Essays on Asset Pricing

Jensen, Christian Skov

Link to publication in CBS Research Portal

ESSAYS ON ASSET PRICING

PhD Series 32-2018

Christian Skov Jensen

ESSAYS ON ASSET PRICING

Ph.D. School in Economics and Management

PhD Series 32.2018

CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Essays on Asset Pricing

Christian Skov Jensen

A thesis presented for the degree of
Doctor of Philosophy

Supervisor: Lasse Heje Pedersen
Ph.D. School in Economics and Management
Copenhagen Business School

Christian Skov Jensen
Essays on Asset Pricing

Preface

The thesis consists of three chapters, which can be read independently. The common theme across all three chapters is the relation between asset prices and investor preferences and beliefs.

The first chapter addresses the question of whether it is possible to recover physical probabilities, marginal utilities, and the discount rate from observed asset prices. We show when such a recovery is possible - and when it isn't - using a simple but general "counting argument". Recovery is possible when the number of states of the economy is no greater than the number of time periods. Our counting argument shows why recovery is impossible in most standard financial models where the state space grows as a multinomial tree. Nevertheless, we provide conditions under which recovery is possible in such an economy. While leaving probabilities fully free, we show that recovery is possible in an economy that evolves as a multinomial tree, if the number of parameters governing the stochastic discount factor is no greater then the number of time periods.

The second chapter addresses the question of how financial market tail risks vary over time and how we can infer such tail risks from asset prices. We show how the market's higher order moments can be estimated ex ante using options written on the market. We find that, the market's higher order moments move together in the sense that skewness becomes more negative when kurtosis becomes more positive. In other words, there are times when higher-moment risk is high, in the sense that the return distribution is both substantially left skewed (due to the large negative skewness) and fat tailed (due to the large positive kurtosis). Interestingly, higher-moment risk tends to be high at times when volatility is low, suggesting that when volatility is low, risk hides in the tails of the market return distribution. We show that this systematic variation

in higher-moment risk has large implications for investors; for example, the tail loss probability of a volatility-targeting investor varies from 3.6% to 9.7%, entirely driven by changes in higher-moment risk. Lastly, we show that times when higher-moment risk is high are characterized by high market and funding liquidity, high turnover, and low expected future returns.

The third chapter addresses the question of how investor risk aversion varies over time and how we can infer this risk aversion from asset prices. Using options written on the market and historical market returns, I present a new method for estimating the market's time-varying risk aversion. Interestingly, I find that the market's risk aversion is negatively related to variance, suggesting that the market became more risk tolerant during the recent 2008-2009 financial crisis. This finding is difficult to reconcile with the leading asset pricing models. Therefore, I discuss two possible explanations for this systematic variation in risk aversion. First, I find that my results are consistent with investors salience. At times of high volatility the expected return on the market is usually high relative to the risk-free return. The relatively high expected return on the market becomes salient for investors which induces heightened risk tolerance among investors. Second, I show how the systematic variation in risk aversion can arise if the stock market is not a perfect proxy for aggregate consumption.

Acknowledgements

This thesis has benefited from discussions with more people than I can mention here. I am grateful to everyone who has helped me, inspired me, and made my time as a doctoral student so much fun. However, a few people deserve special recognitions.

Above all, I am deeply indebted to my advisor Lasse Heje Pedersen who taught me the trade of being an academic within the area of financial economics. Working with Lasse was the most inspiring professional experience of my doctoral studies. I am also indebted to my secondary advisor David Lando who took me under his wings as my master thesis advisor. David encouraged me to do research and shoot for the stars.

In the spring of 2015, I visited Peter Christoffersen at the University of Toronto. Peter quickly became a mentor and a friend and he has my deepest gratitude for his encouragement in my research. I am also deeply grateful to Ian Martin. Visiting Ian at the London School of Economics in the fall of 2017 was a great experience and a perfect preparation for the job market.

Colleagues and friends made it easy for me to enjoy the past five years, for that I am grateful. My co-author Niels challenged everything I said and did. Discussions with Niels taught me a lot.

My partner Pia deserves special thanks. Pia has supported me tremendously throughout my doctoral studies. I have enjoyed every aspect of it: as a partner, a travel companion, and as a professional sparring partner. Finally, I am grateful to my parents for their guidance, support, and encouragement in the choices I have made over the years.

Introduction and Summaries

The central formula in asset pricing relates the price of an Arrow-Debreu security to an investor's preferences and beliefs:

$$\text{Price of an Arrow-Debreu security} = \text{Preferences} \times \text{Beliefs}$$

We observe the prices of Arrow-Debreu securities in the option markets. But we do not directly observe the extent to which these prices are driven by preferences or beliefs. Decomposing and investigating preferences and beliefs is essential for understanding asset prices, and it is therefore the focus of this thesis. In chapter one, my co-authors and I develop a model in which we can disentangle the contribution in asset prices which is driven by preferences and beliefs. In chapter two, my co-author and I estimate investor beliefs and study how these beliefs vary over time. In chapter three, I estimate investor preferences and study how they co-vary with investor beliefs.

1 Summaries in English

Generalized Recovery

Decoding risk preferences and beliefs from asset prices has been viewed as impossible until Ross (2015) provided sufficient conditions for such a recovery. Ross' recovery relies on two critical assumptions: (1) The economy evolves as a time-homogeneous Markov chain. (2) Preferences are time-separable.

In this paper, we generalize Ross' 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, which focuses the attention on the economics of the problem. Specifically, we show that recovery is possible if the number of states of the economy is no greater than the number of time periods with observable option prices. Furthermore, we show that our recovery inversion from prices to probabilities and preferences can be implemented in closed form.

Next, we consider an economy that evolves as a standard multinomial tree. We show that in this economy recovery is impossible because 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. Nevertheless, we show that recovery is possible in a large (continuous) state space model under certain conditions. While maintaining that probabilities are fully general, recovery is possible if we can parameterize the stochastic discount factor by a number of parameters which is no greater than the number of time periods with observable option prices.

Finally, we implement our methodology empirically using a large data set of call and put options written on the S&P 500 stock market index over the time period 1996-2015, testing the predictive power of the recovered expected return and volatility. The recovered expected returns have weak predictive power for the future realized returns, but the predictability is stronger when we exclude the global financial crisis. Recovered volatility has much stronger predictive power for future realized volatility.

Higher-Moment Risk

This paper investigates how financial market tail risk varies over time. Times of financial market distress pose threats to the macroeconomy, as we witnessed in the 2008-2009 financial crisis. For policymakers to act in a timely and preemptive manner in the event of financial market distress, it is important to measure the perceived tail risks in real time.

We estimate the market's higher order moments in real time using a new method based on Martin (2017) and arrive at five main results. First, we show that the moments of the market return, measured ex ante using option prices, predict future realized moments. Specifically,w e show that our ex ante skewness, kurtosis, hyper-

skewness, and hyperkurtosis all have significant predictive power over ex post realized moments.

As our second main result, we find that higher order moments move together in the sense that skewness and hyperskewness are more negative at times when kurtosis and hyperkurtosis are more positive. In other words, there are times when higher-moment risk is high, in the sense that the return distribution is both substantially left skewed (due to the large negative skewness and hyperskewness) and fat tailed (due to the large positive kurtosis and hyperkurtosis). We estimate the principal components of the space spanned by skewness, kurtosis, hyperskewness, and hyperkurtosis. The first principal component explains 90% of the joint variation in the market's higher order moments. We define this first principal component as a higher-moment risk index (HRI) which is meant to capture market tail risk.

As our third main result, we find that higher-moment risk varies systematically with variance. Higher-moment risk tends to be high at times when volatility is low, suggesting that when volatility is low, risk hides in the tail of the market return distribution. In addition, we find that higher-moment risks tend to be high subsequent to market run-ups, which are usually "calm" times as measured by variance.

As our fourth main result, we show that higher-moment risk has important implications for investors; for example, the tail loss probability of a volatility-targeting investor and varies from 3.6% to 9.7%, entirely driven by changes in higher-moment risk.

Finally, as our fifth main result, we investigate what can explain the systematic variation in higher-moment risk. In particular, we test if financial intermediaries are more levered when variance is low, and if such variation in financial intermediary leverage can explain our observed variation in higher moment risk. We find no relation between higher-moment risks and aggregate financial intermediary leverage. Next, we investigate how higher-moment risks are related to previously suggested measures of "bubble" characteristics and market valuation. We consider the "bubble" characteristics: acceleration (Greenwood, Shleifer, and You (2017)), turnover (Chen, Hong, and Stein (2001)), issuance percentage (Pontiff and Woodgate (2008)), and the market valuation measures: CAPE, the dividend-price ratio, and cay (Lettau and Ludvigson (2001)). We find that higher-moment risk is positively related to price acceleration:

there is more higher-moment risk when the recent price path is more convex. Also, higher turnover after market run-ups is associated with more higher-moment risk. Furthermore, there is more higher-moment risk when cay (Lettau and Ludvigson, 2001) is high. We find no conclusive relation between higher-moment risks and CAPE, the dividend-price ratio, or equity issuance.

The Market's Time-Varying Risk Aversion

This paper investigates how the market's risk aversion varies over time. Specifically, I provide a new method for estimating the market's time-varying risk aversion. My methodology allows me to investigate the co-movements between risk-aversion and the physical distribution of market returns.

I find that the market's risk aversion is negatively related to market variance, suggesting that investors are less risk averse during times of financial turmoil, e.g., during the recent 2008-2009 financial crisis. Next, I show that the market's risk aversion is positively related to market skewness, suggesting that investors are more risk averse at times when the normalized market return distribution is more risky.

Finally, I discuss two possible explanations for the systematic variation in market risk aversion. First, I investigate salience theory (Bordalo, Gennaioli, and Shleifer (2012)) as a possible behavioral explanation. Specifically, I follow Lian, Ma, and Wang (2018) who argue that, at times of low interest rates the relatively high expected returns on risky assets are salient, and this salience on the upside of a higher return on the risky asset induces heightened risk tolerance and "reaching for yield" tendencies among investors. I therefore regress the ratio of expected gross returns on the market to gross risk-free returns, $E_t(R_{t,T})/R_{t,T}^f$, onto risk aversion. Consistently with the findings of Lian, Ma, and Wang (2018), I find that investors become more risk tolerant as the ratio of expected returns to risk-free returns increases. Second, I discuss how the systematic variation in risk aversion can arise if the stock market is not a perfect proxy for aggregate consumption.

2 Summaries in Danish

Generalized Recovery

Gendannelse af risiko præferencer og sandsynligheder fra aktivpriser blevet betragtet som umuligt indtil Ross (2015) gav tilstrækkelige betingelser for en sådan gendannelse. Ross' metode afhænger af to kritiske antagelser: (1) Økonomien udvikler sig som en tidshomogen Markov-kæde. (2) Risiko præferencer er tidsseparable.

I dette kapitel generaliserer vi Ross' gendannelsesteorem til at håndtere en generel sandsynlighedsfordeling, der ikke er afhængig af en antagelse om tidshomogenitet eller Markov adfærd. Vi viser hvornår gendannelse er muligt, og hvornår det ikke er, ved hjælp af et simpelt "tælle"-argument. Specifikt viser vi, at gendannelse er mulig, hvis antallet af mulige udfald i økonomien er mindre end antallet af tidsperioder med observerbare optionspriser. Desuden viser vi, at vores gendannelse fra priser til sandsynligheder og præferencer kan implementeres i lukket form.

Dernæst betragter vi en økonomi, der udvikler sig som en standard multinomial træ. Vi viser at, i denne økonomi er gendannelse umulig, fordi antallet af mulige udfald er højere end antallet af tidsperioder. Dermed viser vi at, gendannelse uden yderligere antagelser er umuligt i de fleste standardmodeller for finansiering, hvor økonomien udvikler sig på denne måde.

Herefter viser vi under hvilke forudsætninger, at gendannelse er mulig i et stort (kontinuerligt) tilstandsrum. Samtidig med at vi lader sandsynligheder være helt generelle, er gendannelse mulig, hvis vi kan beskrive prisningskernen ved hjælp af en række parametre, som er mindre end antallet af tidsperioder med observerbare optionspriser.

Til sidst implementerer vi vores metode empirisk ved hjælp af et stort datasæt af optioner skrevet på det amerikanske aktie indeks S&P 500 i perioden 1996-2015. Vi tester hvor godt vores gendannede forventede afkast of varians prædikterer de fremtidige realiserede afkast. Det gendannede forventede afkast har en svag forudsigelseskraft for det fremtidige realiserede afkast, men forudsigeligheden er stærkere, når vi udelukker den globale finanskrise. Den gendannede varians har meget stærkere forudsigende kraft for fremtidig realiseret varians.

Higher-Moment Risk

I dette kapitel undersøger vi hvordan halerisikoen i finanssektoren varierer over tid. Tider hvor den finansielle sektor er i nød udgør trusler mod hele makroøkonomien, som vi oplevede i finanskrisen i 2008-2009. For at regulatorer kan handle rettidigt og forebygge fremtidige kriser i finanssektoren, er det vigtigt at kunne måle hvordan investorerne opfatter hale risici i finanssektoren.

Vi estimerer finanssektorens højere ordens momenter i realtime ved hjælp af en ny metode baseret på resultater i Martin (2017) og kommer frem til de følgende fem hovedresultater. For det første viser vi, at markedets momenter, målt ex ante ved hjælp af optionspriser, forudsiger fremtidige realiserede momenter. Specielt viser vi, at vores ex ante skewness, kurtosis, hyperskewness og hyperkurtosis alle kan prædiktere fremtidige realiserede momenter.

Som vores andet hovedresultat finder vi, at højere ordens momenter bevæger sig sammen i den forstand, at skewness og hyperskewness er mere negative, når kurtosis og hyper kurtosis er mere positive. Med andre ord er der tidspunkter, hvor risikoen i de højere momenter er høj i den forstand, at finansmarkedets afkastfordelingen både er væsentligt venstre skæv (på grund af den store negative skewness og hyperskewness) og har fede haler (på grund af den store positive kurtosis og hyper kurtosis) . Vi estimerer de principale komponenter i rummet udspændt af skewness, kurtosis, hyperskewness og hyper kurtosis. Det første principale komponent forklarer 90% af den fælles variation i markedets højere ordens momenter. Derfor definerer vi det første principale komponent som et højere ordens moment risiko index (HRI).

Som vores tredje hovedresultat finder vi, at risikoen i højere momenter varierer systematisk med finanssektorens varians. Risiko associeret med de højere momenter har tendens til at være høj i tider hvor variansen i finanssektoren er lav, hvilket tyder på, at når variansen er lav, skjuler risikoen sig i halen af afkastfordelingen. Derudover finder vi, at risici associeret med højere momenter har tendens til at være høje efter store opsving i finansmarkedet, hvilket normalt er "rolige" perioder målt ved varians.

Som vores fjerde hovedresultat viser vi, at risici associeret med højere momenter har store konsekvenser for investorer; for eksempel varierer sandsynligheden for et hale udfald i porteføljen for en investor der målrettet har konstant varians i hans portefølje

fra 3,6 % til 9,7 %, denne variation er udelukkende drevet af ændringer i finanssektorens højere ordens momenter.

Som vores femte hovedresultat undersøger vi hvad der kan forklare den systematiske variation i finanssektorens højere ordens momenter. Vi tester om finansielle institutioner gearer sig mere når variansen er lav, og hvis en sådan variation i finansiel institutioners gearing kan forklare vores observerede variation i højere ordens moment risiko. Vi finder ingen sammenhæng mellem højere ordens moment risici og gearingsniveauet for finansielle institutioner. Dernæst undersøger vi, hvordan højere ordens moment risiko er relateret til tidligere foreslåede variable der er associerede med finansielle "bubbler". Vi finder, at risikoen for højere ordens moment risici er positivt relateret til prisaccelerationen: Der er mere risiko for højere øjeblik, når den seneste prisstigning har været mere konveks. Desuden er højere omsætning efter marketsopsving forbundet med mere risiko i finanssektorens højere ordens momenter.

The Market's Time-Varying Risk Aversion

Dette kapitel undersøger hvordan markedets risikoaversion varierer over tid. Specifikt præsenterer jeg en ny metode til at estimere markedets tidsvarierende risikoaversion. Min metode giver mig mulighed for at undersøge samspillet mellem risikoaversion og fordeling af markedsafkast.

Jeg finder at, markedets risikoaversion er negativt relateret til markedets varians, hvilket tyder på, at investorer er mindre risikoaverse i tider med finansiel uro, fx under den seneste finansielle krise i 2008-2009. Dernæst viser jeg, at markedets risikoaversion er positivt relateret til markedets skewness, hvilket tyder på, at investorer er mere risikoaverse i tider hvor den normaliserede fordeling for markedets afkast er mere risikabel.

Til sidst diskuterer jeg to mulige forklaringer for den systematiske variation i markedsrisikoaversion. Først undersøger jeg salience teori (Bordalo, Gennaioli, and Shleifer (2012)) som en mulig adfærdsmæssig forklaring. Specielt følger jeg Lian, Ma, and Wang (2018), som argumenterer for at, i perioder med høj varians er markedets forventede afkast højt i forhold til den risikofrie rente, og at det høje forventede markedsafkast er "salient" (ekstra fremtrædende) i bevidstheden for investorer hvilket

medfører øget risikotolerance blandt investorer. For at teste deres hypotese regresserer jeg forholdet mellem markedets forventede afkast of det risiko-frie afkast på risikoaversion. Konsistent med resultaterne i Lian, Ma, and Wang (2018) finder jeg at, investorer bliver mere risikotolerante når forholdet mellem det forventede markedsafkast og det risikofrie afkast stiger. Dernæst viser jeg hvordan den systematiske variation i risikoaversion kan opstå, hvis aktiemarkedet ikke er en perfekt proxy for det samlet forbrug i økonomien.

Contents

Chapter 1

Generalized Recovery

Co-authored with David Lando and Lasse Heje Pedersen

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 allows a closed-form linearized solution. We implement our model empirically, testing the predictive power of the recovered expected return and other recovered statistics.

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 (2016) (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 (formalized based on Sard's Theorem), which focuses the attention on the economics of the problem. When recovery is possible, we show that our recovery inversion from prices to probabilities and preferences can be implemented in closed form. We implement our method empirically using option data from 1996-2015 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 1.1, 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 1.1, Panel B. In other words, Ross assumes that, if S&P 500 is at the level 2000, then one-period option prices do not depend on the calendar time at which this level is observed.

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 1.1, 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 $S-1$ unknown probabilities for each future time period. 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, for each additional time period 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 and, generically, it is unique.[1] On the other hand, 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 economics of 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

---

[1] Generically means that the result holds for all parameters except on a "small" set of parameters of zero measure. For the measure-zero set of parameters where a certain matrix of prices has less than full rank such that there is a continuum of solutions to our generalized recovery problem, we show that the multi-period version of Ross's problem also has a continuum of solutions.

in this way. In other words, our model provides a fundamentally different way – via our simple counting argument – to understand the critique of Borovicka, Hansen, and Scheinkman (2016) 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 marginal utilities 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 (1985a), Cox, Ross, and Rubinstein (1979), 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 (1985a), 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 Cox-Ross-Rubinstein model (the discrete-time version 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. While the former two models fall in the setting of Borovicka, Hansen, and Scheinkman (2016) (with a non-zero martingale component), we also show how recovery is possible in the non-Markovian setting, which falls outside the framework of Borovicka, Hansen, and Scheinkman (2016) and Ross (2015), illustrating the generality of our framework in terms of the allowed probabilities.

Finally, we implement our methodology empirically using a large data set of call and put options written on the S&P 500 stock market index over the time period 1996-2015. We estimate state price densities for multiple future horizons and 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 using four different methods. The recovered expected returns vary substan-

4

tially across specifications, challenging the empirical robustness of the results. The recovered expected returns have weak predictive power for the future realized returns, but the predictability is stronger when we exclude the global financial crisis. We can also recover ex ante volatilities, which have much stronger predictive power for future realized volatility. We note that a rejection of the recovered distribution is a rejection of the joint hypothesis of the general recovery methodology and the specific empirical choices including the state space and the available options.

The literature on recovery theorems is quickly expanding.[2] Bakshi, Chabi-Yo, and Bakshi (2017) and Audrino, Huitema, and Ludwig (2014) empirically test the restrictions of Ross's Recovery Theorem. Martin and Ross (2013) apply the recovery theorem in a term structure model in which the driving state variable is a stationary Markov chain, illustrating the role played by the (infinitely) long end of the yield curve, a role already recognized in Kazemi (1992). Several papers focus on generalizing the underlying Markov process to a continuous-time process with a continuum of values and an infinite horizon (Carr and Yu (2012), Linetsky and Qin (2016)) and Walden (2017) in particular derive intuitive results on the importance of recurrence. All these papers impose time-homogeneity of the underlying Markov process.[3] Qin and Linetsky (2017) go beyond the Markov assumption, discussing factorization of stochastic discount factors and recovery in a general semimartingale setting.

These approaches require an infinite time horizon while our approach only requires the observed finite-maturity data. Indeed, the martingale decomposition used by Borovicka, Hansen, and Scheinkman (2016) is only defined over an infinite horizon, as is the recurrence condition used by Walden (2017), and the factorization of Qin and Linetsky (2017).[4]

---

[2]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 (e.g., 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)).

[3]See also Schneider and Trojani (2017a) who focus on recovering moments of the physical distribution and Malamud (2016) who 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.

[4]Said differently, if we observe a data from a finite number of time periods from an economy satisfying the conditions on Borovicka, Hansen, and Scheinkman (2016), then there is no unique Markov decomposition. Recurrence means that each state is being visited infinitely often so it can only be defined over an infinite horizon. The factorization of Qin and Linetsky (2017) relies on limits

Our paper contributes to the literature by characterizing recovery of any probability distributions observed over a finite number of periods, by proving a simple solution and its closed-form approximation, and by providing natural empirical tests of our generalized method. Rather than relying on specific probabilistic assumptions (Markov processes and ergodocity) as in Ross (2015) and Borovicka, Hansen, and Scheinkman (2016), 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.[5] Indeed, the martingale decomposition applied by Borovicka, Hansen, and Scheinkman (2016) relies on knowing the infinite-time distribution of Markov processes, which imposes much more structure than needed and removes the focus from the essence of the recovery problem, namely the number of economic variables vs. economic restrictions.

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 pays 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 $i, j = 1, ..., S$

$$\pi^{i,j} = p^{i,j} m^{i,j} \tag{1.1}$$

---

of T-forward measures as T goes to infinity.

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

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 (1.1) 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:

$$m^{i,j} = \delta \frac{u^j}{u^i} \tag{1.2}$$

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.

It is important in Ross's setting as it will be in ours, that a state corresponds to a particular level of the marginal utility of consumption. This level does not depend on calendar time. In our empirical implementation, a state will correspond to a particular level of the S&P500 index.

The most troubling assumption, however, 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 $\pi_{t,t+\tau}^{i,j}$ 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. (1.1) becomes

$$\pi_{t,t+\tau}^{i,j} = p_{t,t+\tau}^{i,j}\, m_{t,t+\tau}^{i,j} \tag{1.3}$$

Similarly, the multi-period analogue to equation (1.2) 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*

$$m_{t,t+\tau}^{i,j} = \delta^\tau \frac{u^j}{u^i} \tag{1.4}$$

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 $\tau > 0$, $p_{t,t+\tau}^{i,j}$ 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.

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 (1.3) consists of four equations:

$$
\begin{aligned}
\pi_{t,t+1}^{1,1} &= p_{t,t+1}^{1,1} & m_{t,t+1}^{1,1} \\
\pi_{t,t+1}^{1,2} &= (1 - p_{t,t+1}^{1,1}) & m_{t,t+1}^{1,2} \\
\pi_{t,t+2}^{1,1} &= p_{t,t+2}^{1,1} & m_{t,t+2}^{1,1} \\
\pi_{t,t+2}^{1,2} &= \underbrace{(1 - p_{t,t+2}^{1,1})}_{2 \text{ unknowns}} & \underbrace{m_{t,t+2}^{1,2}}_{4 \text{ unknowns}}
\end{aligned}
\tag{1.5}
$$

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

9

notation h for the normalized vector of of marginal utilities:

$$h = \left(1, \frac{u^2}{u^1}, \ldots, \frac{u^S}{u^1}\right)' \equiv (1, h_2, \ldots, h_S)'. \tag{1.6}$$

where we normalize by u^1. With this notation and the assumption of time-separable utility, we can rewrite the system (1.5) as follows:

$$\begin{aligned}
\pi_{t,t+1}^{1,1} &= p_{t,t+1}^{1,1}\delta \\
\pi_{t,t+1}^{1,2} &= (1 - p_{t,t+1}^{1,1})\delta h_2 \\
\pi_{t,t+2}^{1,1} &= p_{t,t+2}^{1,1}\delta^2 \\
\pi_{t,t+2}^{1,2} &= (1 - p_{t,t+2}^{1,1})\delta^2 h_2
\end{aligned} \tag{1.7}$$

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

$$\begin{aligned}
\pi_{t,t+1}^{1,1} + \pi_{t,t+1}^{1,2}\frac{1}{h_2} - \delta &= 0 \\
\pi_{t,t+2}^{1,1} + \pi_{t,t+2}^{1,2}\frac{1}{h_2} - \delta^2 &= 0
\end{aligned} \tag{1.8}$$

Solving equation (1.8) 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

$$\pi_{t,t+2}^{1,1} - \frac{\pi_{t,t+2}^{1,2}\pi_{t,t+1}^{1,1}}{\pi_{t,t+1}^{1,2}} + \frac{\pi_{t,t+2}^{1,2}}{\pi_{t,t+1}^{1,2}}\delta - \delta^2 = 0 \tag{1.9}$$

Hence, we can solve the 2-state model by (i) finding δ as a root of the 2nd degree polynomial (1.9); (ii) computing the marginal utility ratio h_2 from (1.8); and (iii) computing the physical probabilities by rearranging (1.7).

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

$$\Pi = \begin{bmatrix} \pi_{11} & ... & \pi_{1S} \\ \vdots & & \vdots \\ \pi_{T1} & ... & \pi_{TS} \end{bmatrix} \tag{1.10}$$

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 (1.6) we define the $S-$dimensional diagonal matrix $H = \mathrm{diag}(h)$. Further, we construct a $T-$dimensional diagonal matrix of discount factors as $D = \mathrm{diag}(\delta, \delta^2, \ldots, \delta^T)$.

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

$$\Pi = DPH \tag{1.11}$$

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:

$$P = D^{-1}\Pi H^{-1} \tag{1.12}$$

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 (1.12) such that it only depends on δ and h:

$$\Pi H^{-1}e = DPe = De = (\delta, \delta^2, \ldots, \delta^T)' \tag{1.13}$$

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

$$\Pi = \begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix} = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} \tag{1.14}$$

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 (1.13) as

$$\Pi_1 + \Pi_2 \begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{bmatrix} \delta \\ \vdots \\ \delta^T \end{bmatrix} \tag{1.15}$$

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

$$\Pi_{11} + \Pi_{12} \begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{bmatrix} \delta \\ \vdots \\ \delta^{S-1} \end{bmatrix} \tag{1.16}$$

with solution[6]

$$
\begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \Pi_{12}^{-1} \left(\begin{bmatrix} \delta \\ \vdots \\ \delta^{S-1} \end{bmatrix} - \begin{bmatrix} \pi_{11} \\ \vdots \\ \pi_{S-1,1} \end{bmatrix} \right)
\tag{1.17}
$$

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

$$
\Pi_{21} + \Pi_{22} \begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{bmatrix} \delta^S \\ \vdots \\ \delta^T \end{bmatrix}
\tag{1.18}
$$

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 generically no solution exists for general Arrow-Debreu prices Π since δ must simultaneously solve several polynomial equations (where "generically" means almost surely as defined just below Proposition 1). 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 by the model" if there exist δ, P, and H such that Π can be found from the right-hand side of (1.11). 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 $S = T$;*

---

[6]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.

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 $S < T$.

The proof of this and all following propositions are in the appendix. The proposition states our results using the notion "generically," which means that they fail to hold at most for a set of measure zero. Said differently, if someone picks parameters "at random," then our results hold almost surely.[7]

Further, since Sard's theorem is not a standard tool in asset pricing theory, some words here on the basic intuition behind our use the theorem are in order. To get started, consider a linear function $f(x) = Ax$ from R^m to R^n given by the $n \times m$ matrix A. We know that if $n = m$ and A has full rank, then the image of A is all of R^n, i.e., every point of R^n is being "hit" by A. If, however, $n > m$, then the image of A is a linear subspace of R^n, which is vanishingly small (has Lebesgue measure 0 in R^n). By Sard's theorem, we can extend this result to a non-linear smooth function f and still conclude that, when $n > m$, the image of f is vanishingly small. Said differently, there exists no solution x to $f(x) = y$ generically (i.e., if you pick a random y then almost surely no solution exists).[8]

4 Generalized Recovery vs. Other Forms of Recovery

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. We now show how our method relates to Ross's method and other recovery results.

4.1 Generalized Recovery in a Ross Economy

We first show that our method generalizes Ross's recovery method in the sense that, if we are in a Ross economy, then any solution to Ross's problem has a corresponding solution to our problem.

---

[7] We note that the fact that our results hold only generically is not a consequence of our solution method – indeed, there exist counter-examples for special sets of parameters as discussed in our examples.

[8] On a more technical note, Sard's theorem in fact states, that if M is the set of critical points of f (i.e., the set of points for which the Jacobian matrix of f has rank strictly smaller than n), then $f(M)$ has Lebesgue measure zero in R^n. When $n > m$ all points are critical points, and therefore in this case $f(M)$ is the same as the image of f, which is what we need for our proof.

It is important to be clear about the terminology here. In Ross's recovery problem, physical transition probabilities are specified in terms of a one-period transition probability matrix $\bar{P}$ which includes transition probabilities from states that we are currently not in ("parallel universes"). Our problem focuses on recovering the matrix P of multi-period transition probabilities as seen from the state we are in at time 0, which we take to be state 1. We say that P *is generated from* $\bar{P}$ if the k'th row of P is equal to the first row of $\bar{P}^k$. The same terminology can be applied to state prices, of course.[9]

Proposition 2 (Generalized Recovery Works in a Ross Economy). *If observed prices* Π *over* $S = T$ *time periods are generated by a Ross economy (i.e., an irreducible matrix* $\bar{\Pi}$ *of one-period state prices and probabilities* P *satisfying Assumptions 1 and 2) then*

1. *The matrix* P *generated from* $\bar{P}$ *is a solution to our generalized recovery problem.*
2. *P is a unique solution to our generalized recovery problem generically in the space of Ross price matrices* $\bar{\Pi}$.
3. *If* Π_{12} *has full rank, then Ross's parallel-universe prices* $\bar{\Pi}$ *can be derived uniquely from multi-period prices* Π *observed from the current state. Otherwise, there may exist a continuum of Ross prices* $\bar{\Pi}$ *consistent with the observed prices. The rank condition is satisfied generically in the space of Ross price matrices.*

Part 1 of the proposition confirms that any solution to Ross's recovery problem corresponds to a solution to our generalized problem. Part 2 of the proposition considers the deeper question of uniqueness. Ross establishes a unique solution while our generalized recovery solution in our earlier Proposition 1 only narrows the solution set down to at most $S = T$ solutions. Interestingly, Proposition 2 shows that our method too yields a unique solution when prices come from a Ross economy, generically. Thus, in this sense, nothing is "lost" by using generalized recovery even when we are in a Ross economy.

One way to understand this result is to note that Ross's problem comes down to solving a characteristic polynomial, and, similarly, our generalized recovery problem

---

[9]The notion of generating P from $\bar{P}$ is based on the fact that, in a Ross economy, the matrix of probabilities of going from state i to state j in k time periods is given by $\bar{P}^k$. Likewise, the k-period state prices are given by $\bar{\Pi}^k$.

can be solved via the polynomial given by (1.18). Even though these polynomials come from different sets of equations, it turns out that they have the same roots when Ross's assumptions are satisfied.

Finally, part 3 of the proposition deals with the issue that some of our results only hold "generically," that is, for almost all parameters. One might ask whether Ross also has a similar problem for the (small set of) remaining parameters. The answer turns out to be "yes," and for a reason that has not yet been discussed in the context of Ross's method. The issue is that Ross finds a unique solution given his parallel universe price matrix $\bar{\Pi}$, but where does this matrix come from? In any real-world application, we start with observed prices Π over time as in our generalized recovery setting. When Ross implements his model empirically, he must first find his $\bar{\Pi}$ from the observed Π and then use his recovery method (but he does not consider the mathematics of the first step, getting $\bar{\Pi}$ from Π). Part 3 of the proposition shows that Ross has the same problem as we do for the small set of parameters where Π_{12} has less than full rank. In other words, his lack of uniqueness arises from the difficulty in finding the price matrix $\bar{\Pi}$. Interestingly, this may have been unnoticed since Ross takes $\bar{\Pi}$ as given in his theoretical analysis (and shows that his recovery is unique for each $\bar{\Pi}$).

This last point is most clearly seen through an example: Consider two different one-period transition probability matrices, that are both irreducible:

$$\bar{P} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \text{ and } \bar{P}' = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}$$

If we assume that the current state is state 1, then since all powers of the matrices $\bar{P}$ and $\bar{P}'$ have the same first row, namely $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, it follows that the matrices P and P' (i.e., the physical transition probabilities as seen from state 1) generated by $\bar{P}$ and $\bar{P}'$ become the same matrix

$$P = P' = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

For given discount factors D and marginal utilities H, $\Pi = DPH$ and $\Pi' = DP'H$

16

are then the same, and hence observing the 3×3 matrix of state prices Π would not allow us to distinguish between the physical transition matrices $\bar{P}$ and $\bar{P}'$. The problem is not mitigated by observing more periods. It is simply impossible in a world where we cannot observe parallel universe prices to distinguish between the two irreducible matrices. In our approach, we do not seek to recover the one period transition probabilities. Rather, we recover the matrix P, and our ability to do so depends on the rank of a submatrix the Π matrix. For example, if we let $\delta = 0.98$, and let $h_1 = 1, h_2 = 0.9, h_3 = 0.8$, then the sub-matrix of state prices Π_{12} has rank 1, and this means that we would not have unique recovery either.

4.2   Ross Recovery in our Generalized Economy

We now establish that our formulation is *strictly* more general, by showing 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.

Proposition 3 (Generalized Recovery is More General)**.** *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.*

This proposition shows that generalized recovery may be useful because it can match a broader class of market prices, in addition to the basic advantage that it starts with the observed multi-period prices (rather than parallel universe prices).

4.3   Recovery in Infinite Horizon

In addition to generalizing Ross's method, our result also provides a simple and intuitive way of understanding why, for example, growth may present a challenge for recovery, cf. the critique of Borovicka, Hansen, and Scheinkman (2016) 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 6.

Our argument is very different from that of Borovicka, Hansen, and Scheinkman (2016) who rely on a martingale decomposition, which requires infinite time horizon. Our counting argument is simple and is based on a finite horizon, consistent with the data observed in practice.

Our finite-horizon recovery theorem is therefore also markedly distinct from the existing approaches that exist in continuous-time models in that we make no reference to, and have no need for, recurrence or stationarity conditions. In a diffusion setting, Walden (2017) shows the fundamental role of recurrence as a necessary condition for recovery in these models. Recurrence essentially means that each state is being visited infinitely often so it can only be defined over an infinite horizon. Recurrence bears some resemblance to Ross' condition of irreducibility in that an infinite time extension of an irreducible chain would be recurrent. The result of Walden (2017) is intuitive since, when states are visited infinitely often, we have a chance to recover probabilities.

Our approach can naturally be used to consider whether recovery is possible in a finite-time version of infinite-horizon process (i.e., even if a process is defined over an infinite horizon, we can ask what happens if we only see it over a couple of years). Further, we can show via some examples that recovery may even be possible for non-recurrent processes or processes with growth.

To give a simple example of this, consider a two period non-homogeneous Markov process with two states defined from the probability transition matrices for each time

$$\bar{P}(0,1) = \begin{pmatrix} 0.4 & 0.6 \\ 0.5 & 0.5 \end{pmatrix} \text{ and, for } t \geq 1, \ \bar{P}(t, t+1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In the first period, the process either stays in its current state or jumps to the other state, but, after that, the process is absorbed in its current state. If we only observe prices for two time periods, then this is clearly the restriction of a non-recurrent process.

Given that $S = T = 2$, our counting argument shows that generalized recovery is feasible.

We could also imagine a process with growth, starting in the "lowest state" 1 and evolving according to a transition matrix specified as an upward drifting process. To give a simple illustration, imagine Assumptions 1 and 2 hold and that the one-period transition matrix of physical probabilities across five states is given as

$$\bar{P} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.1 & 0.5 & 0.4 & 0 & 0 \\ 0 & 0.1 & 0.5 & 0.4 & 0 \\ 0 & 0 & 0.1 & 0.5 & 0.4 \\ 0 & 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

If we observe prices over five time periods, then our counting argument is satisfied $S = T = 5$, and we see that it is not growth per se which makes recovery impossible — it is the expanding state space necessary to accommodate models with growth that may cause problems.

In summary, our results complement those in the literature in two ways. First, generalized recovery may work when other methods don't and vice versa. Second, generalized recovery provides an economic intuition in finite economies while other methods do so in infinite-horizon economies.

4.4 Flat Term Structure and Risk Neutrality

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

$$\frac{1}{(1+r)^\tau} = (\Pi e)_\tau \tag{1.19}$$

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 4 (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.*

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.

5 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 1.2, 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 (1.13) become the following:

$$\begin{pmatrix} \pi_{11} \\ \vdots \\ \pi_{T1} \end{pmatrix} + \begin{pmatrix} \pi_{12} & \dots & \pi_{1S} \\ \vdots & & \vdots \\ \pi_{T2} & \dots & \pi_{TS} \end{pmatrix} \begin{pmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{pmatrix} = \begin{pmatrix} a_1 + b_1\delta \\ \vdots \\ a_T + b_T\delta \end{pmatrix} \tag{1.20}$$

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

$$\begin{pmatrix} -b_1 & \pi_{12} & \dots & \pi_{1S} \\ \vdots & \vdots & & \vdots \\ -b_T & \pi_{T2} & \dots & \pi_{TS} \end{pmatrix} \begin{pmatrix} \delta \\ h_2^{-1} \\ \vdots \\ h_S^{-1} \end{pmatrix} = \begin{pmatrix} a_1 - \pi_{11} \\ \vdots \\ a_T - \pi_{T1} \end{pmatrix} \tag{1.21}$$

Rewriting this equation in matrix form as

$$Bh_\delta = a - \pi_1 \tag{1.22}$$

we immediately see the closed-form solution

$$h_\delta = \begin{cases} B^{-1}(a - \pi_1) & \text{for } S = T \\ (B'B)^{-1}B'(a - \pi_1) & \text{for } S < T \end{cases} \tag{1.23}$$

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 (1.23) 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 5 (Closed-Form Solution)**.** *If prices are generated by the model and B has full rank $S \leq T$ then the closed-form solution (1.23) 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 $\bar{\delta}$.

6 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*

$$m_{0,\tau}^{1,s} = \delta^\tau h_s(\theta) \tag{1.24}$$

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.

6.1   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

$$
\begin{bmatrix} h_2^{-1} \\ \vdots \\ h_S^{-1} \end{bmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_{S-1} \end{pmatrix} + \begin{pmatrix} b_{11} & \ldots & b_{1N} \\ \vdots & & \vdots \\ b_{S-1,1} & \ldots & b_{S-1,N} \end{pmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} = A + B\theta \qquad (1.25)
$$

Combining this equation with the recovery problem (1.15) gives

$$
(\Pi_1 + \Pi_2 A) + \Pi_2 B \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_N \end{pmatrix} = \begin{pmatrix} \delta \\ \vdots \\ \delta^T \end{pmatrix} \qquad (1.26)
$$

This equation has *exactly* the same form as our original recovery problem (1.15), 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

$$m_{0,\tau}^{1,s} = \delta^\tau \left(\frac{c_s}{c_1} \right)^{-\theta} \tag{1.27}$$

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 6 (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.*

As one simple application of the proposition, we can recover preferences from state prices if we know that the pricing kernel is bounded and we have sufficiently many time periods as seen in the following corollary. Said differently, using a simplified or winsorized pricing kernel (or state space) is a special case of Proposition 5.

Corollary 7 (Generalized Recovery with Bounded Kernel). *Suppose that the pricing kernel is bounded in the sense that there exist states $\bar{s} > \underline{s}$ such that $h_s = h_{\bar{s}}$ for $s > \bar{s}$ and $h_s = h_{\underline{s}}$ for $s < \underline{s}$. Then the conclusion of Proposition 5 applies, where N is the number of states from $\underline{s}$ to $\bar{s}$.*

6.2 Continuous State Space

Finally, we note that our framework also easily extends to a continuous state space under Assumption 1* in discrete time (see Walden (2017) for the case of continuous time and continuous state space). 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

$$\pi_\tau(s) = \delta^\tau h_s(\theta) p_\tau(s) \tag{1.28}$$

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

$$\int \pi_\tau(s) h_s^{-1}(\theta) ds = \delta^\tau \tag{1.29}$$

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 6.1, and we have recovery if there are enough time periods as stated in Proposition 6.

As before, the linear case is particularly simple. Suppose that the marginal utilities can be written as[10]

$$h_s^{-1}(\theta) = A(s) + B(s)\theta \tag{1.30}$$

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 (1.29) as a simple equation of the same form as our original recovery problem (1.15):

$$\pi_\tau^A + \pi_\tau^B \ \theta = \delta^\tau \tag{1.31}$$

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.

7 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.

---

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

7.1 Recovery in the Mehra and Prescott (1985) model

The Mehra and Prescott (1985a) 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. This consumption growth between time $t-1$ and t is captured by a process X_t. The aggregate consumption process can be written as

$$Y_t = \prod_{s=1}^{t} X_s \tag{1.32}$$

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

$$\pi_{0,t}^{1,s_t} = \delta^t y_t^{-\gamma} Pr(X_t = x_t, Y_t = y_t) \tag{1.33}$$

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.[11]

Based on this model of Mehra and Prescott (1985a), 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).

We first notice from equation (1.32) 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 6. The simplest way to proceed is to assume that we know the form of the

---

[11]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.

pricing kernel (1.33), 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

$$\Pi h^{-1}(\gamma) = \begin{bmatrix} \delta & \delta^2 & \ldots & \delta^T \end{bmatrix}' \tag{1.34}$$

where h is a one-to-one C^∞ smooth function of the parameter γ based on (1.33), see Appendix B for details.[12] 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 6 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 1.3 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 1.1).[13] 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.

7.2 Cox-Ross-Rubinstein and iid. consumption growth

We can capture the standard binomial model of Cox, Ross, and Rubinstein (1979) (i.e., the discrete time counterpart to Black-Scholes-Merton) as follows. We consider the same model for aggregate 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 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

---

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

[13]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.

Proposition 4, where recovery is impossible.[14] Concretely, the problem is that the price matrix Π from (1.34) is not full rank. Hence, as seen in Figure 1.3 Panel B, the objective of minimizing pricing errors has a continuum of solutions. In other words, recovery is not feasible.

7.3 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 7.1–7.2, 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 1.3 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 (2016)).

8 Empirical Analysis

This section describes our data, empirical methodology, and empirical findings.

8.1 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 last trading day of the month from January 1996 to December 2015. 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

---

[14]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 $E_t(\delta^\tau \frac{u_{t+\tau}}{u_t}) = E_t(\prod_{s=1,\ldots,\tau} \delta \frac{u_{t+s}}{u_{t+s-1}}) =: (\frac{1}{1+r})^\tau$, 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.

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%.

8.2 Recovery Methodology

The Generalized Recovery Theorem relies on the knowledge of Arrow-Debreu 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. Further, given the large number of states, we use the parametric kernel method from Section 6.

To study the robustness of recovery, we consider two different methods for backing out Arrow-Debreu prices and two different specifications of the pricing kernel, for a total of four different recovered distributions and preferences.

More specifically, we apply the following two methods of extracting Arrow-Debreu prices from options: (i) the parametric model of Bates (2000) and (ii) the non-parametric method of Jackwerth (2004). Each of the methods yields Arrow-Debreu prices across multiple time horizons and mutliple index levels for each day t as described in detail in Appendix C.

Given these observed Arrow-Debreu prices, we recover preferences and probabilities based on the two different specifications of the pricing kernel that we denote "piece-wise linear" and "polynomial" pricing kernels, respectively, as described in detail in Appendix D.

8.3 Computing Statistics of the Recovered 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 $(1 - 2.5\text{VIX}_t)S_t$ and $(1 + 4\text{VIX}_t)S_t$. 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 the recovered expected excess return μ_t at time t by summing over the N_t possible states:

$$\mu_t = E_t^{\mathbb{P}}[r_{t,t+1}] - r_{t,t+1}^f = \sum_{\nu=1}^{N_t} p_{t+1,\nu} r_{t+1,\nu} - r_{t,t+1}^f \tag{1.35}$$

where $r_{t,t+1}^f$ is the risk-free rate, $p_{t+1,\nu}$ is the recovered time-t conditional physical probability for the transition to state ν at time $t+1$, $r_{t+1,\nu} = \frac{S_{t+1}(\nu)}{S_t} - 1$ is the return in state ν, and $S_{t+1}(\nu)$ is the value of the index at time $t+1$ if state ν is realized.

We compute the contemporaneous unpredictable innovation in the conditional expected return as

$$\Delta\mu_{t+1} = \mu_{t+1} - E_t[\mu_{t+1}] \tag{1.36}$$

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

$$\mu_{t+1} = \alpha_0 + \alpha_1\mu_t + \varepsilon_{t+1} \tag{1.37}$$

The estimated persistence parameter α_1 is 0.3 at the monthly horizon.

We compute the recovered conditional variance, $\mathrm{VAR}_t^{\mathbb{P}}(r_{t,t+1})$, analogously to how we computed the expected return and we denote the recovered volatility by $\sigma_t = \sqrt{\mathrm{VAR}_t^{\mathbb{P}}(r_{t,t+1})}$.

8.4 Empirical Results

We next investigate the properties of the recovered probabilities based on each of our four methods. We first consider the recovered expected return. Table 1.1 shows the correlation matrix for the recovered expected returns based on each of our four methodologies as well as the VIX volatility index and the SVIX variable of Martin (2017). The good news is that all variables are positively correlated, as we would expect. The ææless good news is that the correlations between the different recovered expected returns are modest in magnitude, with an average pairwise correlation of only 0æ.5. This modest correlation is concerning because all these recovered expected

returns should be measures of the same thing, namely the market's expected return at any given time.

Figure 1.4 shows the time series variation of the recovered expected return based on one of the methodologies (we plot just one time series since it is difficult to look at all four together). These recovered expected returns do not look unreasonable, but we next try to test their predictability of actual realized returns. Specifically, we regress the ex post realized excess return on the ex ante recovered expected excess return, μ_t, and the ex post innovation in expected return, $\Delta\mu_{t+1}$:

$$r_{t,t+1} = \beta_0 + \beta_1\mu_t + \beta_2\Delta\mu_{t+1} + \epsilon_{t,t+1} \tag{1.38}$$

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 ($\beta_1 > 0$), whether an increase in the risk premium is associated with a contemporaneous drop in the price ($\beta_2 < 0$), and whether the intercept as zero ($\beta_0 = 0$).

Table 1.2 reports the results of this regression for each of our four recovery methodologies as well as using VIX and SVIX as the expected return over the full sample from 1997 to 2015. First, the intercept β_0 is insignificantly different from zero in most specifications, but significantly different from zero using method 2 and using VIX, providing evidence against these models. Second, β_1 is positive and marginally significant from 0 in model 1, but otherwise insignificantly different from zero, providing neither evidence in favor or against the models. The coefficient β_2 is highly significant and has the desired negative sign in all models. Further, as expected the absolute value of β_2 is greater than one 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 1.3 reports the result of regression (1.38) over the sub-sample that excludes the global financial crisis (9/2008–7/2009), a sub-sample that has been considered in the literature (e.g., Martin (2017)). The results here are stronger and more consistent with theory. All the key parameters have the expected sign, the estimated coefficient β_0 is small and insignificant in all models, the estimated coefficient β_1 is positive and

31

marginally significant or insignificant, and β_2 is negative and significant.

The reason that the models work better when we exclude the crisis is intuitive: During the crisis, there were several months in which the ex ante recovered expected return was high, but, nevertheless, the ex post realized return was negative and large in magnitude. It seems plausible that investors were scared at that time, which means that it is plausible that the true required return was indeed high, which in turn implies that the negative realized return was a negative surprise. Hence, one could argue that the model gets this period wrong for the "right" reason, but we don't want to push this argument too far as the most compelling evidence is almost always that of using the full sample.

Finally, we consider the recovered physical volatility as plotted in Figure 1.5. This recovered volatility looks reasonable. Further, the recovered volatilities are similar across the different methodologies with an average pairwise correlation of 0.95 and an average correlation to VIX of 0.92. It is not that surprising that volatilities can be recovered, but studying volatility provides a simple and powerful reality check of our method since the true future volatility is known with much less error than the expected return. Hence, we regress the ex post realized volatility on the ex ante recovered conditional volatility, σ_t:

$$\sqrt{\mathrm{VAR}(r_{t,t+1})} = \beta_0 + \beta_1 \sigma_t + \epsilon_{t,t+1} \tag{1.39}$$

where the realized volatility $\sqrt{\mathrm{VAR}(r_{t,t+1})}$ is computed using close-to-close daily data over the 4 weeks from t to $t+1$ by OptionMetrics. We also run the same regression where we replace the recovered volatilities by the VIX volatility index. The theory predicts that $\beta_0 = 0$ and $\beta_1 = 1$.

Table 1.4 reports the results. As seen in Table 1.4, the estimated intercept coefficient β_0 is insignificant for models 1 and 2, but significant for models 3 and 4. However, for all models, the intercept is smaller than that of VIX, suggesting that the recovered volatilities are less biased than VIX.

The estimated slope coefficient β_1 is positive and highly significant for all models. Further, the estimated slope is close to the predicted value of 1, in particular closer than the estimated value for VIX. Lastly, we see that VIX has a slightly higher R^2,

which may reflect that the recovery method introduces some noise in the volatility measures.

In summary, we find substantial differences across the recovered probabilities based on different methodologies, and the predictive power for future returns appears weak in the full sample, but slightly stronger in the sample that excludes the gælobal financial crisis. The recovered volatilities predict well the future volatility in a way that is less biased than VIX, but slightly lower R^2. We are able to reject that the recovered probabilities provide a perfect description of the future evolution of the market based on a Berkowitz (2001) test.[15] This rejection could be due to the details of our implementation. For instance, while the true pricing kernel may depend on multiple factors, we assume that the state space is given by the level of S&P500 since we do not observe option prices depending simultaneously on multiple factors.

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, our counting argument immediately shows that recovery is generally not feasible. While this finding parallels results by Borovicka, Hansen, and Scheinkman (2016), our intuitive counting argument is fundamentally different and does not rely on the assumptions of an infinite-period time-homogeneous Markov setting, but, rather, is based on the general methods pioneered by Debreu (1970) for general equilibrium.

To pursue recovery even in economies with growth, e.g., classical multinomial models, we show how our method can be used when the pricing kernel can be parameterized by a sufficiently low-dimensional parameter vector. When recovery is feasible, our model allows a closed-form linearized solution. We implement our model empiri-

---

[15]The details of this test are not reported for brievity. The idea is that, given the estimated distribution $\hat{F}_t$ of the excess return r_{t+1} at time t, the distribution of the transformed variable $u_{t+1} = \hat{F}_t(r_{t+1})$ should be uniform and the distribution of the further transformed variable $x_{t+1} = \Phi^{-1}(u_{t+1})$ should be standard normal, which is tested by estimated 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 $Var(\epsilon_t) = 1$.

cally using several different specifications, testing the predictive power of the recovered statistics.

Panel A. Ross's Recovery Theorem: one period, two "parallel universes"
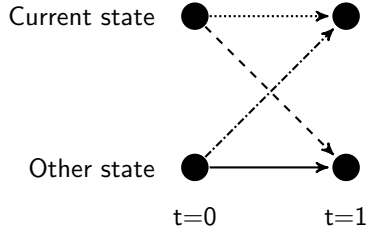
Panel B. Ross's Recovery Theorem: time-homogeneous dynamic setting
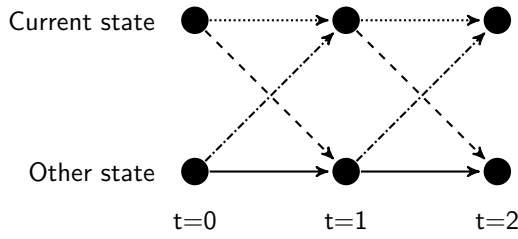
Panel C. Our Generalized Recovery: No assumptions about probabilities
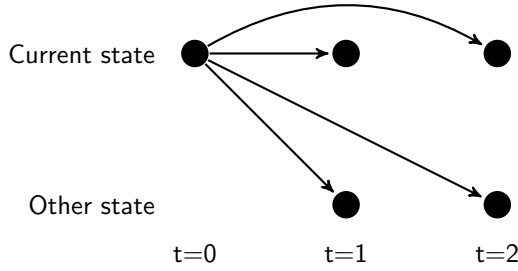
Figure 1.1: **Generalized Recovery Framework.** Panel A illustrates the idea behind Ross's Recovery Theorem, namely that we start with information about all Arrow-Debreu prices in *all* initial states (not just the state we are currently in, but also prices in "parallel universes" where today's state 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.
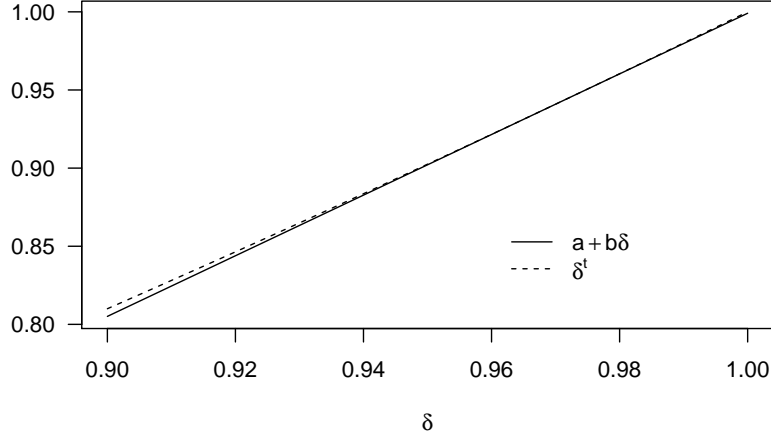
Panel A: $t = 2$ **years**
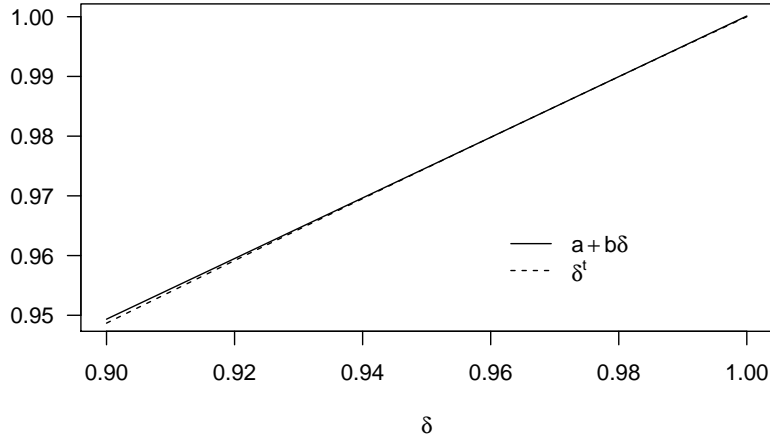
Panel B: $t = 0.5$ **years**

Figure 1.2: **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 1.1: **Correlation Matrix.** This tables shows the pairwise correlations between the recovered conditional expected excess return for different specifications of marginal utilities and method for estimating risk-neutral prices; (i) $\mu_{t,1}$: Bates and polynomial, (ii) $\mu_{t,2}$: Bates and piecewise linear, (iii) $\mu_{t,3}$: Jackwerth and polynomial, (iv) $\mu_{t,4}$: Jackwerth and piecewise linear. We augment the table with pairwise correlations with the VIX_t index and the lower boundary on the equity premium, SVIX_t, due to Martin (2017).

	$\mu_{t,1}$	$\mu_{t,2}$	$\mu_{t,3}$	$\mu_{t,4}$	VIX_t	SVIX_t
$\mu_{t,1}$	1	0.359	0.393	0.392	0.534	0.485
$\mu_{t,2}$		1	0.642	0.523	0.716	0.794
$\mu_{t,3}$			1	0.642	0.784	0.830
$\mu_{t,4}$				1	0.634	0.689
VIX_t					1	0.928
SVIX_t						1
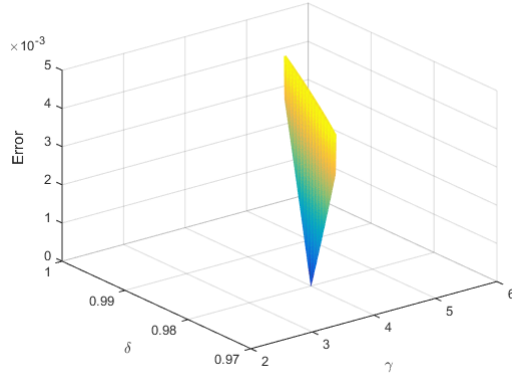
Panel A: Mehra Prescott (1985)
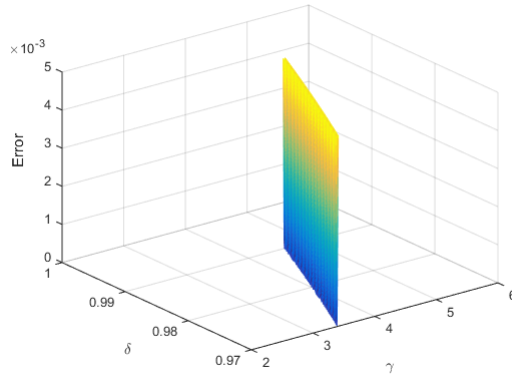
Panel B: Iid. consumption
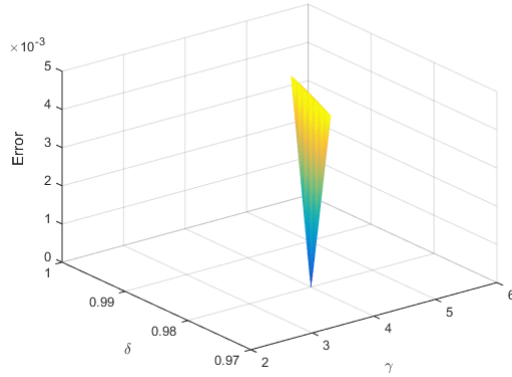
Panel C: Non-Markovian

Figure 1.3: **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 (1.48). 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.
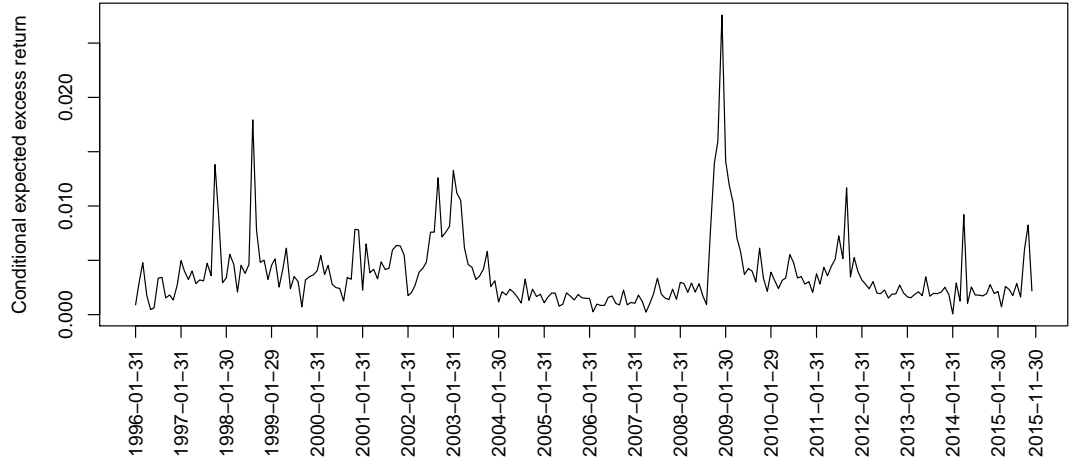
Figure 1.4: **Recovered conditional expected excess return.** The figure plots monthly conditional expected excess market returns, recovered last trading day of each month from 1/1996 to 12/2015. Marginal utilities are piecewise linear and risk-neutral prices are estimated using Jackwerth (2004).
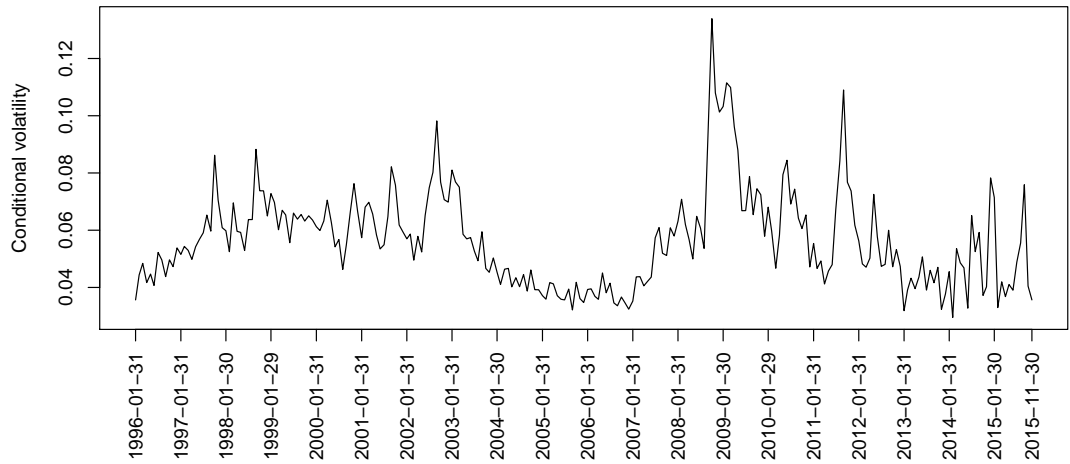
Figure 1.5: **Recovered conditional volatility of excess return.** The figure plots monthly conditional market volatility, recovered last trading day of each month from 1/1996 to 12/2015. Marginal utilities are piecewise linear and risk-neutral prices are estimated using Jackwerth (2004).

39

Table 1.2: **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, $\Delta\mu_{t+1}$, and ex ante the SVIX index of Martin (2017):

$$r_{t,t+1} = \beta_0 + \beta_1\mu_t + \beta_2\Delta\mu_{t+1} + \beta_3 SVIX_t + \beta_4 VIX_t + \epsilon_{t,t+1}$$

The regression uses monthly data over the full sample 1/1996–12/2015, t-statistics are reported in parentheses, and significance at a 10% level is indicated in bold.

Dependent variable	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$
Intercept	-0.00	**0.01**	0.01	0.00	**0.01**	0.01
	(-0.06)	(2.34)	(1.64)	(1.46)	(2.05)	(1.78)
μ_t	1.23	-2.95	-0.22	0.07	-0.00	0.18
	(1.55)	(-1.25)	(-0.18)	(0.09)	(-1.25)	(0.28)
$\Delta\mu_{t+1}$	**-3.66**	**-22.07**	**-14.00**	**-7.83**	**-0.55**	**-16.11**
	(-3.80)	(-7.65)	(-7.93)	(-7.32)	(-10.1)	(-16.01)
Adj. R^2 (%)	5.9	18.1	20.4	17.8	30.0	51.7
Method:						
Expected excess return (μ_t)	Recovered	Recovered	Recovered	Recovered	VIX	SVIX
Q-prices	Bates	Bates	Jackwerth	Jackwerth		
Pricing kernel	Polynomial	Piecewise linear	Polynomial	Piecewise linear		

Table 1.3: **Does the Recovered Expected Return Predict the Future Return - Excluding 8/2008-7/2009**
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, $\Delta\mu_{t+1}$, and ex ante the SVIX index of Martin (2017):

$$r_{t,t+1} = \beta_0 + \beta_1\mu_t + \beta_2\Delta\mu_{t+1} + \beta_3 SVIX_t + \beta_4 VIX_t + \epsilon_{t,t+1}$$

The regression uses monthly data over the full sample 1/1996–12/2015, t-statistics are reported in parentheses, and significance at a 10% level is indicated in bold.

Dependent variable	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$	$r_{t,t+1}$
Intercept	0.00	0.01	-0.00	0.00	0.00	0.00
	(0.06)	(0.98)	(-0.15)	(1.20)	(1.11)	(0.12)
μ_t	**1.37**	0.28	**3.06**	**1.65**	0.00	**1.71**
	(1.79)	(0.07)	(1.82)	(1.72)	(-0.25)	(1.99)
$\Delta\mu_{t+1}$	**-3.04**	**-26.98**	**-12.74**	**-9.00**	**-0.50**	**-17.69**
	(-3.30)	(-7.15)	(-6.29)	(-7.93)	(-8.75)	(-15.53)
Adj. R^2 (%)	5.0	18.1	16.4	23.1	24.6	52.5
Method:						
Expected excess return (μ_t)	Recovered	Recovered	Recovered	Recovered	VIX	SVIX
Q-prices	Bates	Bates	Jackwerth	Jackwerth		
Pricing kernel	Polynomial	Piecewise linear	Polynomial	Piecewise linear		

Table 1.4: **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:

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

The regression uses monthly data over the full sample 1/1996–12/2015, t-statistics are reported in parentheses, and significance at a 10% level is indicated in bold.

Dependent variable	$\sqrt{var(r_{t,t+1})}$	$\sqrt{var(r_{t,t+1})}$	$\sqrt{var(r_{t,t+1})}$	$\sqrt{var(r_{t,t+1})}$	$\sqrt{var(r_{t,t+1})}$
Intercept	-0.00	-0.00	-0.01	-0.01	-0.05
	(-1.45)	(-1.40)	(-2.41)	(-2.83)	(-9.63)
σ_t	**0.89**	**0.86**	**0.95**	**0.98**	**0.71**
	(16.7)	(16.8)	(14.67)	(14.63)	(17.19)
Adj. R^2 (%)	54.0	54.0	47.4	47.3	55.3
Method:					
Volatility (σ_t)	Recovered	Recovered	Recovered	Recovered	VIX
Q-prices	Bates	Bates	Jackwerth	Jackwerth	
Pricing kernel	Polynomial	Piecewise linear	Polynomial	Piecewise linear	

42

A Proofs

Proof of Proposition 1. 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 (1.17) and P is recovered from (1.12). 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 re-

spect 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[16] on $F(X)$. $\square$

Proof of Proposition 2. Let $\bar{\Pi}$ be an $S \times S$ transition matrix corresponding to an irreducible matrix (as in Ross). Without loss of generality we assume that the current state is the first state. Since prices are generated by a Ross economy, the observed matrix Π of multiperiod prices is given as

$$
\Pi := \begin{pmatrix} (\bar{\Pi})_1 \\ (\bar{\Pi}^2)_1 \\ \vdots \\ (\bar{\Pi}^S)_1 \end{pmatrix}
$$

where $(\bar{\Pi})_1$ denotes the first row of $\bar{\Pi}$, $(\bar{\Pi}^2)_1$ is the first row of $\bar{\Pi}^2$, etc. We want to show that all solutions to the eigenvalue problem for $\bar{\Pi}$ give rise to solutions to our system (both the "correct solution" and the ones that, by the Perron-Frobenius theorem, do not generate viable solutions).

Observe that if $z = (z_1, \ldots, z_S)'$ is a (right) eigenvector of $\bar{\Pi}$ with corresponding eigenvalue δ, then

$$
\Pi z = (\delta z_1, \delta^2 z_2, \ldots, \delta^S z_S)'.
$$

If z is the eigenvector corresponding to the maximal eigenvalue of $\bar{\Pi}$, then we know that it is strictly positive. Generically, in the space of matrices $\bar{\Pi}$, the matrix is diagonaliz-

---

[16]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.

able with eigenvectors that contain no zeros and with distinct non-zero eigenvalues – in particular, it has full rank. Therefore, generically, even for the other eigenvectors, we have that the coordinates of z are non-zero, so we can normalize z to have first coordinate 1. Now let the Ross probability matrix be defined (as in Ross)

$$\bar{P} = \frac{1}{\delta} Diag^{-1}(z)\bar{\Pi}Diag(z) \qquad (1.40)$$

with corresponding multi-period probabilities given by

$$P := \begin{pmatrix} (\bar{P})_1 \\ (\bar{P}^2)_1 \\ \vdots \\ (\bar{P}^S)_1 \end{pmatrix}.$$

Note that since the rows of $\bar{P}$ sum to 1, so do rows of P. Further, using (1.40),

$$P = \begin{pmatrix} (\frac{1}{\delta^1} Diag(z)^{-1}\bar{\Pi}^1 Diag(z))_1 \\ \vdots \\ (\frac{1}{\delta^S} Diag(z)^{-1}\bar{\Pi}^S Diag(z))_1 \end{pmatrix} = \begin{pmatrix} (\frac{1}{\delta^1}\bar{\Pi}^1 Diag(z))_1 \\ \vdots \\ (\frac{1}{\delta^S}\bar{\Pi}^S Diag(z))_1 \end{pmatrix} = D^{-1}\Pi Diag(z),$$

where the second equality uses that $z_1 = 1$ and that we only consider the first rows, and the last equation uses our maintained notation $D = Diag(\delta, \ldots, \delta^S)$. We note that this equation is the same as our equation (1.12), which means that all solutions to Ross's eigenvalue problem for the matrix $\bar{\Pi}$ also appear as solutions to our equations. The fact that P generated from the Ross solution $\bar{P}$ is a solution to the generalized problem required no assumptions other than irreducibility, and this proves part 1 of the theorem.

To obtain uniqueness also of our solution, note that, generically, there are S eigenvectors for Ross's matrix from which a matrix P can be generated using (1.40). Each of these solutions can be used to generate a solution P to our problem, as shown above. The $S-1$ solutions are "fake" in the sense that they imply that some marginal utilities (elements in the eigenvector z above) are negative. Hence, these solutions are also fake in the context of the generalized recovery framework. Given that Ross's equations

yield a total of S possible solutions to our problem, of which $S - 1$ are fake, we have a unique viable solution (by Proposition 1) if we can ensure that Π_{12} has full rank.

This follows from the generic property of $\bar{\Pi}$ as being diagnonalizable with distinct, non-zero eigenvalues. In fact, we can show the stronger statement that Π has full rank: Consider the diagonalization of Ross's price matrix as $\bar{\Pi} = VZV'$, where $Z = \mathrm{diag}(z_1, ..., z_S)$ 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 $vZ^kV' = (v_1z_1^k, ..., v_Sz_S^k)V'$ so

$$\Pi = \begin{bmatrix} 1 & ... & 1 \\ \vdots & & \vdots \\ z_1^{T-1} & ... & z_S^{T-1} \end{bmatrix} \begin{bmatrix} v_1z_1 & & 0 \\ & \ddots & \\ 0 & & v_Sz_S \end{bmatrix} V' \tag{1.41}$$

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. Hence our set of equations can have no more than S solutions, and since $S - 1$ of these are "fake", we have unique recovery of the solution corresponding to Ross's solution also, generically.

To see how to derive $\bar{\Pi}$ in an economy where Π arises from a time-homogeneous Ross economy, note that the following equation set must hold:

$$\underbrace{\begin{bmatrix} (\Pi)_2 \\ \vdots \\ (\Pi)_S \end{bmatrix}}_{(S-1) \times S} = \underbrace{\begin{bmatrix} (\Pi)_1 \\ \vdots \\ (\Pi)_{S-1} \end{bmatrix}}_{(S-1) \times S} \bar{\Pi} \tag{1.42}$$

where $(\Pi)_i$ is the i'th row of Π. Further, using the notation from (1.14) for blocks of Π and denoting the first row of $\bar{\Pi}$ by $\bar{\Pi}_1$ and remaining rows by $\bar{\Pi}_2$, we can rewrite

this equation as

$$
\begin{bmatrix} (\Pi)_2 \\ \vdots \\ (\Pi)_S \end{bmatrix} = \begin{bmatrix} \Pi_{11} & \Pi_{12} \end{bmatrix} \begin{bmatrix} \bar{\Pi}_1 \\ \bar{\Pi}_2 \end{bmatrix}
\tag{1.43}
$$

Given that $\bar{\Pi}_1$ is known (because the one-period state prices from state 1 are observed), it is useful to further rewrite this system as

$$
\begin{bmatrix} (\Pi)_2 \\ \vdots \\ (\Pi)_S \end{bmatrix} - \Pi_{11}\bar{\Pi}_1 = \Pi_{12}\bar{\Pi}_2
\tag{1.44}
$$

Hence, when Π_{12} is full rank, the Ross price matrix $\bar{\Pi}_2$ can be derived uniquely and explicitly by pre-multiplying by $(\Pi_{12})^{-1}$. We have already shown in Part 2, that Π_{12} has full rank generically. If Π_{12} does not have full rank, there exists a non-zero vector $v \in R^{S-1}$ for which $\Pi_{12}v = 0$. In this case, if we start from a solution for which $\bar{\Pi}_2$ has strictly positive elements, we can pick $\epsilon > 0$ small enough that adding ϵv to a row of $\bar{\Pi}_2$ yields a perturbed matrix $\bar{\Pi}_2^{\epsilon}$ whose elements are also strictly positive. Clearly, $\bar{\Pi}_2^{\epsilon}$ also satisfies (1.44), and hence the Ross price matrix is not unique, showing part 3. $\quad\square$

Proof of Proposition 3. 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. $\quad\square$

Proof of Proposition 4. 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 $P = Q$. $\quad\square$

Proof of Proposition 5. The result follows from the following lemma. $\quad\square$

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

$$\frac{x - x^*}{\varepsilon} = \frac{\bar{x} - x^*}{\varepsilon} - df^{-1}\frac{f(\bar{x}) - f(x^*)}{\varepsilon} \to \Delta x - df^{-1}df\,\Delta x = 0 \qquad (1.45)$$

$\square$

Proof of Proposition 6. 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, $D =$

diag$(\delta, ..., \delta^T)$ and $H = \text{diag}(h_1(\theta), h_2(\theta), ..., h_S(\theta)))$, 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 (1.12), 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 $F(X)$.
$\square$

49

B   Details on Recovery in Mehra-Prescott

Let

$$
\Pi = \begin{bmatrix}
\pi_{0,1}^{0,d} & \pi_{0,1}^{1,u} & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
0 & 0 & \pi_{0,2}^{0,d} & \pi_{0,2}^{1,d} & \pi_{0,2}^{1,u} & \pi_{0,2}^{2,u} & 0 & \ldots & 0 & 0 & 0 & & 0\ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & \pi_{0,T}^{0,d} & \pi_{0,T}^{1,d} & \pi_{0,T}^{1,u} & \ldots & \pi_{0,T}^{T,u}
\end{bmatrix}
\tag{1.46}
$$

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

$$
h^{-1}(\gamma) = \begin{bmatrix} (y_1^0)^\gamma & (y_1^1)^\gamma & (y_2^0)^\gamma & (y_2^1)^\gamma & (y_2^1)^\gamma & (y_2^2)^\gamma & \ldots & (y_T^T)^\gamma \end{bmatrix}'
\tag{1.47}
$$

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:

$$
\min_{\gamma,\delta} g(\gamma,\delta) := \text{norm}\left(\Pi h^{-1}(\gamma) - \begin{bmatrix} \delta \\ \delta^2 \\ \vdots \\ \delta^T \end{bmatrix} \right)
\tag{1.48}
$$

$$
s.t. \quad \gamma \in \mathbb{R}_+
$$
$$
\delta \in (0,1]
$$

Based on the recovered (γ, δ) that solve this minimizition problem, we can recover the

natural probabilities from (1.33).

C   Computing State Prices Empirically

Before we can recover probabilities, we need to know the 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. We implement two methods; (i) the parametric stochastic volatility model of Bates (2000) and (ii) the non-parametric "Fast and Stable" method of Jackwerth (2004).

C.1   The Bates (2000) Stochastic Volatility Model with Jumps

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

$$dS_t/S_t = (r^f - d - \lambda \bar{k})dt + \sqrt{V_t}dZ_t + kdq_t \tag{1.49}$$

$$dV_t = (\alpha - \beta V_t)dt + \sigma_v \sqrt{V_t}dZ_{vt} \tag{1.50}$$

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:[17] On each

---

[17]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

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:

$$f(\tau; S_\tau) = \frac{1}{\pi} \int_0^\infty \left(\frac{S_\tau}{S_t}\right)^{-iu} \psi(\tau, u) du \qquad (1.51)$$

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 $r^f_{t,t+\tau}$:

$$\pi(\tau; S_\tau) = e^{-r^f_{t,t+\tau}} f(T; S_\tau) \qquad (1.52)$$

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 8.2.

C.2   The Jackwerth (2004) "Fast and Stable" Method

We are interested in converting a (noisy) sparse set of implied volatilities into a full risk-neutral distribution. In section C.1 we imposed a parametric form on the implied volatility surface through a stochastic volatility model with jumps. In this section we refrain from imposing any structure on implied volatilities, that is, we fit a non-parametric method to implied volatilities. The method we have chosen is the "Fast and Stable" method of Jackwerth (2004). This method has a single tuning parameter, λ, which simultaneously controls the smoothness of the function and the fit to observed

_____

option literature of focusing on Wednesdays, where among other reasons option liquidity is high.

implied volatilities. Clearly, there is a trade-off in choosing the value of the tuning parameter, which is: the smoother the function the worse the fit to observations. We therefore control the smoothness of the fit by imposing two conditions; (i) the estimated implied volatilities gives rise to a non-negative risk-neutral distribution, (ii) the risk-neutral distribution is unimodal in the range from 0.8 to 1.2 in moneyness (defined as S_t/S_0, the index level at time t relative to the current index level). Under these conditions we minimize the objective function:

$$\min_{\sigma_s} \frac{1}{2(S+1)} \sum_{s=1}^{S} \left(\sigma_s^{''}\right)^2 + \frac{\lambda}{2I} \sum_{i=1}^{I} (\sigma_i - \bar{\sigma}_i)^2 \tag{1.53}$$

Where S is the number of states. σ_s is the implied volatility associated with state s. $\sigma_s^{''}$ is the second derivative of the implied volatility function with respect to strike prices. $i = 1, ..., I$ is the index for the observed implied volatilities and $\bar{\sigma}_i$ is the i'th observed implied volatility. As seen from (1.53), if λ is high then the fit to observations will be good compared to when λ is low. We therefore choose the highest value of λ which satisfies our two conditions described above. See Jackwerth (2004) for further comments on the method.

Once a smooth function for the implied volatilities is obtained we can back out a risk-neutral distribution by evaluating the Black and Scholes (1973) formula in the estimated implied volatilities and then differentiate the resulting call function twice with respect to strike prices as explained in Breeden and Litzenberger (1978).

The Fast and Stable method estimates a single option maturity at a time. In the period from January 1996 until December 2015 we have at least 7 maturities on any given last trading day of the month. In the framework of Proposition 6 this allows us to parameterize the pricing kernel with up to 6 parameters and still obtain generalized recovery.

D   Pricing Kernels used in Empirical Analysis

Piecewise linear.   The inverse marginal utilities are piecewise linear over states. Given the initial state 1 at time 0 the τ-period inverse marginal utility ratio in state

s is[18]:

$$(h_s^\tau(\theta))^{-1} = B_s\theta \tag{1.54}$$

Here θ is an N-dimensional column vector and B_s is the $s'th$ row of the known $S \times N$ "design matrix" B. In our empirical implementation N is 5[19]. Interpreting the parameters $\theta_1,...,\theta_N$ we let the first parameter θ_1 determine the initial level of the inverse pricing kernel $H^{-1}e = B\theta$. The next parameter, θ_2, determines the initial slope of the first line segment. Similarly, θ_3 is the slope of the next line segment generated by $B\theta$.

We impose that $\theta_1,...,\theta_N \geq 0$ which means that the inverse pricing kernel is monotonically increasing or, equivalently, that the pricing kernel is monotonically decreasing 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 $N-2$ regions. These regions are chosen as follows. The lowest region ranges over states from $(1-2.5\text{VIX}_t)S_t$ to $(1-2\text{VIX}_t)S_t$ where S_t is the current (time t) level of the S&P 500 index. 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 $N-3$ 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 trading day of the month for the period 1/1996 to 12/2015.

With this in place we set up the following minimization problem

$$\min_{\theta,\delta} \text{norm}\left(D^{-1}\Pi B\theta - 1\right) \tag{1.55}$$

$$s.t. \quad \theta > 0$$

$$\delta \in (0,1]$$

Given a state price matrix Π and a design matrix B we estimate the θ and δ that best fit the model in a squared error sense. Once the marginal utilities and discount rate

---

[18]Notice again that $(h_s^\tau(\theta))^{-1} = \frac{1}{h_s^\tau(\theta)}$ and is *not* the inverse function.

[19]The lowest number of maturities with observed option prices in our sample is 7. Therefore, we can impose a structure on the pricing kernel with at most 6 parameters and hence N can at most be 5 because of the sixth parameter δ.

have been recovered, we back out the multi-period physical probabilities as

$$P = D^{-1}\Pi \operatorname{diag}(B\theta) \tag{1.56}$$

where D is a diagonal matrix with elements $D_{ii} = \delta^i$ and $\operatorname{diag}(B\theta)$ is a diagonal matrix with elements $\operatorname{diag}(B\theta)_{jj} = B_j\theta$ where B_j is the j'th row of B. We normalize P to have row sums of one, which is necessary since θ and δ are found from the minimization problem in (1.55) and not solved perfectly.

Polynomial. The inverse marginal utility ratio is a polynomial in the return on the market and time horizon. Given the initial state 1 at time 0 the τ-period inverse marginal utility ratio in state s is:

$$(h_s^\tau(\theta))^{-1} = \beta_0 + \beta_1 r_s + \beta_2 r_s^2 + \beta_3 \tau r_s + \beta_4 \tau r_s^2 \tag{1.57}$$

Here $r_s = S_s/S_1 - 1$ is the return on the market in state s. The parameters of interest are $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)$. In our implementation we impose three conditions on the parameters; (i) $\beta_0 > 0$, ensuring a positive pricing kernel when $r = 0$, (ii) the risk-premium is non-negative and, (iii) the inverse marginal utility ratios are always strictly positive (we set a lower bound on the inverse marginal utility ratio at 0.01.). This means that the parameters $\beta_1, \beta_2, \beta_3, \beta_4$ can move freely (within the space of the conditions) and are all allowed to be either positive or negative.

The polynomial specification of the inverse marginal utility ratios illustrates one possible way of imposing structure on the marginal utilities, not only in the state dimension, but also in the time horizon dimension. This specification allows marginal utilities in a given state, say s, to differ when considering different time horizons, that is, e.g. $h_s^\tau(\theta) \neq h_s^{\tau+1}(\theta)$. The polynomial specification nests the linear specification as a special case when $\beta_2, \beta_3, \beta_4$ are all zero.

The minimization procedure for the polynomial specification is:

$$\min_{\theta,\delta} \sum_{t=1}^{T} \left(\left(\sum_{s=1}^{S} \delta^{-t} \pi_{ts} (h_s^\tau(\theta))^{-1} \right) - 1 \right)^2 \tag{1.58}$$

$$s.t. \quad \beta_0 > 0$$

$$E_0^P(r_t|\theta,\delta) - r_t^f \geq 0 \quad \text{for all } t \in (1, ..., T)$$

$$(h_s^\tau(\theta))^{-1} > 0 \quad \text{for all } s \in (1, ..., S) \text{ and all } \tau \in (1, ..., T)$$

$$\delta \in (0, 1]$$

where π_{ts} is the state price in state s with time horizon t. Here $E_0^P(r_t|\theta,\delta) - r_t^f$ is the excess return given parameter values $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)$, and δ.

Given estimates of $\delta, \beta_0, \beta_1, \beta_2, \beta_3$ and β_4 we can arrive at the t period physical probabilities as

$$P_t = \delta^{-t} \Pi_t \, \text{diag}\left((h_s^\tau(\theta))^{-1}\right) \tag{1.59}$$

where Π_t is the t'th row of the state price matrix Π and r is an $S \times 1$-dimensional vector of returns over states. We normalize P to have row sums of one, this is necessary since θ and δ are found from the minimization problem in (1.58) and not solved perfectly.

Chapter 2

Higher-Moment Risk

Co-authored with Niels Joachim Gormsen

Abstract:
We show how the market's higher order moments can be estimated *ex ante* using methods based on Martin (2017). These ex ante higher order moments predict future realized higher order moments, whereas trailing realized moments have little predictive power. Higher-moment risks move together in the sense that skewness becomes more negative when kurtosis becomes more positive. In addition, higher-moment risk is high when volatility is low, suggesting that risk doesn't go away – it hides in the tails. Higher-moment risk has significant implications for investors; for example, the tail loss probability of a volatility-targeting investor varies from 3.6% to 9.7%, entirely driven by changes in higher-moment risk. We empirically analyze the economic drivers of these risks, such as financial intermediary leverage, market and funding illiquidity, and potential bubbles.

1   Introduction

Times of financial market distress pose threats to the macroeconomy, as we witnessed in the 2008-2009 financial crisis. For policymakers to act in a timely and preemptive manner in the event of financial market distress, it is important to measure the perceived tail risks in real time.

In this paper, we estimate higher-moment risk in real time using a new method and arrive at the following five main results: (1) Moments of the market return, measured ex ante using option prices, predict future realized moments. (2) Higher order moments co-move in the sense that skewness (3rd moment) and hyperskewness (5th moment) become more negative when kurtosis (4th moment) and hyperkurtosis (6th moment) become more positive. In other words, there are times when higher-moment risk is high, in the sense that the return distribution is both substantially left-skewed (due to large negative odd-numbered moments) and fat tailed (due to large positive even-numbered moments). (3) Higher-moment risks tend to be high after market run-ups where the variance is low. (4) Higher-moment risk has important implications for investors; for example, the tail loss probability of a volatility-targeting investor varies from 3.6% to 9.7%, entirely driven by changes in higher-moment risk. (5) The times when higher-moment risks are high are characterized by high market and funding liquidity, high turnover, and low expected future returns.

Our analysis is based on ex ante moments that are estimated from options prices. Using methods based on Martin (2017), we translate risk-neutral moments into physical moments as perceived by an unconstrained power utility investor who wants to hold the market portfolio. Using S&P 500 as a proxy for the market portfolio, we estimate ex ante monthly and quarterly moments. These moments are entirely forward looking and, unlike risk-neutral moments, contain no adjustment for risk, which makes them well suited for studying time-variation in higher-moment risk.

As our first main result, we show that our ex ante moments are positively correlated with ex post realized moments. Consistent with previous research, our ex ante variance predicts ex post realized variance well.[1] More importantly, we show that our ex ante

---

[1]Previous literature has shown that ex post realized variance is well predicted by historical variance or option implied variance, e.g. Bollerslev, Tauchen, and Zhou (2009), Andersen, Fusari, and Todorov (2015), and Bollerslev, Hood, Huss, and Pedersen (2016).

higher order moments also predict ex post higher order moments. We show that our ex ante skewness, kurtosis, hyperskewness, and hyperkurtosis all have significant predictive power over ex post realized moments. We further show that our ex ante moments are better at forecasting ex post realized moments than their trailing (lagged) moments.

Next, we show that these predictability results are robust in several ways. First, we show that our results are not driven by the large price moves that occurred during the financial crisis of 2008 to 2009. Second, we show that our moment prediction holds even when controlling for risk-neutral moments. The latter is important because option-implied risk-neutral skewness has been shown to predict ex post realized skewness, e.g. Neuberger (2012).

As our second main result, we find that higher order moments move together in the sense that skewness and hyperskewness are more negative at times when kurtosis and hyperkurtosis are more positive. Indeed, we find that skewness is negatively correlated with kurtosis with a correlation coefficient of -0.80, a negative correlation of -0.66 with hyperkurtosis, and a positive correlation of 0.79 with hyperskewness. These co-movements in higher order moments are so strong that the first principal component of the space spanned by skewness, kurtosis, hyperskewness, and hyperkurtosis explains 90% of the joint variation in higher order moments.

The first principal component eigenvector has the same signs for skewness and hyperskewness, while the sign is opposite for kurtosis and hyperkurtosis. As shown in Ebert (2013), an investor with power utility has preferences for odd number moments of any order and is averse to even number moments of any order. A high value of the first principal component can therefore be interpreted as times when higher order moment risks are, on average, large (negative for odd moments and positive for even moments). We therefore define the first principal component as a *higher-moment risk index* (HRI).

As our third main result, we find that higher-moment risk varies systematically with variance. Specifically, the correlation between variance and the HRI is -0.53 with 95% bootstrapped confidence bounds of $[-0.60, -0.48]$, which emphasizes that higher-moment risks tend to be high at times when variance is low. In addition, we find that higher-moment risks tend to be high subsequent to market run-ups, which

are usually "calm" times as measured by variance. We find that the HRI is positively related to the past two year return. The relation is statistically significant at a 99% level, showing that the return distribution is more left skewed and fat tailed subsequent to a "good" period where prices have increased significantly.

Fourth, we show that higher-moment risk has large economic implications for investors. To understand the importance of higher-moment risk, we study the portfolio risk of a volatility-targeting investor who holds a portfolio of cash and the market. The investor adjusts the portfolio weights to achieve a constant volatility of $\sigma^{\text{vol target}}$. Despite having constant variance, the riskiness of the portfolio varies substantially over time as higher moment risk varies. Because higher moment risk is high when variance is low, the portfolio is the riskiest when market variance is low.

To understand the economic magnitude of the systematic variation in higher-moment risks, we estimate the probability that the return on the volatility-targeting investor's portfolio is less than $-2\sigma^{\text{vol target}}$. The monthly probability peaked on June 30th 2014 with a probability of 9.7%, almost three times the size of its low, on February 27th 2008, where the probability was 3.6%. Furthermore, the average probability of a $-2\sigma^{\text{vol target}}$ event is 6.6%, which is large compared to the 2.5% that is implied by a normal distribution. Similarly, the probability of a portfolio return that is less than $-3\sigma^{\text{vol target}}$ peaked on November 30th 2006 with a probability of 3.6%, which is four times the size of its low on February 27th 2008, when the probability was 0.76%. These probabilities are also far above what is implied by a normal distribution, which is 0.13%.

Furthermore, we find that the probability of a portfolio return that is less than $-2\sigma^{\text{vol target}}$ for the volatility-targeting investor is negatively correlated with variance with a correlation coefficient of -0.70 and 95% bootstrapped confidence bounds of $[-0.78, -0.65]$. This strong negative correlation further emphasizes the importance of considering higher-moment risks in portfolio choice problems. For example, this finding can help explain why Moreira and Muir (2017b) find that investors can earn high Sharpe-ratios by moving wealth into the market at times when variance is low and moving wealth out of the market when variance increases. The relative (to variance) high expected return in calm times may be compensation for elevated higher-moment risks.

60

Our fifth main result shows how higher-moment risk is associated with several economic drivers. First, our results are closely related to the volatility paradox (Bunnermeier and Sannikov, 2014), which is the notion that systematic risk is high when variance is low. In their model, risk increases when variance is low because specialized investors are more levered. We therefore investigate how the level of financial intermediary leverage is associated with higher-moment risk. In particular, we test if financial intermediaries are more levered when variance is low, and if such variation in financial intermediary leverage can explain our observed variation in higher moment risk. Using the measure of financial intermediary leverage from He, Kelly, and Manela (2016), we find no relation between higher-moment risks and aggregate financial intermediary leverage.

We next investigate how higher-moment risk is related to market illiquidity and funding illiquidity. We find that higher-moment risks are positively associated with both market and funding liquidity. Specifically, using the average value-weighted bid-ask spread of S&P 500 constituents as a proxy for market illiquidity, we find that times when the average bid-ask spread is low are times when higher-moment risks are high. Similarly, using the TED spread as a proxy for funding illiquidity, we find that a low TED spread is associated with high higher-moment risks.

Lastly, we investigate how higher-moment risks are related to previously suggested measures of "bubble" characteristics and market valuation. We consider the "bubble" characteristics: acceleration (Greenwood, Shleifer, and You (2017)), turnover (Chen, Hong, and Stein (2001)), issuance percentage (Pontiff and Woodgate (2008)), and the market valuation measures: CAPE, the dividend-price ratio, and cay (Lettau and Ludvigson (2001)). We find that higher-moment risk is positively related to price acceleration: there is more higher-moment risk when the recent price path is more convex. Also, higher turnover after market run-ups is associated with more higher-moment risk. Furthermore, there is more higher-moment risk when cay (Lettau and Ludvigson, 2001) is high. We find no conclusive relation between higher-moment risks and CAPE, the dividend-price ratio, or equity issuance.

Our paper relates to and extends the existing literature on estimating time-varying market tail risk by integrating two different approaches. Previous research on tail risk is based on either (1) physical moments based on backward looking information

61

or (2) risk-neutral moments based on forward looking option prices. We show that physical higher-moment risks can be estimated in a forward looking manner, and in real time, which complements the existing literature that uses historical (backward looking) returns to estimate tail risks; e.g., using realized returns, Bollerslev and Todorov (2011) suggest using high frequency intraday returns and fit an extreme value distribution to the tails of returns. Also, Kelly and Jiang (2014) estimate market wide tail risks from the cross-section of firm-level returns. Our paper also relates to the literature that studies tail risk using option prices. However, while the existing literature studies tail risk using risk-neutral moments (e.g. Siriwardane (2015), Gao, Gao, and Song (2017), Gao, Lu, and Song (2017), Bates (2000), and Schneider and Trojani (2017b)), we study tail risk using physical moments. Thereby, we can investigate physical tail probabilities and study which economic drivers can explain the time-varying patterns in higher-moment risks.

In summary, higher-moment risks can be measured in real time, and a single factor explains 90% of the joint variation in higher order moments. Furthermore, times when higher-moment risks are high are characterized by: (1) low variance, (2) large (and accelerating) recent price run-ups, (3) low market and funding frictions, (4) high turnover, and (5) low future expected returns.

The paper proceeds as follows: Section 2 covers the theory behind how we estimate higher order moments and tail probabilities. Section 3 covers the data and the empirical implementation. Section 4 investigates the relation between our ex ante moments and ex post realized moments. Section 5 studies the commonalities in higher order moments. Section 6 investigates the systematic patterns in higher-moment risks. Section 7 studies the implications of time-varying higher-moment risks for investors. Section 8 studies the economic drivers of higher-moment risks. Section 9 concludes the paper.

2   Inferring Ex Ante Moments from Asset Prices

We consider an economy where agents can trade two assets, a risk-free asset and a risky asset. The risk-free asset earns a gross risk-free rate of return $R_{t,T}^{f}$ between time t and time T. The risky asset has a price of S and earns a random gross return $R_{t,T}$. The risky asset pays dividends, $D_{t,T}$, between time t and time T such that its gross

return is $R_{t,T} = (S_T + D_{t,T})/S_t$.

Starting from the standard asset pricing formula, we can relate risk-neutral and physical expected values of the time T random payoff, X_T, as

$$E_t[X_T m_{t,T}] = E_t^*[X_T]/R_{t,T}^f \qquad (2.1)$$

where the asterisk denotes risk-neutral expectation and $m_{t,T}$ is a stochastic discount factor. If we define the time T random payoff, $X_{t,T}(n)$, in the following way

$$X_{t,T}(n) = R_{t,T}^n m_{t,T}^{-1} \qquad (2.2)$$

then equation (2.1) implies that the n'th moment of the risky asset's physical return distribution can be expressed in terms of the risk-neutral expectation of $X_{t,T}(n)$:

$$E_t[R_{t,T}^n] = E_t[\underbrace{R_{t,T}^n m_{t,T}^{-1}}_{X_{t,T}(n)} m_{t,T}] = E_t^*[\underbrace{R_{t,T}^n m_{t,T}^{-1}}_{X_{t,T}(n)}]/R_{t,T}^f \qquad (2.3)$$

So if we know the pricing kernel m, then we can derive all moments of $R_{t,T}$ directly from risk-neutral pricing of the claim to $X_{t,T}(n)$. Following Martin (2017), we compute the physical expected value of $R_{t,T}^n$ from the point of view of an unconstrained rational power-utility investor who chooses to be fully invested in the market. This investor has initial wealth W_0 and terminal wealth $W_T = W_0 R_{t,T}$. Given the investor's utility function, $U(x) = x^{1-\gamma}/(1-\gamma)$, with relative risk-aversion, γ, we can determine the investor's stochastic discount factor. Specifically, combining the first order condition from the investor's portfolio choice problem with the fact that the investor holds the market, the stochastic discount factor becomes proportional to $R_{t,T}^{-\gamma}$:

$$m_{t,T} = k R_{t,T}^{-\gamma} \qquad (2.4)$$

for some constant k which is unobservable to us. However, we do not need to learn k to estimate physical moments; we can correct for k by rewriting (2.3) in the following way. First, setting $n = 0$ in (2.2) we get $X_{t,T}(0) = m_{t,T}^{-1}$ and the standard asset pricing

formula (2.1) then implies the relation:

$$E_t^*[m_{t,T}^{-1}] = R_{t,T}^f \qquad (2.5)$$

Then, inserting (2.5) and (2.4) into (2.3), we obtain an expression of the n'th physical moment perceived by an unconstrained rational power utility investor who chooses to be fully invested in the market:

$$E_t[R_{t,T}^n] = \frac{E_t^*[R_{t,T}^n \overbrace{R_{t,T}^{\gamma}/k}^{m_{t,T}^{-1}}]}{E_t^*\underbrace{[R_{t,T}^{\gamma}/k]}_{m_{t,T}^{-1}}} = \frac{E_t^*[R_{t,T}^{n+\gamma}]}{E_t^*[R_{t,T}^{\gamma}]} \qquad (2.6)$$

since k is a constant.

The relation between physical and risk-neutral moments shown in (2.6) is central to our empirical analysis. The key insight is that we can estimate the n'th physical moment directly from risk-neutral pricing of $R_{t,T}^{\gamma}$ and $R_{t,T}^{n+\gamma}$. Furthermore, by pricing claims to the payoffs $R_{t,T}^{m+\gamma}$ for $m \in \{1,...,n\}$, we can then estimate *standardized* moments.

To understand how we estimate standardized moments from (2.6), recall the notion of the n'th standardized moment formula:

$$n\text{'th standardized moment of } R_{t,T} = E_t\left[\left(\frac{R_{t,T} - E_t[R_{t,T}]}{\text{Var}[R_{t,T}]^{1/2}}\right)^n\right] \qquad (2.7)$$

Expanding (2.7) and replacing physical moments with risk-neutral counterparts as presented in equation (2.6), we can arrive at expressions for all physical standardized moments as functions of risk-neutral moments. For example, the third standardized physical moment (skewness) can be expressed in terms of risk-neutral moments by first expanding (2.7) with $n = 3$:

$$\text{Skewness}_{t,T} = \frac{E_t[R_{t,T}^3] - 3E_t[R_{t,T}]E_t[R_{t,T}^2] + 2E_t[R_{t,T}]^3}{(E_t[R_{t,T}^2] - E_t[R_{t,T}]^2)^{3/2}} \qquad (2.8)$$

and then replacing the physical moments in (2.8) with the risk-neutral counterparts using equation (2.6). Similar expressions can be written up for other higher order

moments of interest, as seen in Appendix A. Importantly, the right-hand-side of (2.6) consists of asset prices which can be estimated directly from current and observable call and put options written on the risky asset. Hence, higher order moments can be estimated in real time, without using historical realized returns or accounting data.

2.1 Inferring Ex Ante Market Tail Probabilities

Next, we show how we estimate ex ante tail probabilities from option prices written on the market. To understand our approach, note first that the probability at time t of a market return that is lower than α at time T can be written as the physical expectation of an indicator function in the following way

$$P_t(R_{t,T} < \alpha) = E_t[1_{\{R_{t,T} < \alpha\}}] \tag{2.9}$$

Using the standard asset pricing formula in (2.1), we can rewrite the probability in terms of the risk-neutral measure by adjusting the right hand side of equation (2.9) for the inverse of the stochastic discount factor in (2.4)

$$P_t(R_{t,T} < \alpha) = \frac{E_t^*[R_{t,T}^\gamma 1_{\{R_{t,T} < \alpha\}}]}{E_t^*[R_{t,T}^\gamma]} \tag{2.10}$$

The right hand side of (2.10) is an asset price that has the simple representation presented in Proposition 8, which generalizes Result 2 in Martin (2017) from log-utility to general power utility for any level of relative risk-aversion.

Proposition 8. *For the unconstrained rational power utility investor who wants to hold the market, the conditional physical probability that market return from time t to T is lower than α is:*

$$P_t(R_{t,T} < \alpha) = \frac{R_{t,T}^f}{E_t^*[R_{t,T}^\gamma]} \left[\alpha^\gamma put'_{t,T}(\alpha S_t - D_{t,T}) - \frac{\gamma}{S_t} \alpha^{\gamma-1} put_{t,T}(\alpha S_t - D_{t,T}) \tag{2.11} \right.$$

$$\left. + \int_0^{\alpha S_t - D_{t,T}} \frac{\gamma(\gamma-1)}{S_t^2} \left(\frac{K + D_{t,T}}{S_t} \right)^{\gamma-2} put_{t,T}(K)dK \right] \tag{2.12}$$

where $put'_{t,T}(\alpha S_t - D_{t,T})$ is the first derivative of the put option price with strike $\alpha S_t - D_{t,T}$.

65

Proof. The results of Breeden and Litzenberger (1978) imply the equality

$$E_t^*[R_{t,T}^\gamma 1_{\{R_{t,T}<\alpha\}}] = R_{t,T}^f \int_0^\infty \left(\frac{K+D_{t,T}}{S_t}\right)^\gamma 1_{\{K<\alpha S_t - D_{t,T}\}} \mathrm{put}_{t,T}''(K)dK \qquad (2.13)$$

where $\mathrm{put}_{t,T}''(K)$ is the second derivative of the put option price written on the underlying process S. Splitting the integral at $\alpha S_t - D_{t,T}$ we have

$$E_t^*[R_{t,T}^\gamma 1_{\{R_{t,T}<\alpha\}}] = R_{t,T}^f \int_0^{\alpha S_t - D_{t,T}} \left(\frac{K+D_{t,T}}{S_t}\right)^\gamma \mathrm{put}_{t,T}''(K)dK \qquad (2.14)$$

Proposition 8 then follows from using integration by parts twice. $\qquad\square$

3 Data and Empirical Implementation

We use the Ivy DB database from OptionMetrics to collect information on call options and put options that are written on the S&P 500 index for the last trading day of every month. The data ranges from January 1996 to December 2015. We obtain implied volatilities, strikes, closing bid-prices, closing ask-prices, and maturities. We proxy the risk-free rate with the zero-coupon yield curve from the Ivy DB database, which is derived from the LIBOR rates and settlement prices of CME Eurodollar futures. We also obtain expected dividend payments. We consider options with times to maturity between 10 and 360 calender days, and apply common filters, excluding contracts with zero open interest, zero trading volume, quotes with best bid below \$0.50, and options with implied volatility higher than 100%.

We use daily realized returns to estimate realized daily moments. We also estimate monthly moments from monthly returns. In Appendix A, we discuss the estimation of realized moments in detail.

3.1 Estimating Market Moments

There is a large body of literature devoted to pricing asset derivatives such as those in (2.6), using observable option prices written on the asset. Indeed, Breeden and Litzenberger (1978), Bakshi and Madan (2000), and Bakshi, Kapadia, and Madan (2003) show that the arbitrage free price of a claim on some future (twice differentiable) payoff can be expressed in terms of a continuum of put and call option prices. Specifically

for our purposes, using the results of Breeden and Litzenberger (1978), Martin (2017) shows that we can write the n'th physical moment of $R_{t,T}$ as

$$E_t[R_{t,T}^n] = \frac{E_t^*[R_{t,T}^{n+\gamma}]}{E_t^*[R_{t,T}^\gamma]} = \frac{(R_{t,T}^f)^{n+\gamma} + R_{t,T}^f\,[p(n+\gamma) + c(n+\gamma)]}{(R_{t,T}^f)^\gamma + R_{t,T}^f\,[p(\gamma) + c(\gamma)]} \tag{2.15}$$

with

$$p(\theta) = \int_0^{F_{t,T}} \frac{\theta(\theta-1)}{S_t^\theta}\left(S_t R_{t,T}^f - F_{t,T} + K\right)^{\theta-2} \mathrm{put}_{t,T}(K)dK \tag{2.16}$$

$$c(\theta) = \int_{F_{t,T}}^\infty \frac{\theta(\theta-1)}{S_t^\theta}\left(S_t R_{t,T}^f - F_{t,T} + K\right)^{\theta-2} \mathrm{call}_{t,T}(K)dK \tag{2.17}$$

where $F_{t,T}$ is the forward price and $\mathrm{call}_{t,T}(K)$ and $\mathrm{put}_{t,T}(K)$ are call and put option prices written on the risky asset at time t with horizon $T-t$ and strike K.

In practice, we do not observe a continuum of call and put options and therefore (2.15) must be numerically approximated. Let $F_{t,T}$ be the forward price and, using the notation from Martin (2017), we can write the price, $\Omega_{t,T}(K)$, at time t of an out-of-the money option with strike K and maturity T as

$$\Omega_{t,T}(K) = \begin{cases} \mathrm{call}_{t,T}(K) & \text{if } K \geq F_{t,T} \\ \mathrm{put}_{t,T}(K) & \text{if } K < F_{t,T} \end{cases} \tag{2.18}$$

We let $K_1, ..., K_N$ be the (increasing) sequence of observable strikes for the N out-of-the money put and call options and define $\Delta K_i = \frac{K_{i+1}-K_{i-1}}{2}$ with

$$\Delta K_i = \begin{cases} K_{i+1} - K_i & \text{if } i = 1 \\ K_i - K_{i-1} & \text{if } i = N. \end{cases} \tag{2.19}$$

We approximate the integrals in (2.16) by observable sums such that the n'th physical moment becomes:

$$E_t[R_{t,T}^n] = \frac{(R_{t,T}^f)^{n+\gamma} + R_{t,T}^f\left[\sum_{i=1}^N \frac{(n+\gamma)(n+\gamma-1)}{S^{n+\gamma}}(S_t R_{t,T}^f - F_{t,T} + K_i)^{n+\gamma-2}\Omega_{t,T}(K_i)\Delta K_i\right]}{(R_{t,T}^f)^\gamma + R_{t,T}^f\left[\sum_{i=1}^N \frac{\gamma(\gamma-1)}{S^\gamma}(S_t R_{t,T}^f - F_{t,T} + K_i)^{\gamma-2}\Omega_{t,T}(K_i)\Delta K_i\right]} \tag{2.20}$$

67

In summary, combining equation (2.20) with the standardized moment formula in equation (2.7), we can express standardized physical moments in terms of the derivatives prices written on the risky asset.

When we estimate physical moments for a given horizon, say T, for which we do not observe put and call prices, we linearly interpolate the (standardized) moments between the two closest horizons available in the data. In a few cases, we need to extrapolate to obtain moments for the desired horizon.

Our benchmark investor has power utility and a coefficient of relative risk-aversion of 3, that is, $\gamma = 3$. This level of risk-aversion as the benchmark is motivated by the results of Bliss and Panigirtzoglou (2004), i.e., using our sample we replicate their results and find that 3 is the optimal option-implied level of risk aversion when matching realized returns at the monthly horizon. We also estimate moments for the risk-neutral investor, the log-utility investor, and the power utility investor with a risk-aversion coefficient of 5.

Figure 2.1 shows monthly higher order moments and Table 2.1 shows the moment summary statistics. The average ex ante estimated skewness is negative for both horizons and all levels of risk aversion, suggesting that the physical distributions are left skewed. Consistent with the results of Neuberger (2012), we find that average skewness is not diminishing in the horizon, in the sense that skewness is close to the same on a monthly and quarterly horizon. Similarly, average kurtosis is larger than 3 for both horizons and all levels of risk aversion, which means that the physical distributions are leptokurtic; that is, the tails of the physical return distributions are fatter than what is implied by a normal distribution.

3.2 Estimating Market Tail Probabilities

The main challenge when implementing Proposition 8 is that we are required to estimate the first derivative of the put option price written on the risky asset at strike $\alpha S_t - D_{t,T}$. To handle a sparse and discrete set of observed option prices, we smoothen observed option prices using a Gaussian kernel smoothening procedure. Specifically, we smoothen implied volatilities around the strike $\alpha S_t - D_{t,T}$ and choose the kernel bandwidth to minimize the squared errors between the observed and estimated im-

plied volatilities under the constraint that the estimated option prices do not allow for arbitrage.

Given a smooth set of option prices around the strike $\alpha S_t - D_{t,T}$, we compute the first derivative as the slope between the two adjacent prices:

$$\text{put}'_{t,T}(\alpha S_t - D_{t,T}) = \frac{\text{put}_{t,T}(\alpha S_t - D_{t,T} + h) - \text{put}_{t,T}(\alpha S_t - D_{t,T} - h)}{2h} \qquad (2.21)$$

where h is the chosen grid step size in the discretization.

Let $K_1, ..., K_M$ be the (increasing) sequence of observable strikes for the M out-of-the money put options where K_M is the observed strike that is closest to $\alpha S_t - D_{t,T}$. We approximate the integral in Proposition 8 by the observable sum:

$$\sum_{i=1}^{M} \frac{\gamma(\gamma - 1)}{S_t^2} \left(\frac{K_i + D_{t,T}}{S_t} \right)^{\gamma - 2} \text{put}_{t,T}(K_i) \Delta K_i \qquad (2.22)$$

Inserting (2.21) and (2.22) into Proposition 8, we can estimate physical probabilities.

4   Estimated Moments Predict Realized Moments

In this section, we show that the ex ante higher order moments estimated using the methods described in Sections 2 and 3 predict ex post realized higher order moments.

We start with a simple sorting exercise. For each moment, we first sort ex post realized monthly returns into a "low" or "high" bucket depending on whether the ex ante moment is lower or higher than its median time series value. Next, we estimate the ex post moments for each bucket; for example, we estimate moments using the monthly ex post returns sorted into the "high" bucket. Figure 2.2 shows the monthly ex post realized moments of the two buckets for all moments. The ex post realized returns sorted into the "high" buckets exhibit in-sample higher moment values, suggesting that our ex ante moments predict ex post moments, for example, the "high" bucket for kurtosis has an in-sample kurtosis of 5.93, while the "low" bucket has a kurtosis of 2.85.

Next, we test more formally the relation between ex ante and ex post moments. Specifically, we conduct two tests which differ in the way we estimate ex post realized higher order moments. First, we test if the bucket values following our sorting exercise

are extreme compared to what a random sample would produce. For each moment, we bootstrap a distribution using permutations and then evaluate where in this distribution our observed "low" and "high" bucket values lie. Panel A of Table 2.2 reports the values from the ex ante sorting and significance, which is computed from the bootstrapped distribution in the following way: for the "low" buckets, we estimate the frequency at which a random permutation lies below what we observe. For the "high" buckets, we estimate the frequency at which a random permutation lies above what we observe. For example, the -0.83 value for skewness in the "low ex ante" bucket is not in the lower 10% of the bootstrapped distribution and is therefore insignificant at a 90% level. However, the 5.93 value for kurtosis in the "high ex ante" bucket is in the upper 5% of the bootstrapped distribution for kurtosis and is therefore significant at a 95% level. Importantly, our ex ante moments show statistical significance at a 95% level at least once, for every moment except skewness. Comparing these results to the results we get when sorting the ex post realized returns into two buckets based on the trailing monthly moments (estimated using daily returns), we find that our ex ante moments clearly outperform.

Second, we estimate time-varying ex post monthly (and quarterly) realized moments using daily returns; that is, for a given month, we estimate the in-sample moments for that month using the daily returns during that month. The first two columns of Panel B of Table 2.2 report correlations between our ex ante moments and the ex post realized moments. The latter two columns report correlations between ex post realized moments and their trailing (lagged) moment. We report bootstrapped standard errors in the appendix.

Correlations between our ex ante variances and ex post variances are 49% to 67%, and these correlations are both statistically significant at a 99% level. We also find strong correlations between ex ante and ex post skewness, ranging from 21% to 25%, which are both significantly different from zero at a 99% level. Correlations of our ex ante and ex post hyperskewness are positive and significant at the 99% level. Comparing the correlations of our ex ante moments to those of the trailing moments we find that, on a monthly horizon, trailing moments do not predict either skewness or hyperskewness whereas our ex ante moments do. On a quarterly horizon, trailing moments do predict ex post realized moments, however the correlations are lower than for our

70

ex ante moments.

Neither our ex ante moments nor the trailing moments seem to be able to predict ex post kurtosis or hyperkurtosis. For our ex ante moments, this might be because of the fact that there are fewer available option prices in the right tail of the distribution, that is, deep out-of-the-money call options are traded less frequently than deep out-of-the-money put options. We therefore test if our ex ante kurtosis (and hyperkurtosis) can predict left kurtosis, which is for our purposes the important tail of the distribution to be able to predict. Therefore, we follow Denbee, Julliard, Li, and Yuan (2016), and estimate ex post realized left kurtosis in the following way:

$$\text{Realized left kurtosis}_{t,T} = \frac{\sum_s \left(\frac{\text{Daily return}_s - \text{Realized daily mean}_{t,T}}{\text{Realized daily variance}_{t,T}^{1/2}} \right)^4}{\text{Realized kurtosis}_{t,T}} \tag{2.23}$$

where s is the days in the month where Daily return $<$ Realized daily mean$_{t,T}$. The realized right kurtosis is defined in the obvious way, where daily returns are larger than the realized mean.

Panel C of Table 2.2 shows the correlations between our ex ante kurtosis and the ex post realized left kurtosis and left hyperkurtosis. Both on a monthly and quarterly horizon, our ex ante kurtosis and hyperkurtosis are positively and statistically significantly correlated to ex post realized left kurtosis and left hyperkurtosis. This result should be interpreted in the following way: times when our ex ante kurtosis is high are times when the ex post realized kurtosis can be attributed primarily to the left tail of the return distribution. Comparing the correlations between our ex ante moments and ex post realized left kurtosis and left hyperkurtosis to the correlations between trailing moments and ex post left kurtosis and left hyperkurtosis, we find that while monthly trailing moments do not predict ex post moments, quarterly trailing moments do predict ex post realized left kurtosis and hyperkurtosis, but the correlations are smaller than for our ex ante moments.

Overall, Figure 2.2 and Panel A, B, and C of Table 2.2 show that our ex ante moments predict ex post realized moments. It is natural to worry that the results are driven by the large price moves that occurred during the period of financial distress from 2008 to 2009. To address this concern, Panel A of Table 2.3 shows correlations be-

tween our ex ante moments and ex post realized moments when removing observations that overlap with the period August 1st 2008-July 31st 2009. The results are largely unchanged, suggesting that the financial crisis does not drive the strong predictive results.

Panel B of Table 2.3 shows th correlations for other levels of risk-aversion. The results from the point of view of a log-utility investor or a power-utility investor with a risk-aversion of 5 are not remarkably different from the results presented in Panel B of Table 2.2 for the power-utility investor with a risk-aversion of 3.

As a second robustness test of moment predictability, we ask if physical higher order moments predict ex post realized moments when controlling for risk-neutral moments. Another way to put it is to ask: do we gain anything in terms of predictability for moving from risk-neutral to physical moments? Table 2.3 shows the results of the following two-stage procedure. In the first stage, we run the two regressions:

$$\text{Realized Moment}_{t,T} = \alpha_1 + \beta_1 E_t^*[\text{Moment}_{t,T}] + \epsilon_{t,T} \tag{2.24}$$

$$E_t[\text{Moment}_{t,T}] = \alpha_2 + \beta_2 E_t^*[\text{Moment}_{t,T}] + \eta_{t,T} \tag{2.25}$$

where $E_t[\text{Moment}_{t,T}]$ is our ex ante physical moment and $E_t^*[\text{Moment}_{t,T}]$ is the corresponding risk-neutral moment. The residuals, ϵ and η, are by construction orthogonal to risk-neutral moments, and their correlation therefore determines whether physical moments can explain the variation in realized moments in excess of what is explained by risk-neutral moments. In the second stage we estimate the correlation between ϵ and η. The first two columns of Panel C of Table 2.3 report these correlations, and bootstrapped standard errors that correct for the generated regressor problem we face when estimating the residuals in the first stage regressions are in the appendix.

The correlations between ϵ and η on a monthly horizon range from 0.09 to 0.16 and are statistically significant at a 95% level for kurtosis, hyperskewnes, and hyperkurtosis, implying that our monthly ex ante moments still predict ex post realized moments when controlling for risk-neutral moments. The results are weaker for quarterly moments; only hyperskewness is statistically significant and positive.

As a third robustness test of predictability, we test if our ex ante estimated higher order moments predict ex post realized moments when controlling for trailing (lagged)

moments. We therefore repeat the two-stage procedure described above. In the first stage we run the following two regressions

$$\text{Realized Moment}_{t,T} = \alpha_3 + \beta_3 \text{Realized Moment}_{t-(T-t),t} + \kappa_{t,T} \tag{2.26}$$

$$E_t[\text{Moment}_{t,T}] = \alpha_4 + \beta_4 \text{Realized Moment}_{t-(T-t),t} + \psi_{t,T} \tag{2.27}$$

The residuals, κ and ψ, are by construction orthogonal to the historical moments and their correlation therefore determines whether physical moments can explain the variation in the realized moments in excess of what is explained by historical moments. In the second stage we estimate the correlation between κ and ψ. The last two columns of Panel C of Table 2.3 report these correlations, and bootstrapped standard errors that correct for the generated regressor problem we face when estimating the residuals in the first stage regressions are in the appendix. Controlling for historical (lagged) moments does not change our results. Our ex ante moments have predictive power for ex post realized moments in excess of what is explained by historical moments. Since trailing quarterly moments do predict ex post realized moments, it is particularly important to notice that our quarterly moments add predictability in excess of what the realized trailing moment counterpart can predict.

5 Commonalities in Higher-Moment Risks

Higher order moments exhibit persistent and interesting time-series co-movements, i.e., higher-moment risks move together, in the sense that skewness and hyperskewness are more negative at times when kurtosis and hyperkurtosis are more positive. To see this, Table 2.4 shows monthly (Panel A) and quarterly (Panel B) pairwise correlations between the first six moments of the physical return distribution. The green (lower right) box shows pairwise correlations between higher order moments. We have flipped the signs for skewness and hyperskewness such that a higher (more positive) value can be translated into higher risk — recall that lower (more negative) skewness and hyperskewness implies more mass in the left tail of the return distribution and therefore higher probabilities of large down movements. These correlations are all positive and large, suggesting that risk as measured by individual higher order moments tends to

be simultaneously high or low.

The strong co-movement of higher order moments suggests that the joint variation in higher order moments can be attributed to a single factor. We therefore estimate the principal components of the space spanned by skewness, kurtosis, hyperskewness, and hyperkurtosis. The four principal components are shown in Table 2.5. Interestingly, at both the monthly and quarterly horizon, the first principal components explains about 90% of the joint variation in higher order moments, underlining the strong co-movement in higher-moment risks.

As was expected, the first principal component eigenvectors have the same signs for skewness and hyperskewness, while the sign is opposite for kurtosis and hyperkurtosis. We standardize each moment to make the eigenvector loadings comparable. The size of the loadings for the first principal components are very similar across the moments, namely -0.45 (-0.47 quarterly) for skewness, 0.52 (0.51 quarterly) for kurtosis, -0.52 (-0.52 quarterly) for hyperskewness, and 0.50 (0.50 quarterly) for hyperkurtosis, implying that the first principal component is approximately the average of the standardized higher order moments with the signs flipped for skewness and hyperskewness. As shown in Ebert (2013), an investor with power utility has a preference for odd number moments of any order and is averse to even number moments of any order. A high value of the first principal component can therefore be interpreted as times when higher order moments (the moments that add mass to the lower tail of the return distribution) are on average large. It is therefore natural to define the first principal component as a higher-moment risk index.

Higher-Moment Risk Index: *We define a higher-moment risk index (HRI) as the first principal component of the space spanned by skewness, kurtosis, hyperskewness, and hyperkurtosis.*

6   Systematic Variation in Higher-Moment Risks

Figure 2.1 displays the time-series plot of the monthly HRI which shows clear systematic variation in higher-moment risk. During the period of financial market distress from 2008 to 2009, HRI was low, whereas during the low variance period from 2004 to 2007, leading up to the financial crisis, monthly HRI was high, suggesting that higher-

74

moment risks are high at times when markets are calm. In this section we investigate these systematic patterns.

The blue (upper right) box of Panel A of Table 2.4 shows the pairwise correlations between variance and higher order moments. Variance is negatively correlated to the negative of skewness, kurtosis, the negative of hyperskewness, and hyperkurtosis with correlations ranging from 0.41 to 0.54. This finding is interesting because it reveals the systematic variation in higher-moment risks; that is, higher-moment risks are high at times when the market is perceived to be safe and calm as measured by variance. Said differently, risk doesn't go away – it hides in the tails.

Figure 2.3 shows time-series plots of variance and the HRI. In the years after the high variance period in 2003 (following the dot.com bubble), as the market became more and more safe as measured by variance, higher-moment risks move steadily in the opposite direction, i.e., skewness became more negative, kurtosis became more positive, and overall the HRI increase significantly. Furthermore, as the financial crisis started to reveal itself, following the default of the Bear Sterns hedge funds, then market uncertainty spread through higher variance – as the tail of the distribution diminished, higher-moment risks decreased.

Somewhat surprisingly, the HRI peaked on June 30$^{\text{th}}$ 2014, when monthly ex ante variance was at its lowest point in seven years. This period, which was calm as measured by variance, was associated with high higher-moment risks. The main political and economical uncertainty during this period was associated with the economic sanctions made by the US targeting Russia over Russia's continuing involvement in Crimea.

Panel A of Table 2.6 shows correlations between variances and the HRI. Generally, higher-moment risk as measured through the HRI is high at times when variance is low. On a monthly horizon, the magnitude of the correlation between variance and HRI is -0.53 with 95% bootstrapped confidence bounds of $[-0.60, -0.48]$. The magnitudes and confidence bounds are quantitatively the same for the quarterly HRI.

Related to the co-movements between variance and higher-moment risks, we also find that higher-moment risks tend to be high after recent market run-ups. To show this, Figure 2.4 shows time-series plots of the past two year return and the HRI. Past returns and the HRI are positively correlated with correlations of 0.38 and 0.35 on monthly and quarterly horizons respectively. To further investigate the dependencies

between market run-ups and subsequent higher-moment risks we run a set of regressions of ex ante moments onto the past two year return,[2] $r_{t-24,t} = R_{t-24,t} - 1$:

$$M_{t,T} = \beta_0 + \beta_1 r_{t-24,t} + \epsilon_{t,T} \qquad (2.28)$$

where the moments, $M_{t,T}$, are variance, skewness, kurtosis, hyperkurtosis, hyperskewness, and the higher-moment risk index (HRI). Panel B of Table 2.6 shows the β_1 coefficients of regression (2.28) and in Panel C of Table 2.6 we show β_1 coefficients of regression (2.28) when controlling for the lagged ex ante moment.

We find a negative and significant relation between past returns and variance. This finding is consistent with the intuition that times after market run-ups are "calm" times where risk, as measured by variance, is low. Looking at skewness, we find a statistically significant and negative relation with past returns, implying that the return distribution tilts to the right and leaves more probability mass in the left tail of the return distribution subsequent to market run-ups. Similarly, kurtosis is statistically significant and positive in past returns, hyperskewness is negative in past returns, and hyperkurtosis is positive in past returns. The results are quantitatively similar for monthly and quarterly moments. Panel C of Table 2.6 shows that controlling for lagged risk does not change our results. We still find strong significant systematic variation in higher order moments.

7 Implications for Investors

The results presented in Table 2.4, Table 2.5, and Table 2.6 show that times when variance is low are times when the market's return distribution is highly left skewed (due to large negative skewness and hyperskewness) and fat tailed (due to large positive kurtosis and hyperkurtosis). That higher-moment risks are high at times when variance is low runs counter to the way we usually think about risk, i.e., we often equate risk with variance, saying that risk is high at times when variance is high. To better understand the importance of higher-moment risks, we next investigate portfolio risks for two investors who both hold a portfolio of cash and the market.

The first investor holds a constant notional in the market. The probability that

---

[2]This is similar to the market run-up period of Greenwood, Shleifer, and You (2017).

the investor's portfolio realizes an unexpected return (the shock to the portfolio), $r_{t,T}^{\text{shock}} = R_{t,T} - E_t[R_{t,T}]$, less than α is:

$$P_t(r_{t,T}^{\text{shock}} < \alpha) \tag{2.29}$$

The constant notional investor is exposed to both time-varying variance risk and time-varying higher-moment risks; that is, the probability that the portfolio realizes an unexpected return less than α depends on both conditional variance and conditional higher order moments.

The second investor targets a constant level of portfolio volatility, i.e., the investor moves wealth in and out of the market such that the portfolio has constant volatility. Such volatility-targeting strategies are common practice and have been shown to generate high risk-adjusted returns (e.g. Moskowitz, Ooi, and Pedersen (2012), Asness, Frazzini, and Pedersen (2012), Moreira and Muir (2017a), and Moreira and Muir (2017b)). If $\sigma_{t,T}$ is the market's conditional volatility, $r_{t,T} = R_{t,T} - 1$ is the return on the market, and $r_{t,T}^f$ is the risk-free rate of return, then $r_{t,T}^{\text{vol target}}$ is the return on the volatility-targeting investor's portfolio who targets a constant volatility of $\sigma^{\text{vol target}}$:

$$r_{t,T}^{\text{vol target}} = \underbrace{\frac{\sigma^{\text{vol target}}}{\sigma_{t,T}}}_{\omega_{t,T}} r_{t,T} + \left(1 - \frac{\sigma^{\text{vol target}}}{\sigma_{t,T}}\right) r_{t,T}^f \tag{2.30}$$

where $\omega_{t,T}$ is the fraction of wealth held in the market. If $\omega_{t,T} > 1$, the investor levers up by borrowing cash to invest more than all the initial wealth in the market. We assume for simplicity that the investor is unconstrained. The unexpected return of the volatility-targeting investor's portfolio is $r_{t,T}^{\text{vol target,shock}} = r_{t,T}^{\text{vol target}} - E_t[r_{t,T}^{\text{vol target}}]$ which can be rewritten as:

$$r_{t,T}^{\text{vol target,shock}} = \omega_{t,T} r_{t,T} + (1 - \omega_{t,T}) r_{t,T}^f - \left(\omega_{t,T} E_t[r_{t,T}] + (1 - \omega_{t,T}) r_{t,T}^f\right) \tag{2.31}$$

$$= \omega_{t,T} r_{t,T}^{\text{shock}} \tag{2.32}$$

The probability that the volatility-targeting investor's portfolio realizes an unexpected

return less than α is:

$$P_t(r_{t,T}^{\text{vol target,shock}} < \alpha) = P_t\left(\frac{\sigma^{\text{vol target}}}{\sigma_{t,T}} r_{t,T}^{\text{shock}} < \alpha\right) \qquad (2.33)$$

$$= P_t\left(r_{t,T}^{\text{shock}} < \frac{\alpha}{\sigma^{\text{vol target}}} \sigma_{t,T}\right) \qquad (2.34)$$

For example, if $\sigma^{\text{vol target}} = 5\%$ and $\alpha = -10\%$, then the probability that the volatility-targeting investor's portfolio realizes a return that is 10% lower than expected is $P_t\left(r_{t,T}^{\text{shock}} < -2\sigma_{t,T}\right)$.[3] The volatility-targeting investor's portfolio is only exposed to time-varying higher-moment risks, that is, given a level of $\sigma^{\text{vol target}}$, the probability that the investor's portfolio realizes a return less than α depends only on conditional higher order moments. Time-varying variance risk is eliminated by targeting a constant level of portfolio volatility.

Recall that $\sigma_{t,t+h}$ is the ex ante volatility from time t to $t + h$, and we then define $\bar{\sigma}^h$ as the time series average of $\sigma_{t,t+h}$. For example, the time-series average of monthly volatility for the S&P 500 index is $\bar{\sigma}^h = 5.0\%$. Figure 2.5 shows time-series plots of monthly probabilities, as shown in (2.29) and (2.33), where $\alpha = -2\bar{\sigma}^{\text{month}} = -10.1\%$ and the volatility-target is $\sigma^{\text{vol target}} = 5.0\%$.

The top figure shows the probabilities of $-2\sigma_{t,t+1}$ drops in the market, which are the probabilities of the volatility-targeting investor's portfolio return. The horizontal line shows the probability of a $-2\sigma_{t,t+1}$ drop in the market implied by a normal distribution, which is 2.5%. The shaded area between the two lines is higher-moment risk; that is, the excess probability of a tail event due to negative skewness, excess kurtosis, and all other higher order moments. Interestingly, the probabilities, in excess of what is implied by a normal distribution, range from 1.1% to 7.2%, showing that time-varying higher-moment risks have large economic implications for the risk of the volatility-targeting investor's portfolio. The probability of a $-2\sigma_{t,t+1}$ drop peaked on June 30th 2014 with a probability of 9.7%, almost three times the size of its low on February 27th 2008, where the probability was 3.6%. The systematic variation in the tail probabilities, from 3.6% at high variance times to 9.7% at low variance times, emphasizes that investors who manage risk by managing variance are implicitly imposing more risk

---

[3]Notice that this probability is not necessarily the same as the probability of a portfolio return of -10%. In the example, the probability of a portfolio return of -10% is $P_t\left(r_{t,T}^{\text{shock}} < -2\sigma_{t,T} - E_t[r_{t,T}]\right)$.

into their portfolio when variance is low.

The bottom figure shows the probabilities of $-2\bar{\sigma} = -10.1\%$ drops in the market along with the probabilities implied by a normal distribution. The shaded area between the two lines is higher-moment risk for the constant notional investor. The probability of a -10.1% drop in the market is, as expected, high when variance is high. Importantly, higher-moment risk also contributes to the portfolio risk for the constant notional investor, and the economic magnitude is large. For example, the probabilities, in excess of what is implied by a normal distribution, range from 0.5% on October 31$^{\text{st}}$ 2006 to 4.8% on August 31$^{\text{st}}$ 2015. On August 31$^{\text{st}}$ 2015, the total probability of a -10.1% drop was 9.90%, which means that, on that day, 48% of the probability mass in the left tail of the return distribution beyond -10.1% was due to higher-moment risk.

Figure 2.6 shows time-series plots of monthly probabilities, as shown in (2.29) and (2.33), where $\alpha = -3\bar{\sigma}^{\text{month}} = -15.1\%$. The probability of a portfolio return that is less than $-3\sigma_{t,t+1}$ peaked on November 30$^{\text{th}}$ 2006 with a probability of 3.6%, which is four times the size of its low on February 27$^{\text{th}}$ 2008, where the probability was 0.8%. These probabilities are far from what is implied by a normal distribution, which is 0.13%. Specifically, the average probability of a $-3\sigma_{t,t+1}$ event is 1.8%, which is fourteen times higher than what is implied by the normal distribution. Figure 2.6 shows that higher-moment risk is even more important when evaluating the probability of events further out in the lower tail of the return distribution; that is, the relative amount of probability mass in the lower tail that is due to higher-moment risk increases the further we go out in the tail.

The probabilities co-move in the sense that, when the probability of a $-2\sigma_{t,T}$ event is high, then the probability of a $-2\bar{\sigma}$ event is low. To further investigate these patterns, Panel A of Table 2.7 reports correlations between variance and the probability of a portfolio return that is less than α for the constant notional investor and the volatility-targeting investor. The first column of Panel A shows the correlations between variance and the probability that the market realizes an unexpected return less than $-2\bar{\sigma}^h$ (the probability of a constant notional investor)

$$P_t(r_{t,t+h}^{\text{shock}} < -2\bar{\sigma}^h) \tag{2.35}$$

The correlations range from 0.94 to 0.98 with tight bootstrapped confidence bounds, showing that the conditional probability that the market realizes a return less than -10.1% monthly or -15.1% quarterly is highly correlated with conditional variance, which would be expected just from looking at Figure 2.5.

Panel A of Table 2.7 also reports correlations between variance and the probability of a portfolio return that is less than α for the volatility-targeting investor. Specifically, we estimate the correlations between variance and the probabilities of a $-2\sigma_{t,t+h}$

$$P_t(r_{t,t+h}^{\text{shock}} < -2\sigma_{t,t+h}) \qquad (2.36)$$

This probability is equivalent to the probability in (2.33) with $\frac{\alpha}{\sigma^{\text{vol target}}} = -2$. Interestingly, the correlations in the last two column of Panel A are all negative and range from -0.70 to -0.44, with tight bootstrapped confidence bounds. These high negative correlations show that the portfolio of the volatility-targeting investor is most risky at times when variance is low, even though the investor has eliminated all dependencies on variance in the portfolio.

This finding can help explain why Moreira and Muir (2017a) and Moreira and Muir (2017b) find that investors can earn high Sharpe ratios by moving wealth into the market at times of low variance and moving wealth out of the market when variance increases (in some sense mimicking a volatility targeting strategy). The relatively (to variance) high expected return in calm periods may be compensation for the elevated higher-moment risks.

To better understand the systematic variation in higher-moment risks, we next investigate the relation between tail probabilities and past returns. Specifically, we regress tail probabilities onto past two year returns, e.g. the probability of a $-2\sigma_{t,t+1}$ drop as

$$P_t(r_{t,t+h}^{\text{shock}} < -2\sigma_{t,t+h}) = \beta_0 + \beta_1 r_{t-24,t} + \epsilon_{t,T} \qquad (2.37)$$

Panel B of Table 2.7 reports β_1 coefficients from regressions such as in (2.37). We find that the probability of both a $-2\sigma_{t,t+1}$ and a $-3\sigma_{t,t+1}$ drop in the market is statistically significant and positively related to past returns. The economic magnitude is such that a 50% market run-up over the past two years implies a 1% higher probability of

a monthly $-2\sigma_{t,t+1}$ drop in the market. Furthermore, the monthly probability of a -10% drop in the market is negatively related to past returns, which is to be expected, because this probability is highly correlated to variance, as shown in Table 2.6, and periods after market run-ups are usually associated with low variance. Panel C of Table 2.7 reports β_1 coefficients from regressions such as in (2.37) when controlling for lagged probabilities. Controlling for lagged probabilities does not change our results: high past two year returns imply higher current tail probabilities for the volatility-targeting investor.

Our finding that market run-ups are related to contemporaneously higher higher-moment risks supplements the existing literature that relates market run-ups to subsequent (realized) market "crashes", e.g. Greenwood, Shleifer, and You (2017). Specifically, we find that the probability of an $x\%$ drop in the market decreases in past returns. High past returns means low current volatility and a low probability of a subsequent $x\%$ drop in the market price. However, *conditional on variance*, the probability of an $x\%$ drop in the market increases in past returns.

8 What Explains Higher-Moment Risk?

In this section we investigate three possible explanations for the systematic variation in higher-moment risk. First, we investigate how higher-moment risk is associated with financial intermediary leverage. Second, we study how market and funding liquidity relates to higher-moment risk. Third, we investigate how higher-moment risk is associated with common "bubble" characteristics. Throughout this section, we will focus on monthly horizon ex ante higher-moment risk.

8.1 The Volatility Paradox and Intermediary Leverage

The volatility paradox is the phenomenon that endogenous risk is high even though exogenous risk is low (Brunnermeier and Sannikov (2014)). Loosely speaking, exogenous risk can be seen as variance and endogenous risk can be seen as higher-moment risk. When variance is low, investors take on more risk in their positions, for instance through leverage, which creates endogenous risk. This negative relation between higher-moment risk and variance is closely related to our empirical findings,

we therefore test if our finding can be linked to the economic drivers suggested by Brunnermeier and Sannikov (2014).

One way in which this endogenous risk may arise is through intermediary leverage.[4] We test if financial intermediary leverage can help explain higher-moment risks by running the following regression:

$$M_{t,T} = \beta_0 + \beta_1 \text{Leverage}_t + \epsilon_{t,T} \tag{2.38}$$

where the risk, $M_{t,T}$, is variance, skewness, kurtosis, hyperkurtosis, hyperskewness, and the higher-moment risk index (HRI). Leverage is the financial intermediary leverage ratio of He, Kelly, and Manela (2016). Regression (2.38) relates aggregate financial intermediary leverage to contemporaneous higher-moment risks. Panel A of Table 2.8 shows the results of regression (2.38). We find that leverage is positively associated with contemporaneous ex ante variance, which is consistent with financial intermediary leverage being counter-cyclical, as noted in He, Kelly, and Manela (2016). The first column of Panel A shows that aggregate financial intermediary leverage is not related to the HRI: we find a regression coefficient pf -0.15 which is statistically insignificant. Decomposing higher-moment risks into individual moments, we do not find a significant relation between financial intermediary leverage and individual higher order moments. Overall, aggregate leverage does not help explain higher-moment risks.

Next, we test if conditional (on variance) financial intermediary leverage is associated with higher-moment risks. We run the regression:

$$M_{t,T} = \beta_0 + \beta_1 \text{Leverage}_t + \beta_2 \text{Variance}_{t,T} + \epsilon_{t,T} \tag{2.39}$$

Panel B of Table 2.8 reports the results of regression (2.39). Interestingly, we find that, conditioning on ex ante variance, financial intermediary leverage can help explain contemporaneous higher-moment risks. We find that skewness, kurtosis, hyperskewness, and hyperkurtosis all load statistically significantly on financial intermediary leverage, with negative signs for skewness and hyperskewness and positive signs for kurtosis

---

[4]Several papers have shown that financial intermediary leverage is associated with asset returns, e.g. He, Kelly, and Manela (2016), He and Krishnamurthy (2013), Adrian and Boyarchenko (2012), and Adrian, Etula, and Muir (2014).

and hyperkurtosis. Furthermore, the HRI is positively related to conditional financial intermediary leverage. Given a level of ex ante variance, higher leverage is associated with higher contemporaneous higher-moment risks.

Panel C in Table 2.8 reports regression (2.39) when controlling for lagged risk. Controlling for lagged risk does not change our results. Aggregate financial intermediary leverage is in general not associated with higher-moment risks. Given a level of ex ante variance, and controlling for lagged risk, higher leverage is associated with higher contemporaneous higher-moment risks.

8.2 Market Liquidity and Funding Liquidity.

Several previous papers link market liquidity and funding liquidity to aspects of the stock market's return distribution. Christoffersen, Feunou, Jeon, and Ornthanalai (2016) suggest market illiquidity as an economic factor driving risk-neutral market variance and jump risks, or equivalently, higher order moments. They argue that market illiquidity is the common culprit of market price drops in cases when the price drop happened without news about fundamentals, and it is therefore a reasonable economic driver of market moments. Brunnermeier and Pedersen (2009) show that, from a theoretical point of view, stocks with low market (and funding) liquidity have high variance because they are associated with high margin requirements. Furthermore, Danilova and Julliard (2015) develop a model in which volatility and illiquidity are jointly determined by the same equilibrium forces.

First, we test if high market illiquidity is associated with high contemporaneous ex ante variance. Thereafter, we investigate the relation between market illiquidity and higher-moment risks. When testing the relation between higher-moment risks (or variance) and market illiquidity, we run the regression:

$$\mathrm{M}_{t,T} = \beta_0 + \beta_1 \text{Bid-ask spread}_t + \epsilon_{t,T} \tag{2.40}$$

where the risk, $\mathrm{M}_{t,T}$, is variance, skewness, kurtosis, hyperkurtosis, hyperskewness, and the higher-moment risk index (HRI). As a proxy for market illiquidity, we follow Christoffersen, Feunou, Jeon, and Ornthanalai (2016), and use the average value-

weighted bid-ask spread of constituents of the S&P 500 index.

Panel A of Table 2.9 reports the results of regression (2.40). We find that higher market illiquidity is associated with higher contemporaneous ex ante variance, which is consistent with the model of Brunnermeier and Pedersen (2009). The effect is statistically significant at a 99% level and controlling for lagged variance does not change the result.

The HRI is negatively related to market illiquidity with a regression coefficient of -1.22, which is statistically significant at a 99% level. When we control for lagged HRI, we still get a negative relation between market illiquidity and higher-moment risks, but the relation is insignificant. The negative relation between the HRI and market illiquidity shows that higher-moment risks tend to be high at times when the market is most liquid.

Next, we test the relation between funding illiquidity and higher-moment risks. We run the regression:

$$\mathrm{M}_{t,T} = \beta_0 + \beta_1 \mathrm{TED\ spread}_t + \epsilon_{t,T} \tag{2.41}$$

where the risk, $\mathrm{M}_{t,T}$, is variance, skewness, kurtosis, hyperkurtosis, hyperskewness, and the higher-moment risk index (HRI). The TED spread is a common proxy for funding illiquidity, e.g. Frazzini and Pedersen (2014). The TED spread is the three month LIBOR intrabank interest rate minus the three month T-bill interest rate and it is available from the St. Louis FED.

Panel B of Table 2.9 reports the results of regression (2.41). Contemporaneous ex ante variance is positively related to funding illiquidity, higher TED spread is associated with higher ex ante variance. We find that the HRI is negatively related to funding illiquidity which means that, higher-moment risks are high at times when there is low friction in the funding market. Controlling for the lagged HRI does not change our result.

Figure 2.7 shows time-series plots of market illiquidity and funding illiquidity with the HRI. Consistent with the results presented in Table 2.9, we see that the HRI is negatively correlated with both the bid-ask spread and the TED spread.

8.3 "Bubble" Characteristics

A range of macroeconomic variables have been proposed as possible indicators of increased market "crash" risks, or equivalently, increased higher-moment risks. A partial list of the variables include the suggestion of Chen, Hong, and Stein (2001), who suggest turnover, Pontiff and Woodgate (2008), who use issuance as a characteristic, and Greenwood, Shleifer, and You (2017), who propose price acceleration as a higher-moment risk characteristic.

In this section we investigate the relation between common market "crash" indicators and contemporaneous ex ante higher-moment risks. We therefore run regressions on the form:

$$\text{HRI}_{t,T} = \beta_0 + \beta_1 \text{Characteristic}_t + \epsilon_{t,T} \tag{2.42}$$

where the "bubble" characteristics are: 1) The Greenwood, Shleifer, and You (2017) variable acceleration, which is defined as the annualized past two year return minus the return of the first year of the two year return. Acceleration captures the convexity in the recent price path and a high value of acceleration is intended to be associated with high contemporaneous ex ante higher-moment risks. 2) Issuance as the percentage of firms in the S&P 500 index that issued equity in the past year. We follow Greenwood, Shleifer, and You (2017), and define an equity issuance as the event that a firm's split-adjusted share count increased by five percent or more. 3) Market turnover. The market valuation measures are: 4) CAPE, the Shiller cyclically adjusted price-earnings ratio. 5) The dividend price ratio as the past two year dividends divided by the current market price. 6) Cay, the Lettau and Ludvigson (2001) log consumption - aggregate wealth ratio.

Table 2.10 reports the results of regression (2.42). Marginally, we find that cay is negatively and significantly related to the HRI. Calm times when expected returns, as proxied by cay, are low are times when higher-moment risks are high.

Interacting turnover with the past two year return, we find that turnover is positively related to the HRI. This finding is consistent with the findings in Chen, Hong, and Stein (2001), that is, subsequent to market run-ups, a higher turnover is associated with higher higher-moment risks.

For issuance, we find that, subsequent to market run-ups, a higher level of equity

issuance implies a lower level of contemporaneous higher-moment risks. This finding is counter to the results of Pontiff and Woodgate (2008). Firms have incentives to issue equity when the stock price is higher than its fundamental value, which should be associated with higher contemporaneous higher-moment risks. Other characteristics show no marginal relation with higher-moment risks.

As a last test, we run a horse race including all "bubble" characteristics. The last column of Table 2.10 reports the results of the horse race. Jointly, we find that acceleration is statistically significant and positively related to the HRI, and cay is negatively related to the HRI. Conditional on market run-ups, turnover is positively associated with the HRI. Interestingly, issuance changes sign in the horse race compared to the marginal regressions. Indeed, we find that higher issuance is associated with higher contemporaneous higher-moment risks when controlling for other characteristics.

Figure 2.8 shows time-series plots of the HRI and cay. The two time-series are negatively correlated with an in-sample correlation coefficient of -0.57. Figure 2.9 shows time-series plots of the HRI and turnover times past two year return. The two time-series are positively correlated with a correlation coefficient of 0.43, implying that, after market run-ups when turnover is high, then so are higher-moment risks.

9   Conclusion: when volatility is low, risk hides in the tails

We show that ex ante physical moments estimated using methods based on Martin (2017) are superior to historical moments and risk-neutral moments at predicting ex ante realized moments.

Ex ante higher order moment risks co-move such that the first principal component of the space spanned by skewness, kurtosis, hyperskewness, and hyperkurtosis explains 90% of the joint variation. We define this first principal component as a higher-moment risk index (HRI) which captures the time-variation in higher-moment risks.

Interestingly, the HRI is negatively related to variance. We show that times when variance is low are the times when the physical return distribution is most left skewed (due to large negative skewness and hyperskewness) and fat tailed (due to large positive kurtosis and hyperkurtosis). The economic importance of higher-moment risk is most easily understood from the point of view of a volatility-targeting investor. The portfolio

risk of this investor is high at times when variance is low, even though the investor has eliminated variance risk in the portfolio. For example, the probability that the investor's portfolio realizes a return less than two standard deviations varies from 3.6% during times of financial distress to 9.7% during periods of low variance.

We show empirically how higher-moment risk is associated with market liquidity, funding liquidity, turnover, and the market valuation variable cay. Times with low liquidity frictions, low cay, and high turnover are times when higher-moment risks are high.
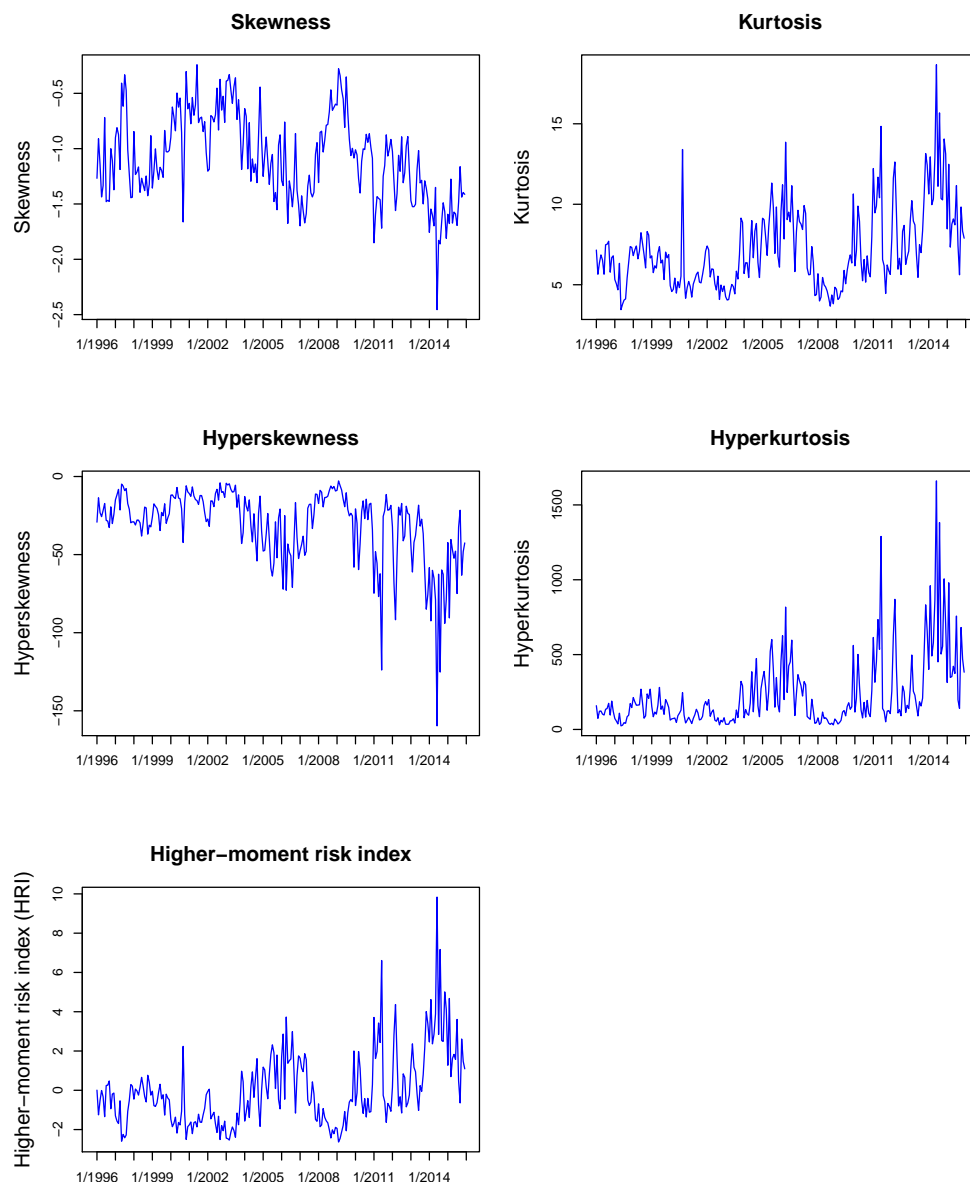
Figure 2.1: **Higher order moments and the higher-moment risk index.** The figures show a time-series plot of monthly higher order moments and the higher-moment risk index (HRI) for the S&P 500 index. HRI is estimated as the first principal component of the space spanned by skewness, kurtosis, hyperskewness, and hyperkurtosis. Times when the HRI is high are times when higher-moment risks are high.
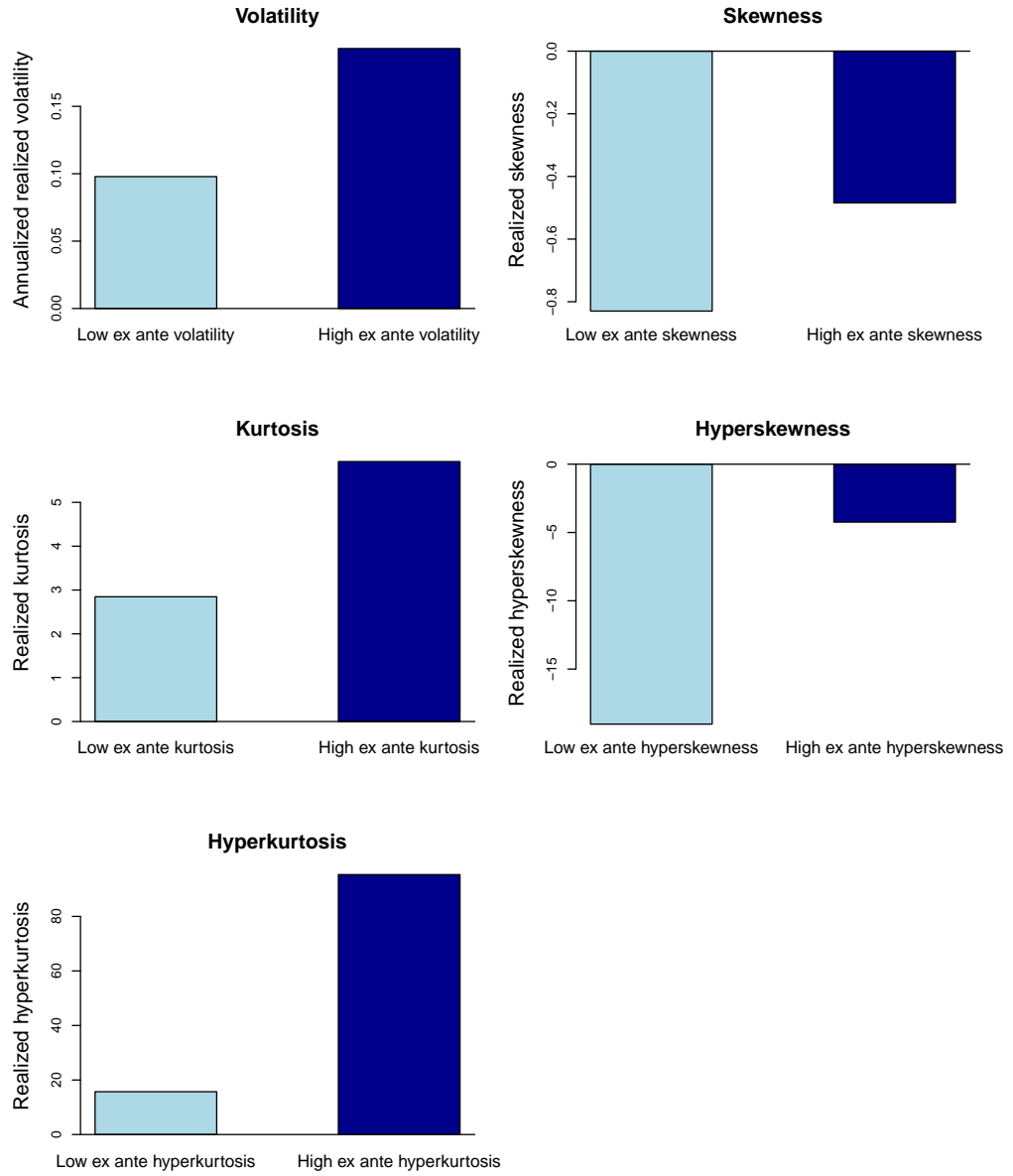
Figure 2.2: **Moments of ex post realized returns sorted on ex ante expected moments.** For each moment, we sort ex post realized monthly returns into a "low" and "high" bucket based on the median ex ante expected value of that moment. Thereafter, we compute the corresponding realized moment for each bucket.
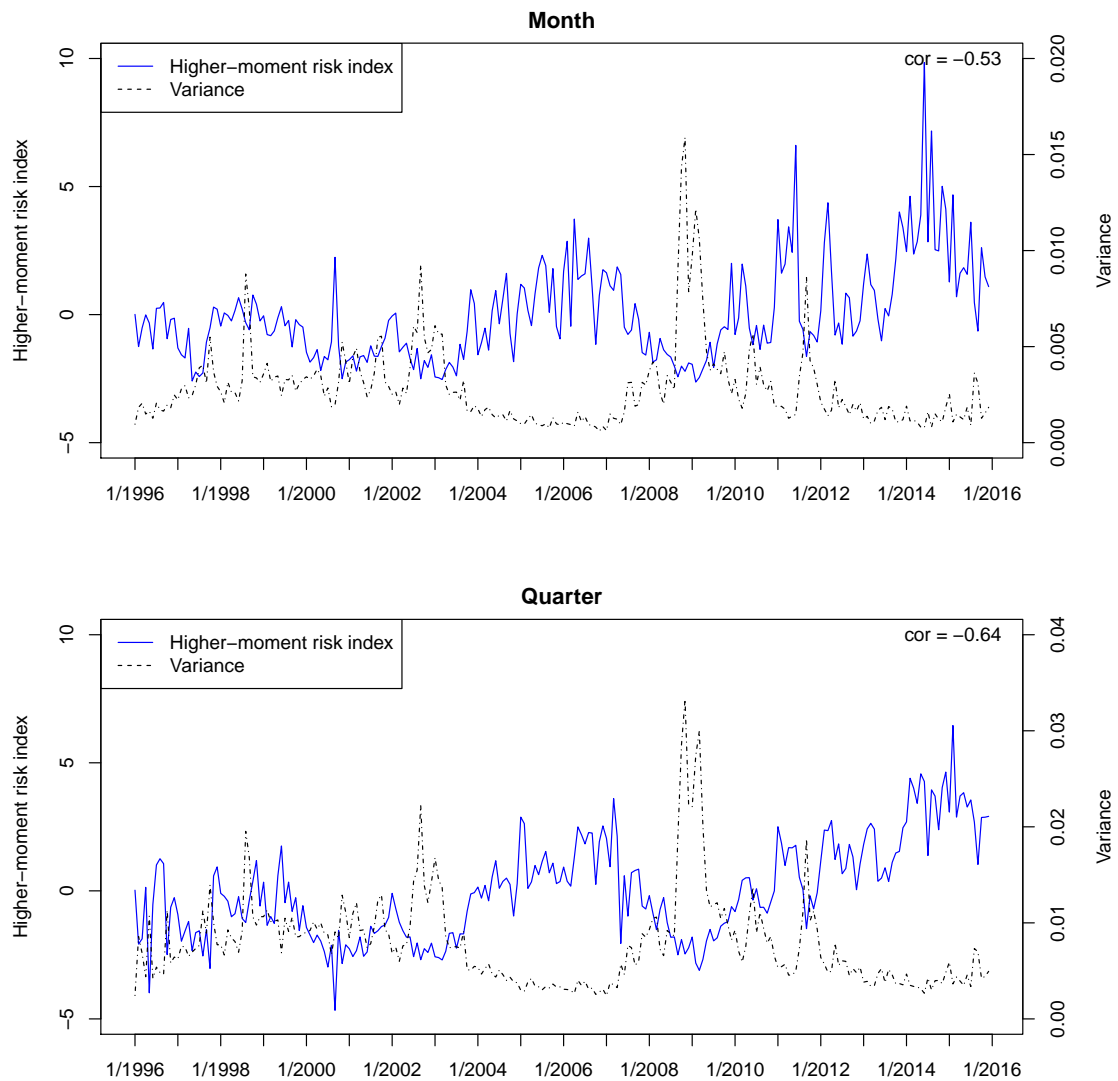
Figure 2.3: **Higher-moment risk index and variance.** This figure shows time-series plots of monthly and quarterly higher-moment risk index (HRI) and variances. The HRI is high at times when variance is low.
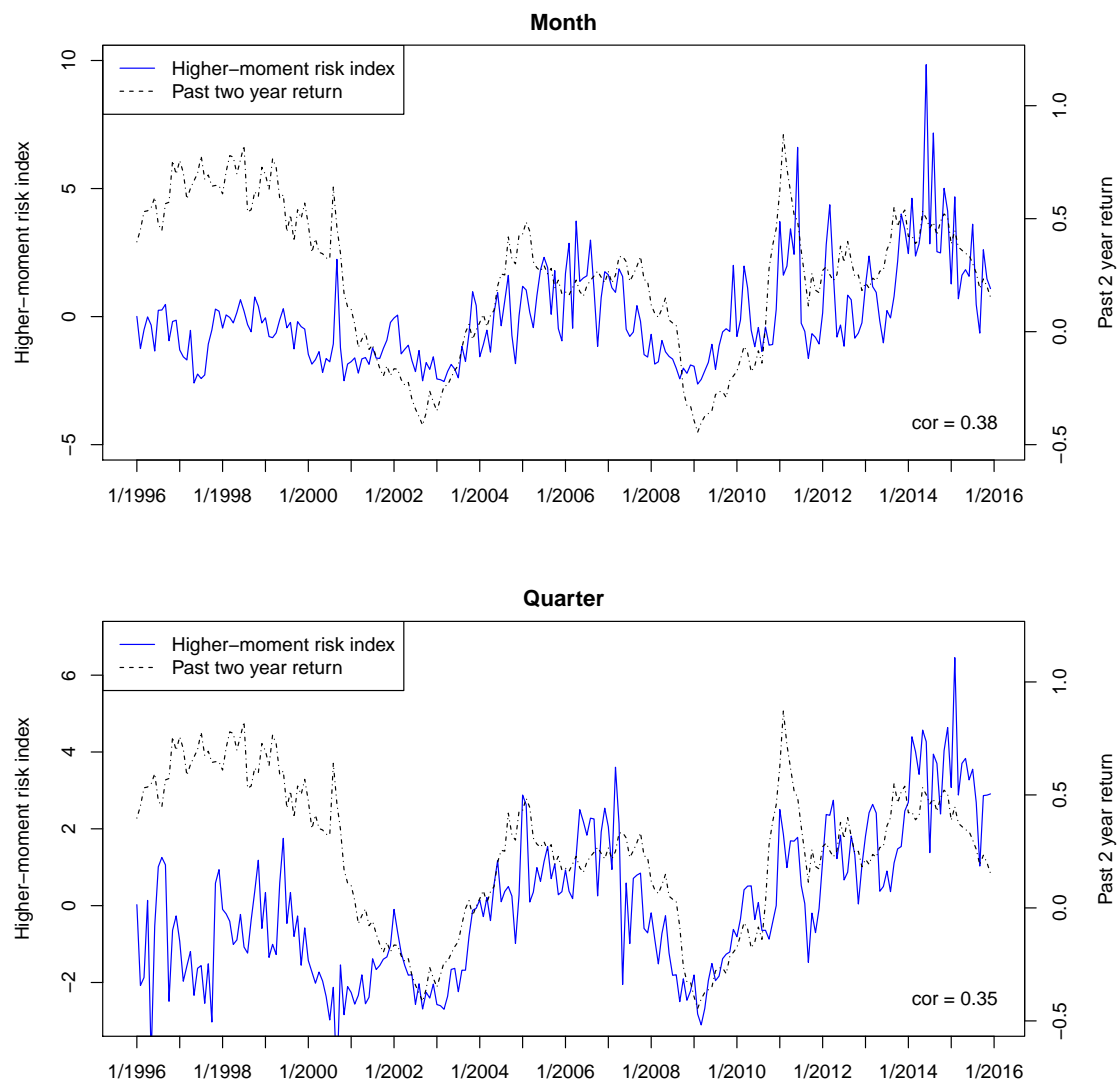
Figure 2.4: **Higher-moment risk index and the past two year return.** These figures show time-series plots of the past two year return and the S&P 500 higher-moment risk index (HRI). Past return and the HRI are positively correlated, implying that higher-moment risk is high subsequent to market run-ups.
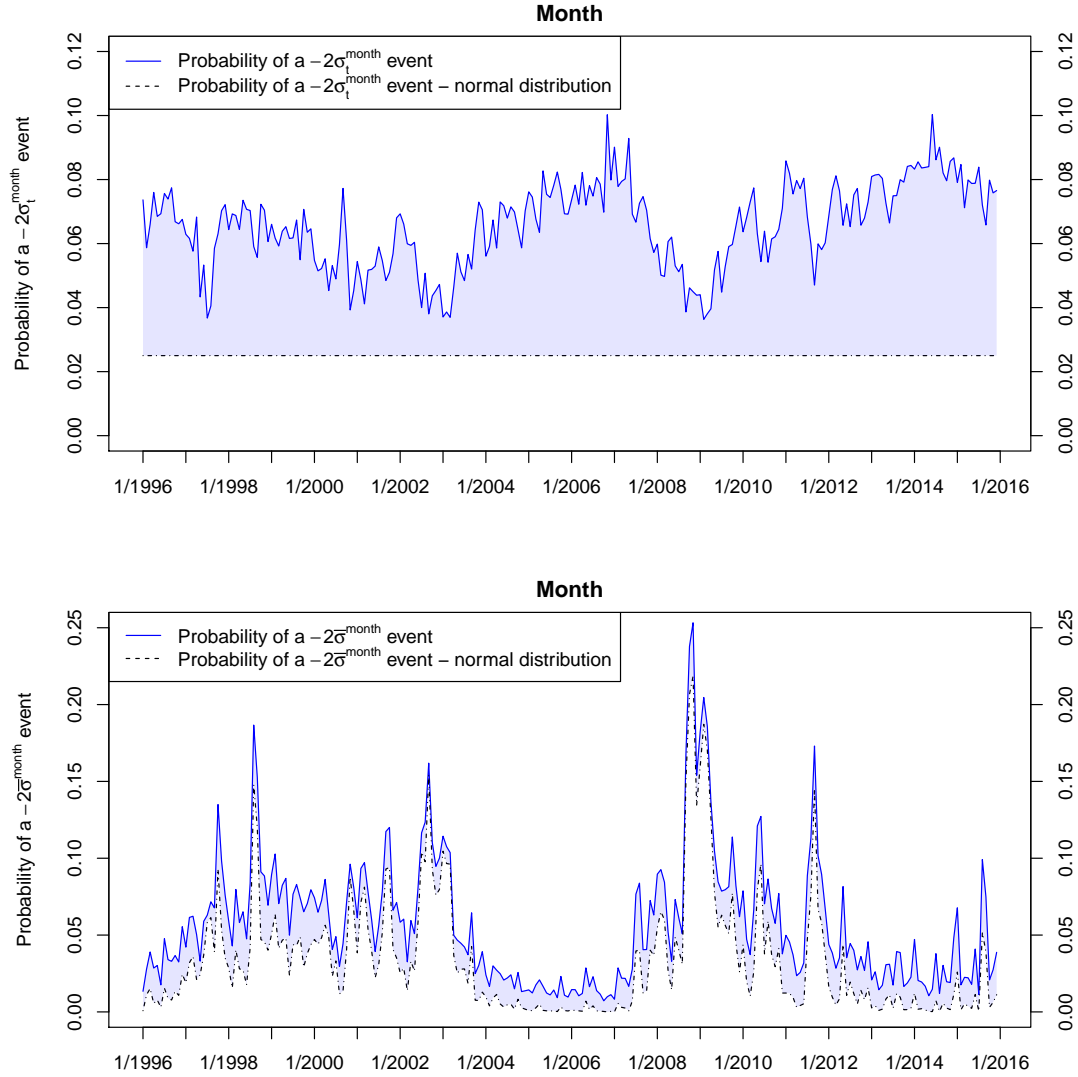
Figure 2.5: **Market tail loss probabilities – two sigma.** The top figure shows portfolio tail loss probabilities for the volatility-targeting investor; that is, the probability of an unexpected return lower than $-2\sigma_t^{\text{month}}$. The dashed line is the tail loss probabilities implied by a normal distribution. The shaded area between the lines is higher-moment risk, that is, the part of the tail loss probability that is entirely driven by changes in higher order moments. The bottom figure shows portfolio tail loss probabilities for the constant notional investor. Here, σ_t^{month} is the conditional monthly ex ante variance and $\bar{\sigma}^{\text{month}}$ is the time-series average of σ_t^{month}. In our sample, $\bar{\sigma}^{\text{month}} = 5.0\%$.

Figure 2.6: **Market tail loss probabilities – three sigma.** The top figure shows portfolio tail loss probabilities for the volatility-targeting investor; that is, the probability of an unexpected return lower than $-3\sigma_t^{\mathrm{month}}$. The dashed line is the tail loss probabilities implied by a normal distribution. The shaded area between the lines is higher-moment risk, that is, the part of the tail loss probability that is entirely driven by changes in higher order moments. The bottom figure shows portfolio tail loss probabilities for the constant notional investor. Here, $\sigma_t^{\mathrm{month}}$ is the conditional monthly ex ante variance and $\bar{\sigma}^{\mathrm{month}}$ is the time-series average of $\sigma_t^{\mathrm{month}}$. In our sample, $\bar{\sigma}^{\mathrm{month}} = 5.0\%$.
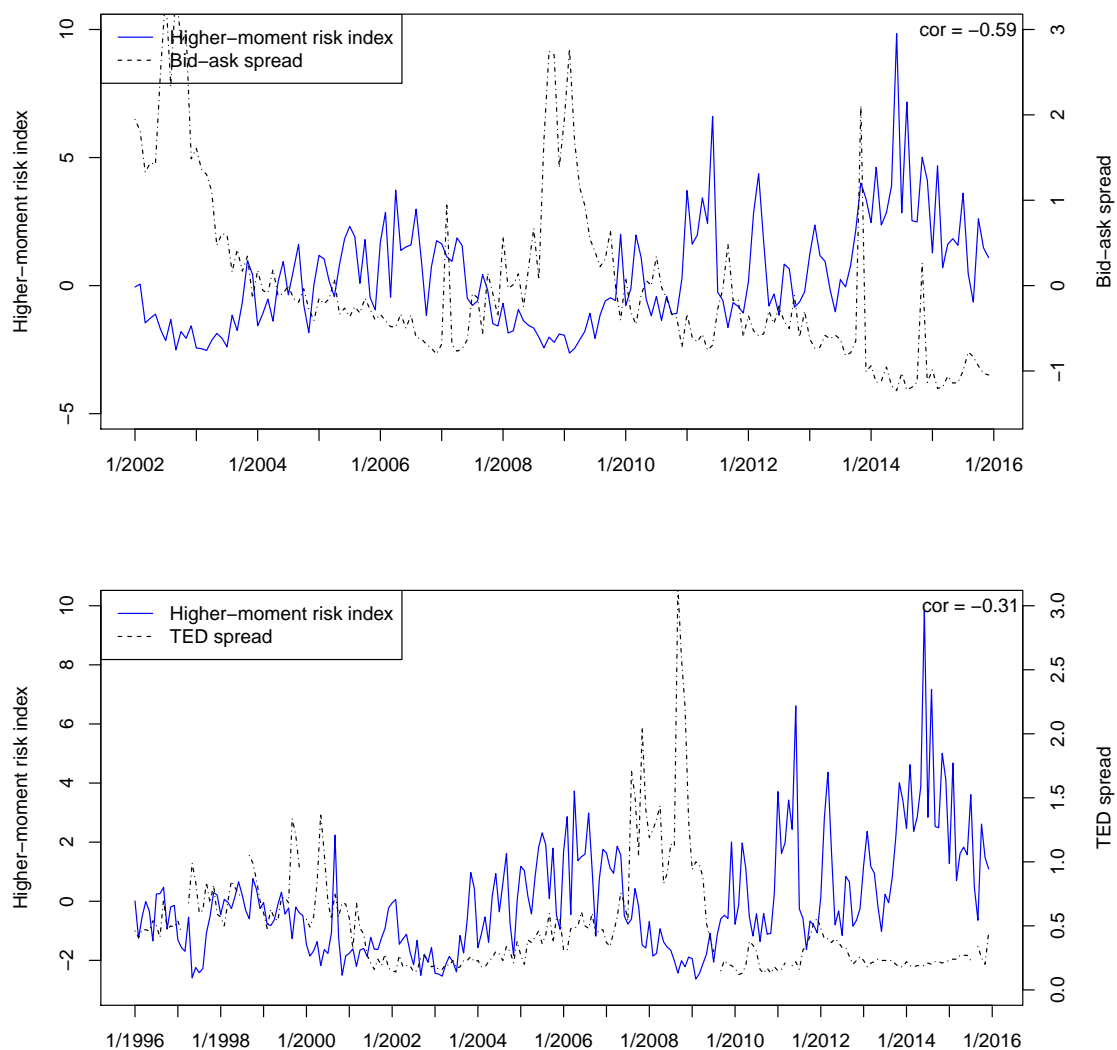
Figure 2.7: **Higher-moment risk index and market and funding illiquidity.**
The top figure shows time series plots of the HRI and market illiquidity (proxied by
the average value-weighted bid-ask spread of S&P 500 constituents). The bottom
figure shows time series plots of the HRI and funding illiquidity (proxied by the TED
spread).
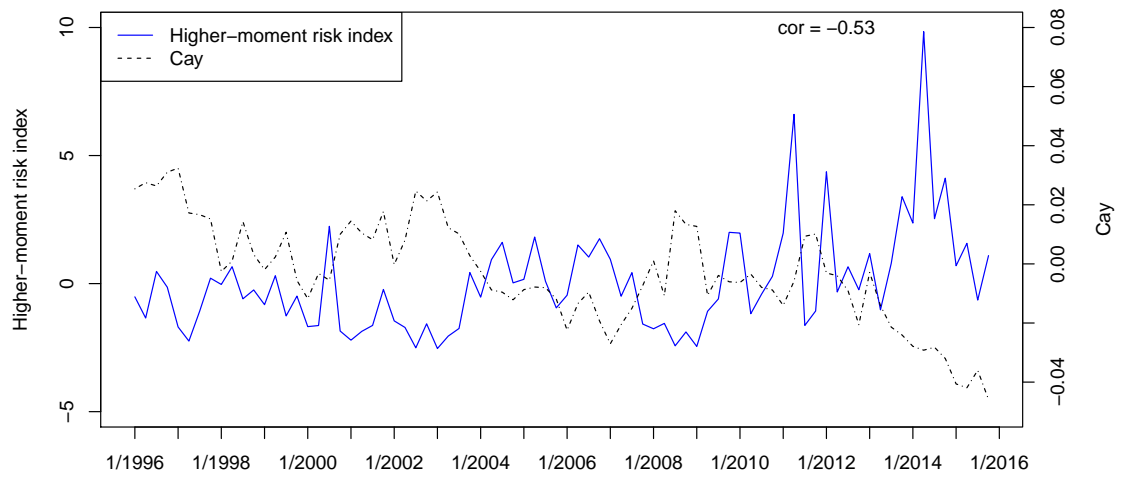
Figure 2.8: **Higher-moment risk index and cay.** The figure shows time series plots of the HRI and cay.
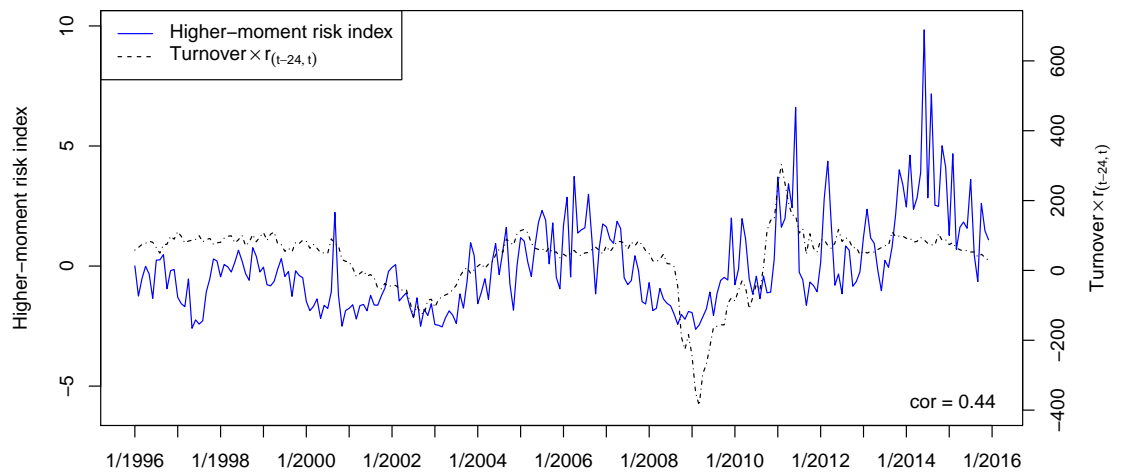
Figure 2.9: **Higher-moment risk index and turnover.** The figure shows time series plots of the HRI and turnover times the past two year return.

Table 2.1: **Moment Summary Statistics.** In this table we report the average time-series values for ex ante estimated moments: excess return (ER−Rf), standard deviation (St. dev.), skewness (Skew), kurtosis (Kurt), hyperskewness (Hskew), and hyperkurtosis (Hkurt). We estimate ex ante moments from the point of view of a risk-neutral investor ($\gamma = 0$), a log-utility investor ($\gamma = 1$), and two power-utility investors ($\gamma = 3$, $\gamma = 5$).

Horizon	Risk-aversion	Annualized (%)		Skew	Kurt	Hskew	Hkurt
		ER−Rf	St. dev.				
Month	$\gamma = 0$	0	21.07	-1.45	8.90	-46.58	347.41
Month	$\gamma = 1$	4.44	19.89	-1.31	8.25	-41.01	307.44
Month	$\gamma = 3$	12.00	18.33	-1.08	7.20	-31.16	233.07
Month	$\gamma = 5$	18.36	17.32	-0.89	6.43	-23.47	175.88
Quarter	$\gamma = 0$	0	21.07	-1.17	5.78	-20.58	110.36
Quarter	$\gamma = 1$	4.44	19.59	-1.09	5.57	-18.64	97.45
Quarter	$\gamma = 3$	11.48	17.55	-0.95	5.23	-15.56	80.69
Quarter	$\gamma = 5$	17.12	16.25	-0.81	4.94	-12.93	68.19

Table 2.2: **Ex Ante Conditional Moments Predict Ex Post Realized Moments.** Panel A reports ex post moments for monthly returns sorted into a low or high bucket based on the ex ante moment. Panel B reports correlations between our ex ante moments and ex post realized moments. Panel C reports correlations between our ex ante kurtosis and hyperkurtosis with ex post left kurtosis and left hyperkurtosis. We also report correlations between historical moments and ex post moments. We report bootstrapped standard errors in the appendix and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$.

Panel A: Sorting on ex ante monthly moments

	Our moments		Historical moments	
	Low ex ante	High ex ante	Low ex ante	High ex ante
Variance (%)	0.08***	0.31***	0.07***	0.28***
Skewness	−0.83	−0.48	−0.49	−0.75
Kurtosis	2.85	5.93**	3.35	4.46
Hyperskewness	−19.03**	−4.24	−5.40	−9.01
Hyperkurtosis	15.70	95.24***	15.34	40.10

Panel B: Correlation between ex ante moments and ex post realized moments

	Our moments		Historical moments	
	Month	Quarter	Month	Quarter
Variance	0.67***	0.49***	0.72***	0.46***
Skewness	0.21***	0.25***	0.07	0.24***
Kurtosis	−0.01	0.00	0.06	0.04
Hyperskewness	0.17***	0.20***	0.07	0.15***
Hyperkurtosis	−0.03	0.06	0.03	0.01

Panel C: Left kurtosis and left hyperkurtosis

	Our moments		Historical moments	
	Month	Quarter	Month	Quarter
Left kurtosis	0.19***	0.26***	0.02	0.14**
Left hyperkurtosis	0.17***	0.21***	0.05	0.13**

Table 2.3: **Ex Ante Conditional Moments Predict Ex Post Realized Moments — Robustness.** Panel A reports correlations between our ex ante moments and ex post realized moments when we remove observations that overlap with the period from August 1st 2008 to July 31st 2009. Panel B reports correlations between our ex ante moments (estimated with different levels of relative risk aversion) and ex post realized moments. Panel C reports correlations when controlling for risk-neutral moments or historical moments. We report bootstrapped standard errors in the appendix and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$..

Panel A: Excluding August 1st 2008 to July 31st 2009

	Our moments		Historical moments	
	Month	Quarter	Month	Quarter
Variance	0.52^{***}	0.49^{***}	0.51^{***}	0.42^{***}
Skewness	0.23^{***}	0.23^{***}	0.11	0.25^{***}
Kurtosis	-0.01	0.00	0.07	0.06
Hyperskewness	0.18^{***}	0.18^{***}	0.09	0.13^{***}
Hyperkurtosis	-0.03	0.04	0.04	0.01

Panel B: Other levels of risk-aversion

	$\gamma = 1$		$\gamma = 5$	
	Month	Quarter	Month	Quarter
Variance	0.67^{***}	0.48^{***}	0.67^{***}	0.48^{***}
Skewness	0.20^{***}	0.25^{***}	0.21^{***}	0.23^{***}
Kurtosis	-0.03	0.00	0.01	0.02
Hyperskewness	0.13^{***}	0.17^{***}	0.20^{***}	0.22^{***}
Hyperkurtosis	-0.06	0.04	-0.01	0.07

Panel C: Marginal correlations

	Controlling for risk-neutral moments		Controlling for historical moments	
	Month	Quarter	Month	Quarter
Variance	0.09	0.12	0.18^{**}	0.20^{***}
Skewness	0.09	0.01	0.20^{***}	0.17^{***}
Kurtosis	0.11^{**}	-0.02	-0.02	0.01
Hyperskewness	0.16^{***}	0.14^{***}	0.16^{***}	0.17^{**}
Hyperkurtosis	0.09^{**}	0.07	-0.03	0.05

Table 2.4: **Correlations Between S&P 500 Moments.** Panel A reports pairwise correlations between monthly S&P 500 moments. Expected return (Er), variance (Var), skewness (Skew), kurtosis (Kurt), hyperskewness (Hskew), and hyperkurtosis (Hkurt). Panel B shows the correlation between quarterly horizon moments. We report 95% bootstrapped confidence bounds in brackets.

Panel A: Month

	Er	Var	−Skew	Kurt	−Hskew	Hkurt
Er	1	0.99	−0.46	−0.50	−0.48	−0.41
		[0.99,1]	[−0.56,−0.37]	[−0.57,−0.46]	[−0.55,−0.43]	[−0.48,−0.37]
Var		1	−0.52	−0.54	−0.51	−0.44
			[−0.60,−0.43]	[−0.60,−0.50]	[−0.58,−0.47]	[−0.51,−0.40]
-Skew			1	0.80	0.78	0.66
				[0.76,0.84]	[0.74,0.82]	[0.60,0.72]
Kurt				1	0.97	0.93
					[0.95,0.98]	[0.90,0.95]
-Hskew					1	0.98
						[0.97,0.98]
Hkurt						1

Panel A: Quarter

	Er	Var	−Skew	Kurt	−Hskew	Hkurt
Er	1	0.99	−0.46	−0.54	−0.58	−0.56
		[0.99,0.99]	[−0.55,−0.37]	[−0.60,−0.49]	[−0.64,−0.54]	[−0.62,−0.52]
Var		1	−0.54	−0.62	−0.65	−0.62
			[−0.62,−0.47]	[−0.67,−0.57]	[−0.71,−0.61]	[−0.68,−0.57]
-Skew			1	0.83	0.86	0.74
				[0.79,0.86]	[0.82,0.90]	[0.68,0.79]
Kurt				1	0.96	0.94
					[0.95,0.97]	[0.92,0.96]
-Hskew					1	0.97
						[0.96,0.98]
Hkurt						1

Table 2.5: **Principal Components of Higher-Moment Risks.** We estimate the four principal components (PC) spanning the space of monthly (Panel A) and quarterly (Panel B) skewness (Skew), kurtosis (Kurt), hyperskewness (Hskew), and hyperkurtosis (Hkurt). Panel A reports the loadings on each of the monthly moments. Panel B reports the loadings on each of the quarterly moments. The last column of Panel A shows that the first principal component (PC 1) explains 89% of the variation in monthly higher order moments. Similarly, the last column of Panel B shows that 91% of the variation in quarterly higher order moments is captured by the first principal component.

Panel A: Month

	Skew	Kurt	Hskew	Hkurt	Variation explained
PC 1 eigenvector	−0.45	0.52	−0.52	0.50	89%
PC 2 eigenvector	0.85	0.07	−0.20	0.48	10%
PC 3 eigenvector	−0.23	−0.83	−0.16	0.48	1%
PC 4 eigenvector	−0.13	0.19	0.81	0.54	0%
PC 1 correlation	−0.85	0.98	−0.99	0.95	

Panel B: Quarter

	Skew	Kurt	Hskew	Hkurt	Variation explained
PC 1 eigenvector	−0.47	0.51	−0.52	0.50	91%
PC 2 eigenvector	−0.84	−0.16	0.11	−0.51	7%
PC 3 eigenvector	−0.13	−0.84	−0.32	0.40	2%
PC 4 eigenvector	−0.25	0.01	0.78	0.57	0%
PC 1 correlation	−0.89	0.98	−0.99	0.96	

Table 2.6: **Cyclicality in Higher-Moment Risks.** Panel A reports correlations between ex ante variance and the higher-moment risk index (HRI). We report bootstrapped 95% confidence intervals in brackets. Panel B reports β_1 coefficients when regressing physical moments onto the past two year returns:

$$M_{t,T} = \beta_0 + \beta_1 r_{t-24,t} + \epsilon_{t,T}$$

where the moment $M_{t,T}$ is variance (Var), skewness (Skew), kurtosis (Kurt), hyperskewness (Hskew), hyperkurtosis (Hkurt), and the higher-moment risk index. Panel C reports the regression when controlling for lagged moments. We report t-statistics in parentheses and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$. We correct standard errors for autocorrelation using Newey and West (1987).

Panel A: Variance and the higher-moment risk index

Horizon	HRI
Month	−0.53
95% CI	[−0.60,−0.48]
Quarter	−0.64
95% CI	[−0.69,−0.59]

Panel B: Past return and higher-moment risks

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	2.26***	−0.30*	−0.61***	2.75***	−24.57***	201.65**
(sd)	(0.79)	(0.16)	(0.14)	(1.01)	(9.41)	(98.06)
Quarter	2.13**	−0.68*	−0.45***	1.02*	−9.62**	50.63**
(sd)	(0.88)	(0.39)	(0.12)	(0.56)	(3.83)	(25.23)

Panel C: Past return and higher-moment risks — controlling for lagged risk

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	0.94***	−0.07**	−0.23***	1.16***	−11.11**	109.07*
(sd)	(0.34)	(0.03)	(0.06)	(0.39)	(4.53)	(57.91)
Quarter	0.47**	−0.14**	−0.15***	0.25*	−2.04**	11.98**
(sd)	(0.19)	(0.06)	(0.04)	(0.14)	(0.86)	(5.33)

Table 2.7: **Constant Notional, Volatility-Targeting, and Higher-Moment Risks.** Panel A reports correlations between ex ante variance and tail loss probabilities. The probabilities are $P(r_t^h < -2\bar{\sigma}^h)$ and $P(r_t^h < -2\sigma_t^h)$ where $r_t^h = R_{t,t+h} - E_t[R_{t,t+h}]$, σ_t^h is the ex ante volatility from time t to $t+h$, and we define $\bar{\sigma}^h$ as the time series average of σ_t^h. We report bootstrapped 95% confidence intervals in brackets. Panel B reports regression slope coefficients when regressing physical tail loss probabilities onto the past two year returns. Panel C reports coefficients when controlling for lagged probabilities. We report t-statistics in parentheses and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$. We correct standard errors for autocorrelation using Newey and West (1987)

Panel A: Correlations between variance and tail probabilities

Horizon	$P(r_t^h < -2\bar{\sigma}^h)$	$P(r_t^h < -3\bar{\sigma}^h)$	$P(r_t^h < -2\sigma_t^h)$	$P(r_t^h < -3\sigma_t^h)$
Month	0.97	0.98	−0.70	−0.54
95% CI	[0.96,0.98]	[0.96,0.99]	[−0.78, −0.65]	[−0.62,−0.47]
Quarter	0.97	0.94	−0.58	−0.44
95% CI	[0.96,0.98]	[0.92,0.96]	[−0.66,−0.51]	[−0.53,−0.35]

Panel B: Tail probabilities (%) and past return

Horizon	$P(r_t^h < -2\bar{\sigma}^h)$	$P(r_t^h < -3\bar{\sigma}^h)$	$P(r_t^h < -2\sigma_t^h)$	$P(r_t^h < -3\sigma_t^h)$
Month	−5.28*	−2.23	2.02***	0.65***
(sd)	(3.03)	(1.44)	(0.63)	(0.17)
Quarter	−5.67	−2.09	1.24**	0.38**
(sd)	(3.61)	(1.46)	(0.57)	(0.18)

Panel C: Tail probabilities (%) and past return - controlling for lagged probabilities

Horizon	$P(r_t^h < -2\bar{\sigma}^h)$	$P(r_t^h < -3\bar{\sigma}^h)$	$P(r_t^h < -2\sigma_t^h)$	$P(r_t^h < -3\sigma_t^h)$
Month	−1.17*	−0.56*	0.55***	0.26***
(sd)	(0.60)	(0.30)	(0.20)	(0.09)
Quarter	−0.96*	−0.44*	0.38**	0.11**
(sd)	(0.54)	(0.26)	(0.16)	(0.05)

Table 2.8: **Financial Intermediary Leverage and Higher-Moment Risk.** Panel A reports regression slope coefficients when regressing higher-moment risks onto the financial intermediary leverage of He, Kelly, and Manela (2016). Panel B reports coefficients when conditioning on ex ante variance. Panel C reports coefficients when conditioning on ex ante variance and controlling for lagged risk. We report standard errors in parentheses and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$. We correct standard errors for autocorrelation using Newey and West (1987).

Panel A: Leverage and higher-moment risks

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	−0.14	0.11***	0.05	−0.22	1.27	−4.65
(sd)	(0.29)	(0.03)	(0.07)	(0.30)	(3.41)	(28.54)
Quarter	−0.03	0.24**	0.01	−0.00	0.23	0.37
(sd)	(0.51)	(0.10)	(0.06)	(0.39)	(2.00)	(15.64)

Panel B: Conditional leverage and higher-moment risks

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	0.51**	—	−0.07	0.65***	−6.59***	68.43***
(sd)	(0.20)	—	(0.05)	(0.24)	(2.49)	(24.94)
Quarter	0.76***	—	−0.10***	0.51***	−3.28***	22.29***
(sd)	(0.19)	—	(0.03)	(0.08)	(0.80)	(2.60)

Panel C: Conditional leverage and higher-moment risks — controlling for lagged risk

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	0.29***	—	−0.03*	0.38***	−3.93***	45.12***
(sd)	(0.10)	—	(0.02)	(0.13)	(1.35)	(15.42)
Quarter	0.34***	—	−0.05**	0.26***	−1.44***	10.16***
(sd)	(0.10)	—	(0.02)	(0.08)	(0.37)	(2.37)

Table 2.9: **Market Liquidity, Funding Liquidity, and Higher-Moment Risks.**
This table reports the results when regressing higher-moment risks onto market liquidity (Panel A) and funding liquidity (Panel B). We use the value-weighted bid-ask spread of S&P 500 constituents as a proxy for market illiquidity. We use the TED spread as a proxy for funding illiquidity. Panel C and Panel D report results when controlling for lagged market or funding illiquidity respectively. We report t-statistics in parentheses and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$. We correct standard errors for autocorrelation using Newey and West (1987).

Panel A: Market illiquidity

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	−1.22***	0.17***	0.27***	−1.53***	14.18***	−128.26***
(sd)	(0.20)	(0.05)	(0.04)	(0.25)	(2.47)	(28.29)
Quarter	−1.44***	0.40***	0.24***	−0.85***	6.30***	−37.05***
(sd)	(0.19)	(0.11)	(0.03)	(0.12)	(0.81)	(5.85)

Panel B: Funding illiquidity

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	−1.41***	0.25***	0.18	−2.12***	14.19***	−128.26***
(sd)	(0.39)	(0.11)	(0.11)	(0.52)	(2.47)	(28.28)
Quarter	−1.37***	0.54**	0.08	−1.07***	6.30***	−37.05***
(sd)	(0.51)	(0.22)	(0.11)	(0.32)	(0.81)	(5.85)

Panel C: Market illiquidity — controlling for lagged risk

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	−0.68***	0.08***	0.14***	−0.78***	8.37***	−83.77***
(sd)	(0.17)	(0.02)	(0.02)	(0.20)	(2.15)	(24.75)
Quarter	−0.55***	0.15***	0.11***	−0.31***	2.51***	−14.96***
(sd)	(0.11)	(0.05)	(0.02)	(0.07)	(0.51)	(3.93)

Panel D: Funding illiquidity — controlling for lagged risk

Horizon	HRI	Var (%)	Skew	Kurt	Hskew	Hkurt
Month	−0.57***	0.10***	0.08***	−0.88***	8.37***	−83.77***
(sd)	(0.18)	(0.02)	(0.03)	(0.23)	(2.15)	(24.75)
Quarter	−0.38***	0.20***	0.05*	−0.31***	2.51***	−14.96***
(sd)	(0.12)	(0.07)	(0.03)	(0.08)	(0.51)	(3.93)

Table 2.10: **"Bubble" Characteristics and Higher-Moment Risks.** This table reports the results when regressing the higher-moment risk index onto "bubble" characteristics. These are: acceleration, CAPE, dividend-price ratio, cay, turnover, and issuance. We report t-statistics in parentheses and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$. We correct standard errors for autocorrelation using Newey and West (1987).

Dependent variable: Monthly HRI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Acceleration	2.90								3.86***
(sd)	(1.62)								(3.52)
CAPE		-0.03							-0.01
(sd)		(-0.76)							(-0.39)
Dividend-price ratio			-23.73						162.08
(sd)			(-0.29)						(1.45)
Cay				-64.10***					-45.80***
(sd)				(-3.63)					(-3.12)
Turnover					-3.52***	1.58			-0.00
(sd)					(-2.81)	(1.33)			(-1.07)
Turnover$\times r_{t-24,t}$						9.90***			0.01**
(sd)						(3.94)			(2.11)
Issuance							-2.12	-0.72	6.20**
(sd)							(-0.54)	(-0.18)	(1.97)
Issuance$\times r_{t-24,t}$								-27.27***	-4.97
(sd)								(-3.19)	(-0.67)
$r_{t-24,t}$						0.28		10.77	0.90
(sd)						(0.40)		(4.00)	(0.30)
No. obs.	240	238	240	80	240	240	240	240	80
Adj. R^2	0.02	0.00	-0.00	0.32	0.06	0.18	0.00	0.27	0.46

105

A   Ex Ante Physical Moments, Risk-Neutral Pricing, and Realized Moments

Ex ante physical moments and risk-neutral pricing

Using equation (2.6) we can represent physical ex ante moments in terms of asset prices:

$$E_t[R_{t,T}^i] = \frac{E_t^*[R_{t,T}^{i+\gamma}]}{E_t^*[R_{t,T}^\gamma]} \tag{2.43}$$

for $i \in \{1, ..., 6\}$. These asset prices can be used to estimate ex ante physical moments by expanding the standardized moment formula in equation (2.7). We estimate kurtosis, hyperskewness, and hyperkurtosis in the following way:

$$\text{Kurtosis}_{t,T} = \frac{E_t[R_{t,T}^4] - 3E_t[R_{t,T}]^4 + 6E_t[R_{t,T}]^2 E_t[R_{t,T}^2] - 4E_t[R_{t,T}]E_t[R_{t,T}^3]}{(E_t[R_{t,T}^2] - E_t[R_{t,T}]^2)^2} \tag{2.44}$$

$$\tag{2.45}$$

$$\text{Hyperskewness}_{t,T} = \frac{E_t[R_{t,T}^5] + 4E_t[R_{t,T}]^5 + 10E_t[R_{t,T}]^2 E_t[R_{t,T}^3] - 10E_t[R_{t,T}]^3 E_t[R_{t,T}^2] - 5E_t[R_{t,T}]E_t[R_{t,T}^4]}{(E_t[R_{t,T}^2] - E_t[R_{t,T}]^2)^{5/2}} \tag{2.46}$$

$$\tag{2.47}$$

$$\text{Hyperkurtosis}_{t,T} = \frac{E_t[R_{t,T}^6] - 5E_t[R_{t,T}]^6 + 15E_t[R_{t,T}]^4 E_t[R_{t,T}^2] - 20E_t[R_{t,T}]^3 E_t[R_{t,T}^3] + 15E_t[R_{t,T}]^2 E_t[R_{t,T}^4]}{(E_t[R_{t,T}^2] - E_t[R_{t,T}]^2)^3} \tag{2.48}$$

$$- \frac{6E_t[R_{t,T}]E_t[R_{t,T}^5]}{(E_t[R_{t,T}^2] - E_t[R_{t,T}]^2)^3} \tag{2.49}$$

Estimating ex post realized moments

Let N be the number of daily realized returns between time t and T and denote the daily return between day s and $s+1$ as $r_{s,s+1}$. The realized moments between time t and T are estimated from daily realizations in the following way:

$$\mu_{t,T} = \frac{1}{N} \sum_{i=1}^{N} r_{i-1,i} \tag{2.50}$$

$$\tag{2.51}$$

$$\sigma_{t,T}^2 = \frac{N \sum_{i=1}^{N} (r_{i-1,i} - \mu_{t,T})^2}{N-1} \tag{2.52}$$

$$\tag{2.53}$$

$$\text{Realized Skewness}_{t,T} = \frac{N^{1/2} \sum_{i=1}^{N} \left(r_{i-1,i} - \mu_{t,T}\right)^3}{\sigma_{t,T}^3} \tag{2.54}$$

$$\tag{2.55}$$

$$\text{Realized Kurtosis}_{t,T} = \frac{N \sum_{i=1}^{N} \left(r_{i-1,i} - \mu_{t,T}\right)^4}{\sigma_{t,T}^4} \tag{2.56}$$

$$\tag{2.57}$$

$$\text{Realized Hyperskewness}_{t,T} = \frac{N^{3/2} \sum_{i=1}^{N} \left(r_{i-1,i} - \mu_{t,T}\right)^5}{\sigma_{t,T}^5} \tag{2.58}$$

$$\tag{2.59}$$

$$\text{Realized Hyperkurtosis}_{t,T} = \frac{N^2 \sum_{i=1}^{N} \left(r_{i-1,i} - \mu_{t,T}\right)^6}{\sigma_{t,T}^6} \tag{2.60}$$

This is similar to the methods used by Amaya, Christoffersen, Jacobs, and Vasquez (2015).

B Appendix Tables

Table AI: **Ex Ante Conditional Moments Predict Ex Post Realized Moments (Test statistics).** This Table reports test statistics for the results reported in Table 2.2. Panel A reports p-values from the bootstrapped distribution. Panel B reports bootstrapped standard errors for the correlation coefficient between ex ante moments and ex post realized moments. Panel C reports bootstrapped standard errors for the correlation coefficient between ex ante kurtosis and ex post realized left kurtosis. Panel C reports also reports correlations for hyperkurtosis.

Panel A: Sorting on ex ante monthly moments

	Our moments		Historical moments	
	Low ex ante	High ex ante	Low ex ante	High ex ante
Variance	0.00	0.00	0.00	0.01
Skewness	0.19	0.46	0.54	0.74
Kurtosis	0.21	0.03	0.46	0.20
Hyperskewness	0.02	0.45	0.47	0.75
Hyperkurtosis	0.39	0.01	0.38	0.20

Panel B: Correlation between ex ante moments and ex post realized moments

	Our moments		Historical moments	
	Month	Quarter	Month	Quarter
Variance	0.07	0.09	0.09	0.10
Skewness	0.06	0.06	0.07	0.06
Kurtosis	0.06	0.07	0.05	0.04
Hyperskewness	0.05	0.06	0.06	0.04
Hyperkurtosis	0.05	0.06	0.04	0.03

Panel C: Left kurtosis and left hyperkurtosis

	Our moments		Historical moments	
	Month	Quarter	Month	Quarter
Left kurtosis	0.06	0.07	0.07	0.06
Left hyperkurtosis	0.06	0.06	0.07	0.06

Table AII: **Ex Ante Conditional Moments Predict Ex Post Realized Moments — Robustness (Test statistics).** Panel A reports bootstrapped standard errors for the correlations between our ex ante moments and ex post realized moments when we remove observations that overlap with the period from August 1, 2008 to July 31, 2009. Panel B reports bootstrapped standard errors for the correlations between our ex ante moments (estimated with different levels of relative risk aversion) and ex post realized moments. Panel C reports bootstrapped standard errors for the correlations when controlling for risk-neutral moments or historical moments. We report bootstrapped standard errors in the appendix and significance as; * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$..

Panel A: Excluding August 1, 2008 to July 31, 2009

| | Our moments | | Historical moments | |
	Month	Quarter	Month	Quarter
Variance	0.06	0.05	0.07	0.08
Skewness	0.06	0.05	0.07	0.06
Kurtosis	0.06	0.08	0.06	0.07
Hyperskewness	0.05	0.06	0.07	0.07
Hyperkurtosis	0.05	0.07	0.06	0.05

Panel B: Other levels of risk-aversion

| | $\gamma = 1$ | | $\gamma = 5$ | |
	Month	Quarter	Month	Quarter
Variance	0.07	0.10	0.06	0.09
Skewness	0.06	0.06	0.06	0.06
Kurtosis	0.06	0.06	0.05	0.06
Hyperskewness	0.05	0.05	0.05	0.07
Hyperkurtosis	0.05	0.05	0.05	0.06

Panel C: Marginal correlations

| | Controlling for risk-neutral moments | | Controlling for historical moments | |
	Month	Quarter	Month	Quarter
Variance	0.17	0.10	0.12	0.06
Skewness	0.06	0.06	0.06	0.06
Kurtosis	0.05	0.06	0.06	0.07
Hyperskewness	0.05	0.06	0.05	0.06
Hyperkurtosis	0.04	0.06	0.05	0.06

Chapter 3

The Market's Time-Varying Risk Aversion

Abstract:

I present a new method for estimating the market's time-varying risk aversion using historical market returns and option prices written on the S&P 500 index. Market risk aversion varies in a systematic way; it tends to be *low* during times of financial market distress, e.g., during the 2008-2009 financial crisis, and it tends to be *high* at times when the market is considered to be calm as measured by variance. This systematic variation in market risk aversion is difficult to reconcile with the leading asset pricing models. I discuss several possible explanations for these time-varying patterns in risk aversion including investor salience and time-varying correlations between aggregate consumption and the market.

1 Introduction

Risk averse market participants require compensation for taking on market risk. The amount of compensation required depends on the level of investor risk aversion and the expected distribution of future market returns. Understanding how risk aversion and the expected distribution of future market returns co-vary over time is therefore important in understanding market prices.

In this paper, I provide a new method for estimating the market's time-varying risk aversion. My methodology allows me to investigate the co-movements between market risk aversion and the physical distribution of market returns. I arrive at the following two main results: (i) Market risk aversion varies over time and tends to be low when volatility is high, e.g., during the recent 2008-2009 financial crisis. (ii) Market risk aversion tends to be high at times when market tail risk is high. This systematic variation in market risk aversion is difficult to reconcile with the leading asset pricing models. In the last part of the paper, I discuss possible explanations for the systematic variation in risk aversion. Specifically, I consider salience theory as a behavioral explanation and time-varying correlation between the market and aggregate consumption as a rational explanation.

Before I go into the details on my results, it is instructive to understand how I estimate the market's time-varying risk aversion. As in Martin (2017) and Gormsen and Jensen (2017b), I consider the preferences of a power utility investor who chooses to be fully invested in the market. I estimate option implied risk neutral distributions for multiple horizons corresponding to the last trading day of the month from January 1996 until December 2015. Using these risk neutral distributions, I estimate the time series of risk aversion coefficients which best match the historical realized monthly returns on the market while still being consistent with observable option prices. To do so, I rely on the generalized recovery methodology developed in Jensen, Lando, and Pedersen (2017) combined with the Berkowitz test as used in Bliss and Panigirtzoglou (2004). Specifically, I estimate the power utility investor's time preference parameter by minimizing the Berkowitz test statistic under the constraint that, for a given value of the time preference parameter, the time-varying risk aversion coefficients solves the generalized recovery equation set.

As my first main result, I show that the market's risk aversion varies systematically with variance. Specifically, risk aversion is negatively correlated with market variance with a correlation coefficient of -0.65 and tight bootstrapped 95% confidence bounds of $[-0.70, -0.62]$.[1] The negative correlation suggests that market participants are more risk tolerant at times when the market is generally considered highly risky, i.e., during times of high market volatility. This finding is hard to reconcile with the leading asset pricing models. For example, Campbell and Cochrane (1993) explain asset prices with a risk aversion that is countercyclical.

To understand why it is reasonable that market risk aversion is low at times when variance is high, consider the following heuristic example. Under power utility and log-normality of the market return distribution then the expected excess return on the market is equal to $\gamma\sigma^2$, where γ is market risk aversion and σ^2 is market variance. Suppose risk aversion is equal to 3 and constant, which is a commonly chosen value in the financial economics literature, see e.g. Bliss and Panigirtzoglou (2004) and Gormsen and Jensen (2017b). Then, during the peak of the financial crisis in 2008-2009 where annualized realized market variance reached almost 60%, the implied annual expected return on the market was about 180%. It is hard to believe that the market participants believed that, in expectation, the market was going to bounce back with an increase of 180% from the peak of the crisis. Therefore, if annualized expected returns were truly lower than 180% during the financial crisis, then either risk aversion is lower than 3 and constant or it fluctuates and was low during the crisis.

I find that risk aversion on average is 2.77 and that it falls to 1.49 during the peak of the financial crisis. Using the heuristic argument above, this level of risk aversion implies an expected excess return on the market during the peak of the crisis of about 90% annually and 7.5% on a monthly horizon.

As my second main result, I show that the market's risk aversion varies systematically with market tail risk. Specifically, I find that risk aversion is: (i) negatively correlated with market skewness with a correlation of -0.42 with bootstrapped 95% confidence bounds of $[-0.51, -0.31]$ and (ii) positively correlated with market kurtosis with a correlation of 0.62 and bootstrapped confidence bounds of $[0.54, 0.70]$. This

---

[1] Bliss and Panigirtzoglou (2004) also find evidence of a risk aversion which is negatively related to variance.

finding suggests that the investor becomes more risk averse as the market's higher order moments become more risky, i.e., there is more probability mass in the tail of the market return distribution.

Next, I discuss evidence of this systematic variation in risk aversion that is already in the finance literature. For example, in a recent paper by Moreira and Muir (2017b), the authors show that investors can earn large alphas by timing market volatility. They find that investors should exit the market when volatility increases and enter the market when it drops. Their result arises because expected excess returns are high relative to variance during low variance periods, i.e., the ratio $\frac{ER_t}{\sigma_t^2}$ varies over time and becomes high when variance is low. Their findings imply a time variation in market risk aversion which is consistent with the results I present. To see why, recall that we can express the expected excess return on the market as $ER_t = \gamma_t \sigma_t^2$ which implies a market risk aversion of $\gamma_t = \frac{ER_t}{\sigma_t^2}$. Their results are therefore reminiscent of a risk aversion that is low at times when volatility is high. Similarly, Gormsen and Jensen (2017a) also find evidence that the market price of risk is high during low volatility periods.

In the final part of the paper, I discuss two possible explanations for the systematic variation in market risk aversion. First, I show that my results are consistent with aspects of salience theory by Bordalo, Gennaioli, and Shleifer (2012). Specifically, I follow Lian, Ma, and Wang (2018) who argue that, at times of low interest rates the relatively high expected returns on risky assets are salient, and this salience on the upside of a higher return on the risky asset induces heightened risk tolerance and "reaching for yield" tendencies among investors. I therefore regress the ratio of expected gross returns on the market to gross risk-free returns, $E_t(R_{t,T})/R_{t,T}^f$, onto risk aversion. Consistently with the findings of Lian, Ma, and Wang (2018), I find that investors become more risk tolerant as the ratio of expected returns to risk-free returns increases.

Lastly, I discuss how the systematic variation in risk aversion can arise if the stock market is not a perfect proxy for aggregate consumption. Specifically, for an investor with power utility who cares about aggregate consumption, then the stock market implied risk aversion can be expressed as $\alpha_t = \rho_t \gamma_t^c \sigma_t^c / \sigma_t$ where ρ_t is the correlation between the stock market and aggregate consumption, σ_t^c is consumption volatility,

γ_t^c is the investor's risk aversion, and σ_t is the volatility of the stock market. As an illustrative example, I consider the case where risk aversion and consumption volatility are constant, i.e., $\gamma_t^c = \gamma^c$ and $\sigma_t^c = \sigma^c$. In this case, the stock market implied risk aversion is proportional to the ratio of the correlation between the stock market and aggregate consumption and the volatility of the stock market. During times of financial distress, stock market volatility increases, if the correlation between the stock market and aggregate consumption does not increase enough to offset the increase in stock market volatility then α_t will decrease.

My paper relates to and extends the existing literature on estimating market risk aversion. I present a new method for estimating the market's risk aversion by integrating the Berkowitz test with the generalized recovery method of Jensen, Lando, and Pedersen (2017). Previous research which use option prises to estimate time-varying risk aversion is based either (1) a full identification (e.g. a parameterization through a GARCH model) of the market's physical return distribution from historical returns or (2) parameterization of investor preferences, e.g. Ross (2015) and Jensen, Lando, and Pedersen (2017). My method allows for the identification of time-varying risk aversion while leaving the physical probability distribution fully free, which compliments the existing literature that imposes structure on the physical return distribution, e.g. Barone-Adesi, Engle, and Mancini (2008), Jackwerth (2004), Jackwerth (2000), Ait-Sahalia and Lo (2000). My paper also relates to Bliss and Panigirtzoglou (2004) who use the Berkowitz test, but they use it to estimate a constant risk aversion.

The paper proceeds as follows: Section 2 describes how I estimate the market's time-varying risk aversion. Section 3 covers data and details on the empirical implementation. Section 4 studies the empirical results of the time-varying market risk aversion. Section 5 studies the relation between market risk aversion and the market's physical moments. Section 6 investigates what might explain the systematic variation in the market's time-varying risk aversion. Section 7 concludes the paper.

2 Inferring Financial Market Risk Aversion

I consider the preferences and beliefs of a power utility investor who optimally chooses to be fully invested in the market. The economy consists of a risk-free asset with gross

risk-free return of $R_{t,T}^f$ and a risky asset, the stock market, with price S, dividends $D_{t,T}$, and gross returns

$$R_{t,T} = \frac{S_T + D_{t,T}}{S_t} \qquad (3.1)$$

The power utility investor's utility function at time t is $U_t(x) = x^{1-\gamma(t)}/(1-\gamma(t))$ where $\gamma(t)$ is the investor's (possibly time-varying) risk aversion coefficient. The investor has initial wealth W_0 and terminal wealth $W_0 R_{t,T}$. Given that the power utility investor chooses to be fully invested in the stock market, I can express the investor's stochastic discount factor in the following way:

$$m_{t,T}(R_{t,T}) = \delta(t,T) R_{t,T}^{-\gamma(t)} \qquad (3.2)$$

where $\delta(t,T)$ is a time-preference parameter and $\gamma(t)$ is the market's risk aversion[2]. Table 3.1 reports the five different specifications of the stochastic discount factor in (3.2) which I investigate in this paper.

Next, I show how I infer the market's time-varying risk aversion. I combine the results of two related papers, the generalized recovery methodology of Jensen, Lando, and Pedersen (2017) and the Berkowitz test as used in Bliss and Panigirtzoglou (2004). The main methodological objective of both these papers is to backward engineer preference parameters from observable asset prices,[3] in this paper I exploit these methods in a joint setting.

I start by fixing some notation. I write the standard asset pricing formula in the following common way:

$$\pi_{t,T}(r) = p_{t,T}(r) m_{t,T}(r) \qquad (3.3)$$

where $\pi_{t,T}$ is the time t and $T-t$ horizon known state price density, $p_{t,T}$ is the corresponding unknown physical probability density, and $m_{t,T}$ is the unknown stochastic discount factor. In Section 3, I discuss how I estimate state price densities using option prices written on the market. Given the standard asset pricing formula, I can write

---

[2]I assume that the market's risk aversion is independent of the horizon. There are several papers that address horizon dependent risk aversion, see e.g., Bliss and Panigirtzoglou (2004) and Lazarus (2018). The methodology I use in this paper can accommodate a horizon specific risk aversion, however, I choose to focus on the variation in the time series of risk aversion rather than the term structure of risk aversion.

[3]

the market's probability distribution function, say F, as

$$F_{t,T}(r) = \int_{-\infty}^{r} p_{t,T}(x)dx = \int_{-\infty}^{r} \frac{\pi_{t,T}(x)}{m_{t,T}(x)}dx \tag{3.4}$$

If I insert the power utility investor's stochastic discount factor from (3.2) into (3.4), then I can rewrite the market's distribution function as follows

$$F_{t,T}(r) = \int_{-\infty}^{r} \frac{\pi_{t,T}(x)x^{\gamma(t)}}{\delta(t,T)}dx \tag{3.5}$$

Given values of the parameters $\delta(t,T)$ and $\gamma(t)$, I can estimate the market's probability distribution as perceived by the power utility investor who chooses to invest everything in the market. The objective now is figuring out what the true values of $\delta(t,T)$ and $\gamma(t)$ are.

To estimate these true values of the preference parameters, I follow Bliss and Panigirtzoglou (2004) and use the so-called Berkowitz test, cf. Berkowitz (2001). The idea behind the Berkowitz test is that, for the true values $\hat{\delta}(t,T)$ and $\hat{\gamma}(t)$, the distribution of $u_{t,T} = \hat{F}_{t,T}(R_{t,T})$ is uniform and the distribution $y_{t,T} = \Phi^{-1}(u_{t,T})$ is standard normal. Here $\hat{F}_{t,T}(R_{t,T})$ denotes the distribution function in (3.5) with the true values, $\hat{\delta}(t,T)$ and $\hat{\gamma}(t)$, inserted. In the Berkowitz test, I estimate the coefficients in the regression model:

$$y_{t,T} = \hat{a} + \hat{\beta}y_{t-1,T-1} + \epsilon_{t,T}, \qquad \epsilon_{t,T} \sim N(0, \hat{\sigma}) \tag{3.6}$$

and perform a likelihood ratio test of the joint hypothesis that $a = \beta = 0$ and $\text{Var}(\epsilon_{t,T}) = 1$. The hypothesis that $b = 0$ is natural when considering non-overlapping returns, for overlapping returns see e.g. Bliss and Panigirtzoglou (2004) for a thorough discussion of the test. It is also worth noticing that, even though there might be momentum effects in returns, then we will still want $b = 0$ because the true distribution should take these momentum effects into account.

The Berkowitz likelihood ratio test for non-overlapping returns is then:

$$LR = -2(LL(0,0,1) - LL(\hat{a}, \hat{\beta}, \hat{\sigma})) \sim \chi_3^2 \tag{3.7}$$

117

where $LL(\hat{a}, \hat{\beta}, \hat{\sigma})$ is the log likelihood of (3.6).

I find the true values of $\delta(t, T)$ and $\gamma(t)$ by minimizing the Berkowitz test statistic in (3.7) under the constraint that, for all dates t and horizons $T - t$, the equation $\int_{-\infty}^{\infty} \frac{\pi_{t,T}(x)x^{\gamma(t)}}{\delta(t,T)}dx = 1$ must hold. This constraint ensures that the resulting physical return distribution integrates to one. The optimization problem is therefore:

$$\min_{\delta(t,T)} - 2\left(LL(0, 0, 1) - LL(\hat{a}, \hat{\beta}, \hat{\sigma})\right) \tag{3.8}$$

$$s.t. \quad \gamma(t) \text{ solves } \int_{-\infty}^{\infty} \frac{\pi_{t,T}(x)x^{\gamma(t)}}{\delta(t,T)}dx = 1, \quad \text{for all } t, T \tag{3.9}$$

For a given level of $\delta(t, T)$,[4] the constraints provide enough equations to solve for the time-varying risk aversion, $\gamma(t)$. Specifically, for a given level of $\delta(t, T)$, at any point in time, I only have to solve for the risk aversion coefficient, $\gamma(t)$. If $\gamma(t)$ were linear in the constraint, then solving for the parameter would be straightforward. However, $\gamma(t)$ enters non-linearly in the equation $\int_{-\infty}^{\infty} \frac{\pi_{t,T}(x)x^{\gamma(t)}}{\delta(t,T)}dx = 1$ and I need to address this non-linearity. Luckily, the generalized recovery methodology of Jensen, Lando, and Pedersen (2017) provides me with a way of recovering $\gamma(t)$. The authors show that, at any point in time, if we know the state price densities for multiple horizons, say M different horizons, then recovery is possible if the stochastic discount factor is characterized by a number of of parameters, say N, that is less than the number of horizons with known state price densities, i.e., if $M > N$. At any point in time and for a given level of $\delta(t, T)$, I have $N = 1$ (the risk aversion parameter) and can therefore solve for $\gamma(t)$ if I know state price densities for two or more horizons.

My methodology is closely related to the methodology used in Bliss and Panigirtzoglou (2004), specification (3) in Table 3.1, and it is therefore instructive to understand how my methodology differs from theirs. In short, their methodology implies a constant level of risk aversion but time-varying time preferences. My methodology on the other hand implies structure on time preferences and allows for time-varying risk aversion. To understand their approach in my setting, notice that for a constant level of risk aversion, $\gamma(t) = \gamma$,[5] I can write the Bliss and Panigirtzoglou (2004) distribution

---

[4]I need to add structure on the time preference parameter in order to minimize (3.8). Specifically, I need to make $\delta(t, T)$ a function of just one free parameter. Table 3.1 shows the five specifications I investigate in this paper.

[5]Note that Bliss and Panigirtzoglou (2004) minimize (3.8) over γ and not over $\delta(t, T)$. Specifically,

function, $F_{t,T}^{(3)}$[6] as

$$F_{t,T}^{(3)}(r) = \frac{\int_{-\infty}^{r} \pi_{t,T}(x)x^{\gamma}dx}{\int_{-\infty}^{\infty} \pi_{t,T}(x)x^{\gamma}dx} \tag{3.12}$$

The denominator is time-varying due to time-varying state price densities. I can restate the denominator in terms of time preferences as $\delta(t,T) = \int_{-\infty}^{\infty} \pi_{t,T}(x)x^{\gamma}dx$. Essentially, at any point in time, for a given (constant) level of risk aversion, they force the time-varying time preference parameter to solve the equation $\int_{-\infty}^{\infty} \frac{\pi_{t,T}(x)x^{\gamma}}{\delta(t,T)}dx = 1$, simply because the resulting physical probabilities must integrate to one. This condition is similar to the one I impose, but, as noted above, I impose structure on time preferences and allow for time-varying risk aversion. Using one method over the other is simply a question of whether you find it more reasonable to have time-varying risk aversion or time-varying time preferences.

3 Data and Risk Neutral Distributions

I use the Ivy DB database from OptionMetrics to gather information on call and put options written on the S&P 500 index for the last trading day of every month. The data are from January 1996 to December 2015. I obtain implied volatilities, strikes, closing bid-prices, closing ask-prices, and maturities. As a proxy for the risk- free rate, I use the zero-coupon yield curve from the Ivy DB database. I also obtain expected dividend payments. I apply standard filters, excluding contracts with zero open interest, zero trading volume, quotes with best bid below \$0.50, and options with implied volatility higher than 100%.

Using prices of options written on the market, I estimate risk neutral distributions using the "Fast and Stable" method proposed by Jackwerth (2004). I follow the implementation procedure described in Jensen, Lando, and Pedersen (2017). For each last trading day of the month in the period from January 1996 until December 2015,

their minimization problem is:

$$\min_{\gamma} - 2\left(LL(0,0,1) - LL(\hat{a}, \hat{\beta}, \hat{\sigma})\right) \tag{3.10}$$

$$s.t. \quad \delta(t,T) \text{ solves } \int_{-\infty}^{\infty} \frac{\pi_{t,T}(x)x^{\gamma}}{\delta(t,T)}dx = 1, \quad \forall t \tag{3.11}$$

[6]Here the superscript in $F_{t,T}^{(3)}$ denotes the third specification in Table 3.1.

I estimate risk neutral distributions for two different horizons. I choose the horizon which is closest to 30 calender days along with the horizon with the shortest maturity that is longer than 30 calender days and which has observable options written on the market.

4 Testing the Estimated Distributions

Next, I test how the estimated physical distributions match the historical return distributions. I have estimated the parameters for each specification of the stochastic discount factor shown in Table 3.1 using the methods described in Section 2. Then I convert the risk neutral distributions into the corresponding physical return distributions using the estimated parameters of the stochastic discount factor.

For example, for specification (4) in Table 3.1, on the last trading day of January 1996, I estimate the monthly horizon risk neutral distribution, say $q_{t,t+1}(R_{t,t+1})$,[7] using options written on that date and convert this risk neutral distribution into a physical distribution:

$$p^{(4)}_{t,t+1}(R_{t,t+1}) = \frac{q_{t,t+1}(R_{t,t+1})}{m^{(4)}_{t,t+1}(R_{t,t+1})} \tag{3.13}$$

$$= \frac{q_{t,t+1}(R_{t,t+1})}{1.0056 \times R_{t,t+1}^{-3.17}} \tag{3.14}$$

where $\gamma(t) = 3.17$ was the market's risk aversion on the last trading day of January 1996 and $\delta(t,T) = 1.0056$ is the time preference parameter. Also, the superscript of $p^{(4)}_{t,t+1}(R_{t,t+1})$ denotes that I am talking about the 4'th specification of the stochastic discount factor in Table 3.1. I then compute the cumulative distribution function and evaluate the realized return over the subsequent month, $\tilde{R}_{t,t+1}$:

$$F^{(4)}_{t,t+1}(\tilde{R}_{t,t+1}) = \int_{-\infty}^{\tilde{R}_{t,t+1}} p^{(4)}_{t,t+1}(x)dx \tag{3.15}$$

For each last trading of the month from January 1996 until December 2015, I estimate these probabilities, giving me 240 non-overlapping periods with cumulative proba-

---

[7]If there are no observable options with exactly one month maturity, then I estimate risk neutral distributions for maturities closest to one month and linearly interpolate the cumulative probabilities estimated in (3.15) between these maturities.

bilities $F_{s,s+1}^{(4)}(\tilde{R}_{s,s+1})$ for $s \in \{1, ..., 240\}$. Given this time series of probabilities, I perform a Berkowitz test on both the full distribution and parts of the distribution. If the estimated distributions reflect the 'true' realized distribution, then $F_{t,t+1}^{(4)}(\tilde{R}_{t,t+1})$ is uniformly distributed and any partition of the distribution should not be rejected by the Berkowitz test. Therefore, I test if the estimated distribution reflects the 'true' distribution using multiple partitions of the uniform distribution.

Table 3.2 shows the result of Berkowitz tests of the five specifications of the stochastic discount factors as shown in Table 3.1. Panel A shows the results when testing the full distribution. Panel B reports results when testing partitions of the distribution. The second column reports the results when testing the risk neutral specification, that is, specification (1) in Table 3.1. The first row of Panel B reports the results when testing the distribution of $F_{t,t+1}^{(1)}(\tilde{R}_{t,t+1})$ on the interval from 0 to 0.5. This means that, when looking strictly at those cumulative probabilities that are lower than 0.5, I test whether these probabilities are uniform on 0 to 0.5. The test statistic is 5.20 which gives a p-value of 0.16 when evaluated in a χ^2-distribution with 3 degrees of freedom. Therefore, I cannot reject that the realized returns that fall within the lowest 50% of the risk neutral distributions are actually drawn from these distributions. Looking at the remaining rows, I can reject that the risk neutral distributions match the realized returns, that is, the test statistics for the intervals from 0 to 0.7 or higher are all statistically significant. The risk neutral distribution is therefore a poor proxy for the historical 'true' distribution. This result is not surprising as we expect investors to be risk averse and thereby adjust their physical beliefs accordingly when pricing financial assets.

The third column of Table 3.2 reports the results of the Berkowitz test when evaluating the second specification of the stochastic discount factor in Table 3.1. The distribution of $F_{t,t+1}^{(2)}(\tilde{R}_{t,t+1})$ is rejected on the interval from 0 to 0.8 and higher, which means that this specification is a poor match for the realized return distribution.

The fourth column of Table 3.2 reports the results of the Berkowitz test when evaluating the Bliss and Panigirtzoglou (2004) parameterization of the stochastic discount factor (3.2), that is, specification (3) in Table 3.1. Their specification of the stochastic discount factor performs very well. The estimated physical distributions, $F_{t,t+1}^{(3)}(\tilde{R}_{t,t+1})$, arising from a constant level of risk aversion, $\gamma = 2.44$, and time-varying time pref-

erences match realized returns very well. However, I can reject the full distribution estimated using the Bliss and Panigirtzoglou (2004) methodology simply because there are too few realized observations in the far right tail (upper 2.5%) of the estimated distributions.

The fifth column of Table 3.2 reports the result of the Berkowitz test when risk aversion is time-varying and time preferences are constant, that is, specification (4) in Table 3.1. The distribution $F_{t,t+1}^{(4)}(\tilde{R}_{t,t+1})$ cannot be rejected by the Berkowitz test on the interval from 0 to 0.975 and any partition within these boundaries. However, as for specification (3), the full distribution is rejected because of too few realizations in the far right tail of the estimated distribution.

The last column of Table 3.2 reports the result of the Berkowitz test when risk aversion is time-varying and time preferences are affine in the inverse of the gross risk-free return, that is, specification (5) in Table 3.1. As with specifications (3) and (4), the distribution $F_{t,t+1}^{(5)}(\tilde{R}_{t,t+1})$ is not rejected by the Berkowitz test on the interval from 0 to 0.975 and any partition within these boundaries but it is rejected on the full distribution.

Table 3.2 shows that the transformation from the risk neutral distributions to the physical distributions using constant relative risk aversion preferences with parameterization (3), (4), or (5) of the stochastic discount factor as shown in Table 3.1 can help explain historical realized returns. According to the Berkowitz tests, the three specifications perform equally well. This result is not surprising since the number of free parameters in each specification is the same. Specification (3) has 1 degree of freedom in the risk aversion parameter and 240 degrees of freedom in the time preference parameters (one for each month in my sample). Specification (4) has 240 free variables in the risk aversion parameters and 1 time preference parameter. Specification (5) has 240 free variables in the risk aversion parameters and 1 time preference parameter, however, this free variable in the time preference parameter measures the (constant) difference between the inverse of the gross risk-free return and the time preference parameter.

4.1 Empirical Risk Aversion Estimates

In this section, I focus on three of the specifications of the stochastic discount factor as shown in Table 3.1, that is, specifications (3), (4), and (5). Figure 3.1 shows empirical estimates of market risk aversion, that is, the market's time-varying risk aversion, γ_t, along with the constant risk aversion implied by the Bliss and Panigirtzoglou (2004) methodology.

As seen from the figure, the time-varying risk aversion implied by specifications (4) and (5) are almost identical. Therefore, for the remainder of the paper, I will focus on one just of them. My choice falls on specification (5), simply because it allows time preferences to be dependent on the risk-free rate.

Panel A of Table 3.3 reports summary statistics for the estimated risk aversion parameter. The market's time-varying risk aversion implied by specification (5) varies from 1.49 during the peak of the financial crisis in 2008-2009 to 5.03 in the post crisis period of 2014-2015. The average risk aversion is 2.77, slightly higher than the constant risk aversion estimate implied by the Bliss and Panigirtzoglou (2004) methodology which is 2.44.

4.2 Empirical Time Preference Estimates

Panel B of Table 3.3 reports summary statistics for the estimated time preference parameter. The time-varying time preferences implied by the Bliss and Panigirtzoglou (2004) methodology, specification (3), vary from 0.9985 to 1.0239. Similarly, the time preferences implied by specification (5) vary from 0.9982 to 1.0076 with an average value of 1.0057. The constant time preference parameter implied by specification (4) is 1.0056. A value of the time preference parameter which is larger than one implies that investors prefer future over immediate consumption. It is therefore surprising that the estimates of the time preference parameter are on average above one. However, given the form of the stochastic discount factor in (3.2), a value of the time preference parameter above one fits the data best.

To better understand the magnitude of the time preference parameter, I first investigate how a commonly used value for the time preference parameter fits the data.

Therefore, in specification (2) from Table 3.1, I fix $\delta(t,T) = 0.98^{T-t}$.[8] The second column of Table 3.2 shows the Berkowitz test for specification (2). The constant time preference parameter with an annualized value of 0.98 is strongly rejected by the data. This result means that, if $\delta(t,T) = 0.98^{T-t}$, there is no time series of $\gamma(t)$ which can simultaneously match the historical returns on the market and the observable option prices. From the point of view of a power utility investor who wants to hold the market, if we want to match both option prices and returns, then we need to accept that time preferences are on average above one.

It is worth noting that, even though $\delta(t,T)$ is above one, the setting I adopt will always (and by constriction) match interest rates. To understand how, recall that the law of one price gives us the following relation:

$$E_t \left[m_{t,T}(R_{t,T}) \right] = \frac{1}{R_{t,T}^f} \tag{3.16}$$

which for the power utility investor is equivalent to:

$$E_t \left[\delta(t,T) R_{t,T}^{-\gamma(t)} \right] = \int_{-\infty}^{\infty} p_{t,T}(x) \delta(t,T) x^{-\gamma(t)} dx \tag{3.17}$$

From the standard asset pricing formula in (3.3), I am given the relation $p_{t,T}(r) = \pi_{t,T}(r)/m_{t,T}(r)$ which is equal to $\pi_{t,T}(r) r^{\gamma(t)}/\delta(t,T)$ for the power utility investor. Inserting this expression of physical probabilities into (3.17), I get:

$$\int_{-\infty}^{\infty} \frac{\pi_{t,T}(x) x^{\gamma(t)}}{\delta(t,T)} \delta(t,T) x^{-\gamma(t)} dx = \int_{-\infty}^{\infty} \pi_{t,T}(x) dx = \frac{1}{R_{t,T}^f} \int_{-\infty}^{\infty} q_{t,T}(x) dx \tag{3.18}$$

where $q_{t,T}(r)$ is the risk-neutral distribution. The integral on the rhs. of (3.18) is equal to one since $q_{t,T}(r)$ is a probability distribution. Therefore, for any level of $\delta(t,T)$, my setting will by construction match interest rates.

Next, I want to better understand the scenarios in which $\delta(t,T) > 1$. Therefore, I consider the following equation which again arises from the law of one price and a

---

[8]E.g. Campbell and Cochrane (1993) derive an annualized value of the time preference parameter close to 0.98.

power utility investor who choses to hold the market

$$\delta(t,T) \int_{-\infty}^{\infty} p_{t,T}(x) x^{-\gamma(t)} dx = \frac{1}{R_{t,T}^f} \tag{3.19}$$

Now, if $\int_{-\infty}^{\infty} p_{t,T}(x) x^{-\gamma(t)} dx < \frac{1}{R_{t,T}^f}$ then $\delta(t,T)$ must, by the law of one price, be above one. This scenario can for example arise if the physical return distribution has most of its probability mass in the states where the market increases in value, that is, when $\int_1^{\infty} p_{t,T}(x) dx$ is large (close to one). In this case, the integral $\int_{-\infty}^{\infty} p_{t,T}(x) x^{-\gamma(t)} dx$ could potentially be lower than $1/R_{t,T}^f$ because $x^{-\gamma(t)} < 1$ for values of $x > 1$ and $\gamma(t) > 0$.

Therefore, a value of $\delta(t,T)$ which is larger than one can arise in a scenario where the probability of a positive return on the market is large. That is, the investor is more willing to postpone consumption into the future when the probability of a future good state (a state with high consumption) is high.[9] This result is highly unintuitive. I expected the exact opposite result, that the investor is willing to postpone (smooth) consumption if the probability of a bad state in the future is high.

The result may arise for several reasons. For example, it might be a consequence of a misspecified functional form of the stochastic discount factor. Also, it can be because the state variable is wrong, that is, market returns might not be a perfect proxy for aggregate consumption. There are of course several other possible explanations.

5   Systematic Variation in Market Risk Aversion

Figure 3.1 shows the market's time-varying risk aversion. As seen from the figure, risk aversion was low during the 2008-2009 financial crisis and high both in the years leading up to the crisis and in the post crisis period of 2010-2015. This variation in risk aversion suggests that the market became *less* risk averse during the peak of the financial crisis and more risk averse after the crisis. In this section, I investigate these time-varying patterns of market risk aversion.

First, using methods developed in Martin (2017) and Gormsen and Jensen (2017b), I estimate the market's physical return distribution. Table 3.4 shows the summary

---

[9]Similarly, a value of $\delta(t,T)$ which is lower than one can arise in a scenario where the probability of a negative return on the market is large. That is, the investor is more willing to consume today if the probability of a future bad state (a state with low consumption) is high.

statistics for the first four moments of the physical return distribution: expected excess return, variance, skewness, and kurtosis. Looking at the last row of Table 3.4 we see the market's time-varying moments implied by specification (5) of Table 3.1. The physical return distribution implied by this specification of the stochastic discount factor is on average negatively skewed and exhibit excess kurtosis. Results are similar for specification (3) and (4).

Figure 3.2 shows a time-series plot of the market's time-varying risk aversion and its physical conditional variance. Clearly, risk aversion tend to be low at turbulent times when variance is high. Panel A of Table 3.5 shows the correlations between the market's time-varying risk aversion and the market's physical moments. The correlation between market risk aversion and variance is -0.65 with bootstrapped 95% confidence bounds of $[-0.70, -0.62]$.

Figure 3.3 shows the market's risk aversion and its physical conditional skewness. Risk aversion and skewness are negatively correlated with a correlation of -0.42 and 95% bootstrapped confidence bounds of $[-0.51, -0.31]$ as shown in Table 3.5. The negative correlation between risk aversion and skewness suggests that the market is more risk averse at times when market tail risk is high.

The expected return on the market is negatively related to risk aversion with a correlation coefficient of -0.56 and bootstrapped 95% confidence bounds of $[-0.61, -0.52]$. This negative relation suggests that the expected return on the market is high during times of financial market distress when risk aversion is low. Similarly, the market's Sharpe ratio tends to be high during periods of high volatility. To better understand why the market's Sharpe ratio is high at times when risk aversion is low, notice that, under power utility and log-normal market returns, the expected excess return on the market is $ER_t = \gamma_t \sigma_t^2$. Rewriting this equation leads to the following expression of risk aversion:

$$\gamma_t = SR_t / \sigma_t \tag{3.20}$$

where $SR_t = ER_t / \sigma_t$ is the market's Sharpe ratio. Even though the market's Sharpe ratio might increase at times of financial market distress when volatility also increases, then the market's risk aversion can decrease if the increase in the Sharpe ratio is too low to offset the higher market volatility.

6 Understanding the Market's Time-Varying Risk Aversion

As shown in the previous section, market risk aversion varies in a systematic way, that is, it tends to be high at times when volatility is low and vice versa. A natural question to ask is; does it make sense at all that market risk aversion is *low* during times of financial distress when volatility is high? It is definitely difficult to reconcile this systematic variation in market risk aversion with the leading asset pricing models. For example, Campbell and Cochrane (1993) use habit formation as a mechanism for time-varying risk aversion and find that a countercyclical risk aversion can help explain asset prices and the behavior of the stock market.

As a start to understanding this systematic variation in market risk aversion, it is instructive to think about the following heuristic argument: standard textbook asset pricing tells us that, under power utility and a log-normal assumption on market returns, the expected excess return on the market is:

$$ER_t = \gamma_t \sigma_t^2 \tag{3.21}$$

During the financial crisis of 2008-2009, annualized monthly variance reached 60%. If I assume that risk aversion is constant and at a level of 3, which is a common choice in the financial literature[10], then (3.21) implies an annualized expected excess return of 180% and a monthly expected excess return of 15%. It is hard to believe that the market expected the excess returns on the monthly horizon was as high as 15% during the peak of financial crisis, not to mention the extremely high annualized expected excess returns. Therefore, since we can estimate market variance rather accurately, then if we want expected excess returns to be at a realistic level during times of high volatility, then market risk aversion should go down when volatility spikes.

As noted above, this systematic variation in market risk aversion is difficult to reconcile with the leading asset pricing models. However, when looking at the intersection between option prices and market returns, this variation seems to be a persistent pattern and it is therefore interesting to understand how it can arise in our data.

---

[10]See e.g. Bliss and Panigirtzoglou (2004) and Gormsen and Jensen (2017b)

6.1 Evidence of a Procyclical Market Risk Aversion

In a recent paper, Moreira and Muir (2017b) show that investors can earn large alphas by timing market volatility. They find that investors should exit the market when volatility increases and enter the market when it drops. Their result arises because expected excess returns are high relative to variance during low variance periods, i.e., the ratio $\frac{ER_t}{\sigma_t^2}$ varies over time and becomes high when variance is low. Similarly, Gormsen and Jensen (2017a) also find evidence that the market price of risk is high during low volatility periods.

These findings have implications for how we should think about the time variation in market risk aversion. To see why, consider again equation (3.21) which tells us that the expected excess return to variance ratio is equivalent to market risk aversion and their results are therefore reminiscent of a risk aversion that is high at times when volatility is low.

The results of Moreira and Muir (2017b) can also in part be explained by the fact that market tail risk is higher at times of low volatility, see e.g. Gormsen and Jensen (2017b). The higher expected excess return to variance during low volatility periods could simply be compensation for higher tail risk (or a combination of higher tail risk and higher risk aversion). Nevertheless, it is an artifact of the data that risk aversion tends to be high at times when the market is generally considered to be calm. In the following two subsections, I discuss two possible explanations for these time-varying patterns. First, I investigate a behavioral explanation through investor salience. Second, I consider a rational explanation through time-varying correlations between the market and aggregate consumption.

6.2 Salience Theory and Market Risk Aversion

In this subsection, I discuss investor salience as a possible explanation for the systematic variation in market risk aversion which I document in this paper. Bordalo, Gennaioli, and Shleifer (2012) develop a model in which the focus of investors are drawn to unusual, different or odd events. Specifically, an investor is risk-seeking if his attention is drawn to the "upside" of a lottery and risk-averse if his attention is drawn to the "downside" of a lottery. They refer to the state which has drawn the investor's

attention as the salient state or salient outcome.

A practical example of salience theory can be found in Lian, Ma, and Wang (2018). The authors show that, investors have greater appetite for risk taking when interest rates are low. They argue that, at times of low interest rates the relatively high expected returns on risky assets become salient, and this salience on the upside of a higher return on the risky asset induces heightened risk tolerance and "reaching for yield" tendencies. Also, they refer to Weber's law and argue that investors tend to evaluate assets in terms of proportional returns rather than by differences. In my setting both the risk-free rate and the expected return on the market fluctuate. Therefore, to test how their hypothesis fits into my setting, I regress my measure of risk aversion onto the ratio of gross expected returns to gross risk-free returns:[11]

$$\gamma(t) = \beta_0 + \beta_1 \frac{E_t(R_{t,T})}{R_{t,T}^f} + \epsilon_{t,T} \tag{3.22}$$

The idea is that, as expected returns on the risky asset increase relative to the safe return on the risk-free asset, investors will find the higher return on the risky asset salient and therefore increase their risk tolerance which results in a lower $\gamma(t)$.

Table 3.6 presents the results of regression (3.22). The coefficient β_1 is -6.97 with 95% bootstrapped confidence bounds of $[-8.63, -5.92]$. This results suggests that an increase in the ratio of gross expected return on the market to risk-free gross return of, say 0.1, results in a decrease in risk aversion of almost 0.7, that is, as expected returns on the market increase relative to the return on the risk-free asset, the investor becomes more risk tolerant.

6.3 Consumption and Stock Market Correlation

A natural critique of the method I apply in this paper is that the power utility investor only cares about how the stock market develops and gains all utility from market movements. In this subsection, I discuss what can drive my results if I in fact have

---

[11]Lian, Ma, and Wang (2018) argue that, when looking at gross returns the effect of salience is lower than when considering 'normal' returns because the ratio of, say 1.1/1.05, is much smaller than the ratio 0.1/0.05. Nevertheless, they find evidence that investors still exhibit salient preferences when considering gross returns. I do not consider 'normal' returns because they, at the end of my sample, are very close to zero.

postulated the wrong state variable.

Suppose for now that the true state variable is aggregate consumption and that the stock market is not a prefect proxy for aggregate consumption. Also, suppose that risk aversion is actually constant, say γ_c, the power utility investor's 'true' stochastic discount factor is then:

$$m_{t,T} = \delta^{T-t} \left(\frac{C_T}{C_t} \right)^{\gamma_c} \qquad (3.23)$$

where C_t is aggregate consumption at time t. If I assume that consumption is log-normally distributed with constant variance, σ_c^2, then in the standard Merton (1973) model, I can express expected return on the stock market in the following way:

$$ER_t = \gamma_c \sigma_t \sigma_c \rho_t \qquad (3.24)$$

where ρ_t is the time-varying correlation between the stock market and aggregate consumption. Keeping the stock market volatility constant, then as the correlation between the stock market and aggregate consumption increases, so does the required expected return on the stock market.

Now, if I define the time-varying parameter $\alpha_t = \frac{\rho_t \gamma_c \sigma_c}{\sigma_t}$[12] then I can rewrite (3.24) as

$$ER_t = \alpha_t \sigma_t^2 \qquad (3.25)$$

which is equivalent to imposing that the stock market is the state variable, log-normally distributed, and that the power utility investor has time-varying risk aversion α_t. Therefore, even though risk aversion is constant at γ_c, the stock market implied risk aversion can be time-varying because the correlation between consumption and the stock market is time-varying.

As a consequence of the assumptions made in this subsection, the time-varying correlation between consumption and the stock market is proportional to the market's Sharpe ratio, $\rho_t \propto \frac{ER_t}{\sigma_t}$, and the stock market implied time-varying risk aversion is proportional to the consumption to stock market correlation over stock market volatility,

---

[12]In a (more realistic) setting where risk aversion (γ_t^c), consumption volatility, and the correlation between the market and aggregate consumption are time-varying, then $\alpha_t = \frac{\rho_t \gamma_t^c \sigma_t^c}{\sigma_t}$ is the market's implied risk aversion as seen from the point of view of a power utility investor who holds the market. Clearly, in this setting, I cannot distinguish between the contributions in the time variation of α_t which comes from either of the time-varying parameters.

$\alpha_t \propto \frac{\rho_t}{\sigma_t}$.

The correlation between consumption and the stock market is proportional to the Sharpe ratio of the market and is therefore high during high volatility periods. For instance, during the recent financial crisis when stock market volatility spiked, then so did the correlation between consumption and the stock market. However, the increase in correlation between consumption and the stock market was not enough to offset the increase in stock market volatility and as a consequence stock market risk aversion decreased. Figure 3.4 shows the time series of implied correlations between consumption and the stock market when consumption risk aversion is constant at 38 as implied by the equity premium puzzle, cf. Mehra and Prescott (1985b), and monthly consumption volatility is constant at 0.0075.

Clearly, a setting in which both risk aversion and consumption volatility are constant is too simplified to fully capture the real world. Many of the conclusions which I draw in this subsection can be altered simply by assuming that, for example, risk aversion (γ_c) is time-varying or consumption volatility is time-varying. Nevertheless, the discussion in this subsection highlights that the time variation in market risk aversion, which I document in this paper, can be a consequence of a misinterpretation of the implied risk aversion. That is, the risk aversion implied by the market is time-varying and tends to be low at times when volatility is high because the mapping from the true state variable, aggregate consumption, onto the market is time-varying, e.g., through time-varying correlations.

7 Conclusion

I present a new method for estimating the market's time-varying risk aversion directly from option written on the market and historical market returns. My method combines the generalized recovery method of Jensen, Lando, and Pedersen (2017) and the Berkowitz test as used in Bliss and Panigirtzoglou (2004).

The estimated market risk aversion coefficients are on average 2.77 and vary from 1.49 to 5.03 over the period from 1996 to 2016. Interestingly, market risk aversion is negatively related to variance. During the peak of the financial crisis in 2008-2009 when market volatility was high, risk aversion was at its all-time low. These co-movements

between risk aversion and variance are difficult to reconcile with the leading asset pricing models.

I discuss two possible explanations for the negative relation between risk aversion and variance. First, I show that my results are consistent with salience theory. Specifically, during periods of high volatility when expected returns on the market are high relative to the risk-free returns, investors find the relatively high expected return on the market salient and therefore become more risk tolerant. Secondly, I show that my results can arise if the stock market is a poor proxy for aggregate consumption.

Table 3.1: **Specifications of the stochastic discount factor.** This table reports five different specifications of the CRRA stochastic discount factor that I use in this paper. The stochastic discount factors take the form:

$$m_{t,T}(R_{t,T}) = \delta(t,T)R_{t,T}^{-\gamma(t)} \tag{3.26}$$

Here $\delta(t,T)$ is a time preference parameter and $\gamma(t)$ is the coefficient of relative risk aversion.

| # | Comments | Parameterization | |
		$\delta(t,T)$	$\gamma(t)$
(1)	Risk neutral distribution	$1/R_{t,T}^f$	0
(2)	Fixed time preferences Time-varying risk aversion	0.98^{T-t}	γ_t
(3)	Time-varying time preferences (Bliss and Panigirtzoglou (2004)) Constant risk aversion	δ_t^{T-t}	γ
(4)	Constant time preferences Time-varying risk aversion	δ^{T-t}	γ_t
(5)	Time preferences affine in the inverse of the gross risk-free return Time-varying risk aversion	$1/R_{t,T}^f + \delta^{T-t} - 1$	γ_t

Table 3.2: **Berkowitz tests.** This table reports Berkowitz test statistics for five different specifications of the stochastic discount factor (SDF) as shown in Table 3.1. Specifically, I test if the distribution of $F_{t,T}(R_{t,T}) = \int_{-\infty}^{R_{t,T}} \frac{\pi_{t,T}(x)x^{\gamma(t)}}{\delta(t,T)}dx$ is uniform on the interval from 0 to 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.975, and 1. If the specification of the stochastic discount factor perfectly captures the 'true' historical distribution of realized returns, then any partition of the interval from 0 to 1 should not be rejected by the Berkowitz test. I report significance as: * when $p < 0.1$, ** when $p < 0.05$, and *** when $p < 0.01$. P-values are found by evaluating the test statistic in a χ^2-distribution with three degrees of freedom.

Panel A: The full distribution

	Specification of SDF (#)				
Uniform on	(1)	(2)	(3)	(4)	(5)
0 to 1	15.84***	14.08***	12.08***	12.33***	12.25***

Panel B: Partitions of the distribution

	Specification of SDF (#)				
Uniform on	(1)	(2)	(3)	(4)	(5)
0 to 0.5	5.20	6.14	3.38	0.35	1.26
0 to 0.6	5.53	3.49	1.49	1.41	1.84
0 to 0.7	8.06**	5.21	4.77	0.81	0.83
0 to 0.8	7.83**	6.94*	3.06	2.20	2.15
0 to 0.9	10.99**	7.83**	5.38	2.80	3.87
0 to 0.95	12.32***	7.24*	2.87	2.62	2.70
0 to 0.975	12.28***	9.76**	6.93*	6.98*	6.92*

Table 3.3: **Parameter summary statistics.** This table reports the average time-series values for the estimated preference parameters, $\delta(t,T)$ and $\gamma(t)$. The number (#) refers to the specification of the stochastic discount factor from Table 3.1. I do not show results for specification (1) and (2) since they were strongly rejected in the data as seen from Table 3.2.

Panel A: Risk aversion

#	$\gamma(t)$	Mean	St. dev.	Min	Median	Max
(3)	γ	2.44	-	-	-	-
(4)	γ_t	2.76	0.83	1.54	2.58	5.38
(5)	γ_t	2.77	0.71	1.49	2.62	5.03

Panel B: Time preferences

#	$\delta(t,T)$	Mean	St. dev.	Min	Median	Max
(3)	δ_t^{T-t}	1.0055	0.0046	0.9985	1.0047	1.0239
(4)	δ^{T-t}	1.0056	-	-	-	-
(5)	$1/R_{t,T}^f + \delta^{T-t} - 1$	1.0057	0.0019	0.9982	1.0063	1.0076

Table 3.4: **Moment summary statistics.** This table reports the average time-series values for estimated physical moments: expected excess return (ER−Rf), standard deviation (St. dev.), skewness, and kurtosis. Moments are estimated using methods developed in Gormsen and Jensen (2017b). The number (#) refers to the specification of the stochastic discount factor from Table 3.1. I do not show results for specification (1) and (2) since they were strongly rejected in the data as seen from Table 3.2.

	Parameterization		Annualized (%)			
#	$\delta(t,T)$	$\gamma(t)$	ER−Rf	St. dev.	Skewness	Kurtosis
(3)	δ_t^{T-t}	γ	9.98	18.81	−1.14	7.46
(4)	δ^{T-t}	γ_t	9.78	17.82	−1.09	7.09
(5)	$1/R_{t,T}^f + \delta^{T-t} - 1$	γ_t	9.83	17.81	−1.09	7.12

Table 3.5: **Correlations.** Panel A reports correlations between the market's time-varying risk aversion (estimated using specification (5) from Table 3.1) and the market's physical moments estimated using methods developed in Gormsen and Jensen (2017b): expected excess return (ER−Rf), variance, skewness , and kurtosis. Panel B reports correlations between the market's physical moments. The number (#) refers to the specification of the stochastic discount factor from Table 3.1.

Panel A: Risk aversion and moment correlations

#		ER−Rf	Variance	Skewness	Kurtosis
(5)	γ_t	−0.56	−0.65	−0.42	0.63
		$[-0.61, -0.52]$	$[-0.70, -0.62]$	$[-0.51, -0.31]$	$[0.54, 0.70]$

Panel B: Moment correlations

#		ER-Rf	Variance	Skewness	Kurtosis
(5)	ER−Rf	1	0.97	0.36	−0.50
			$[0.96, 0.98]$	$[0.25, 0.46]$	$[-0.56, -0.44]$
(5)	Variance		1	0.41	−0.53
				$[0.32, 0.50]$	$[-0.60, -0.49]$
(5)	Skewness			1	−0.75
					$[-0.80, -0.69]$
(5)	Kurtosis				1

Table 3.6: **Salience Theory.** This table reports the regression result when regressing the market's risk aversion onto the ratio of gross expected return on the market to gross risk-free returns:

$$\gamma(t) = \beta_0 + \beta_1 \frac{E_t(R_{t,T})}{R_{t,T}^f} + \epsilon_{t,T} \qquad (3.27)$$

$\gamma(t)$ is estimated using specification (5) in Table 3.1. The expected return on the market is estimated using methods developed in Gormsen and Jensen (2017b). As the ratio of expected return on the risky asset to the risk-free return increases, then the high expected return on the risky asset becomes salient making the investor more risk tolerant. I report 95% bootstrapped confidence bounds on the estimates.

	β_0	β_1	Adj. R^2
Coefficient	10.41	-6.97	0.29
95% CI	$[8.89, 11.93]$	$[-8.63, -5.92]$	
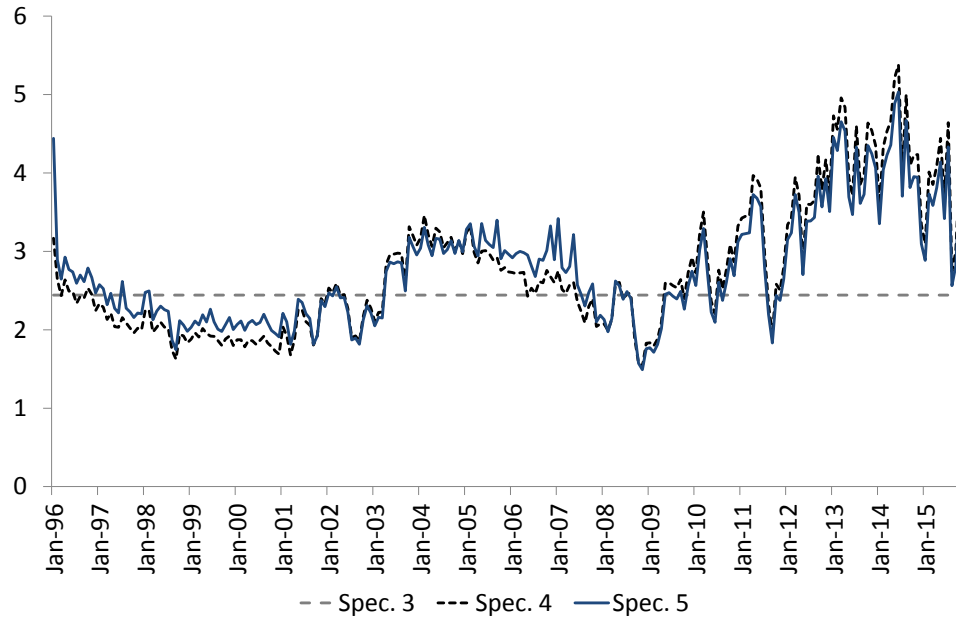
Figure 3.1: **Risk aversion estimates.** This figures shows risk aversion estimates for specification (3), (4), and (5) from Table 3.1. The solid blue line is when risk aversion is time-varying and time preferences are proportional to the gross risk-free return. The dashed black line is when risk aversion is time-varying and time preferences are constant. The dashed gray line is the constant risk aversion implied by the Bliss and Panigirtzoglou (2004) methodology. I do not show results for specification (1) and (2) since they were strongly rejected in the data as seen from Table 3.2.
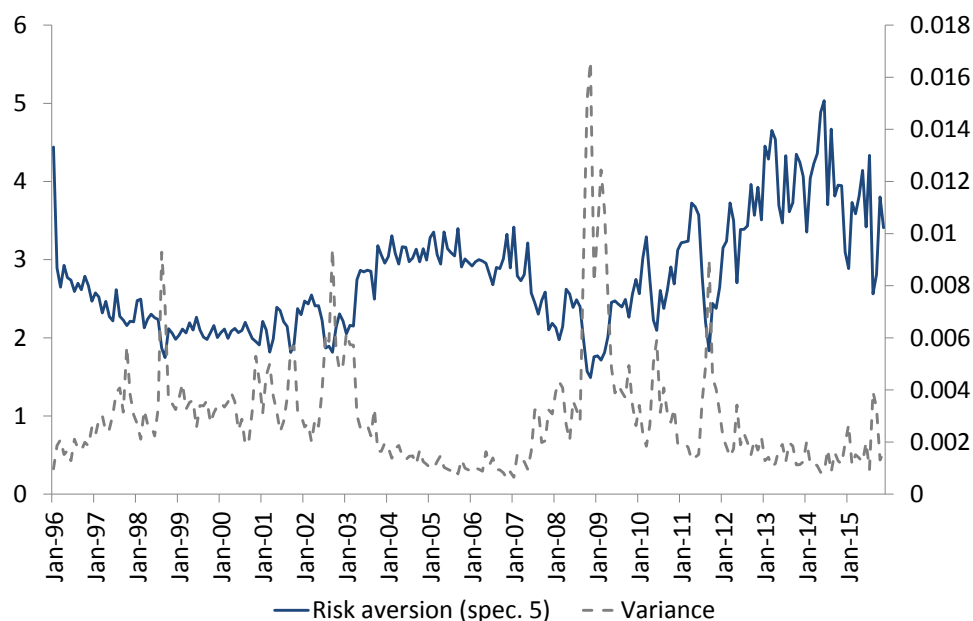
Figure 3.2: **Risk aversion and market variance.** This figure shows the market's time-varying risk aversion (left axis) and the physical conditional variance of the market (right axis). Conditional variance is estimated using methods developed in Martin (2017) and Gormsen and Jensen (2017b). Risk aversion is estimated using specification (5) from Table 3.1. Correlation: −0.65.
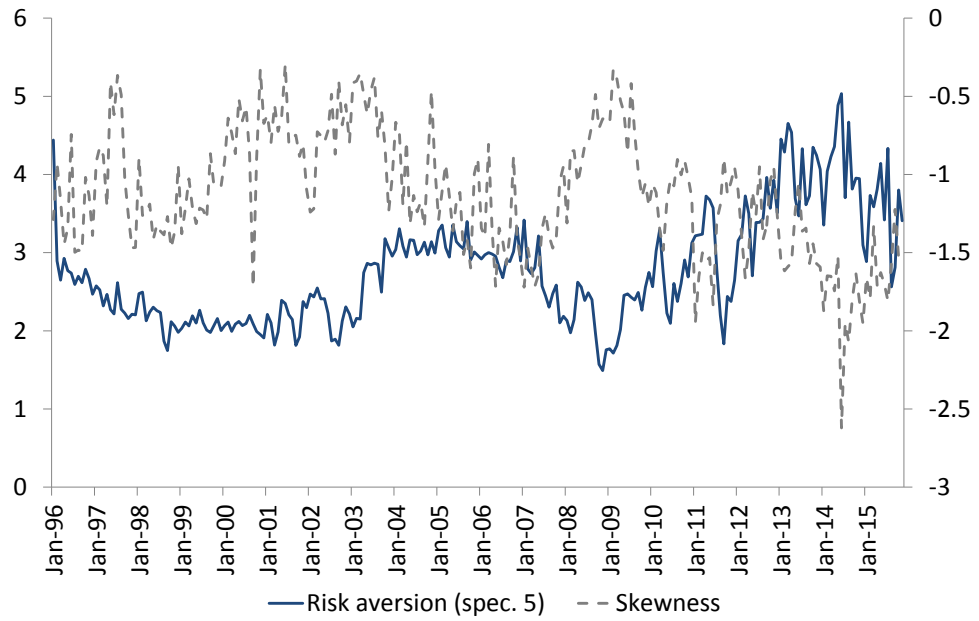
Figure 3.3: **Risk aversion and market skewness.** This figure shows the market's time-varying risk aversion (left axis) and the physical conditional skewness of the market (right axis). Conditional skewness is estimated using methods developed in Martin (2017) and Gormsen and Jensen (2017b). Risk aversion is estimated using specification (5) from Table 3.1. Correlation: -0.42.
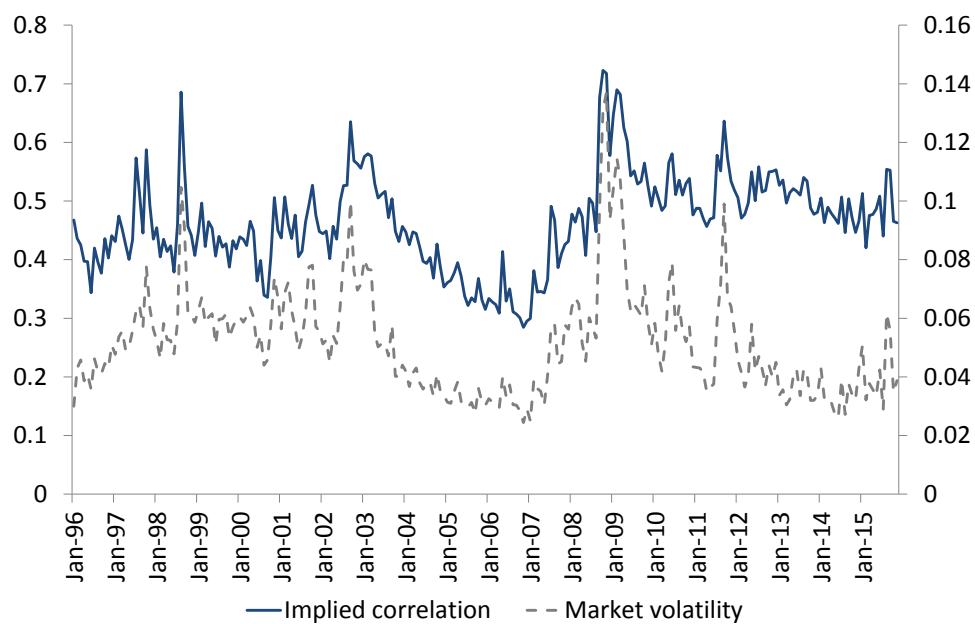
Figure 3.4: **Implied consumption and stock market correlation.** This figure shows the implied correlation between the stock market and aggregate consumption (left axis) and the markets condition monthly horizon volatility (right axis). Conditional variance is estimated using methods developed in Martin (2017) and Gormsen and Jensen (2017b)

Bibliography

Adrian, Tobias, and Nina Boyarchenko, 2012, Intermediary leverage cycles and financial stability, *Working paper*.

Adrian, Tobias, Erkko Etula, and Tyler Muir, 2014, Financial intermediaries and the cross-section of asset returns, *The Journal of Finance* 69, 2557—-2596.

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

Amaya, Diego, Peter Christoffersen, Kris Jacobs, and Aurelio Vasquez, 2015, Does realized skewness predict the cross-section of equity returns?, *Journal of Financial Economics* 118, 135–167.

Andersen, Torben, Nicola Fusari, and Viktor Todorov, 2015, The risk premia imbedded in index options, *Journal of Financial Economics* 117(3), 558–584.

Asness, Cliff, Andrea Frazzini, and Lasse Heje Pedersen, 2012, Leverage aversion and risk parity, *Financial Analysts Journal* 68, 47—-59.

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 Bakshi, 2017, A recovery that we can trust? deducing and testing the restrictions of the recovery theorem, *Review of Financial Studies, forthcoming*.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies*.

Bakshi, Gurdip, and Dilip Madan, 2000, Spanning and derivative-security valuation, *Journal of Financial Economics.*

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, Benjamin Hood, John Huss, and Lasse Heje Pedersen, 2016, Risk everywhere: Modeling and managing volatility, *Duke, working paper.*

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *The Review of Financial Studies* 22, 4463–4492.

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

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2012, Salience theory of choice under risk, *The Quarterly Journal of Economics* pp. 1243–1285.

Borovicka, Jaroslav, Lars Peter Hansen, and Jose A. Scheinkman, 2016, Misspecified recovery, *Journal of Finance* 71, 2493–2544.

Breeden, Douglas T., and Robert H. Litzenberger, 1978, Prices of state-contingent claims implicit in option prices, *The Journal of Business* 51, 621—-651.

Brunnermeier, Markus, and Yuliy Sannikov, 2014, A macroeconomic model with a financial sector, *American Economic Review* 104, 379–421.

Brunnermeier, Markus .K., and Lasse H. Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Campbell, John Y., and John H. Cochrane, 1993, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

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

Chen, Joseph, Harrison Hong, and Jeremy C Stein, 2001, Forecasting crashes: Trading volume, past returns and conditional skewness in stock prices, *Journal of Financial Economics* 61, 345–391.

Christoffersen, Peter, Bruno Feunou, Yoontae Jeon, and Chayawat Ornthanalai, 2016, Time-varying crash risk: The role of stock market liquidity, *Working paper*.

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

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

Danilova, Albina, and Christian Julliard, 2015, Information asymmetries, volatility, liquidity, and the tobin tax, *Working paper*.

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

Denbee, Edward, Christian Julliard, Ye Li, and Kathy Yuan, 2016, Network risk and key players: A structural analysis of interbank liquidity, *Working paper*.

Ebert, Sebastian, 2013, Moment characterization of higher-order risk preferences, *Theory and Decision* 74, 267—-284.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Gao, George, Pengjie Gao, and Zhaogang Song, 2017, Do hedge funds exploit rare disaster concerns?, *Review of Financial Studies forthcoming.*

Gao, George, Xiaomeng Lu, and Zhaogang Song, 2017, Tail risk concerns everywhere, *Management Science, forthcoming.*

Gormsen, Niels J., and Christian S. Jensen, 2017a, Conditional risk, *Copenhagen Business School working paper.*

Gormsen, Niels J., and Christian S. Jensen, 2017b, Higher-moment risk, *Copenhagen Business School working paper.*

Greenwood, Robin, Andrei Shleifer, and Yang You, 2017, Bubbles for Fama, *Journal of Financial Economics, forthcoming.*

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

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2016, Intermediary asset pricing: New evidence from many asset classes, *forthcoming in Journal of Financial Economic.*

He, Zhiguo, and Arvind Krishnamurthy, 2013, Intermediary asset pricing, *American Economic Review* 103, 732—-770.

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

Jackwerth, Jens C., 2004, Option-implied risk-neutral distributions and risk aversion, *CFA Institute Research Foundation of AIMR Publications* pp. 1–86.

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

Jensen, Christian Skov, David Lando, and Lasse Heje Pedersen, 2017, Generalized recovery, *forthcoming Journal of Financial Economics.*

Kazemi, Hossein, 1992, An intertemporal model of asset prices in a markov economy with a limiting stationary distribution, *Review of Financial Studies* 5, 85–104.

Kelly, Bryan, and Hao Jiang, 2014, Tail risk and asset prices, *The Review of Financial Studies* 27, 2841—-2871.

Lazarus, Eben, 2018, Horizon dependent risk pricing: Evidence from short-dated options, *Working paper*.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *The Journal of Finance* 56, 815—-849.

Lian, Chen, Yueran Ma, and Carmen Wang, 2018, Low interest rates and risk taking: Evidence from individual decisions, *Working paper*.

Linetsky, Vadim, and Likuan Qin, 2016, Positive eigenfunctions of markovian pricing operators: Hansen-Scheinkman factorization, Ross recovery and long-term pricing, *Operations Research* 64, 99–117.

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

Martin, Ian, 2017, What is the expected return on the market?, *Quarterly Journal og Economics* 132(1), 367—433.

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

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

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

Merton, R., 1973, Theory of rational option pricing, *The Bell Journal of Economics and Management Science* 4, 141–183.

Moreira, Alan, and Tyler Muir, 2017a, Should long-term investors time volatility?, *Journal of Financial Economics, forthcoming*.

Moreira, Alan, and Tyler Muir, 2017b, Volatility-managed portfolios, *The Journal of Finance* 72, 1611—-1644.

Moskowitz, Tobias, Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228—-250.

Neuberger, Anthony, 2012, Realized skewness, *The Review of Financial Studies* 25, 3423–3455.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *The Journal of Finance* 63, 921—-945.

Qin, Likuan, and Vadim Linetsky, 2017, Long term risk: A martingale approach, *Econometrica* 85, 299–312.

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, 2017a, (almost) model-free recovery, *Journal of Finance, Forthcoming.*

Schneider, Paul, and Fabio Trojani, 2017b, Fear trading, *Working paper.*

Siriwardane, Emil, 2015, The probability of rare disasters: Estimation and implications, *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, 2017, Recovery with unbounded diffusion processes, *Review of Finance* 21, 1403–1444.

TITLER I PH.D.SERIEN:

2004

1. Martin Grieger
   Internet-based Electronic Marketplaces and Supply Chain Management

2. Thomas Basbøll
   LIKENESS
   A Philosophical Investigation

3. Morten Knudsen
   Beslutningens vaklen
   En systemteoretisk analyse of moderniseringen af et amtskommunalt sundhedsvæsen 1980-2000

4. Lars Bo Jeppesen
   Organizing Consumer Innovation
   A product development strategy that is based on online communities and allows some firms to benefit from a distributed process of innovation by consumers

5. Barbara Dragsted
   SEGMENTATION IN TRANSLATION AND TRANSLATION MEMORY SYSTEMS
   An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process

6. Jeanet Hardis
   Sociale partnerskaber
   Et socialkonstruktivistisk casestudie af partnerskabsaktørers virkelighedsopfattelse mellem identitet og legitimitet

7. Henriette Hallberg Thygesen
   System Dynamics in Action

8. Carsten Mejer Plath
   Strategisk Økonomistyring

9. Annemette Kjærgaard
   Knowledge Management as Internal Corporate Venturing

   – a Field Study of the Rise and Fall of a Bottom-Up Process

10. Knut Arne Hovdal
    De profesjonelle i endring
    Norsk ph.d., ej til salg gennem Samfundslitteratur

11. Søren Jeppesen
    Environmental Practices and Greening Strategies in Small Manufacturing Enterprises in South Africa
    – A Critical Realist Approach

12. Lars Frode Frederiksen
    Industriel forskningsledelse
    – på sporet af mønstre og samarbejde i danske forskningsintensive virksomheder

13. Martin Jes Iversen
    The Governance of GN Great Nordic
    – in an age of strategic and structural transitions 1939-1988

14. Lars Pynt Andersen
    The Rhetorical Strategies of Danish TV Advertising
    A study of the first fifteen years with special emphasis on genre and irony

15. Jakob Rasmussen
    Business Perspectives on E-learning

16. Sof Thrane
    The Social and Economic Dynamics of Networks
    – a Weberian Analysis of Three Formalised Horizontal Networks

17. Lene Nielsen
    Engaging Personas and Narrative Scenarios – a study on how a user-centered approach influenced the perception of the design process in the e-business group at AstraZeneca

18. S.J Valstad
    Organisationsidentitet
    Norsk ph.d., ej til salg gennem Samfundslitteratur

Effectiveness of Grocer Media Advertising
Measuring Ad Recall and Recognition, Purchase Intentions and Short-Term Sales

11. Allan Mortensen
Essays on the Pricing of Corporate Bonds and Credit Derivatives

12. Remo Stefano Chiari
*Figure che fanno conoscere
Itinerario sull'idea del valore cognitivo e espressivo della metafora e di altri tropi da Aristotele e da Vico fino al cognitivismo contemporaneo*

13. Anders McIlquham-Schmidt
*Strategic Planning and Corporate Performance
An integrative research review and a meta-analysis of the strategic planning and corporate performance literature from 1956 to 2003*

14. Jens Geersbro
*The TDF – PMI Case
Making Sense of the Dynamics of Business Relationships and Networks*

15. Mette Andersen
*Corporate Social Responsibility in Global Supply Chains
Understanding the uniqueness of firm behaviour*

16. Eva Boxenbaum
Institutional Genesis: Micro – Dynamic Foundations of Institutional Change

17. Peter Lund-Thomsen
Capacity Development, Environmental Justice NGOs, and Governance: The Case of South Africa

18. Signe Jarlov
Konstruktioner af offentlig ledelse

19. Lars Stæhr Jensen
Vocabulary Knowledge and Listening Comprehension in English as a Foreign Language

An empirical study employing data elicited from Danish EFL learners

20. Christian Nielsen
*Essays on Business Reporting
Production and consumption of strategic information in the market for information*

21. Marianne Thejls Fischer
Egos and Ethics of Management Consultants

22. Annie Bekke Kjær
*Performance management i Proces-innovation
– belyst i et social-konstruktivistisk perspektiv*

23. Suzanne Dee Pedersen
*GENTAGELSENS METAMORFOSE
Om organisering af den kreative gøren i den kunstneriske arbejdspraksis*

24. Benedikte Dorte Rosenbrink
*Revenue Management
Økonomiske, konkurrencemæssige & organisatoriske konsekvenser*

25. Thomas Riise Johansen
*Written Accounts and Verbal Accounts
The Danish Case of Accounting and Accountability to Employees*

26. Ann Fogelgren-Pedersen
The Mobile Internet: Pioneering Users' Adoption Decisions

27. Birgitte Rasmussen
Ledelse i fællesskab – de tillidsvalgtes fornyende rolle

28. Gitte Thit Nielsen
*Remerger
– skabende ledelseskræfter i fusion og opkøb*

29. Carmine Gioia
A MICROECONOMETRIC ANALYSIS OF MERGERS AND ACQUISITIONS

*A case study of the Fashion and
Design Branch of the Industrial District
of Montebelluna, NE Italy*

12. Mikkel Flyverbom
*Making the Global Information Society
Governable*
*On the Governmentality of Multi-
Stakeholder Networks*

13. Anette Grønning
Personen bag
*Tilstedevær i e-mail som inter-
aktionsform mellem kunde og med-
arbejder i dansk forsikringskontekst*

14. Jørn Helder
One Company – One Language?
The NN-case

15. Lars Bjerregaard Mikkelsen
*Differing perceptions of customer
value*
*Development and application of a tool
for mapping perceptions of customer
value at both ends of customer-suppli-
er dyads in industrial markets*

16. Lise Granerud
Exploring Learning
*Technological learning within small
manufacturers in South Africa*

17. Esben Rahbek Pedersen
*Between Hopes and Realities:
Reflections on the Promises and
Practices of Corporate Social
Responsibility (CSR)*

18. Ramona Samson
*The Cultural Integration Model and
European Transformation.*
The Case of Romania

2007
1. Jakob Vestergaard
*Discipline in The Global Economy
Panopticism and the Post-Washington
Consensus*

2. Heidi Lund Hansen
*Spaces for learning and working
A qualitative study of change of work,
management, vehicles of power and
social practices in open offices*

3. Sudhanshu Rai
*Exploring the internal dynamics of
software development teams during
user analysis*
*A tension enabled Institutionalization
Model; "Where process becomes the
objective"*

4. Norsk ph.d.
Ej til salg gennem Samfundslitteratur

5. Serden Ozcan
*EXPLORING HETEROGENEITY IN
ORGANIZATIONAL ACTIONS AND
OUTCOMES*
A Behavioural Perspective

6. Kim Sundtoft Hald
*Inter-organizational Performance
Measurement and Management in
Action*
*– An Ethnography on the Construction
of Management, Identity and
Relationships*

7. Tobias Lindeberg
Evaluative Technologies
*Quality and the Multiplicity of
Performance*

8. Merete Wedell-Wedellsborg
*Den globale soldat
Identitetsdannelse og identitetsledelse
i multinationale militære organisatio-
ner*

9. Lars Frederiksen
*Open Innovation Business Models
Innovation in firm-hosted online user
communities and inter-firm project
ventures in the music industry
– A collection of essays*

10. Jonas Gabrielsen
*Retorisk toposlære – fra statisk 'sted'
til persuasiv aktivitet*

36. Annegrete Juul Nielsen
Traveling technologies and transformations in health care

37. Athur Mühlen-Schulte
Organising Development
Power and Organisational Reform in the United Nations Development Programme

38. Louise Rygaard Jonas
Branding på butiksgulvet
Et case-studie af kultur- og identitets- arbejdet i Kvickly

2011

1. Stefan Fraenkel
Key Success Factors for Sales Force Readiness during New Product Launch A Study of Product Launches in the Swedish Pharmaceutical Industry

2. Christian Plesner Rossing
International Transfer Pricing in Theory and Practice

3. Tobias Dam Hede
Samtalekunst og ledelsesdisciplin – en analyse af coachingsdiskursens genealogi og governmentality

4. Kim Pettersson
Essays on Audit Quality, Auditor Choi- ce, and Equity Valuation

5. Henrik Merkelsen
The expert-lay controversy in risk research and management. Effects of institutional distances. Studies of risk definitions, perceptions, management and communication

6. Simon S. Torp
Employee Stock Ownership: Effect on Strategic Management and Performance

7. Mie Harder
Internal Antecedents of Management Innovation

8. Ole Helby Petersen
Public-Private Partnerships: Policy and Regulation – With Comparative and Multi-level Case Studies from Denmark and Ireland

9. Morten Krogh Petersen
'Good' Outcomes. Handling Multipli- city in Government Communication

10. Kristian Tangsgaard Hvelplund
Allocation of cognitive resources in translation - an eye-tracking and key- logging study

11. Moshe Yonatany
The Internationalization Process of Digital Service Providers

12. Anne Vestergaard
Distance and Suffering Humanitarian Discourse in the age of Mediatization

13. Thorsten Mikkelsen
Personligheds indflydelse på forret- ningsrelationer

14. Jane Thostrup Jagd
Hvorfor fortsætter fusionsbølgen ud- over "the tipping point"? – en empirisk analyse af information og kognitioner om fusioner

15. Gregory Gimpel
Value-driven Adoption and Consump- tion of Technology: Understanding Technology Decision Making

16. Thomas Stengade Sønderskov
Den nye mulighed Social innovation i en forretningsmæs- sig kontekst

17. Jeppe Christoffersen
Donor supported strategic alliances in developing countries

18. Vibeke Vad Baunsgaard
Dominant Ideological Modes of Rationality: Cross functional

integration in the process of product innovation

19. Throstur Olaf Sigurjonsson
    Governance Failure and Icelands's Financial Collapse

20. Allan Sall Tang Andersen
    Essays on the modeling of risks in interest-rate and inflation markets

21. Heidi Tscherning
    Mobile Devices in Social Contexts

22. Birgitte Gorm Hansen
    Adapting in the Knowledge Economy Lateral Strategies for Scientists and Those Who Study Them

23. Kristina Vaarst Andersen
    Optimal Levels of Embeddedness The Contingent Value of Networked Collaboration

24. Justine Grønbæk Pors
    Noisy Management A History of Danish School Governing from 1970-2010

25. Stefan Linder
    Micro-foundations of Strategic Entrepreneurship Essays on Autonomous Strategic Action

26. Xin Li
    Toward an Integrative Framework of National Competitiveness An application to China

27. Rune Thorbjørn Clausen
    Værdifuld arkitektur Et eksplorativt studie af bygningers rolle i virksomheders værdiskabelse

28. Monica Viken
    Markedsundersøkelser som bevis i varemerke- og markedsføringsrett

29. Christian Wymann
    Tattooing The Economic and Artistic Constitution of a Social Phenomenon

30. Sanne Frandsen
    Productive Incoherence A Case Study of Branding and Identity Struggles in a Low-Prestige Organization

31. Mads Stenbo Nielsen
    Essays on Correlation Modelling

32. Ivan Häuser
    Følelse og sprog Etablering af en ekspressiv kategori, eksemplificeret på russisk

33. Sebastian Schwenen
    Security of Supply in Electricity Markets

2012
1. Peter Holm Andreasen
    The Dynamics of Procurement Management - A Complexity Approach

2. Martin Haulrich
    Data-Driven Bitext Dependency Parsing and Alignment

3. Line Kirkegaard
    Konsulenten i den anden nat En undersøgelse af det intense arbejdsliv

4. Tonny Stenheim
    Decision usefulness of goodwill under IFRS

5. Morten Lind Larsen
    Produktivitet, vækst og velfærd Industrirådet og efterkrigstidens Danmark 1945 - 1958

6. Petter Berg
    Cartel Damages and Cost Asymmetries

7. Lynn Kahle
    Experiential Discourse in Marketing A methodical inquiry into practice and theory

8. Anne Roelsgaard Obling
    Management of Emotions in Accelerated Medical Relationships

31. Fumiko Kano Glückstad
Bridging Remote Cultures: Cross-lingual concept mapping based on the information receiver's prior-knowledge

32. Henrik Barslund Fosse
Empirical Essays in International Trade

33. Peter Alexander Albrecht
Foundational hybridity and its reproduction
Security sector reform in Sierra Leone

34. Maja Rosenstock
CSR - hvor svært kan det være?
Kulturanalytisk casestudie om udfordringer og dilemmaer med at forankre Coops CSR-strategi

35. Jeanette Rasmussen
Tweens, medier og forbrug
Et studie af 10-12 årige danske børns brug af internettet, opfattelse og for-ståelse af markedsføring og forbrug

36. Ib Tunby Gulbrandsen
'This page is not intended for a US Audience'
A five-act spectacle on online communication, collaboration & organization.

37. Kasper Aalling Teilmann
Interactive Approaches to Rural Development

38. Mette Mogensen
The Organization(s) of Well-being and Productivity
(Re)assembling work in the Danish Post

39. Søren Friis Møller
From Disinterestedness to Engagement Towards Relational Leadership In the Cultural Sector

40. Nico Peter Berhausen
Management Control, Innovation and Strategic Objectives – Interactions and Convergence in Product Development Networks

41. Balder Onarheim
Creativity under Constraints
Creativity as Balancing 'Constrainedness'

42. Haoyong Zhou
Essays on Family Firms

43. Elisabeth Naima Mikkelsen
Making sense of organisational conflict
An empirical study of enacted sense-making in everyday conflict at work

2013
1. Jacob Lyngsie
Entrepreneurship in an Organizational Context

2. Signe Groth-Brodersen
Fra ledelse til selvet
En socialpsykologisk analyse af forholdet imellem selvledelse, ledelse og stress i det moderne arbejdsliv

3. Nis Høyrup Christensen
Shaping Markets: A Neoinstitutional Analysis of the Emerging Organizational Field of Renewable Energy in China

4. Christian Edelvold Berg
As a matter of size
THE IMPORTANCE OF CRITICAL MASS AND THE CONSEQUENCES OF SCARCITY FOR TELEVISION MARKETS

5. Christine D. Isakson
Coworker Influence and Labor Mobility Essays on Turnover, Entrepreneurship and Location Choice in the Danish Maritime Industry

6. Niels Joseph Jerne Lennon
Accounting Qualities in Practice Rhizomatic stories of representational faithfulness, decision making and control

7. Shannon O'Donnell
Making Ensemble Possible
How special groups organize for collaborative creativity in conditions of spatial variability and distance

TITLER I ATV PH.D.-SERIEN

1992
1. Niels Kornum
    Servicesamkørsel – organisation, øko-
    nomi og planlægningsmetode

1995
2. Verner Worm
    Nordiske virksomheder i Kina
    Kulturspecifikke interaktionsrelationer
    ved nordiske virksomhedsetableringer i
    Kina

1999
3. Mogens Bjerre
    Key Account Management of Complex
    Strategic Relationships
    An Empirical Study of the Fast Moving
    Consumer Goods Industry

2000
4. Lotte Darsø
    Innovation in the Making
    Interaction Research with heteroge-
    neous Groups of Knowledge Workers
    creating new Knowledge and new
    Leads

2001
5. Peter Hobolt Jensen
    Managing Strategic Design Identities
    The case of the Lego Developer Net-
    work

2002
6. Peter Lohmann
    The Deleuzian Other of Organizational
    Change – Moving Perspectives of the
    Human

7. Anne Marie Jess Hansen
    To lead from a distance: The dynamic
    interplay between strategy and strate-
    gizing – A case study of the strategic
    management process

2003
8. Lotte Henriksen
    Videndeling
    – om organisatoriske og ledelsesmæs-
    sige udfordringer ved videndeling i
    praksis

9. Niels Christian Nickelsen
    Arrangements of Knowing: Coordi-
    nating Procedures Tools and Bodies in
    Industrial Production – a case study of
    the collective making of new products

2005
10. Carsten Ørts Hansen
    Konstruktion af ledelsesteknologier og
    effektivitet

TITLER I DBA PH.D.-SERIEN

2007
1. Peter Kastrup-Misir
    Endeavoring to Understand Market
    Orientation – and the concomitant
    co-mutation of the researched, the
    re searcher, the research itself and the
    truth

2009
1. Torkild Leo Thellefsen
    Fundamental Signs and Significance
    effects
    A Semeiotic outline of Fundamental
    Signs, Significance-effects, Knowledge
    Profiling and their use in Knowledge
    Organization and Branding

2. Daniel Ronzani
    When Bits Learn to Walk Don't Make
    Them Trip. Technological Innovation
    and the Role of Regulation by Law
    in Information Systems Research: the
    Case of Radio Frequency Identification
    (RFID)

2010
1. Alexander Carnera
    Magten over livet og livet som magt
    Studier i den biopolitiske ambivalens