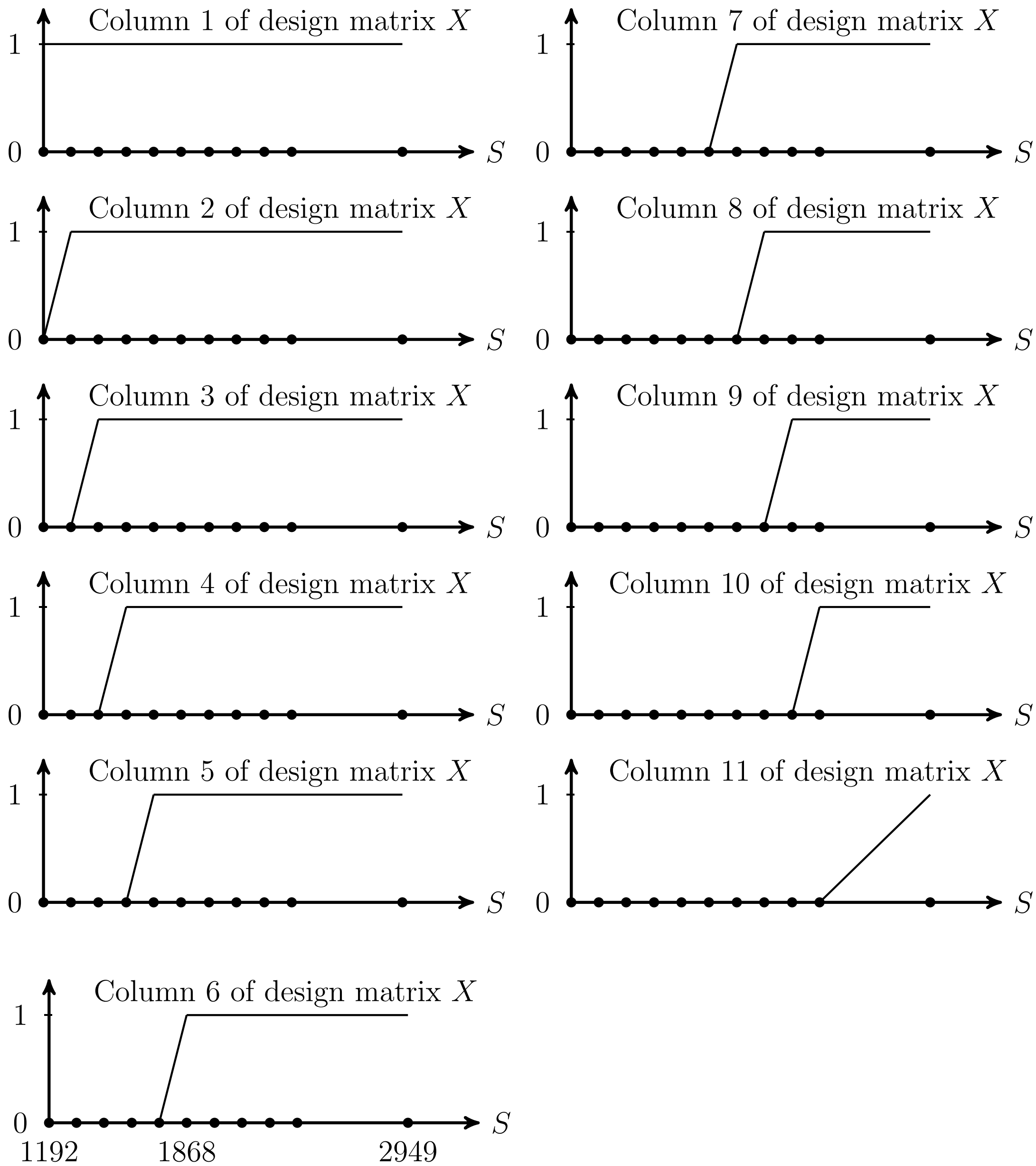


Panel B: Iid. consumption



Panel C: Non-Markovian

Figure 3: **Generalized Recovery: Objective Function in Specific Economic Models.** This figure shows the objective function used for the generalized recovery method, the squared pricing errors in (B.3). Panel A shows that the objective function for the Mehra Prescott (1985) model has a unique minimum, making the generalized recovery feasible. Panel B shows that generalized recovery is not feasible in the Black-Scholes-Merton model with iid. consumption as the objective has a continuum of solutions. Panel C shows that generalized recovery is feasible in the non-Markovian model.

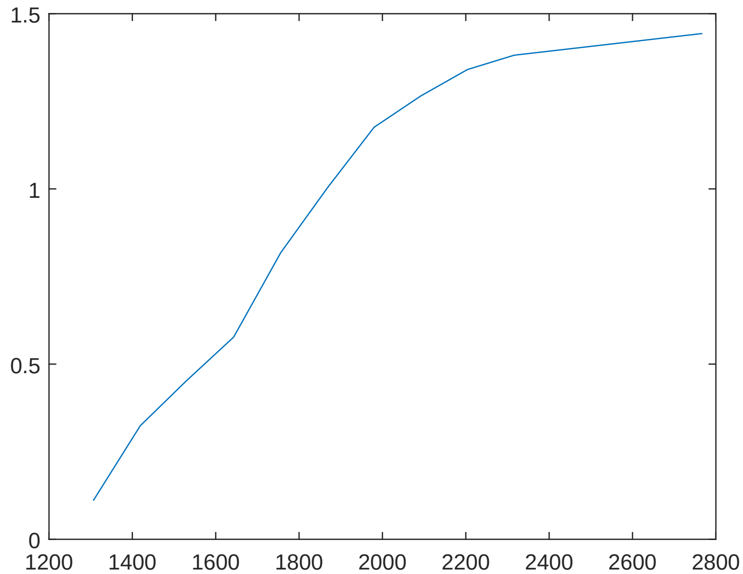


Panel A: **Design Matrix: Illustration.**

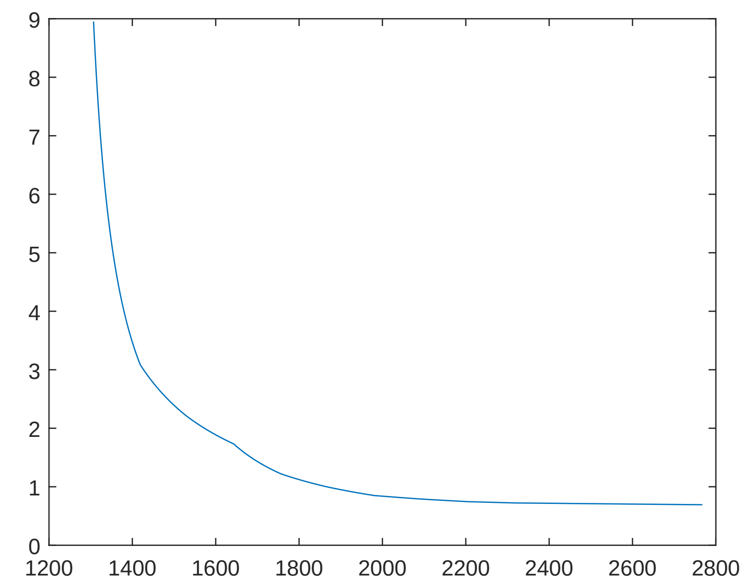
$$B = \begin{bmatrix} 1 & 0 & ... & 0 & 0 \\ 1 & \frac{1}{n_2} & ... & 0 & 0 \\ 1 & \frac{2}{n_2} & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{n_2-1}{n_2} & ... & 0 & 0 \\ 1 & 1 & ... & 0 & 0 \\ 1 & 1 & ... & 0 & 0 \\ 1 & 1 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & ... & 0 & 0 \\ 1 & 1 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & ... & 0 & 0 \\ 1 & 1 & ... & \frac{1}{n_{10}} & 0 \\ 1 & 1 & ... & \frac{2}{n_{10}} & 0 \\ 1 & 1 & ... & \vdots & \vdots \\ 1 & 1 & ... & \frac{n_{10}-1}{n_{10}} & 0 \\ 1 & 1 & ... & 1 & 0 \\ 1 & 1 & ... & 1 & \frac{1}{n_{11}} \\ 1 & 1 & ... & 1 & \frac{2}{n_{11}} \\ 1 & 1 & ... & 1 & \vdots \\ 1 & 1 & ... & 1 & \frac{n_{11}-1}{n_{11}} \\ 1 & 1 & ... & 1 & 1 \end{bmatrix}$$

Panel B: **Design Matrix: Equation.**

Figure 4: **Design Matrix.** This figure illustrates the design matrix B used to span the inverse pricing kernel as a function of the state, which is the level of the S&P500. Panel A depicts the columns of the design matrix graphically while Panel B shows the matrix mathematically. We see that each column is piece-wise linear and increasing, ensuring that the inverse pricing kernel enherits the same properties. The first column controls the level, and each of the next columns control the slope of the successive line segments.



Panel A: **Inverse Pricing Kernel is Piecewise Linear.**



Panel B: **Pricing Kernel has Piecewise Constant Curvature.**

Figure 5: **Pricing Kernel.** Panel A shows the estimated inverse pricing kernel on March 12, 2014. We note that it consists of ten linear pieces governed by the columns of the design matrix shown in Figure 4 weighted by the estimated parameters θ. Panel B shows the corresponding pricing kernel, which has piece-wise constant curvature.

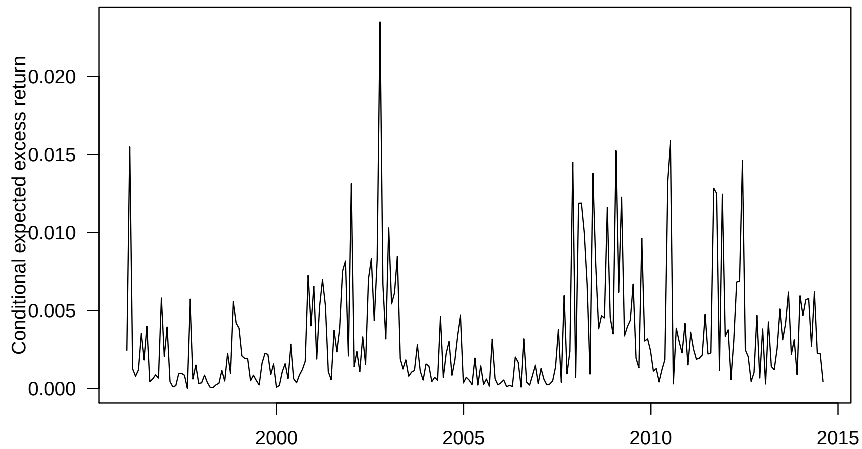


Figure 6: **Recovered conditional expected excess return.** The figure plots monthly conditional expected excess market returns, recovered last Wednesday of each month from 1/1996 to 7/2014.

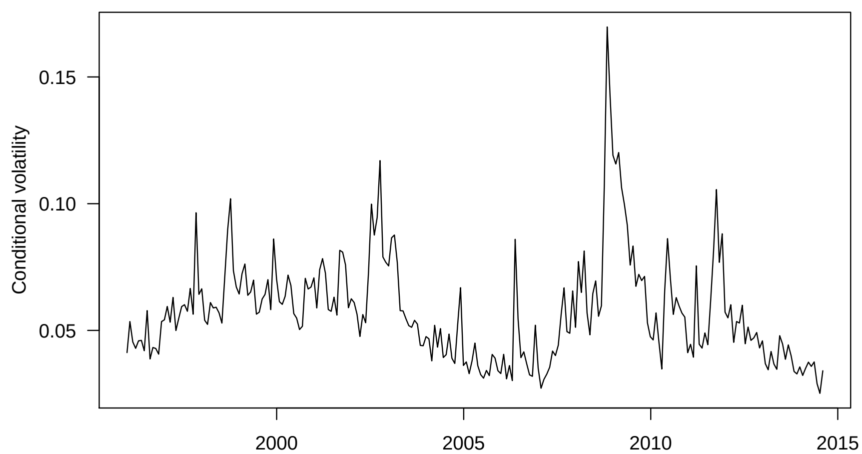


Figure 7: **Recovered conditional volatility of excess return.** The figure plots monthly conditional market volatility, recovered last Wednesday of each month from 1/1996 to 7/2014.

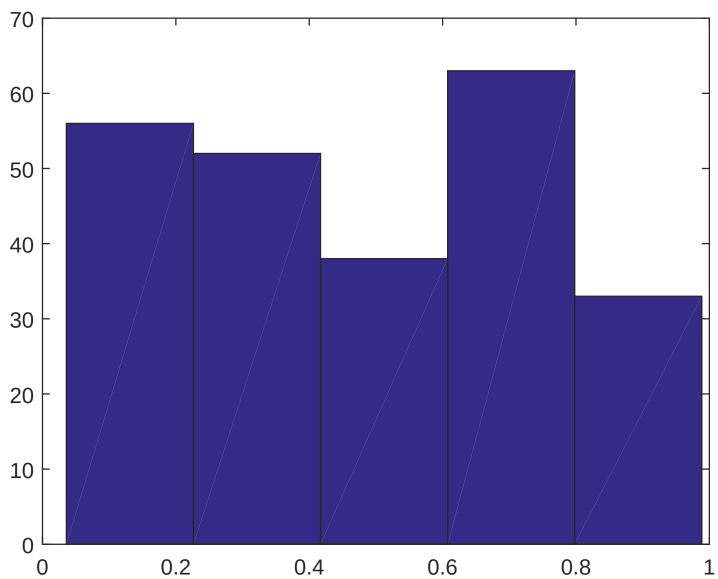


Figure 8: **Frequency histogram.** The figure plots the frequency at which the realized excess return, r_{t+1}, ended up in the estimated time t conditional excess return distribution bucket. The null hypothesis is that each bar has equal height, but based on a Berkowitz test we can rejects the hypothesis that the recovered distribution fully captures the realized excess return distribution.

Table 2: **Does the Recovered Expected Return Predict the Future Return?** This table reports results of the regression of the ex post realized excess return r_{t+1} on the ex ante recovered expected excess return, μ_t, the ex post innovation in expected return, $\Delta\mu_{t+1}$, the ex ante recovered volatility, σ_t, and ex ante the VIX volatility index:

$$r_{t,t+1} = \beta_0 + \beta_1\mu_t + \beta_2\Delta\mu_{t+1} + \beta_3\sigma_t + \beta_4\mathrm{VIX}_t + \epsilon_{t,t+1}$$

The regression uses monthly data over the full sample 1/1997–6/2014 , t-statistics are reported in parentheses, and significance is indicated as * for $p < 0.1$, ** for $p < 0.05$, and *** for $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-0.00	0.00	0.00	0.00	-0.00	0.00	0.01	-0.00
	(-0.30)	(1.33)	(0.28)	(0.39)	(-0.28)	(0.70)	(0.75)	(-0.59)
μ_t	1.44*				1.41*	1.87**	1.89**	1.18
	(1.82)				(1.95)	(2.03)	(2.08)	(1.39)
$\Delta\mu_{t+1}$		-5.09***			-5.09***			-5.18***
		(-6.64)			(-6.67)			(-6.60)
σ_t			0.01			-0.04		0.02
			(0.13)			(-0.92)		(0.52)
VIX_t				0.00			-0.04	
				(0.03)			(-1.00)	
Adj. R^2 (%)	0.95	15.16	-0.40	-0.42	16.14	0.89	0.95	15.88

Table 3: **Does the Recovered Expected Return Predict the Future Return in Sub-Samples?** This table reports the result of the same regressions as in Table 2 over two sub-samples. Panel A reports regression results for the pre crisis period (1/1996–8/2008) and Panel B reports the results outside the financial crisis (full sample excluding 9/2008–1/2010). The tables report t-statistics in parentheses and significance is indicated as * for $p < 0.1$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Pre crisis (1/1996–8/2008)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-0.00	0.00	-0.01	-0.01	-0.00	-0.01	-0.01	-0.02*
	(-0.58)	(0.13)	(-1.12)	(-1.05)	(-0.52)	(-0.86)	(-0.84)	(-1.84)
μ_t	1.96*				1.51*	1.11	1.18	0.47
	(1.96)				(1.66)	(0.97)	(1.06)	(0.44)
$\Delta\mu_{t+1}$		-4.53***			-4.66***			-5.00***
		(-4.46)			(-4.82)			(-5.10)
σ_t			0.08			0.05		0.11*
			(1.36)			(0.74)		(1.77)
VIX_t				0.07			0.04	
				(1.31)			(0.71)	
Adj. R^2 (%)	1.74	10.68	0.54	0.45	12.88	0.74	0.71	14.04

Panel B: Full sample except the financial crisis (1/1997–6/2014 excluding 9/2008–7/2009)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-0.00	0.00*	-0.01	-0.01	-0.00	-0.01	-0.01	-0.02*
	(-0.22)	(1.65)	(-1.26)	(-1.18)	(-0.42)	(-0.78)	(-0.77)	(-1.97)
μ_t	1.94**				1.90***	1.58*	1.60*	1.08
	(2.47)				(2.61)	(1.74)	(1.79)	(1.29)
$\Delta\mu_{t+1}$		-4.83***			-4.82***			-5.08***
		(-6.26)			(-6.33)			(-6.62)
σ_t			0.09*			0.04		0.10*
			(1.82)			(0.73)		(1.94)
VIX_t				0.08*			0.04	
				(1.78)			(0.72)	
Adj. R^2 (%)	2.20	14.30	1.01	0.93	16.43	1.89	1.88	17.44

Table 4: **Does the Recovered Volatility Predict the Future Volatility?** This table reports results of a monthly regression of the ex post realized volatility on the ex ante recovered return volatility, σ_t, and the VIX volatility index:

$$\sqrt{\mathrm{VAR}(r_{t,t+1})} = \beta_0 + \beta_1\sigma_t + \beta_2\mathrm{VIX}_t + \epsilon_{t,t+1}$$

using the full sample 1/1997–6/2014. The realized volatility in computed using close-to-close daily data over the month by OptionMetrics. The table reports t-statistics in parentheses and significance is indicated as * for $p < 0.1$, ** for $p < 0.05$, and *** for $p < 0.01$.

	(1)	(2)
Intercept	-0.03**	-0.02
	(-2.27)	(-1.63)
σ_t	1.01***	
	(16.69)	
VIX_t		0.90***
		(17.27)
Adj. R^2 (%)	53.43	55.14