

# Generalized Recovery\*

Christian Skov Jensen, David Lando, and Lasse Heje Pedersen<sup>†</sup>

First version March, 2015. This version December 14, 2015

*Preliminary*

## Abstract

We characterize when physical probabilities, marginal utilities, and the discount rate can be recovered from observed state prices for several future time periods. Our characterization makes no assumptions of the probability distribution, thus generalizing the time-homogeneous stationary model of Ross (2015). Our characterization is simple and intuitive, linking recovery to the relation between the number of time periods on the number of states. 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, crash risk, and other recovered statistics.

---

\*We are grateful for helpful comments from Jens Jackwerth, Horatio Cuesdeanu, Peter Christoffersen, Chayawat Ornthanalai, Pia Mølgaard as well as from seminar participants at University of Konstanz, University of Toronto, HEC-McGill winter finance workshop 2015, and Copenhagen Business School, and especially to Dan Petersen for guiding us to Sard's theorem. All three authors gratefully acknowledges support from the FRIC Center for Financial Frictions (grant no. DNRF102) and Lasse Heje Pedersen from the European Research Council (ERC grant no. 312417).

<sup>†</sup>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](http://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).

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 need not be time homogeneous. We show when recovery is possible and when it is not using a simple “counting” argument (based on Sard’s Theorem). 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 the 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. Said simply, 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 distribu-

tion 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, 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  $T \geq S$ .

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.

Our result echoes Ross's result under more general conditions. We show that, when Ross' time homogeneity conditions are met, then our solution naturally is the same as his. 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 critique of Borovicka, Hansen, and Scheinkman (2015) in a simpler way, consider what happens if the economy evolves in a 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, we see that recovery is impossible because of the number of states must be higher than the number of time periods ( $S > T$ ).

Nevertheless, we show that our method can also be useful when the state space is unbounded and the probability distribution is unrestricted. 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  $N$  is no greater than the number of time periods minus one. Further, we show that the method also works for path-dependent preferences such as habit formation and Epstein-Zin utility.

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. Each week, we estimate state price densities for multiple future horizons and apply our closed-form method to recover probabilities and preferences. Based on the recovered probabilities, we derive the risk and expected return over the future month, quarter, and year, among other statistics from the physical distribution of returns. Our preliminary results suggest that the recovered statistics may have predictive power for the distribution of future realized returns, although we caution that these results are preliminary and subject to revision.

The literature on recovery is old and 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. 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' use of Perron-Frobenius 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. 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, Jacobs, and Heston (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.

The remainder of the paper is structured as follows. Section 2 briefly reviews Ross' 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 describes our data and empirical methodology and Section 7 provides our empirical results.

## 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 when 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  $1, \dots, S$ . 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  $i, j = 1, \dots, S$

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

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 a bit of additional structure on the pricing kernel, we can recover physical probabilities. Indeed, there exists a unique set of physical probabilities  $p^{i,j}$  for all  $i, j = 1, \dots, S$  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  $u = (u_1, \dots, u_S)$  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  $\delta$  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' 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  $s = 1, \dots, S$ , and  $\pi_{t,t+\tau}^{i,j}$  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

$$\pi_{t,t+\tau}^{i,j} = p_{t,t+\tau}^{i,j} m_{t,t+\tau}^{i,j} \tag{3}$$

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  $\tau$ , the pricing kernel can be written as*

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

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  $\tau > 0$ ,  $p_{i,t+\tau}^{i,j}$  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 with preference specifications such as habit formation or Epstein-Zin utility.

### 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 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:

$$\begin{aligned}
 \pi_{t,t+1}^{1,1} &= p_{t,t+1}^{1,1} m_{t,t+1}^{1,1} \\
 \pi_{t,t+1}^{1,2} &= (1 - p_{t,t+1}^{1,1}) m_{t,t+1}^{1,2} \\
 \pi_{t,t+2}^{1,1} &= p_{t,t+2}^{1,1} m_{t,t+2}^{1,1} \\
 \pi_{t,t+2}^{1,2} &= \underbrace{(1 - p_{t,t+2}^{1,1})}_{2 \text{ unknowns}} \underbrace{m_{t,t+2}^{1,2}}_{4 \text{ unknowns}}
 \end{aligned} \tag{5}$$

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 marginal utilities:

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

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

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

This system now has 4 equations with 4 unknowns, so there is hope to recover the physical probabilities  $p$ , the discount rate  $\delta$ , 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

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

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

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

Hence, we can solve the 2-state model by (i) finding  $\delta$  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  $\tau$  if the realized state

is  $s$  and collect the set of observed state prices in a  $T \times S$  matrix  $\Pi$  defined as

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

Similarly, letting  $p_{\tau s}$  denote the physical transition probabilities of going from the current state 1 to state  $s$  in  $\tau$  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  $\tau$  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

$$\Pi = DPH \quad (11)$$

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

$$P = D^{-1}\Pi H^{-1} \quad (12)$$

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

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

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

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

Here,  $\Pi_1$  is a column vector of dimension  $T$ , where the first  $S - 1$  elements are denoted by  $\Pi_{11}$  and the rest of the vector is denoted  $\Pi_{21}$ . Similarly,  $\Pi_2$  is a  $T \times (S - 1)$  matrix, where the first  $S - 1$  rows are called  $\Pi_{12}$  and the last rows are called  $\Pi_{22}$ . With this notation and the fact that  $H(1, 1) = h(1) = 1$ , we can write (13) as

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

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

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

with solution<sup>1</sup>

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

Hence, if  $\delta$  were known, we would be done! And, by the way,  $\delta$  is “almost known” since the discount rate is the least mysterious parameter. We later use this insight to derive a closed-form approximation which is accurate as long as we have a sense of the ballpark of  $\delta$ , here we proceed for general unknown  $\delta$ .

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

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

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 polynomial in  $\delta$  of degree at most  $T$ . If  $T = S$ , then  $\delta$  is a root to a single polynomial so at most a finite number of solutions exist. If  $T > S$ , then typically no solution exists for general Arrow-Debreu prices  $\Pi$  since  $\delta$  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  $\Pi$  has been “generated by the model” if there exist  $\delta$ ,  $P$ , and  $H$  such that  $\Pi$  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*

---

<sup>1</sup>Of course, to invert  $\Pi_{12}$  it must have full rank. As long as  $\Pi_2$  has full rank, we can re-order the rows to ensure that  $\Pi_{12}$  also has full rank.

recovery problem has

1. a continuum of solutions if  $S > T$ ;
2. at most  $S$  solutions if the submatrix  $\Pi_2$  has full rank and  $S = T$ ;
3. no solution generically in terms of an arbitrary  $\Pi$  matrix of positive elements and  $S < T$ ;
4. a unique solution generically if  $\Pi$  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  $(\delta, h, P)$  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,  $D = \text{diag}(\delta, \dots, \delta^T)$  and  $H = \text{diag}(1, h_2, \dots, h_S)$ , and  $F$  is  $C^\infty$ . If  $S < T$ , the image  $F(X)$  has Lebesgue measure zero in  $R^{T \times S}$  by Sard's theorem, proving 3.

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

For two different choices of the discount rate  $(\delta_1, \delta_2)$  and a single set of prices  $\Pi$ , we consider the triplet  $(\delta_1, \delta_2, \Pi)$ . 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  $\Pi$  is of full rank and both discount rates are consistent with the prices, that is, there exists (unique)  $P_i$  and  $H_i$  ( $i=1,2$ ) such that  $D_1 P_1 H_1 = D_2 P_2 H_2 = \Pi$ .

Given that probabilities and marginal utilities can be uniquely recovered from prices and a discount rate (as explained above), we have a smooth map from  $M$  to  $X$  by mapping any triplet  $(\delta_1, \delta_2, \Pi)$  to  $(\delta_1, h_1, P_1)$ , where  $(h_1, P_1)$  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 small enough.

To study the dimension of  $M$ , we note that we can think of  $M$  as the space of triplets such that the span of  $\Pi$  contains both the points  $(\delta_1, \delta_1^2, \dots, \delta_1^T)'$  and  $(\delta_2, \delta_2^2, \dots, \delta_2^T)'$ . The span of  $\Pi$  is given by  $V_\Pi := \{\Pi \cdot (1, h_2, h_3, \dots, h_S)' | h_s > 0\}$ , which is an affine  $(S-1)$ -dimensional subspace of  $\mathbb{R}^T$  for  $\Pi$  of full rank. The set of all those  $\Pi \in \mathbb{R}^{T \times S}$  such that  $V_\Pi$  passes through 2 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  $(\delta_1, \delta_2)$  depends on 2 parameters, and, for a given pair, there is a  $(ST - 2T + 2S - 2)$ -dimensional subspace of possible  $\Pi$  (any two distinct points are always in general position). ■

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  $S$ .

Proposition 1 also helps us understand the intuition behind the critique of Borovicka, Hansen, and Scheinkman (2015) with 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). If Assumption 2 also applies (time homogeneity) then a solution to Ross' Recovery problem produces a solution to our generalized recovery problem as well; a solution to the generalized recovery problem is not in general a solution to Ross's recovery problem without Assumption 2.*

**Proof.** Let  $\bar{P}$  denote an  $S \times S$  matrix of one-period state prices as considered in Ross (2015), i.e.,  $\bar{p}_{ij}$  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

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

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

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

If the starting state is  $i$ , then the equations of our generalized recovery problem are the subset obtained by considering the  $i$ 'th 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  $i$ 'th row of  $F^k$ , we have a solution to the equations for our generalized recovery problem. Naturally, Ross's solution based on



time-homogeneity cannot work if the true probabilities are not time-homogeneous as we show by specific examples below. ■

To see that our problem is, in fact, a strict generalization, consider first the case where the number of time-periods  $T$  is strictly larger than the number of states. In Ross' setting, the unique solution can be found from the first  $S$  time periods, and all subsequent periods add new equations but no new unknowns. Indeed, Ross's assumption of time-homogeneity allows one to compute all transition probabilities from the transition matrix determined on the basis of the first  $S$  equations. Therefore, Ross's framework imposes significant over-identifying assumptions that are likely to be violated empirically. In contrast, our setting allows any probability distribution at any time period, so we add  $S - 1$  new unknowns for each time-period we add (recall that probabilities must add up to one). In other words, our framework allows much more flexibility.

Our framework is also more general in the the case with the same number  $S$  of time periods and states. Seeing this is a bit more tricky. Recalling that  $p_{\tau i}$  denotes the probability of going from the current state 1 to state  $i$  in  $\tau$  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. Needless to say, the bound is much stricter once we account for other possible paths as well. However, such a bound may not apply for the true probabilities if the real world is not time-homogeneous. Indeed, the probability of going from state 1 to state 2 tomorrow could be very different from the current probability of this transition and, therefore, the probability of going from state 1 to 2 in two periods could be anything if the world is not homogeneous. In other words, the true probabilities may have  $p_{22} < p_{11}p_{12}$ , in which case Ross's model can never recover the true probabilities. In contrast, our model allows any probability distributions, so we can still achieve correct recovery in

this case. Actually, violations of time-homogeneity in observed data is an important reason why implementing Ross's method is difficult.<sup>2</sup>

We finally note that, as in the time-homogeneous-case considered by Ross, 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.

To see this result, note that the price of a zero coupon bond with maturity  $\tau$  is equal to the sum of the  $\tau$ 'th row of  $\Pi$ , 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

$$\frac{1}{(1+r)^\tau} = (\Pi e)_\tau \tag{21}$$

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  $\tau = 1, \dots, T$ . 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  $\Pi$  of state prices as seen from a

---

<sup>2</sup>More formally, a solution to our recovery problem (when  $S = T$  and the starting state is state 1) consists of  $S$  probability vectors  $(p_{\tau 1}, \dots, p_{\tau S})$ , where  $\tau$  runs from 1 to  $S$ . Our lower bound argument shows that this solution cannot in general be obtained from a time-homogeneous solution. Indeed, let  $\bar{P}$  be a Ross probability matrix where  $p_{ij}$  is the probability of a one-period transition from state  $i$  to state  $j$ . The first row must be  $(p_{1,1}, \dots, p_{1,S})$ . For any choice of the other rows, the first row in each of the matrices  $(\bar{P})^\tau$  where  $\tau$  runs from 1 to  $S$  cannot coincide with our  $P$ .

particular starting state can be written as

$$\Pi = RQ$$

Clearly, by letting  $\delta = r$  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 pricing kernel  $\Pi$  with low rank, which implies that a continuum of solutions may exist and the representative agent may well be risk averse (as one would expect). Intuitively, a flat term structure may be generated by a  $\Pi$  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  $\delta$  which enter into the problem as a polynomial. In practical implementations over the time horizons where options are liquidly traded, a linear approximation provides an accurate approximation given that  $\delta$  is close to one. For instance, we know from the literature that  $\delta$  is close to 0.97 at an annual horizon.

The linear approximation is straightforward. To linearize the discounting of  $\delta^\tau$  around a point  $\delta_0$  (say,  $\delta_0 = 0.97$ ), we write  $\delta^\tau \approx a_\tau + b_\tau \delta$  for known constants  $a_\tau$  and  $b_\tau$ . 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 relatively accurate for  $\delta \in [0.94, 1]$  for time horizons less than 2 years.

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

$$\begin{pmatrix} \pi_{11} \\ \vdots \\ \pi_{T1} \end{pmatrix} + \begin{pmatrix} \pi_{12} & \dots & \pi_{1S} \\ \vdots & & \vdots \\ \pi_{T2} & \dots & \pi_{TS} \end{pmatrix} \begin{pmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{pmatrix} = \begin{pmatrix} a_1 + b_1\delta \\ \vdots \\ a_T + b_T\delta \end{pmatrix} \quad (22)$$

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

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

Rewriting this equation in matrix form as

$$Bh_\delta = a - \pi_1 \quad (24)$$

we immediately see the closed-form solution

$$h_\delta = \begin{cases} B^{-1}(a - \pi_1) & \text{for } S = T \\ (B'B)^{-1}B'(a - \pi_1) & \text{for } S < T \end{cases} \quad (25)$$

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  $\Pi$  that are not exactly generated by the model (e.g., because of noise in the data), then (25) provides the values of  $\delta$  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  $\delta_0$  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 (25) approximates the true solution in the following sense: The distance between the true solution  $(\bar{\delta}, \bar{h}, \bar{P})$  and the approximate solution  $(\delta, h, P)$  approaches 0 faster than  $(\delta_0 - \bar{\delta})$  as  $\delta_0$  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 really 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. The difficulty in handling expanding state spaces is an important element of the critique in Borovicka, Hansen, and Scheinkman (2015). We now 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. In doing so, we can allow a broader class of utility functions such that path dependence may be allowed. Hence, this specification can be used for Epstein-Zin preferences or habit formation.

**Assumption 1\* (General utility with  $N$  parameters)** *The pricing kernel at*

time  $\tau$  given the history of states  $\bar{s}_\tau = (s_0, \dots, s_\tau)$  can be written as

$$m_\tau^{\bar{s}_\tau} = h_\tau(\bar{s}_\tau, \theta) \quad (26)$$

where  $h(\cdot)$  is a one-to-one function of the observed states and the parameter  $\theta \in \Theta$ , and  $\Theta$  is a subset of  $R^N$  whose interior is non-empty.

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  $\Theta$ . Let us give two simple examples of how this can be done. 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  $\Theta$  to marginal utilities).

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

$$\begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_{S-1} \end{bmatrix} + \begin{bmatrix} b_{11} & \dots & b_{1N} \\ \vdots & & \vdots \\ b_{S-1,1} & \dots & b_{S-1,N} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} = A + B\theta \quad (27)$$

Combining this equation with the recovery problem (15) gives

$$(\Pi_1 + \Pi_2 A) + \Pi_2 B \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} = \begin{bmatrix} \delta \\ \vdots \\ \delta^T \end{bmatrix} \quad (28)$$

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

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 where we have price data.

Assumption 1\* also allows that the marginal utilities can be non-linear function of the risk aversion parameters  $\theta$ . 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  $\tau$  (given the current state 1 at time 0) is

$$m_{0,\tau}^{1,s} = \delta^\tau \left( \frac{c_s}{c_1} \right)^{-\theta} \tag{29}$$

where  $c_s$  is consumption in state  $s$  of the representative agent and  $\theta$  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 (Large State Space)** *Consider an economy satisfying Assumption 1\* with Arrow-Debreu prices for each of the  $T$  time periods and  $S$  states. Then the conclusions of Proposition 1 apply if the conditions for  $S$  are replaced by  $N + 1$ .*

**Proof.** The proof consistent of verifying that recovery in the large state space with  $N$  parameters corresponds to a small state space with  $S - 1$  free marginal utilities.

■

## 5.1 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  $\tau = 1, \dots, T$  (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  $\tau$  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  $\tau$  periods. The fundamental recovery equations now become

$$\pi_\tau(s) = \delta^\tau h_s(\theta) p_\tau(s) \quad (30)$$

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

$$\int \pi_\tau(s) h_s^{-1}(\theta) ds = \delta^\tau \quad (31)$$

For each time period  $\tau$ , this gives an equation to help us recover the  $N + 1$  unknowns (discount rate  $\delta$  and  $\theta \in R^N$ ).

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

$$h_s^{-1}(\theta) = A(s) + B(s)\theta \quad (32)$$

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 (31) as a simple equation of the same form as our original recovery problem (15):

$$\pi_\tau^A + \pi_\tau^B \theta = \delta^\tau \quad (33)$$

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 Data and Methodology

This section describes our data and empirical methodology.



## 6.1 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 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 30 and 750 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%.

## 6.2 Computing State Prices

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, to fully characterize the risk-neutral distribution, and back out state prices, we use the Pan (2002) specification of the model in Bates (2000) to match market option prices. We choose a parametric model to match prices in order to obtain prices that are consistent with a no-arbitrage argument in both the strike and maturity dimension. Furthermore the parametric approach puts structure on the tails of the risk-neutral density which allows us to extrapolate outside the range of observable option quotes.

In this model, the risk-neutral process for the underlying asset,  $S_t$ , and the in-

stantaneous variance,  $V_t$ , are assumed to be of the form

$$dS_t/S_t = (r - d - \lambda V_t \bar{k})dt + \sqrt{V_t}dZ_t + kdq_t \quad (34)$$

$$dV_t = (\alpha - \beta V_t)dt + \sigma_v \sqrt{V_t}dZ_{vt} \quad (35)$$

where  $Z_t$  and  $Z_{vt}$  are Brownian motions with correlation  $\rho$ ,  $q_t$  is a Poisson counting process with state-dependent jump intensity  $\lambda V_t$  and  $k$  is the random percentage jump size, with a Gaussian distribution  $\ln(1 + k) \sim N(\ln(1 + \bar{k})\frac{1}{2}\delta^2, \delta^2)$  conditional upon a jump. The term  $r - d$  is the cost of carry which is the risk-free rate minus the dividend. We use risk-neutral structural parameters from Table 1 in Pan (2002). To estimate the instantaneous variance we use the call/put option with the highest volume for each available maturity and ensure that we choose options with different strikes. This approach gives us between six and ten options per day.

Once we have obtained model estimates, we compute the risk-neutral density using the characteristic function for the Bates model as found in Pan (2002). The relation between the  $T$ -period characteristic function  $\psi_S(T; \Phi)$  and the  $T$ -period risk-neutral density is

$$f(T; x) = \frac{1}{\pi} \int_0^\infty e^{-i\Phi x} \psi_S(T; \Phi) d\Phi \quad (36)$$

where  $x = \log(S_T/S_0)$ . We use the Gauss-Laguerre quadrature method to numerically approximate the integral of equation (36). Knowing the risk-neutral density we time discount with the risk-free rate and estimate the corresponding  $T$ -period state price density as

$$\pi(T; x) = e^{-r_T T} f(T; x) \quad (37)$$

where  $r_T$  is the annualized  $T$ -period risk-free rate. We estimate a state price density for 34 different horizons, namely 30, 40, 50, ..., 350, 360 days. For each horizon we

estimate a state price density that ranges from 30% to 200% moneyness.

### 6.3 Recovery Methodology

We apply the closed-form approximation of Proposition 4. Our state space will be very large since we let each integer index value ranging from 30% to 200% moneyness be its own state, so clearly with  $T = 34$  we have  $S \gg T$ . We therefore impose a linear, lower-dimensional structure on  $H^{-1}e$  in the simplified recovery equation

$$\Pi H^{-1}e = De. \tag{38}$$

We do this by assuming that  $H^{-1}e = X\theta$ , where  $\theta$  is an  $N$ -dimensional column vector and  $X$  is a known  $S \times N$  matrix. Our choice of  $X$  corresponds to having a pricing kernel that has piecewise constant curvatures:

$$X = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & (n_{10} - 1)/n_{10} \\ 1 & 1 & \dots & 1 & (n_{10} - 2)/n_{10} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 1/n_{10} \\ 1 & 1 & \dots & (n_9 - 1)/n_9 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 0 & 0 \\ 1 & (n_2 - 1)/n_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 2/n_2 & \dots & 0 & 0 \\ 1 & 1/n_2 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} \tag{39}$$

where  $n_i$  is the number of index values covered by the curve controlled by  $\theta_i$ . In our implementation, we let  $N = 10$ . There are of course alternative relevant specifications of  $X$  which will be investigated in future versions. The appearance of  $X$  determines how the parameters in  $\theta$  should be interpreted. In our specification we can interpret  $\theta_1$  as a level parameter for the first curve piece of the pricing kernel and  $\theta_2$  is the slope. For the second curve piece  $\theta_1 + \theta_2$  is the level and  $\theta_3$  is the slope, for the third curve piece  $\theta_1 + \theta_2 + \theta_3$  is the level and  $\theta_4$  is the slope, and so on for the remaining curve pieces.

In choosing which states that should be covered by each  $\theta_i$  we do the following; we let  $\theta_2$  control the slope of the pricing kernel over states ranging from  $0.3S_0$  to  $(1 - \sigma)S_0$  where  $S_0$  is the current index value and  $\sigma$  is the implied one year risk-neutral standard deviation. We let  $\theta_{10}$  govern the slope of the pricing kernel over states ranging from  $(1 + 2\sigma)S_0$  to  $2S_0$ . The remaining seven parameters  $\theta_3, \dots, \theta_9$  cover equally large intervals in the range  $(1 - \sigma)S_0$  to  $(1 + 2\sigma)S_0$ . Letting the index values at which the parameters cover depend on volatility is simply an attempt to ensure that  $\Pi X$  has full rank. When using this specification of  $X$  we obtain a  $\Pi X$  with full rank for every Wednesday from 1/1996 to 7/2014.

Imposing a first order Taylor expansion of  $\delta^\tau$  in  $De$  around some guess  $\bar{\delta}$  we end up in the setting of Proposition 4 and parameter estimates of  $\theta, \delta$  are straightforward to obtain. Once the parameters are known we back out the multiperiod physical probability distribution as

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

where  $D$  is a diagonal matrix with elements  $D_{ii} = \delta^i$  and  $\text{diag}(X\theta)$  is a diagonal matrix with elements  $\text{diag}(X\theta)_{jj} = X_j\theta$  where  $X_j$  is the  $j$ 'th row of  $X$ .

## 6.4 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, of course. If  $r_{t,t+\tau}$  is the return on the index in period  $t$  to  $t + \tau$  and  $rf_{t,t+\tau}$  is the  $\tau$  period risk free rate, then we compute

$$\mu_{t,t+\tau} = E_t^{\mathbb{P}}[r_{t,t+\tau}] - rf_{t,t+\tau} \quad (41)$$

as the conditional expected excess return over from time  $t$  to  $\tau$ . Furthermore, we let

$$\sigma_{t,t+\tau} = \sqrt{\text{VAR}_t^{\mathbb{P}}(r_{t,t+\tau})} \quad (42)$$

be time  $t$  conditional  $\tau$  period volatility. Likewise, we can compute tail risk measures by summing the probabilities of suitably chosen extreme outcomes.

## 7 Empirical Results

We now investigate the predictive power of the estimated physical probabilities, noting that these results are preliminary and should be interpreted with caution. Figure 3 shows annualized monthly, quarterly and annual conditional expected excess returns. We see that these conditional expected excess returns vary over time and are closely related. The average annualized conditional expected excess return for monthly, quarterly and annual horizons are respectively 6.24%, 5.26% and 3.9%. Generally, the shorter the horizon, the higher the conditional expected excess return. This is consistent with van Binsbergen, Brandt, and Kojen (2012) who suggest that the term structure of equity premium and Sharpe Ratio is downward sloping.

The estimated physical volatility is plotted in Figure 4. We note that it is not

surprising that volatilities can be recovered, so we report these as a simple reality check on our method.

Figure 5 shows our estimated annualized conditional Sharpe Ratios, the ratio of the expected excess return to the volatility as discussed above. Here we also find evidence broadly consistent with the findings of van Binsbergen, Brandt, and Koijen (2012).

Figure 6 shows the estimated conditional probability of a fall in the S&P 500 index of 20% or more over a month, a quarter, or a year. We see that this measure of tail risk naturally varies significantly over time.

Table 1 reports results for a regression of ex post realized excess return on ex ante excess returns as well as the change in expected returns:

$$r_{t,t+\Delta t} = \beta_0 + \beta_1 \mu_{t,t+\Delta t} + \beta_2 \Delta \mu_{t+\Delta t} + \epsilon_{t+\Delta t} \quad (43)$$

where  $\epsilon_{t+\Delta t}$  is a noise term,  $\mu_{t,t+\Delta t}$  is the ex ante conditional expected return

$$\mu_{t,t+\tau} = E_t^{\mathbb{P}}[r_{t,t+\tau}] - r f_{t,t+\tau} \quad (44)$$

and  $\Delta \mu_{t+\Delta t}$  is the contemporaneous unpredictable change in the conditional expected return:

$$\Delta \mu_{t+\Delta t} = \mu_{t+\Delta t,t+2\Delta t} - E_t[\mu_{t+\Delta t,t+2\Delta t}] \quad (45)$$

To understand this regression, note that we are interested in testing whether the recovered probabilities give rise to reasonable expected returns, that is, 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

(43) should satisfy

$$\beta_0 = 0 \quad \text{and} \quad \beta_1 = 1 \quad \text{and} \quad \beta_2 < 0 \quad (46)$$

To estimate the innovation to the risk premium, we first estimate an AR(1)-process:

$$\mu_{t+\Delta t, t+2\Delta t} = \underbrace{\alpha_0 + \alpha_1 \mu_{t, t+\Delta t}}_{E_t[\mu_{t+\Delta t, t+2\Delta t}]} + \zeta_{t+\Delta t} \quad (47)$$

where the first part is the predicted future risk premium and  $\zeta_{t, t+\Delta t}$  is the unpredictable shock to the risk premium,  $\Delta \mu_{t+\Delta t}$ . The estimated persistence parameter  $\alpha_1$  is 0.58 for both a monthly and quarterly horizon and 0.2 annually.

Panel A in Table 1 reports the result of regression (43) on the full sample from 1/1997 to 7/2014. Here we see that  $\beta_1$  is positive but insignificant in every case, even if we control for the change in expectation. The coefficient  $\beta_2$  is highly significant on both the monthly and quarterly horizon and has the desired negative sign. Panel B reports the result of regression (43) on the sample period that excludes the financial crisis from 1/2008 to 7/2009. In this subsample, the estimated  $\beta_1$  is significant for both the monthly and quarterly horizon. The estimated  $\beta_1$  coefficient ranges from 1.47 to 3.62 which means that the estimated expected excess returns are on average a bit too low, but the estimate is not statistically significantly different from 1 ( $t$ -statistics values of 1.68 and 1.07, respectively, for monthly and quarterly horizons).

Table 2 reports the results of regressing ex post realized volatility on the ex ante recovered conditional volatility. We use squared daily returns to estimate monthly and quarterly realized variance and volatility. As seen in Table 2, the estimated slope coefficients are close to 1, with values of 0.94 monthly and 0.89 quarterly. These estimates are significantly different from zero ( $t$ -statistics above 5) and not significantly different from one ( $t$ -statistics of  $-0.34$  and  $-0.72$ ).

## 8 Conclusion

We provide a recovery result which makes no assumptions on the physical probability distribution, thus generalizing Ross (2015) who relies on strong time-homogeneity assumptions. In other words, we provide a simple method to recover physical probabilities and preference parameters from observed state prices. Our framework can handle classical multinomial models, models with an infinite state space and models with non-Markovian behavior. Our framework is also rich enough to handle general preference specifications. 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 preliminary findings indicate that ex ante expected returns based on recovered physical probabilities may help predicting returns although further empirical tests are needed.



## A Appendix

**Lemma 1** *Suppose that  $x^* \in \mathbb{R}^n$  is defined by  $f(x^*) = 0$  for a differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  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  $\varepsilon \rightarrow 0$ .*

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

$$\frac{x - x^*}{\varepsilon} = \frac{\bar{x} - x^*}{\varepsilon} - df^{-1} \frac{f(\bar{x}) - f(x^*)}{\varepsilon} \rightarrow \Delta x - df^{-1}df \Delta x = 0 \quad (\text{A.1})$$

■

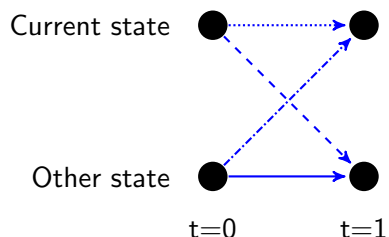
## References

- 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*.
- Bates, D. S., 2000, Post 87 crash fears in the S&P 500 futures options market, *Journal of Econometrics* 94, 181–238.
- 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* forthcoming.
- Carr, Peter, and Jiming Yu, 2012, Risk, return, and Ross recovery, *Working paper*.
- Christoffersen, Peter, Kris Jacobs, and Steven L. Heston, 2013, Capturing option anomalies with a variance-dependent pricing kernel, *Review of Financial Studies* 26, 1963–2006.
- 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.

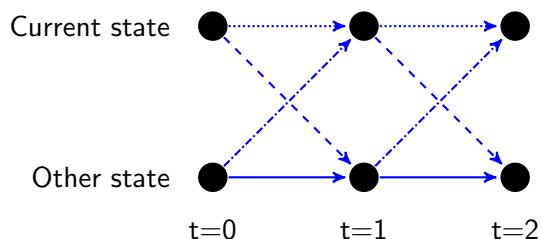
- Jackwerth, Jens C., and Mark Rubinstein, 1996, Recovering probability distributions from option prices, *Journal of Finance* 51, 1611–1631.
- Linetsky, Vadim, and Likuan Qin, 2015, Positive eigenfunctions of markovian pricing operators: Hansen-scheinkman factorization, Ross recovery and long-term pricing, *Working paper*.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.
- Pan, Jun, 2002, The jump-risk premia implicit in options: evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.
- 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*.
- van Binsbergen, Jules, Michael Brandt, and Ralph Koijen, 2012, On the timing and pricing of dividends, *American Economic Review* 102, 1596–1618.
- Walden, Johan, 2013, Recovery with unbounded diffusion processes, *Working paper*.

## B 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

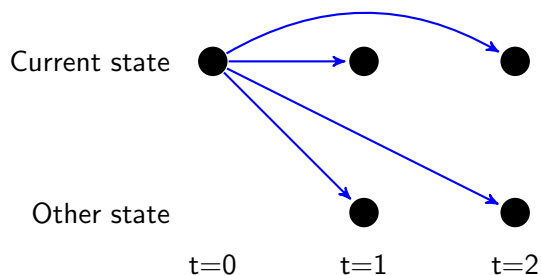
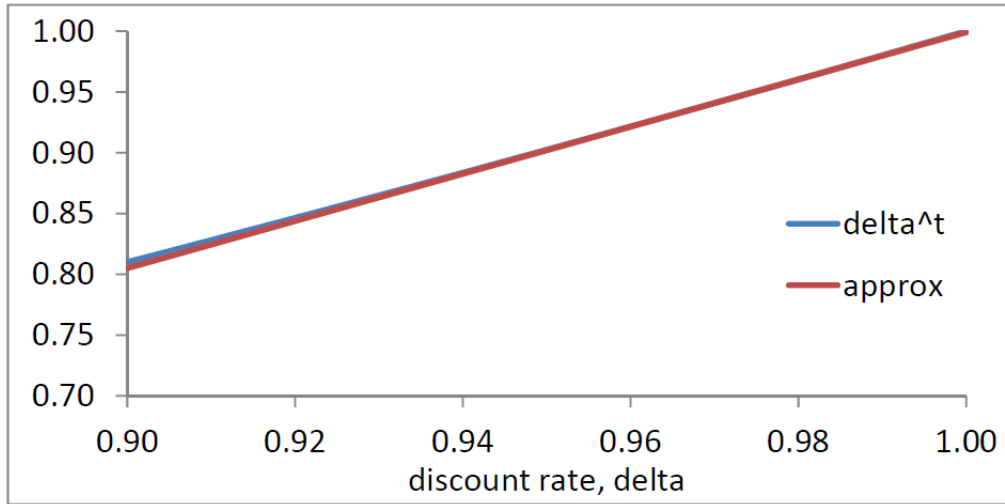
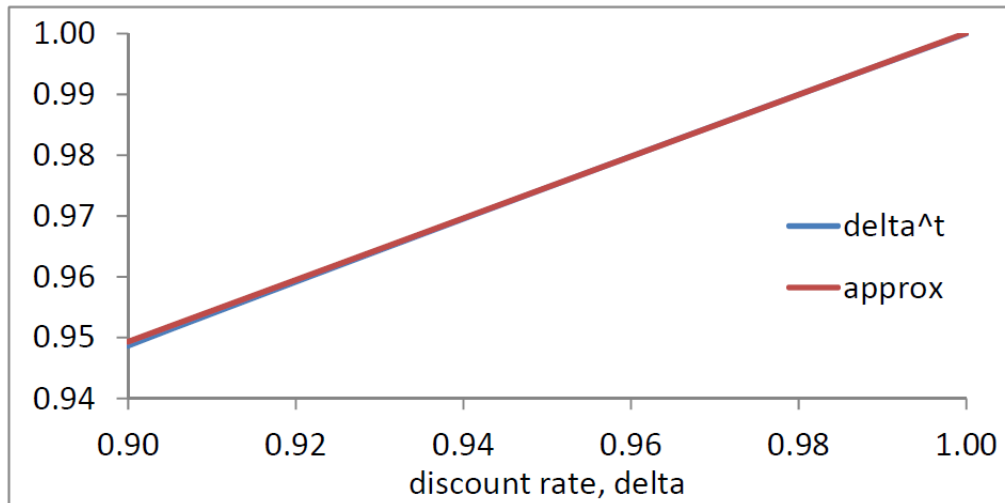


Figure 1: **Generalized Recovery Framework.** Panel A illustrates the idea behind Ross's Recovery Theorem, namely that we start with information about all Arrow-Debreu prices in *all* initial states (not just the state we are currently in, but also prices in "parallel universes" where today's state is 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  $\delta$  due to the powers of time,  $\delta^t$ . However, the function  $\delta \rightarrow \delta^t$  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.

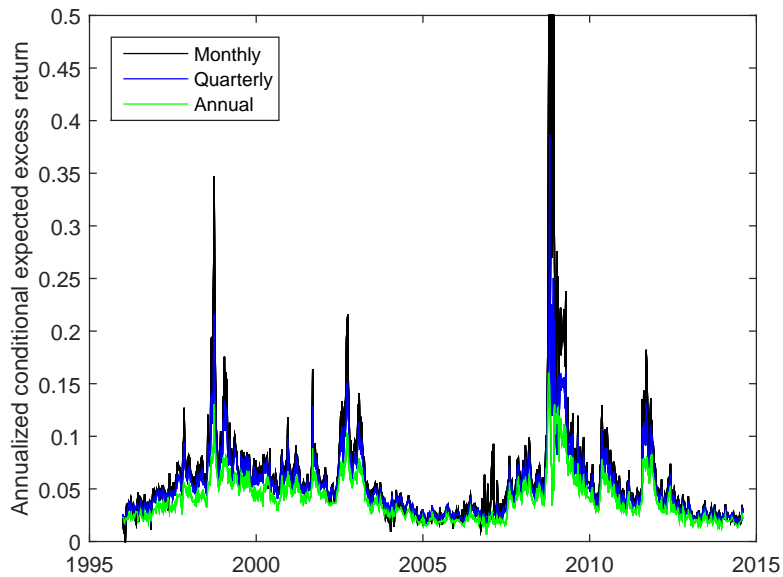


Figure 3: **Estimated annualized conditional expected excess return.** The figure plots annualized monthly, quarterly and annual condition expected excess returns, estimated each Wednesday from 1/1996 to 7/2014.

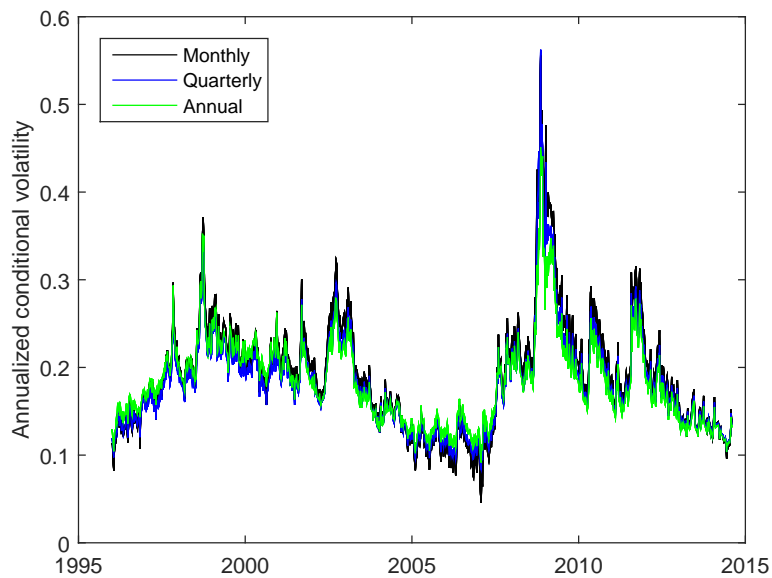


Figure 4: **Estimated conditional volatility of excess return.** The figure plots annualized monthly, quarterly and annual condition volatility, estimated each Wednesday from 1/1996 to 7/2014.

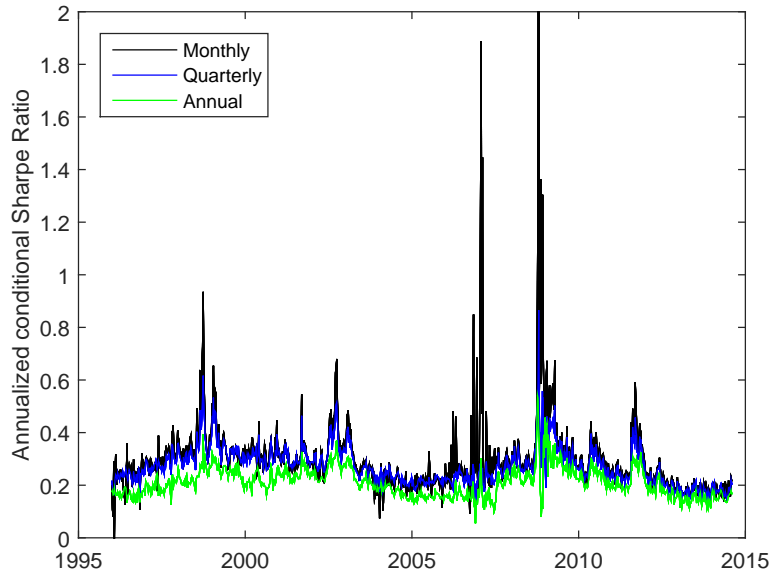


Figure 5: **Annualized Sharpe Ratio.** The figure plots annualized monthly, quarterly and annual Sharpe Ratio, estimated each Wednesday from 1/1996 to 7/2014.

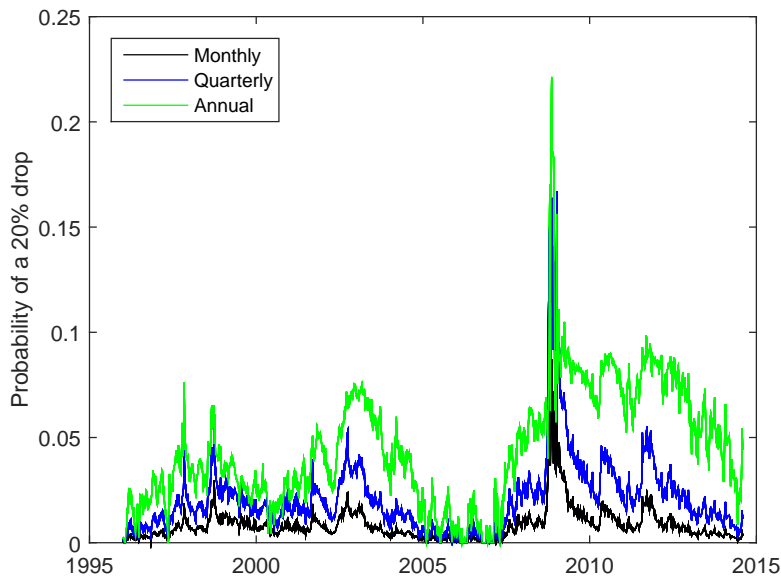


Figure 6: **Probability of a larger than 20% drop in the stock market index.** The figure plots monthly, quarterly and annual probability of a drop in the index price of at least 20%, estimated each Wednesday from 1/1996 to 7/2014.

Table 1: **Does the Recovered Expected Return Predict the Future Return?**

This table regresses the ex post realized return on the ex ante recovered expected return ( $\mu_{t,t+\Delta t}$ ) as well as the contemporaneous innovation in the expected return ( $\Delta\mu_{t+\Delta t}$ ):

$$r_{t,t+\Delta t} = \beta_0 + \beta_1\mu_{t,t+\Delta t} + \beta_2\Delta\mu_{t+\Delta t} + \epsilon_{t,t+\Delta t}$$

We perform this regression at the 1 and 3-month horizon. Panel A reports the results for the full sample, 1/1997 to 6/2014. Panel B reports the results of the subsample where the financial crisis (8/2008 to 7/2009) is eliminated. All of the regressions are based on weekly observations. Robust  $t$ -statistics accounting for the overlap following Newey and West (1987) are reported in parentheses and significance is indicated as \* for  $p < 0.1$ , \*\* for  $p < 0.05$  and \*\*\* for  $p < 0.01$ .

A: Full sample				
Horizon	1M	1M	3M	3M
Constant	0.00 (0.28)	0.00 (0.26)	-0.01 (-0.46)	-0.04** (-2.49)
$\mu_{t,t+\Delta t}$	0.33 (0.55)	0.32 (0.47)	1.09 (0.98)	0.33 (0.31)
$\Delta\mu_{t+\Delta t}$		-4.37*** (-3.49)		-7.32*** (-10.42)
Adj. $R^2$ (%)	0.05	19.05	1.12	43.41

B: Excluding the financial crisis 8/08-7/09				
Horizon	1M	1M	3M	3M
Constant	-0.01 (-1.28)	-0.02*** (-3.48)	-0.02 (-1.21)	-0.05*** (-2.49)
$\mu_{t,t+\Delta t}$	2.15** (3.14)	3.62*** (5.86)	2.15*** (2.00)	1.47** (2.06)
$\Delta\mu_{t+\Delta t}$		-11.93*** (-4.93)		-8.06*** (-10.42)
Adj. $R^2$ (%)	2	34.27	3.84	42.94



Table 2: **Does the Recovered Volatility Predict the Future Volatility?** This table regresses the ex post realized volatility on the ex ante recovered return volatility.

$$\sqrt{\text{VAR}[r_{t,t+\Delta t}]} = \gamma_0 + \gamma_1 \sigma_{t,t+\Delta t} + \nu_{t,t+\Delta t}$$

We perform this regression at the 1 and 3-month horizons. The table reports the results for the full sample, 1/1997 to 6/2014 based on weekly observations. Robust  $t$ -statistics accounting for the overlap following Newey and West (1987) are reported in parentheses and significance is indicated as \* for  $p < 0.1$ , \*\* for  $p < 0.05$  and \*\*\* for  $p < 0.01$ .

Horizon	1M	3M
Constant	-0.01 (-1.18)	-0.01 (-0.83)
$\sigma_{t,t+\Delta t}$	0.94*** (5.47)	0.89*** (5.87)
Adj. $R^2$ (%)	49.72	42.38