Tails, Fears, and Equilibrium Option Prices*

David Schreindorfer†

July 23, 2014

Abstract

Recent empirical evidence suggests that the compensation for rare events accounts for a large fraction of the average equity and variance premia. I replicate this fact in a parsimonious consumption-based asset pricing model based on a (generalized) disappointment averse investor and conditionally Gaussian fundamentals. In the model, regime-switches in endowment volatility interact with the investor's tail aversion to produce large endogenous return jumps and a realistic implied volatility smirk. The presence of multiple shock frequencies in volatility gives the variance premium the ability to predict returns over short horizons and the price-dividend ratio the ability to predict returns over long horizons.

*First draft: June 1, 2013. I am grateful to many people who have given me advise, including Ravi Bansal, Sreedhar Bharath, Darrell Duffie, Anisha Ghosh, Burton Hollifield, Emilio Osambela, Ioanid Rosu, Alex Schiller, Fallaw Sowell, Tan Wang, and seminar participants at Arizona State University, University of British Columbia, Carnegie Mellon University, Duke University, Georgetown University, HEC Paris, Northwestern University, and Penn State University. I am especially thankful to Lars-Alexander Kuehn, Bryan Routledge, and Duane Seppi for invaluable advice and guidance. The latest version of the paper is available at www.davidschreindorfer.com

†Department of Finance, W. P. Carey School of Business, Arizona State University, PO Box 873906, Tempe, AZ 85287, david.schreindorfer@asu.edu

1 Introduction

Prices of equity index options provide significant information about the composition of risk premia in financial markets. A key observations is that deep out-of-the-money (OTM) index put options appear overpriced from the perspective of standard models. This suggests that investors are more concerned with large declines in the aggregate stock market than typically assumed. Indeed, recent nonparametric estimates in Bollerslev and Todorov (2011) show that about two thirds of the average equity premium represent compensation for extreme left tail events, defined as returns of -10% or less over the horizon of a few weeks. This finding is puzzling because such returns occur very infrequently in the data.

I show that investors' fear of extreme events can account for average option prices in a consumption-based asset pricing model with fundamentals that are conditionally Gaussian. The representative agent has Epstein-Zin (1989) utility with generalized disappointment aversion (GDA) risk preferences (Routledge and Zin 2010). Relative to the more commonly-used recursive utility function with expected utility (EU) risk aggregation, GDA can overweight lower-tail outcomes in the distribution over future aggregate wealth and produce a high price of tail risk. Cash flows (consumption and dividend growth) have constant means and time-varying volatility. Volatility follows the multifractal process of Calvet and Fisher (2001, 2004), which allows for shocks with different persistence levels and generates substantial volatility feedback in equilibrium, i.e. large endogenous return jumps.[1] The combination of tail risk resulting from volatility feedback and the aversion to such risks implied by GDA preferences allows the model to replicate the steep implied volatility smirk implied by option prices.

The model successfully reproduces a number of additional stylized facts of asset markets and the equity index option market in particular, including (1) the high average equity premium, (2) the low risk-free rate, (3) the large excess volatility of returns relative to fundamentals, (4) the predictive ability of the price-dividend ratio for excess returns over long horizons, (5) the high excess kurtosis of monthly returns and the low excess

[1] Calvet and Fisher (2007) have previously shown that a model with recursive utility with EU risk preferences generates substantial volatility feedback. Different from the present paper, these authors focus on a model where multifractal volatility affects dividends but not consumption. In this case, volatility risks are not priced because they do not affect the pricing kernel.

kurtosis of annual returns (6) the high average variance premium[2] and its ability to predict excess returns over short horizons and (7) the time-series moments of option prices, which are reflected in the term structure of variance swap rates. At the same time, the model remains tightly parameterized.

In the model, the probability of disappointments increases in times of high macroeconomic uncertainty, during which extreme events are more likely. When disappointments occur, they tend to be accompanied by large negative returns that trigger payoffs to out-of-the-money put options. Puts thus provide a hedge against disappointments and the investor is willing to pay a large insurance premium for them in states of *high* uncertainty. In times of *low* uncertainty, disappointments are rare and insurance premia only account for a small part of put option prices. In these states the IV smirk instead arises from mean reversion in the conditional volatility of cash flows. Specifically, when volatility is low and therefore expected to increase, the price-dividend ratio is expected to *decrease* because volatility carries a negative price of risk in the model. As reductions in the price-dividend ratio translate into negative returns, the conditional return distribution is left-skewed when volatility is low.[3] In these states, the model produces a steep implied volatility curve due to the relatively high likelihood of negative return jumps rather than due to high insurance premia.

I show that the magnitude of the conditional return skewness depends crucially on the nature of endowment volatility. When volatility is modeled as a persistent AR(1) process, mean reversion occurs too slowly to have a strong impact on the conditional skewness of short horizon returns. Instead, it affects the conditional skewness at longer horizons. A model with AR(1) volatility (and GDA preferences) therefore produces an IV curve that is steep at long maturities but too flat at short maturities. In contrast, multifractal volatility allows for mean reversion at different time scales. In particular, the shocks with low persistence levels mean revert quick enough to produce substantial return skewness at short horizons, and they allow the model to match the steep IV

[2]The variance premium equals the difference between the risk-neutral and statistical expectations of future market variance. Details are discussed in Section 2.

[3]For similar reasons, the conditional return distribution (under the statistical measure) is right-skewed when volatility is high. All else equal, this makes the IV curve flat or even upward sloping. However, I show that the effect of GDA preferences is strong enough to induce strong left-skewness in the *risk neutral* distribution when volatility is high.

curve at both short and long maturities.

To summarize the *dynamics* of option prices I turn to variance swaps, i.e. forward contracts on realized stock market variance (see Section 2 for more details). These instruments can be replicated from portfolios of equity index options and they display some revealing patterns about the shocks that drive option prices. Specifically, swaps with a monthly maturity are considerably more volatile and less persistent than swaps with an annual maturity. Additionally, the autocorrelation functions of swap rates display pronounced long-memory behavior – a well-known feature of return volatility (Ding, Granger, and Engle 1993). The model is able to capture these complex time series features despite relying on a very parsimonious endowment specification. In particular, the multifractal volatility process depends on only four parameters, regardless of the number of volatility components. I show that these parameters can be calibrated such that the model provides almost an exact match for the term structure of variance swap rates. In contrast, the dynamics of option prices and time series properties of the conditional return volatility are counterfactual when cash flow volatility is modeled as a persistent AR(1) process. The present paper is the first, to my knowledge, to replicate the term structure of variance swap rates in an endowment economy.

Return predictability arises in the model from the interaction between time-varying risk (stochastic endowment volatility) and time-varying risk aversion. As previously emphasized by Routledge and Zin (2010), GDA preferences are capable of producing time-variation in effective risk aversion when combined with a persistent state variable. In the present economy, endowment volatility (the state variable) equals the product of several volatility components with different persistence levels. All else equal, when a component with a given frequency is currently in its high state, the conditional endowment volatility is higher and there is a greater chance of disappointing tail events. The overweighting of tail events in the GDA utility function raises the agent's effective risk aversion when these outcomes are more likely to occur, which results in higher expected returns. However, volatility components with different frequencies have a different effect on expected returns. Specifically, less persistent shocks alter expected returns over short horizons whereas more persistent shocks alter expected returns over long horizons. As time-variation in the price-dividend ratio is mainly driven by persistent shocks, it is

a good predictor of long horizon returns. On the other hand, the (1-month) variance premium has considerably more exposure to transient shocks, making it a better predictor of short horizon returns. These implications for the term structure of risk premia agree with the empirical results of Martin (2013), who derives a lower bound on the equity premium from option prices and concludes that the equity premium has both a business cycle component and a higher-frequency component.

An important question is whether GDA preferences imply a reasonable level of risk aversion. To compare level of effective risk aversion implied by different risk preference calibrations, I conduct a welfare analysis in the spirit of Lucas (1987). Specifically, I hold the parameters controlling time preference and the endowment calibration constant, and consider various risk preference calibrations that all match the historical equity premium. The set of alternatives includes the (pure) disappointment aversion (DA) model of Gul (1991) as well as the EU model, i.e. the Epstein-Zin specification used by Bansal and Yaron (2004). For each economy, I then compute the welfare costs of heteroscedasticity risk as well as the welfare costs of total endowment risk. The results show that the GDA agent is less risk averse overall, but that he is considerably more averse to stochastic volatility than both EU and DA agents. The reason is that for a constant endowment variance, the extreme tail outcomes constituting a disappointment for the GDA agent are very rare and risk premia are close to zero. In line with the empirical findings of Bollerslev and Todorov (2011) discussed above, risk premia in the model therefore primarily arise from the aversion to extreme tail risks.

The rest of the paper is structured as follows. Section 1.1 points to connections with the existing literature. In Section 2, I discuss the option dataset, define option-related statistics, and present the set of stylized option market facts that serves as the empirical target. The model and the associated solution technique are shown in Section 3, whereas Section 4 shows calibration results and discusses the model mechanism. Section 5 illustrates how option pricing implications and effective risk aversion change for nested preference specifications, thereby shedding additional light on the mechanism. The appendix contains details on both the data and the model solution technique.

1.1 Connections with Prior Literature

Rare disasters and option prices. An alternative mechanism for increasing the importance of tail risks is the rare disaster framework of Rietz (1988) and Barro (2006). While the assumption of a Peso problem greatly improves the asset pricing implications of simple representative agent models, Backus, Chernov, and Martin (2011) point out that the implied volatility smirk in these models is far steeper and lower than in the data. These authors argue that, contrary to the large and rare disasters assumed in Barro's calibration, equity index options imply relatively small and frequent consumption disasters. A recent paper by Seo and Wachter (2013) shows that a model with *stochastic* disaster intensity and recursive references can reconcile this conflicting evidence and generate a more realistic smirk. Du (2011) shows that a combination of rare disasters and external habit formation in preferences also produces a smirk, but his results only focus on options on the consumption claim.

In order to generate realistic prices of OTM puts, the above papers rely on the strong assumption that there is no default on option payoffs in the case of a macroeconomic disaster. In contrast, the current model increases the importance of tail risk by increasing its price rather than its quantity, i.e. it does not assume a Peso problem. As a consequence, the most severe drops in consumption are far less extreme and the no-default assumption is less restrictive. A second important difference is that my model matches not only the implied volatility smirk but also the variance premium. While these two phenomena are intimately linked, the question of whether disaster models are consistent with the high premium for variance risk has not been addressed yet.

Long run risks and option prices. A number of prior papers also build on Bansal and Yaron (2004) to study the implied volatility smirk or the variance premium. These studies extend the cash flow dynamics of the basic long run risks model by allowing for stochastic volatility-of-volatility (Bollerslev, Tauchen, and Zhou (2009)) or jumps in the conditional moments of consumption and dividend growth (Benzoni, Collin-Dufresne, and Goldstein (2011), Drechsler and Yaron (2011), Drechsler (2013)). Incorporating jumps into the state variable processes increases the quantity of tail risk because they endogenously map into jumps in returns. For example, a negative jump in the expected growth rate of dividends induces a discrete reduction in the price-dividend ratio

and therefore a negative jump in returns. Similarly, I model endowment volatility as a Markov chain, which allows for discrete changes in volatility that result in endogenous return jumps. Different from the long run risks model, I do not rely on persistent variation in the mean of consumption growth to generate high average risk premia. Additionally, I am able to rely on much more modest jumps in the conditional volatility than the aforementioned papers because GDA preferences amplify the pricing effect of such shocks.

The only prior model that explains the smirk as well as the variance premium and its predictive power for excess returns is Drechsler (2013), and I discuss his paper in a bit more detail. Drechsler extends the long run risks model by adding jumps with time-varying intensity to the first two conditional cash flow moments. The precise dynamics of fundamentals are assumed to be uncertain, and the representative agent is averse to this ambiguity. Variance swaps and out-of-the-money put options pay off when large jumps materialize, which is most likely if the benchmark model underestimates the true jump intensity. These assets thus provide insurance against model misspecification and the representative agent is willing to pay a large premium for them. Lastly, because the degree of ambiguity aversion is modelled as a stochastic process, the economy features time-variation in effective risk aversion. This channel induces strong co-movement between the equity premium and the variance premium and allows the model to match the predictive power of the variance premium.

There are several important differences between Drechsler's study and mine. First, my model is considerably more parsimonious, relying on just over half as many free parameters (11 vs. 20). While the recent quantitative asset pricing literature often ignores parsimony as a model evaluation criterion, it lies at the heart of the Lucas critique and has long been considered an important feature of a successful structural economic model.[4] Second, my model matches a number of additional data features that are intimately linked with index option prices. Specifically, it accounts for the very high kurtosis of monthly returns, the time series moments of implied volatilities (in the form of the term structure of variance swap rates), and the ability of the variance premium and the price-dividend ratio to predict excess returns over different horizons. Lastly,

[4]See Wachter (2002) for a forceful discussion of this point in the asset pricing context.

whereas Drechsler's model generates a smirk for a particular value of its state variables, I illustrate that my model is capable of generating a data-like smirk in finite samples. To my knowledge, my model is the first to do so.

Nature and pricing of consumption volatility shocks. Nakamura, Sergeyev, and Steinsson (2012) estimate a long run risks model based on the international consumption data of Barro and Ursua (2008) (without using any asset market data) and find strong evidence in favor of priced consumption volatility shocks. Boguth and Kuehn (2013) find similar evidence based on the cross-section of U.S. consumption data alone and they do not assume a particular asset pricing model. Their evidence additionally suggests that time-variation in the conditional mean of consumption growth is not priced in the cross-section. Tamoni (2011) provides empirical evidence for the presence of consumption volatility shocks with highly heterogeneous persistence levels, and he shows that models with a single shock frequency have counterfactual implications for the long-run relationship between consumption growth volatility and expected returns in the data. In agreement with this empirical evidence, time-variation in risk premia in the present model is driven by consumption volatility shocks with different persistence levels, and I not not model time-variation in the conditional mean.

2 Stylized Facts of Equity Index Options

This section discusses the option dataset and presents the set of stylized option market facts that serves as a target for the model. Specifically, I focus on moments of equity index options and variance swap rates with different maturities, as well as the variance premium and its predictive power for excess returns.

2.1 Data Sources

The option dataset, which spans the period from January 2, 1990 to December 31, 2012, was obtained from Market Data Express, a subsidiary of the Chicago Board Options Exchange (CBOE). It contains end-of-day information for all option contracts traded on the CBOE for which the S&P 500 index is the underlying asset. Variables include trading volume, open interest, and the daily open, high, low, and last sales prices. I approximate the option price by the mid quote. The exercise style of the options is European. On average, the dataset contains observations on 852 different option contracts per day, which amounts to approximately $5m$ total observations.[5] I apply standard filters to the data (see Appendix A for details). For estimating the realized variance of market returns, I use tick-by-tick transaction prices of S&P500 futures for the same sample period, which were obtained from TICKDATA.

2.2 Variance Swap Rates

A variance swap is a forward contract on the underlying's future variance. At maturity, the seller (floating leg) pays the asset's realized variance, defined as the sum of squared daily log returns over the term of the contract, i.e.

$$RV_{t:t+\tau} = \sum_{i=1}^{\tau} r_{t+i}^2, \tag{1}$$

where τ denotes the number of trading days. The buyer (fixed leg) pays the variance swap rate $\mathcal{S}_t(\tau)$, which is agreed upon at contract initiation.[6] The variance swap rate can therefore be interpreted as the (forward) price of the underlying's realized variance. In the absence of arbitrage opportunities, the swap rate equals the risk neutral conditional

[5] The average number of daily observations increased from 270 in 1990 to 2600 in 2012.

[6] In reality, payments are netted. Also, the difference between RV and $\mathcal{S}$ is typically multiplied by a factor that converts variances to annual units, as well as a notional. These details are irrelevant for the purposes of this paper.

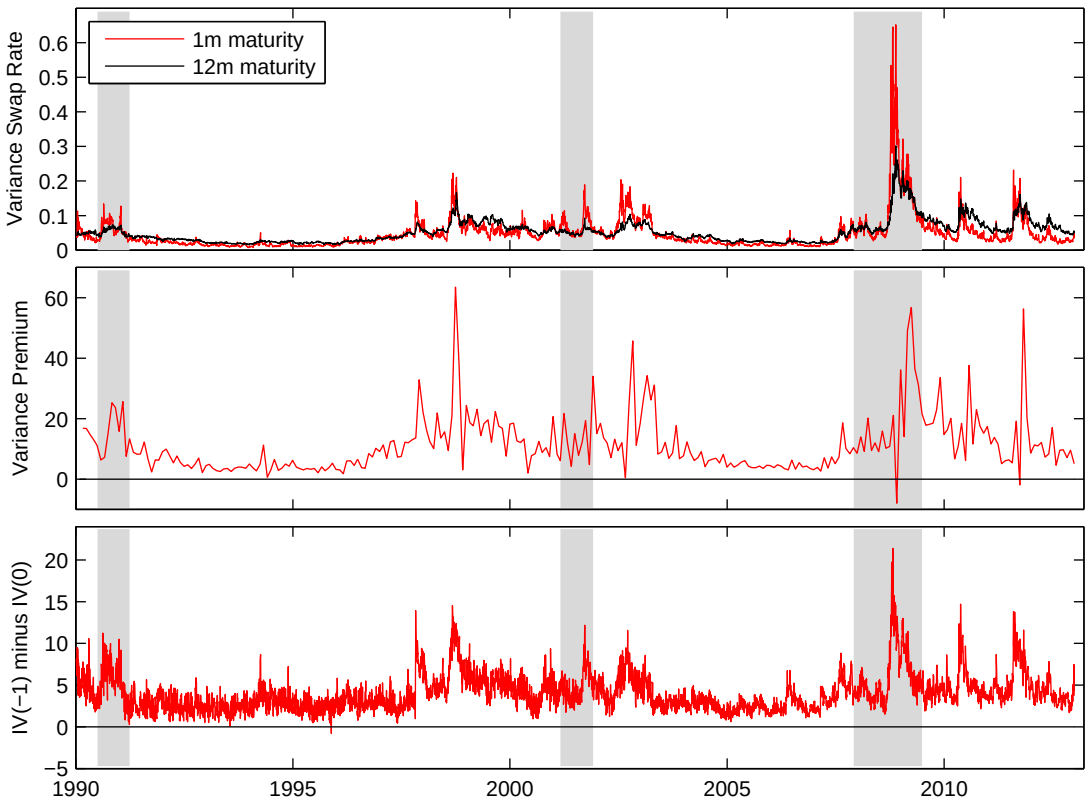

Figure 1: Equity Index Option Markets, 1990-2012

Figure 1 shows time series of option-related statistics. The top panel shows variance swap rates of maturities 1 and 12 months, expressed in annualized variance units. The middle panel shows the 1-month variance premium, expressed in monthly variance units. The lower panel shows the slope of the implied volatility (IV) curve, defined as the difference between the IV for a standardized moneyness of -1 and the IV for an at-the-money option. IVs are expressed in annualized standard deviation units. The time series for swaps and IVs are daily, whereas the time series for the variance premium is monthly. Shaded regions represent NBER recessions.

expectation of future variance, i.e.

$$\mathcal{S}_t(\tau) = E_t^{\mathbb{Q}}[RV_{t:t+\tau}]. \tag{2}$$

The swap payoff can be replicated with a static portfolio of European options and a dynamic, self-financing position in the underlying and a bond. It follows that the price of the option portfolio equals the variance swap rate. The proof is an extension of the classic Breeden and Litzenberger (1978) result, which asserts that the second derivative of the call price with respect to the strike price X equals the risk neutral density

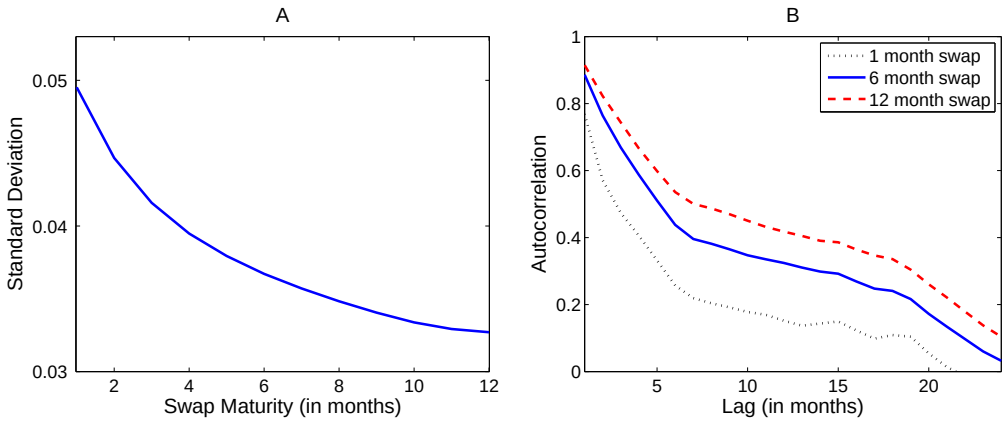

Figure 2: The volatility and persistence of variance swap rates

Figure 2 shows moments of variance swap rates, expressed in annualized variance units. The sample is daily and spans 1990-2012. Swap rates are expressed in annualized variance units. To compute the autocorrelation function, I create multiple overlapping monthly samples, compute the autocorrelation function for each sample, and then average across estimates.

evaluated at X and multiplied by the price of a risk-free bond[7] (see Britten-Jones and Neuberger (2000), Jiang and Tian (2005) and Carr and Wu (2009)). Using this replication result, which also underlies the CBOE's volatility index VIX, I compute synthetic variance swap rates for various maturities (τ's in the notation above). Specifically, for each day of the sample, I compute swap rates for all available option maturities and linearly interpolate them to constant maturities from 1 to 12 months. Further details of the implementation are discussed in Appendix A.

The top panel of Figure 1 shows the daily time series of variance swap rates of maturities of one and twelve months. Both series and very countercyclical and peak in the Fall of 2008 after the collapse of Lehman Brothers. Further, the one-month swap rate is both more volatile and less persistent than the 12-months swap rate. This feature is illustrated further in Figure 2. Panel A shows standard deviation of variance swap rates for maturities of $1, 2, ..., 12$ months. The standard deviation falls monotonically in the swap horizon. The high volatility of short maturity swaps indicates the presence of some high-frequency shocks that partially average out at the longer horizons. Panel B shows the autocorrelation function for maturities 1, 6, and 12 months, and for monthly lags up to two years. The autocorrelation of all lags increases monotonically in the swap

[7] In symbols, $\frac{\partial^2 C(k)}{\partial k^2}|_{k=X} = e^{-rT} f^Q(X)$, where $C(k)$ denotes the price of a call with strike k, r denotes the risk-free rate, and T denotes the maturity of the option.

Table 1: The Variance Premium and Expected Returns

Horizon (in m)	1	3	6
$\hat{\beta}$	0.89	0.88	0.65
t-statistic	2.10	4.10	3.29
R^2 (%)	2.53	6.49	6.52

Table 1 presents return predictability regressions. Excess returns of horizons 1, 3, and 6 months are regressed on the one-month variance premium. Regressions with horizons > 1 month use overlapping data. T-statistics are Newey and West (1987) (HAC) adjusted using $2 * (h - 1)$ lags.

horizon.[8] Further, the autocorrelation of all swap maturities declines very slowly with the lag length, which reflects the well-known long-memory property of return variances (Ding, Granger, and Engle (1993)). The slow decay indicates the presence of some low-frequency shocks with very long-lasting effects.

2.3 The Variance Premium

Using the Euler equation, the τ-period variance swap rate can be decomposed into the statistical expectation of realized variance and the variance premium, i.e.

$$\mathcal{S}_t(\tau) = E_t[RV_{t:t+\tau}] + \underbrace{\frac{Cov_t[M_{t:t+\tau}, RV_{t:t+\tau}]}{E_t[M_{t:t+\tau}]}}_{\text{variance premium}}, \tag{3}$$

where $M_{t:t+\tau}$ denotes a τ-period pricing kernel.[9] Measuring the (conditional) variance premium empirically requires an estimate of both the variance swap rate and the conditional expectation of realized variance. For the latter, I first compute a time series of monthly realized variance estimates from tick-by-tick transaction data, and then I estimate a simple time series model based on the realized variance series (see, e.g. Andersen, Bollerslev, Diebold, and Ebens (2001) and Andersen, Bollerslev, Diebold,

[8] Using option data for 12 major international equity indices, Foresi and Wu (2005) document that short term option prices (expressed as IVs) are both more volatile and less persistent than long term option prices. These patterns are therefore a robust empirical feature of equity index options in that they hold for both (synthetic) variance swap rates and simple implied volatilities.

[9] Note that, while the equity premium equals the difference between the physical and risk neutral expectations of future returns, i.e. $E_t[R_{t+1}] - R_t^f = E_t[R_{t+1}] - E_t^{\mathbb{Q}}[R_{t+1}]$, Equation 3 defines the variance premium as the difference between the risk neutral and the physical expectations of future realized variance, i.e. $\mathcal{S}_t(1) - E_t[RV_{t:t+1}] = E_t^{\mathbb{Q}}[RV_{t:t+1}] - E_t[RV_{t:t+1}]$. I use this "reversed" definition because it makes the variance premium positive and because it corresponds to the convention in most of the previous literature.

and Labys (2003)). The difference between the one-month variance swap rate and the one-step-ahead forecast from the time-series model serves as a proxy for the 1-month variance premium. Details are contained in Appendix A. For comparability with previous studies, I express the variance premium in *monthly* variance units.

The middle panel of Figure 1 shows the monthly time series of the variance premium, which is positive in all but 2 out of 275 months in the sample. The mean variance premium equals 11.29, which amounts to almost 40% of the average realized variance. The large magnitude of the premium suggests that variance swaps provide a hedge for macroeconomic risks. In other words, the market variance correlates positively with investors' marginal utility (see Equation 3).

Table 1 show predictability regressions for excess returns, which I measure as the difference between the value-weighted CRSP return and the yield of a 30 day Treasury bill. The variance premium can account for 2.5% of the return variation at the monthly horizon, and for about 6.5% at horizons of 3 and 6 months. Compared to other known predictors, these magnitudes are quite large for the short time horizons.[10] The equity premium and the variance premium thus share common factors at high frequencies.

2.4 The Implied Volatility Smirk

Throughout the paper, I express option prices in terms of Black-Scholes implied volatilities (IVs) and I graph them against standardized moneyness. I measure standardized moneyness as

$$\text{standardized moneyness } = \frac{\ln(X/S_t)}{\sqrt{\mathcal{S}_t(\tau)/\tau}},$$

where X denotes the option's strike price and S_t the underlying's price. The division by $\sqrt{\mathcal{S}_t(\tau)/\tau}$ converts moneyness, $\ln(X/S_t)$, to standard deviation units. This standardization allows for an easy comparison of the IV curve across option maturities.[11]

[10] The predictive ability of the variance premium was first documented by Bollerslev, Tauchen, and Zhou (2009).

[11] Many papers graph IVs against simple moneyness ($\ln(X/S_t)$). This is helpful for comparing the IV curve across different points in time. On the other hand, it makes it more challenging to compare it across option maturities. For example, if one considers a moneyness range of $[-10\%, +10\%]$, the endpoints of this range correspond to fairly extreme price moves at the monthly horizon, but to much more common moves at the annual horizon. I follow Carr and Wu (2003) and Foresi and Wu (2005) in using a standardized moneyness measure.

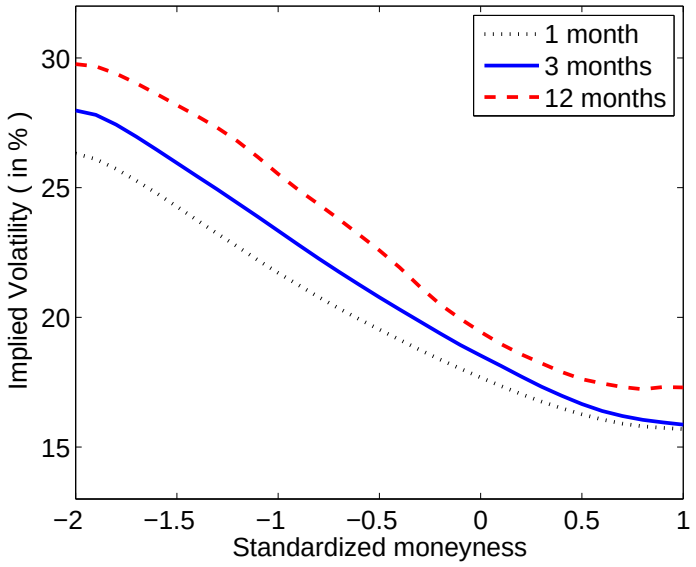

Figure 3: The Implied Volatility Smirk, 1990-2012

Figure 3 shows average (Black-Scholes) implied volatilities as a function of the options' relative moneyness. The sample is daily and spans 1990-2012.

A standardized moneyness of -1 is equivalent to a one standard deviation drop in the stock price. I interpolate observed IVs to a fixed grid of standardized moneyness from -2 to 1 and maturities from 1 to 12 months. The bounds of the grids are chosen such that it is covered by liquid options on most days in the sample. As before, details are discussed in Appendix A.

Figure 3 shows the IV curve for option maturities of 1, 6, and 12 months. According to the Black-Scholes model, IVs across all strikes and maturities should be equal to the volatility of the underlying asset. In the data, IVs are considerably higher for low strikes (for low values of standardized moneyness), a fact first documented by Rubinstein (1994). Additionally, the figure shows that both the level and the slope of the IV smirk are increasing in maturity. This fact is well-known – see e.g. Foresi and Wu (2005), who show that the same pattern holds for 12 major international equity markets. Deviations from log-normality in the risk-neutral distribution are therefore more severe at longer horizons. In particular, the large negative skewness exists even at long horizons, whereas long-horizon log returns are close to normally distributed under the statistical measure. Long maturity options are therefore particularly informative about plausible preference and endowment specifications in asset pricing models.

3 Model

3.1 The Economy

Preferences. Following Epstein and Zin (1989), the representative agent's time t utility, V_t, is given by the constant elasticity of substitution recursion

$$V_t = [(1-\beta)C_t^{\rho} + \beta\mu_t^{\rho}]^{\frac{1}{\rho}}. \tag{4}$$

The parameters β and ρ capture time preferences, whereas the function $\mu_t \equiv \mu_t(V_{t+1})$ captures risk preferences. μ_t equals the certainty equivalent of random future utility using the time t conditional probability distribution. I omit the argument for notational convenience. The certainty equivalent features Generalized Disappointment Aversion (GDA) as in Routledge and Zin (2010) (hereafter RZ), and it is defined by the implicit function

$$u(\mu_t) = E_t\Big[u(V_{t+1})\Big] - \theta E_t\Big[\Big(u(\delta\mu_t) - u(V_{t+1})\Big)\mathbf{1}\{V_{t+1} \leq \delta\mu_t\}\Big], \tag{5}$$

where

$$u(x) = \begin{cases} \frac{x^{\alpha}}{\alpha} & \text{for } \alpha \leq 1, \alpha \neq 0 \\ \log(x) & \text{for } \alpha = 0 \end{cases}, \tag{6}$$

and where $\mathbf{1}\{\cdot\}$ denotes the indicator function. Equation 5 nests two well-known preference specifications as special cases. First, for $\theta = 0$ the second term drops out and risk preferences simplify to expected utility (hereafter EU). In this case, the certainty equivalent is given by the *explicit* function $\mu_t = (E_t[V_{t+1}^{\alpha}])^{\frac{1}{\alpha}}$ (or by $\mu_t = e^{\log(V_{t+1})}$ for $\alpha = 0$) and the utility function equals the Epstein-Zin specification used in Bansal and Yaron (2004). Second, for $\delta = 1$ and $\theta \neq 0$ risk preferences simplify to Gul's (1991) model of disappointment aversion (hereafter DA).[12] In this case, all outcomes that fall below the certainty equivalent are considered disappointing and receive a penalty. The magnitude of the penalty is governed by the parameter θ. GDA preferences, which represent the most general version of Equation 5 ($\theta \neq 0$ and $\delta \neq 0$), place the disappointment threshold further into the tail of the (conditional) distribution of V_{t+1}. In particular, only realizations of V_{t+1} that fall below a fraction δ of the certainty equivalent μ_t are

[12] DA preferences were developed to resolve the Allais (1979) paradox. The asset pricing implications of recursive utility with DA risk preferences have been analyzed by Epstein and Zin (2001) in an endowment economy and by Campanale, Castro, and Clementi (2010) in a production economy. Neither paper analyzes option prices. I show in Section 5.1 that DA risk preferences are not capable of generating a realistic implied volatility smirk or variance premium in my framework.

considered disappointing. In Sections 4 and 5, I use the nested cases to highlight why GDA is needed for realistic option prices.

RZ show that the solution to the representative agent's portfolio optimization problem yields the pricing kernel

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{\rho-1} \left(\frac{V_{t+1}}{\mu_t}\right)^{\alpha-\rho} \left(\frac{1+\theta \mathbf{1}\left\{V_{t+1} \leq \delta\mu_t\right\}}{1+\delta^{\alpha}\theta \mathbf{1} E_t\left[\mathbf{1}\left\{V_{t+1} \leq \delta\mu_t\right\}\right]}\right). \tag{7}$$

Notice that for $\theta = 0$, the last term cancels and the pricing kernel simplifies to EU, i.e. the most common form of Epstein-Zin. Relative to that simpler case, the GDA pricing kernel overweighs left tail outcomes. These outcomes (for which $V_{t+1} \leq \delta\mu_t$) receive a weight that is $(1+\theta)$ times as large as the weight for other outcomes (for which $V_{t+1} > \delta\mu_t$). I return to a more detailed discussion of disappointing outcomes after presenting the dynamics of economic fundamentals.

Cash Flows. The log growth rates of consumption and dividends are given by

$$\begin{aligned} \Delta c_{t+1} &= \mu + \sigma_t \varepsilon^c_{t+1} \\ \Delta d_{t+1} &= \mu + \sigma_t \varphi \varepsilon^d_{t+1} \end{aligned} \tag{8}$$

where μ is the mean growth rate, σ_t is the conditional volatility, ε^c and ε^d are standard normals with correlation ϱ, and φ is a scaling factor that allows dividends to be more volatile than consumption. Note that, in contrast to the long run risks model, mean growth rates are assumed to be constant. Endowment variance follows the Markov Switching Multifractal (MSM) process of Calvet and Fisher (2001, 2004), which allows for a large number of states while remaining tightly parameterized. Specifically, σ_t^2 equals the product of several variance components (and a constant), given by

$$\sigma_t^2 = \overline{\sigma}^2 \prod_{k=1}^{K} \mathcal{M}_{k,t}. \tag{9}$$

For tractability, components are assumed to be mutually independent. Each of the K variance components $\mathcal{M}_{k,t}$ ($\mathcal{M}$ for multiplier) follows a two state Markov chain with state space $\{1-\nu, 1+\nu\}$, which is identical for all components. The parameter $\nu \in (0,1)$ determines the high and low state. Since each component has two states, the Markov chain for σ_t^2 has $N = 2^K$ states. However, due to the fact that all components share the same state space, σ_t^2 can only take on $K+1$ different values.[13] In the benchmark

[13]The possible values of σ_t^2 are given by $\overline{\sigma}^2(1-\nu)^0(1+\nu)^K$, $\overline{\sigma}^2(1-\nu)^1(1+\nu)^{K-1}$, ...,$\overline{\sigma}^2(1-\nu)^K(1+\nu)^0$, i.e. between 0 and K components in the low state and the others in the high state

calibration I set $K = 6$, so that there are 64 states and 7 possible variance values.

The heterogeneity between components lies in their persistence levels. The transition matrix for the k^{th} component is given by

$$P_k = \begin{bmatrix} 1-\gamma_k/2 & \gamma_k/2 \\ \gamma_k/2 & 1-\gamma_k/2 \end{bmatrix}, \tag{10}$$

i.e. it is completely characterized by one parameter. To prevent the number of parameters from growing with the number of variance components (K), the parameters are modeled via the recursion

$$\gamma_k = 1 - (1-\gamma_{k-1})^b, \tag{11}$$

where $\gamma_1 \in (0,1)$ and $b \in (1,\infty)$. Variance components with a higher index have higher γ's and are therefore less persistent. The three parameters (ν, γ_1, b) control the volatility and persistence levels of all components, regardless of how many components there are.[14] Since the transition matrices are symmetric, each component is equally likely to be in the high and low state in the long run, so that $E[\mathcal{M}_{k,t}] = \frac{1}{2}(1-\nu)+\frac{1}{2}(1+\nu) = 1$ for all k. The assumed independence between components therefore implies that $E[\sigma_t^2] = \overline{\sigma}^2 \prod_{k=1}^{K} E[\mathcal{M}_{k,t}] = \overline{\sigma}^2$, so that the parameter $\overline{\sigma}^2$ controls the mean of the process.

Determinants of a disappointment. Note that the disappointment event $V_{t+1} \leq \delta\mu_t$ can be written as

$$\Delta c_{t+1} + \log\left(\frac{\lambda_{t+1}^V}{\lambda_t^V}\right) \leq \log\left(\frac{\delta\mu_t}{V_t}\right), \tag{12}$$

where $\lambda_t^V \equiv V_t/C_t$ equals the utility-consumption ratio. The RHS is known at time t. Disappointments are therefore caused by a combination of shocks to contemporaneous consumption growth (ε_{t+1}^c) and changes in the persistent volatility components ($\mathcal{M}_{k,t}$), which in turn induce changes in λ_t^V.

The state of the volatility process affects the two terms on the LHS in opposite ways. All else equal, when volatility component $\mathcal{M}_{k,t}$ is currently in its high state, the conditional consumption growth volatility σ_t is higher. This *increases* the disappointment

[14] All else equal, increasing ν makes the high and low states more heterogeneous, which increases the volatility of σ_t^2. Increasing γ_1 makes each component less persistent (by increasing γ_k for al k), thereby making σ_t^2 less persistent. Increasing b leads to a faster growth rate between γ_k's, i.e. the persistence level declines faster when moving to a higher k. This makes components more heterogeneous, and it also makes σ_t^2 less persistent.

probability because it increases the likelihood of low Δc_{t+1}-values. On the other hand, $\mathcal{M}_{k,t}$ being in its high state also implies the possibility that it switches to its low state, which in turn results in a positive jump in $\log\left(\frac{\lambda^V_{t+1}}{\lambda^V_t}\right)$ because volatility carries a negative price of risk. This possibility *decreases* the disappointment probability. Clearly, which effect dominates is a quantitative question. I show in Section 4.5 that the former channel prevails in my calibration, so that the conditional disappointment probability is increasing in the level of endowment volatility. Lastly, fluctuations in σ_t may also induce time-variation in the relative disappointment threshold on the RHS. However, I find that this variation is quantitatively negligible for various preference parameter combinations and endowment processes.[15]

3.2 Solution

The model solution is exact, i.e. it does not rely on any log-linear approximations. It is characterized by a set of $N \times 1$ vectors that contain the values of the endogenous objects for each of the model's N states. With the exception of multi-period option prices, which I compute via Monte Carlo simulation, all asset prices are available in closed-form. In the remainder of this section, I define asset prices and outline the solution technique. Derivation details are contained in Appendix B.

The first step of the solution is to solve for the endogenous objects in the pricing kernel. I do so by extending the method of Bonomo, Garcia, Meddahi, and Tedongap (2011) to the case $\alpha = 0$ (log utility), which is used in my benchmark calibration. Denote the utility-consumption and certainty equivalent-consumption ratios by $\lambda^V_t \equiv V_t/C_t$ and $\lambda^\mu_t \equiv \mu_t/C_t$. Like all endogenous objects in the model, each of these ratios can take on N different values. The appendix shows how to express the value function (Equation 4) and the certainty equivalent (Equation 5) in terms of λ^V_t, λ^μ_t, and ε^c_{t+1}. After integrating out ε^c_{t+1}, the remaining system consists of $2N$ (nonlinear) equations in $2N$ unknowns. The solution has to be found numerically. However, this is fast, even for large values of N, and a solution is guaranteed to exist as long as the period utility function $u(\cdot)$ is

[15] In the benchmark calibration of Section 4, the RHS is equal to -0.0369 up to four decimals in all states. Numerically, I find that the relative threshold is also close to being constant for different calibrations of the model in Routledge and Zin (2010) (who model consumption growth as a 2-state Markov chain), as well as a version of the present model for which endowment variance is described by an AR(1) in logs.

continuous.

Having obtained the pricing kernel, one can solve for asset prices. While the associated algebra is somewhat involved, the basic idea is straightforward: After integrating over the normal innovations in the Euler equation, all remaining randomness is associated with the Markovian state and expectations can therefore be evaluated as simple matrix products. The main challenge in integrating over $(\varepsilon_{t+1}^c, \varepsilon_{t+1}^d)$ stems from the fact that they occur within the indicators in the disappointment term in the pricing kernel and the max operator in the option payoff. Derivations therefore rely heavily on change of variables and results for truncated normal random variables. The following Theorem defines asset prices and shows the corresponding closed-form solutions.

Theorem 1 (Closed-Form Asset Prices) *Denote the price of the dividend claim (equity) at time t by S_t and the associated log ex-dividend return by $r_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right)$. Let P denote the transition matrix of the Markov chain for σ_t, let $\odot$ denote the Hadamard (element-wise) matrix product, and let ι_N and $\mathbf{1}_N$ denote a $N \times 1$ vector of ones and a $N \times N$ matrix of ones respectively. Denote the multi-period pricing kernel by $M_{t:t+\tau} \equiv \prod_{h=1}^{\tau} M_{t+h}$. The $N \times N$ matrices A^b, A^d, $A^c(K)$, A^ν and A^p, which occur below, can be computed in closed-form and are defined in Appendix B.*

A. Risk-free Bonds. *The price of a risk-free τ-period zero coupon bond, given by $\mathcal{B}_t(\tau) \equiv E_t[M_{t:t+\tau}]$, can be computed recursively in vector form as*

$$\mathcal{B}(\tau) = (P \odot A^b) \cdot \mathcal{B}(\tau - 1),$$

where $\mathcal{B}(1) = \iota_N$.

B. Equity. *The price-dividend ratio, given by $\mathcal{D}_t \equiv \frac{S_t}{D_t}$, can be computed in vector form as*

$$\mathcal{D} = (I_N - P \odot A^d)^{-1} \cdot (P \odot A^d) \cdot \iota_N.$$

C. Call Options. *The (relative) price of a 1-period call option with strike price X and moneyness $K \equiv \frac{X}{S_t}$, given by*

$$\mathcal{C}_t(1, K) \equiv \frac{1}{S_t} \times E_t\left[M_{t+1} \max\left(0, S_{t+1} - X\right)\right] = E_t\left[M_{t+1} \max\left(0, e^{r_{t+1}} - K\right)\right],$$

can be computed in vector form as

$$\mathcal{C}(1,K) = (P \odot A^c(K)) \cdot \iota_N.$$

D. Variance Swaps. *The τ-period variance swap rate, given by*

$$\mathcal{V}_t(\tau) \equiv E_t^Q \left[\sum_{h=1}^{\tau} r_{t+h}^2 \right] = E_t \left[M_{t+\tau} \sum_{h=1}^{\tau} r_{t+h}^2 \right] \mathcal{B}_t(\tau)^{-1},$$

can be computed in vector form as

$$\mathcal{V}(\tau) = \begin{cases} (P \odot A^\nu) \cdot \iota_N & , \tau = 1 \\ \mathcal{B}(\tau)^{-1} \odot \sum_{h=1}^{\tau} \mathcal{V}(h,\tau) & , \tau > 1 \end{cases}$$

where

- *with a slight abuse of notation, $\mathcal{B}(\tau)^{-1}$ denotes the element-wise inverse of the vector containing τ-period bond prices.*
- *$\mathcal{V}(1,1) \equiv \mathcal{V}(1)$ equals the 1-period swap rate.*
- *$\mathcal{V}(1,\tau) = (P \odot A^\nu) \cdot \mathcal{B}(\tau - 1)$ for $\tau > 1$.*
- *$\mathcal{V}(h,\tau) = (P \odot A^b) \cdot \mathcal{V}(h-1,\tau-1)$ for $h > 1$ and $\tau > 1$.*

E. Variance Premium. *The (1-month) variance premium, defined as $VP_t = E_t^Q \left[\sum_{h=1}^{\tau} r_{t+h}^2\right] - E_t^P \left[\sum_{h=1}^{\tau} r_{t+h}^2\right]$, can be evaluated in matrix form as*

$$VP = \mathcal{V}(1) - (P \odot A^p) \cdot \iota_N$$

.

Proof: See Appendix B.

Note that both options and variance swaps are written on the *ex dividend* market return, as in the data. The price of a τ-period call is given by

$$\mathcal{C}_t(\tau,K) \equiv E_t \left[\left(\prod_{h=1}^{\tau} M_{t+h} \right) \max \left(0, \exp \left(\sum_{h=1}^{\tau} r_{t+h} \right) - K \right) \right].$$

Because the call payoff cannot be factored into τ single-period terms (as it was the case for the variance swap payoff), the expectation cannot be evaluated recursively. More importantly, the expectation for a τ-period option involves N^τ possible paths for the

Markov chain, so that an analytical solution becomes quickly intractable as τ grows. I therefore compute multi-period option prices via Monte Carlo simulation. Specifically, starting from each of the N states, I simulate 100 million paths of both the Markov chain and $(\varepsilon^c, \varepsilon^d)$, use them to compute the pricing kernel and the option payoff, and evaluate the expectation in the associated Euler equation as the average across paths.[16] To convert model-based option prices into Black-Scholes implied volatilities, I use the model-implied interest rate and dividend-yield.

3.3 Small Sample Statistics

In the remainder of the paper, I produce small sample statistics for different model calibrations and compare them to the data. Specifically, I simulate $100,000$ samples of the same length as the data, compute the statistic of interest in each sample, and report the median value. For statistics that appear in tables (rather than figures), a model-based 90% confidence interval is reported in addition. The longest available dataset for cash flow and standard asset pricing moments is annual and spans 83 years (= 996 months). The option data spans 23 years (= 276 months). All model-based calibration results are based on small samples whose lengths match these empirical counterparts.

In constructing annualized moments, I closely follow Beeler and Campbell (2012) and Bansal, Kiku, and Yaron (2012). Consumption and dividend growth rates are computed by adding twelve monthly consumption and dividend levels, and then taking the growth rate of the sum. Annual log stock returns are the sum of monthly values, while log price-dividend ratios use prices measured from the last month of the year. Because the price-dividend ratio in the data divides by the previous years dividends, I multiply the price-dividend ratio in the model by the dividend in that month and divide by the dividends over the previous year. The annual risk-free rate is the sum of the four quarterly risk-free rates within a year.

[16] I repeated this procedure for different seeds of the random number generator to ensure that the Monte Carlo error is negligible given the number of paths.

4 Results

4.1 Calibration

I calibrate the model at the monthly frequency. All calibration targets equal small sample medians. The parameters μ and $\overline{\sigma}^2$ are chosen to match the mean and volatility of annual consumption growth, φ is set to match the volatility of annual dividend growth, and ϱ is set to match the correlation between the two endowments. The parameters governing the volatility and persistence of the endowment variance (ν, γ_K, b) naturally have a large effect on the dynamics of the conditional return variance, which are in turn reflected in the dynamics of variance swap rates. I calibrate them to match as good as possible the volatilities and autocorrelation functions of variance swaps with different maturities.[17] I use $K = 6$ variance components in the MSM process. The calibration implies a first-order autocorrelation of 0.82 for σ_t, with persistence levels that range from 0.5 to 0.994 for the individual components. For comparison, the autoregressive variance process in Bansal and Yaron (2004) has a much higher first-order autocorrelation of 0.987. As a result, the basic long run risks model vastly overstates the persistence of market variance.

Preference parameters are calibrated as follows. Given a time-discount factor of $\beta = 0.96^{1/12}$, I set ρ to match the mean risk-free rate. The implied elasticity of intertemporal substitution equals $\rho^{-1} = 0.49$. The risk preference parameters θ and δ are chosen jointly to match the equity premium and the variance premium – the implied volatility smirk does not serve as a calibration target! I refer to the full model as the GDA-MSM model.

In order to illustrate the role of the two main model components, I also show results for two alternative economies. First, The EU-MSM model consists of the MSM endowment (calibrated as in the benchmark model) and recursive utility with expected utility risk preferences. This nested preference specification obtains for $\theta = 0$. I set $\beta = 0.96^{1/12}$ and choose ρ to match the mean risk-free rate. The curvature parameter α is chosen

[17] I find that preference parameters have a negligible effect on the autocorrelations of variance swap rates (see Section 4.6). This theoretical finding agrees with the empirical fact that both the physical return variance and variance swap rates (which represent a form of risk-neutral variance) display similar long-memory behavior.

Table 2: Calibrations

GDA-MSM	β	ρ	θ	δ	α		
	$0.96^{\frac{1}{12}}$	0.49^{-1}	43.2	0.9625	0		
	$\overline{\sigma}^2$	ν	γ_K	b	μ	ϱ	φ
	0.008^2	0.33	0.5	2.6	0.015	0.53	5.2
EU-MSM	β	ρ	θ	δ	α		
	$0.96^{\frac{1}{12}}$	0.353^{-1}	0	–	-18.38		
	$\overline{\sigma}^2$	ν	γ_K	b	μ	ϱ	φ
	0.008^2	0.33	0.5	2.6	0.015	0.53	5.2
GDA-AR1	β	ρ	θ	δ	α		
	$0.96^{\frac{1}{12}}$	0.687^{-1}	13.44	0.927	0		
	$E[\sigma_t^2]$	$std[\sigma_t^2]$	$AC1[\sigma_t^2]$		μ	ϱ	φ
	6.30×10^{-3}	5.78×10^{-3}	0.98		0.015	0.53	5.2

Table 2 reports the configuration of investors preferences and the time-series parameters that describe the endowment process. The model is calibrated at a monthly decision interval.

to match the average equity premium. Second, the GDA-AR1 model combines GDA preferences with an AR(1) process for log endowment variance.[18] To fit the process into the Markov switching environment, I discretize it with the method of Rouwenhorst (1995). I use 51 states and calibrate the process to a first-order autocorrelation of 0.98 to mimic the high persistence typically used in the long run risks literature. The mean of σ_t^2 is chosen to match the mean volatility of annual consumption growth and the volatility of σ_t^2 is set to the same value as in the other two economies. All three calibrations are summarized in Table 2.

[18]Note that the long run risks literature typically models cash flow volatility as an AR(1) in levels rather than in logs. The log specification has the advantage that volatility cannot become negative. Additionally, it allows for a cleaner comparison with the multifractal process, whose unconditional distribution is right-skewed and resembles that of a log normal distribution.

Table 3: Quantities and Prices

		Data	GDA-MSM			EU-MSM			GDA-AR1		
			5%	50%	95%	5%	50%	95%	5%	50%	95%
Cash Flows	$E[\Delta c]$	1.82	1.30	1.80	2.30	1.30	1.80	2.30	1.30	1.80	2.30
	$\sigma[\Delta c]$	2.18	1.69	2.21	2.84	1.69	2.21	2.84	1.62	2.17	2.92
	$AC1[\Delta c]$	0.51	0.04	0.23	0.41	0.04	0.23	0.41	0.03	0.23	0.42
	$E[\Delta d]$	1.26	-0.81	1.80	4.40	-0.81	1.80	4.40	-0.78	1.80	4.39
	$\sigma[\Delta d]$	11.32	8.81	11.50	14.78	8.81	11.50	14.78	8.43	11.28	15.21
	$AC1[\Delta d]$	0.20	0.04	0.23	0.41	0.04	0.23	0.41	0.03	0.23	0.41
	$corr[\Delta c, \Delta d]$	0.53	0.35	0.53	0.68	0.35	0.53	0.68	0.34	0.53	0.68
Prices and Returns	$E[R - r^f]$	5.24	2.53	5.26	8.38	2.61	5.23	7.90	1.63	5.30	11.56
	$\sigma[R]$	19.85	13.17	17.03	21.32	11.97	15.55	19.51	13.77	20.33	27.84
	$AC1[R]$	-0.01	-0.25	-0.06	0.14	-0.22	-0.03	0.17	-0.28	-0.06	0.16
	$kurt[R]$	3.54	2.74	3.77	6.33	2.71	3.73	6.37	3.12	4.65	8.61
	$kurt[R]$(monthly)	9.45	5.90	8.37	13.29	4.93	6.98	11.09	5.40	7.41	11.62
	$E[r^f]$	0.49	-0.39	0.52	1.69	0.23	0.49	0.80	-4.32	0.51	2.72
	$\sigma[r^f]$	2.87	1.23	1.88	3.47	0.14	0.36	0.51	1.49	5.61	11.79
	$AC1[r^f]$	0.70	0.36	0.55	0.74	0.49	0.75	0.85	0.44	0.74	0.89
	$E[p - d]$	3.39	3.09	3.20	3.32	3.11	3.22	3.33	3.08	3.24	3.35
	$\sigma[p - d]$	0.45	0.10	0.14	0.18	0.06	0.12	0.15	0.11	0.21	0.34
	$AC1[p - d]$	0.88	0.73	0.86	0.93	0.78	0.90	0.95	0.70	0.83	0.92
Variance Premium	$E[VP]$	11.29	4.45	11.30	20.67	-1.49	1.79	4.62	0.14	10.98	45.55
	$\sigma[VP]$	9.65	7.04	15.27	21.45	0.65	1.61	3.36	0.63	22.42	37.75
	$AC1[VP]$	0.46	0.60	0.79	0.91	0.90	0.96	0.99	0.69	0.93	0.98

Table 3 presents moments for annual cash flows and asset prices. Also shown is the kurtosis of monthly returns and the (monthly) variance premium. R denotes the log cum dividend return on the market, r^f the log risk-free rate, $p - d$ the log price-dividend ratio, and VP the variance premium. Model equivalents are computed from 100,000 samples of the same length as the data as described in Section 3.3. Details on the data are contained in Appendix A. The sample for moments in the first two panels spans 1930-2012, whereas the sample for the variance premium spans 1990-2012.

4.2 Cash Flows

The top panel of Table 3 shows moments of annual cash flows. Except for the mean of dividend growth, the models are calibrated to match the first two endowment moments exactly.[19] I follow the convention in the previous literature of assuming that consumption and dividend growth have the same mean, which implies balanced growth. The fact that cash flows are modeled without time-variation in conditional means implies that their first-order autocorrelations are close to the ones of a time-aggregated continuous-time random walk, which equals 0.25 (see Working (1960)). The Table shows this is a good approximation for dividend growth rates, whose empirical autocorrelation of 0.2 is close to the median model estimate of 0.23 in all three economies. Only the autocorrelation of consumption growth falls slightly out of the model-implied 90% confidence interval.

4.3 Basic Asset Prices

The middle panel of Table 3 presents the model implications for annual asset pricing moments. All three models are calibrated to match the equity premium and the risk-free rate exactly. For both GDA models, the volatility of returns falls inside the model-implied 90% confidence interval, whereas the volatility is too low in the EU model. The reason is that the time-varying risk aversion generated by GDA results in more variability in the equity premium and therefore more volatile returns than in the EU model. The effect of GDA shows up even stronger in the volatility of the risk-free rate. With EU preferences, there is little time-variation in the conditional mean of the pricing kernel and r_t^f displays much less volatility than in the data. The upper end of the EU-MSM confidence interval equals 0.51, which is less than a fifth of the data value. In contrast, the risk free rate volatility of 1.88% in the benchmark model is close to its data counterpart, which comfortably falls in the model-based confidence interval. In the GDA-AR1 model, the risk-free rate is twice as volatile as in the data, which shows that the slow-moving nature of the AR(1) process results in too much time-variation in the conditional mean of M_t when combined with GDA preferences. Lastly, all three models are successful at replicating the high mean and persistence of the log price-dividend ratio, but fall somewhat short of matching its volatility.

Because the main goal of my paper is to explain features of index options, it appears important to ensure that the model's implications for the quantity of tail risk are not

[19] Negligible differences between data and model moments remain due to the time-intensive nature of computing small sample statistics.

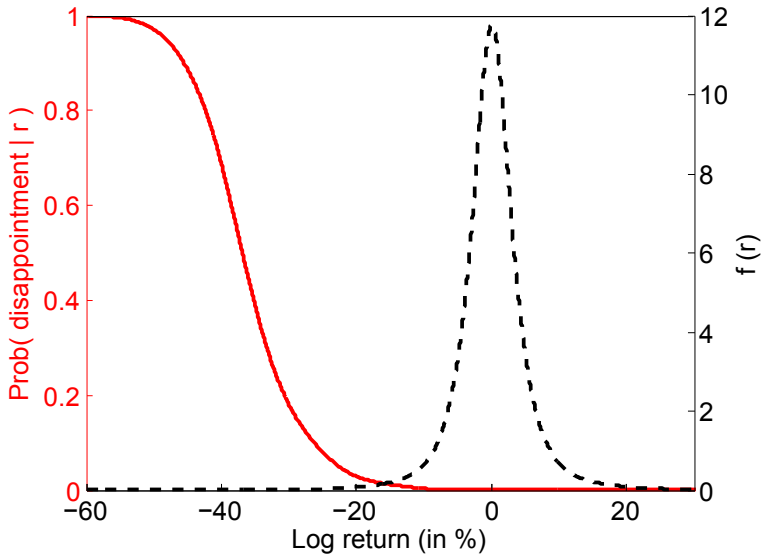

Figure 4: Returns and Disappointments

Figure 4 shows the unconditional density of monthly ex-dividend log returns (dashed line, right axis), as well the the probability of disappointment conditional on a given return (solid line, left axis).

counterfactual. The Table shows that the kurtosis of both annual and monthly returns in the GDA-MSM model are very close to their data counterparts. At the 1 month horizon, the kurtosis equals 8.37 in the model and 9.45 in the data. At the 12 month horizon, the value falls to 3.77 in the model and 3.54 in the data. The model therefore successfully replicates the non-normality of short-horizon returns and the fact that this feature considerably weakens at longer horizons. To my knowledge, the model is the first to successfully match this well-known empirical fact. A comparison with the EU-MSM model shows that part of the kurtosis in monthly returns stems from the large weight that GDA places on tail outcomes. In particular, GDA generates larger price jumps in response to changes in endowment volatility, i.e. more volatility feedback than EU. I discuss this channel in more detail in Section 4.5, where I explicitly compute tail probabilities in all three models.

4.4 Variance Premium

The bottom panel of Table 3 shows moments of the 1-month variance premium. Both GDA models are calibrated to match the average variance premium exactly. The same is not possible for the EU model,[20] whose sole risk preference parameter (α) was chosen

[20]Drechsler and Yaron (2011) present a model with EU risk preferences that *does* match both the equity premium and the variance premium, despite the fact that they use an even lower curvature parameter than in my EU calibration. This is possible because their calibration implies considerably larger jumps than those implied by the Markov switching process in the present paper. Additionally,

to match the equity premium. To illustrate the source of the variance premium in the GDA models, it is helpful to consider the relationship between returns and the probability of disappointments. Equation 12 showed that disappointments are caused by a combination of negative innovations in consumption growth and positive innovations in endowment volatility. Because consumption and dividend growth share the same volatility process as well as (imperfectly) correlated innovations, disappointments tend to coincide with negative returns. This is illustrated in Figure 4, which shows the unconditional distribution of monthly returns (dashed line, right axis) along with the probability of a disappointment conditional on a given return (solid line, left axis), both for the GDA-MSM model. It is clear that the lower a given return, the higher the chance that it is associated with a disappointment. Because strongly negative returns also result in a large realized variance (large payoffs to variance swaps), variance swaps embed a large insurance premium.

4.5 Option Prices

Figure 5 shows the empirical and model-based implied volatility (IV) curve for 1-month (left column) and 12-month (right column) maturities. The full model (top rows) provides a very good match for the empirical IV curve. At the 1-month maturity, options with a relative moneyness of -2 have an IV of about 26.0%, both in the model and in the data. At-the-money options have an IV of 17% in the model, close to the data value of 18%. At the 12-month maturity, the GDA-MSM model provides a similarly good match for the data, replicating the higher level of the curve relative to the shorter maturity. It is worth emphasizing that the IV curve did not serve as a calibration target for the model, i.e. risk preferences were calibrated to match the equity and variance premia only.

The bottom row of Figure 5 shows the IV curve for the two alternative economies. The EU-MSM model (dotted lines) implies a slight smirk, but its level and slope fall short of the data counterparts for both maturities. Nevertheless, it is noteworthy that the EU model can produce a smirk at all because Benzoni, Collin-Dufresne, and Goldstein (2011) have shown that the IV curve is almost exactly horizontal in the basic long run risks (LRR) model. The LRR model is based on the same utility function but an autoregressive process for endowment variance. Average option prices therefore seem to support the presence of multifractal variance risks in cash flows. The GDA-AR1 model (dash-dotted lines) provides a good match for the level of the IV curve at both maturi-

their economy features jumps not only in the conditional variance of endowment growth rates, but additionally in their conditional means.

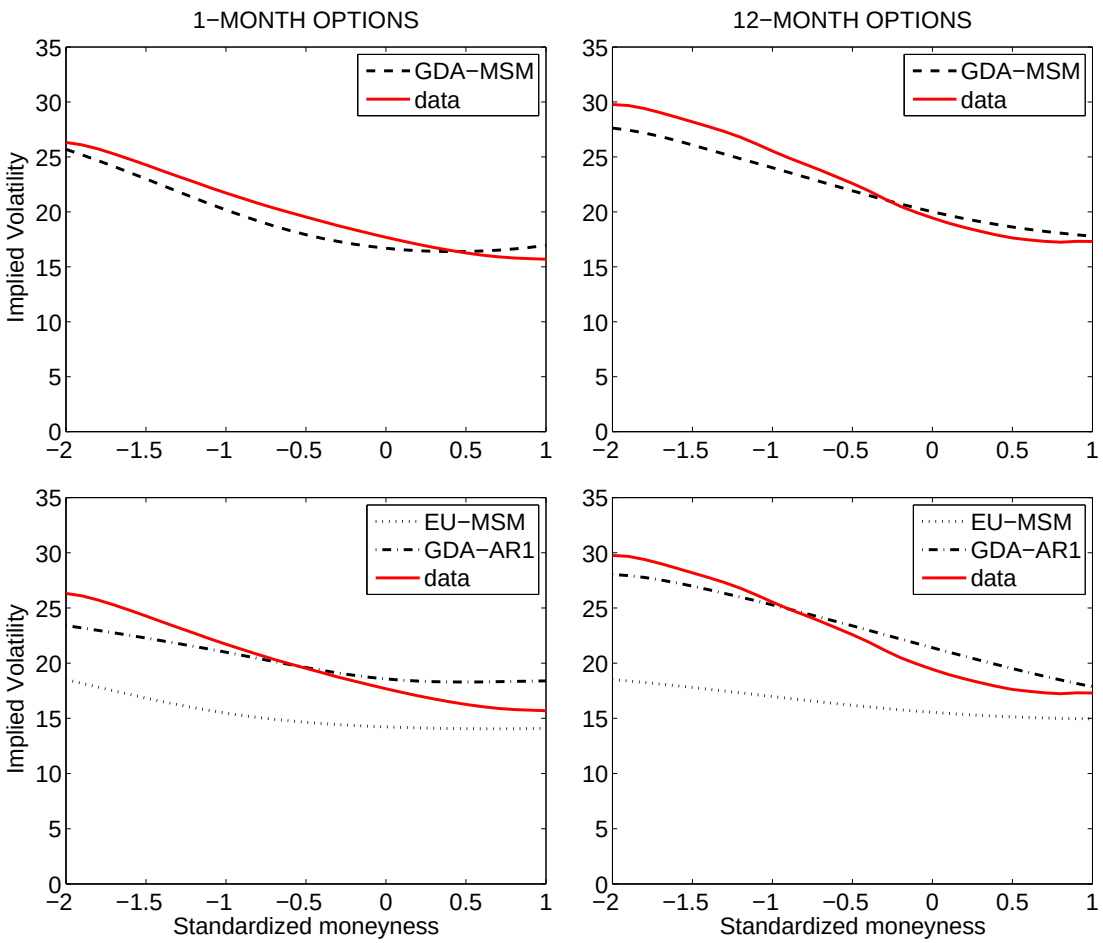

Figure 5: Implied Volatility Smirk

Figure 5 shows average (Black-Scholes) implied volatilities for option maturities of 1 and 12 months, and relative moneyness of -2 to 1. Volatilities are expressed in annualized percentage units. The top row shows results for the benchmark model (GDA-MSM), whereas the bottom panel shows results for the EU-MSM and GDA-AR1 models. Small-sample model statistics are computed from 100,000 samples whose length equals that of the data – see Section 3.3. The sample spans 1990-2012.

ties, which was expected from the fact that the model matched the return volatility and was calibrated to match the average variance premium. The model is further capable of matching the steep slope of the IV curve at the annual maturity, but it produces a curve that is too flat at the 1-month maturity. Additionally, I illustrate below that the GDA-AR1 model has very counterfactual implications for the *conditional* smirk.

To understand why the full model is able to replicate the smirk, I plot conditional IV curves. Figure 6 shows 1-month implied volatilities, conditional on different percentiles of the endowment volatility σ_t.[21] In the GDA-MSM model (top-left panel), the curve

[21] In the MSM models, σ_t only takes on $K+1=7$ different values (with different probabilities). The

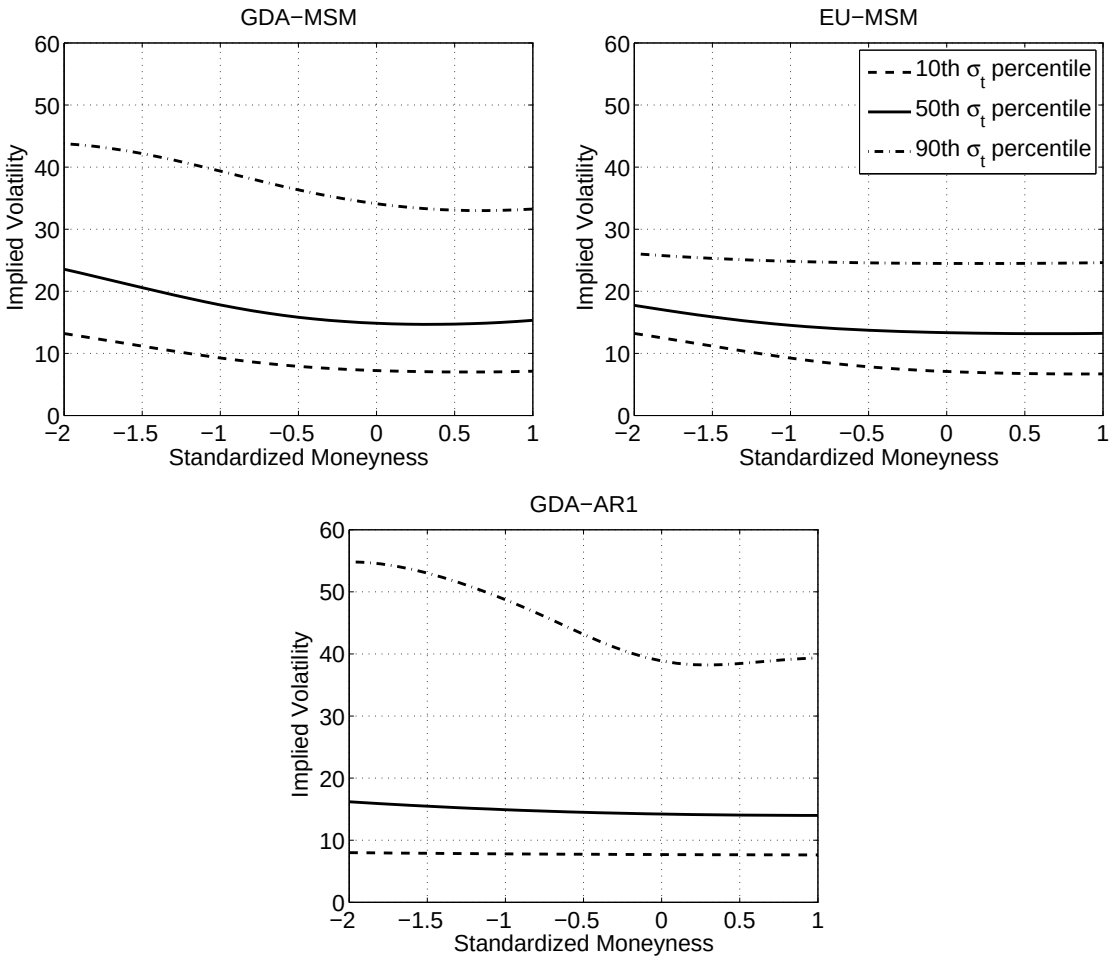

Figure 6: Conditional 1-Month Implied Volatility Smirk

Figure 6 shows the implied volatility smirk conditional on values of the endowment variance σ_t. Volatilities are expressed in annualized percentage units. Each plot shows the IV curve at the 10^{th}, 50^{th}, and 90^{th} percentile of σ_t for a different model.

has a noticeably negative slope at the 10^{th}, 50^{th}, and 90^{th} percentiles of σ_t. The steep unconditional IV curve therefore results from the fact that the curve is relatively steep in most states. This is an attractive feature of the model because the slope of the curve is nearly always significantly positive in the data (see Figure 1). In contrast to the benchmark model, the slope of the IV curve becomes flat for certain regions of σ_t in both alternative economies. In the EU-MSM model (top-right panel), the slope is negative when endowment volatility is low but close to zero when it is high. In the GDA-AR1 model (bottom panel), the conditional slopes behave in the exact opposite way, being strongly negative when endowment volatility is high but close to zero when

10^{th} percentile is given by the states in which exactly one component is in its high state. The IV curve at the 10^{th} percentile of σ_t equals the mean curve across these states. The curve is computed similarly at the other two percentiles.

it is low. These differences can be explained by time-variation in the conditional disappointment probability as well as time-variation in the distribution of (endogenous) return jumps, which in turn result from changes in σ_t.

In what follows, I show that (1) in the two models with GDA preferences, the smirk is steep for high values of σ_t due to a high conditional disappointment probability and (2) in the two models with MSM volatility, the smirk is steep when volatility is low due to a high probability of large negative returns. To quantify the effect of the two channels, Table 4 shows conditional return moments and conditional tail probabilities under both the statistical measure ($\mathbb{P}$) and the risk-neutral measure ($\mathbb{Q}$).[22] As in Figure 6, all numbers are conditional on percentiles of σ_t (I omit the 50^{th} percentile to save space). For comparability with the IV plots, I express tail thresholds in standard deviation units.

States of low endowment volatility. Notice from Table 4 that when volatility is at its 10^{th} percentile, the return distribution in the two MSM models is strongly left-skewed and leptokurtic under $\mathbb{P}$. In the benchmark model, the conditional skewness (kurtosis) equals -2.57 (23.51) whereas it equals -2.05 (23.39) in the EU-MSM model. The higher moments are also reflected in a considerably larger probability of left tail outcomes relative to right tail outcomes in both models. For example, in the GDA-MSM (EU-MSM) model, there is a chance of 1.78% (1.22%) of observing a return of less than *minus* two standard deviations but only a chance of 0.34% (0.28%) of observing a return of more than *plus* two standard deviations. In contrast, the conditional return distribution in the GDA-AR1 model is close to normal with a skewness of -0.08 and a kurtosis of 3.13. The left-skewness in the MSM models results from mean reversion in endowment volatility: When σ_t increases as expected, it induces a decrease in the price dividend ratio and therefore a negative return. In particular, the less persistent volatility components in the MSM models mean-revert fast enough to induce large negative return jumps and a conditional return distribution that is strongly left-skewed at the 1-month horizon. In contrast, mean reversion occurs too slowly in the GDA-AR1 model to have a significant effect on returns over short horizons.[23]

[22] The risk neutral density can be computed from call prices using the result of Breeden and Litzenberger (1978), who show that it equals the second derivative of the call price with respect to the strike price, divided by the price of a risk-free bond of equal maturity. I compute call prices on a fine grid of strikes and compute the second derivative via finite differences.

[23] At the *annual* return horizon, mean reversion has a stronger effect in the GDA-AR1 model, resulting in conditional return skewness (kurtosis) of -0.44 (3.95). The higher skewness contributes to the model's ability to produce a reasonable IV curve for options with a 12-month maturity (see Figure 5). For comparison, the skewness (kurtosis) of annual returns equals -0.79 (5.02) in the GDA-MSM model and

Table 4: Conditional Moments and Tail Probabilities

Model	σ_t-pct	Dist.	Moments			Tail Probabilities			
			std	skew	kurt	< -3	< -2	> 2	> 3
GDA-MSM	10	$\mathbb{P}$	2.63	−2.57	23.51	1.29	1.78	0.34	0.17
		$\mathbb{Q}$	2.91	−3.57	34.25	1.61	2.13	0.30	0.15
	90	$\mathbb{P}$	8.31	0.16	4.05	0.04	0.60	1.48	0.20
		$\mathbb{Q}$	11.12	−0.51	3.68	1.35	6.94	0.96	0.11
EU-MSM	10	$\mathbb{P}$	2.33	−2.05	23.39	0.75	1.22	0.28	0.11
		$\mathbb{Q}$	2.93	−3.43	25.40	1.79	2.44	0.18	0.05
	90	$\mathbb{P}$	7.39	0.12	3.64	0.14	1.68	2.90	0.44
		$\mathbb{Q}$	7.43	−0.11	3.64	0.49	3.18	1.52	0.14
GDA-AR1	10	$\mathbb{P}$	2.23	−0.08	3.13	0.22	2.56	1.92	0.13
		$\mathbb{Q}$	2.23	−0.08	3.14	0.24	2.74	1.82	0.12
	90	$\mathbb{P}$	9.93	−0.05	3.90	0.06	0.80	0.89	0.02
		$\mathbb{Q}$	13.57	−0.74	3.70	1.34	9.43	0.58	0.01

Table 4 shows conditional moments and tail probabilities of 1-month ex-dividend log returns for the physical ($\mathbb{P}$) and risk-neutral ($\mathbb{Q}$) distributions. All quantities condition on a given percentile of the endowment volatility σ_t. Returns are expressed in monthly percent, i.e. a standard deviation of 2.63 represents 2.63% per month. Return thresholds are measured in standard deviation units, where the square root of the 1-month variance swap rate is used to measure the conditional standard deviation. Probabilities are expressed in percent, i.e. 1.29 stands for 1.29%.

For the risk neutral distribution ($\mathbb{Q}$), the moments and tail probabilities at the 10^{th} percentile of σ_t are similar to those under $\mathbb{P}$ in both GDA models. As one may guess from this result, disappointment aversions plays a minor role in states with low endowment volatility. The conditional disappointment probability only equals 0.0001% in the GDA-MSM model and it essentially equals zero in the GDA-AR1 model. The fact that the two MSM models produce a large smirk in the low volatility states is therefore primarily a consequence of the MSM process rather than the preference specification. Put differently, OTM puts are expensive because there is relatively more left tail risk than right tail risk when volatility is low. Insurance premia play a minor role in these states.

States of high endowment volatility. Mean-reversion in volatility has the opposite effect when σ_t is at its 90^{th} percentile, resulting in positive rather than negative return

-0.58 (4.64) in the EU-MSM model, both conditional on σ_t being at its 10^{th} percentile.

jumps on average. In both MSM models, the conditional return distribution is therefore slightly right-skewed under $\mathbb{P}$. All else equal, right-skewness makes the smirk flat or even increasing. However, in the two models with GDA preferences, high volatility states are also associated with an increased disappointment probability. This probability equals 0.33% in the GDA-MSM model and 0.51% in the GDA-AR1 model.[24] As a consequence, the conditional return distribution is significantly left-skewed under the risk-neutral measure in both GDA models. Table 4 shows that the amplification of left tail probabilities under $\mathbb{Q}$ is especially pronounced for the most extreme tails. For example, returns of -3 standard deviations or less are $\frac{1.35}{0.04} \approx 34$ times as likely under $\mathbb{Q}$ than under $\mathbb{P}$ in the GDA-MSM model and $\frac{1.34}{0.06} \approx 22$ times as likely in the GDA-AR1 model. In contrast, the ratio only equals $\frac{0.49}{0.14} \approx 4$ in the EU-MSM model because an EU investor is much less focused on tail events. The fact that the two GDA models produce a large smirk in the high volatility states is therefore a consequence of the utility function rather than the process for endowment variance. In other words, OTM puts are expensive because they embed a large insurance premium when volatility is high.

Lastly, it is interesting to note that the results in Table 4 are broadly consistent with the empirical findings of Bollerslev and Todorov (2011). These authors look at nonparametric estimates of tail probabilities under both measures and find that large negative returns have a much higher probability under the risk neutral measure. GDA preferences represent a possible explanation for this finding.

4.6 The Term Structure of Variance Swap Rates

Figure 7 presents moments of variance swap rates for maturities from 1 to 12 months. Panel A shows the mean of swap rates as a function of the swap maturity. Longer-term swaps have a slightly higher mean compared to short-term swaps, indicating the presence of a term premium in swap rates.[25] All three models can replicate this increasing pattern in mean swap rates, but only the GDA models matches their level. The level is too low in the EU-MSM model because both the variance premium (see Section 4.4) and the return volatility (see Table 3) fall below their respective data counterparts, whereas

[24] A comparison with the disappointment probabilities at the low σ_t percentile shows that the conditional disappointment probability is increasing in the endowment variance. Recall from the discussion following Equation 12 on page 17 that this conclusion was not clear ex ante because endowment volatility has two opposing effects on the conditional disappointment probability.

[25] The (annualized) variance of realized returns is very close to constant across return horizons, both in the data and in the two models. The term premium in swap rates therefore arises due to higher variance premia at longer return horizons.

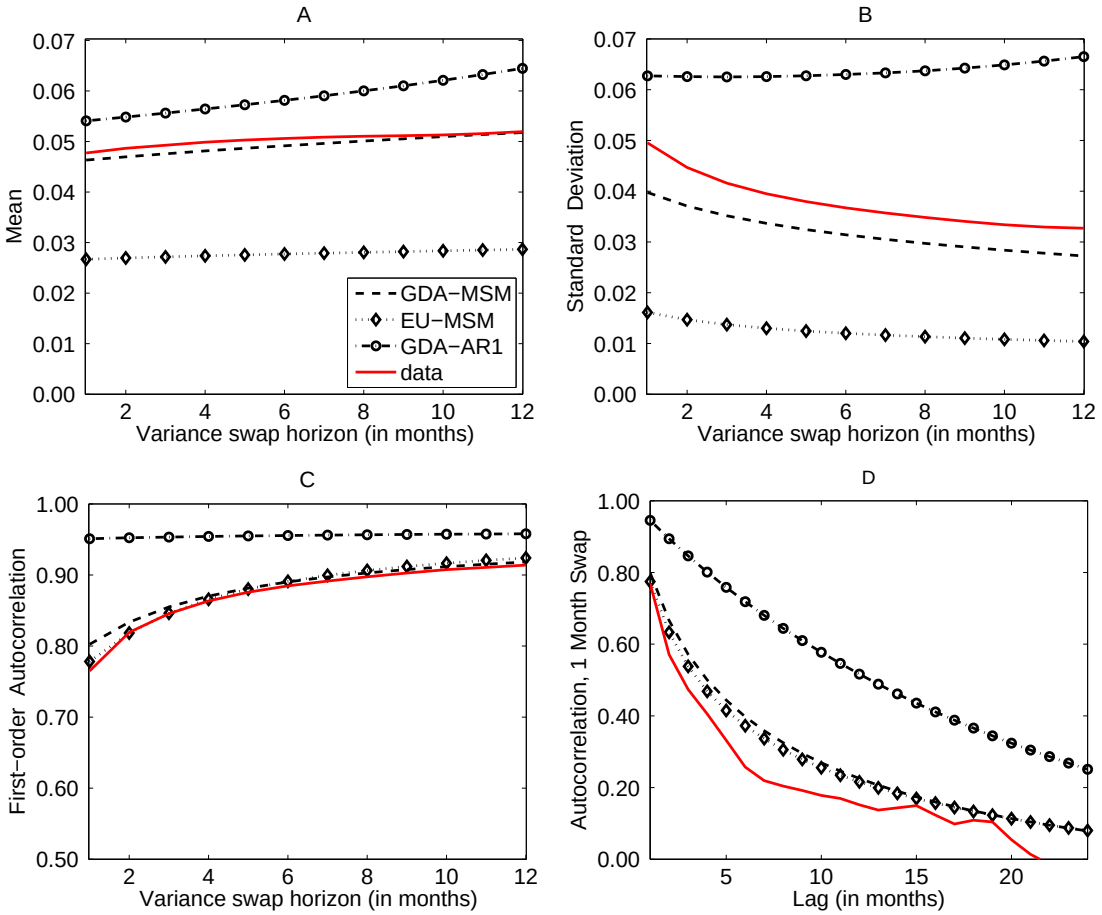

Figure 7: Moments of variance swap rates

Figure 7 shows moments of variance swap rates with maturities between 1 to 12 months. Small-sample model statistics are computed as discussed in Section 3.3. The sample spans 1990-2012.

these moments are matched well in the other two models.

The remaining panels of Figure 7 illustrate the volatility and persistence of swaps with different maturities. Panels B and C show the volatility and first-order autocorrelation of swap rates as a function of the swap maturity, whereas Panel D shows the autocorrelation *function* of 1-month swap rates. These moments served as a calibration target for the parameters controlling the volatility and persistence of the MSM endowment variance, i.e. (ν, γ_K, b). Both MSM models are able to replicate the higher volatility and lower persistence of short term swaps, as well as the long-memory behavior reflected in the autocorrelation function. The GDA-MSM model is additionally able to replicate the high level of the volatility because its state-dependent risk aversion induces time-variation in the variance premium, which makes swap rates more volatile. The same is not true for the EU model. The next section shows that this feature of the GDA model

also endows the variance premium with large predictive power for excess returns.

The success of the MSM models in replicating moments of the term structure of swap rates is due to the presence of both high-frequency and low-frequency shocks in the variance process. The less persistent shocks increase the volatility of variance over short horizons but they average out over longer horizons, making the long end of the term structure less volatile and more persistent. Calvet and Fisher (2004) have previously shown that the MSM process exhibits an autocorrelation structure that mimics the one of a long memory process. While this is a feature of the endowment variance, it is inherited by the endogenously determined variance swap rates in the present model. In contrast, the GDA-AR1 model relies on a single shock frequency and therefore it cannot match the dynamics of variance swap rates.

4.7 Return Predictability

The data section highlighted that the variance premium has high predictive power for excess returns over horizons of a few month. The top panel of Table 5 shows the results of running the same regressions on simulated model data. I report the slope coefficients and R^2s for predictive regressions of log excess returns on the (one-month) variance premium for return horizons of one, three and six months. As in the data, the GDA-MSM model-based slope coefficients are falling in the return horizon and the R^2s are quite large for the short horizons. Further, all median model estimates are close to their data counterparts. In the EU-MSM model, the median slope estimates and R^2s fall considerably below the data values because the model does not generate time-variation in effective risk aversion, resulting in less co-movement among risk premia. On the other hand, the GDA-AR1 model implies considerably too much predictability, producing R^2s that exceed their data counterparts by a factor of 3 to 5. The large degree of predictability results from too much time-variation in the conditional disappointment probability, which also showed up in the very high volatility of the risk-free rate (see Table 3) and the an IV smirk that is extremely steep in high volatility states but essentially flat in low volatility states (see Figure 6). Thus, the model relies on both GDA preferences and the low persistence shocks of the multifractal process to match the predictive power of the variance premium.

In addition to the short-horizon predictive power of the variance premium, it is well-known that the price-dividend ratio has high predictive power over horizons that span several years. In the bottom panel of Table 5, I report the results of regressing excess returns of horizons one, three and five years on the log price-dividend ratio. As in the

Table 5: Return Predictability

	Data	GDA-MSM			EU-MSM			GDA-AR1		
		5%	50%	95%	5%	50%	95%	5%	50%	95%
Regressor: VP										
$\hat{\beta}(1m)$	0.89	0.16	0.72	1.48	-4.13	0.55	5.70	0.54	1.03	7.27
$R^2(1m)$	2.53	0.18	3.04	9.03	0.00	0.40	3.76	0.66	7.72	16.18
$\hat{\beta}(3m)$	0.88	0.13	0.60	1.19	-4.39	0.36	4.93	0.52	0.98	5.73
$R^2(3m)$	6.49	0.03	6.40	19.16	0.01	1.13	9.58	1.46	19.63	36.97
$\hat{\beta}(6m)$	0.65	0.09	0.49	0.96	-4.45	0.24	4.40	0.46	0.91	4.43
$R^2(6m)$	6.52	0.04	8.67	27.40	0.02	2.13	16.22	1.93	31.65	54.74
Regressor: p-d										
$\hat{\beta}(1y)$	-0.08	-0.59	-0.31	1.48	-0.48	-0.16	0.10	-0.69	-0.48	-0.26
$R^2(1y)$	3.67	0.58	6.31	17.01	0.02	1.53	9.18	5.56	21.82	41.69
$\hat{\beta}(3y)$	-0.27	-1.36	-0.71	1.19	-1.23	-0.40	0.32	-1.59	-1.10	-0.54
$R^2(3y)$	17.90	0.58	12.24	33.20	0.04	3.70	21.54	8.44	37.88	64.99
$\hat{\beta}(5y)$	-0.42	-1.89	-0.96	0.96	-1.81	-0.59	0.56	-2.18	-1.43	-0.61
$R^2(5y)$	29.87	0.42	14.50	40.99	0.05	5.24	29.71	7.37	39.21	68.38

Table 5 shows predictive regressions for excess log returns. The regressions take the form $r_{t:t+h} - r_t^f = \alpha(h) + \beta(h)x_t + \varepsilon_{t+h}(h)$, where h stands for the predictive horizon and x_t for the predictor. The top panel shows the results of regressing excess returns of horizons 1, 3, and 6 months on the one month variance premium. The regression is run on overlapping monthly data. The bottom panel shows the results of regressing excess returns of horizons 1, 3, and 5 years on the log price dividend ratio. The regression is run on overlapping annual data. In each case, I report the slope coefficient as well as the R^2. Small-sample model statistics are computed as discussed in Section 3.3. The sample for the first set of regressions spans 1990-2012, whereas the sample for the second set of regressions spans 1930-2012.

data, the price dividend ratio in the GDA-MSM model has large predictive power for excess returns. The median R^2s in the model rise from 6.3% at the annual horizon to 14.5% at the 5-year horizon. The EU-MSM model produces substantially less predictability. In particular, the median R^2 equals 1.5% at the 1-year horizon and it only rises to 5.2% at the 5-year horizon. These numbers are very similar to the ones implied by the long run risks model of Bansal, Kiku, and Yaron (2012), which is based on the same utility function. On the other hand, the GDA-AR1 model once again produces considerably too much predictability. For example, the median R^2 of 21.82% at the 1-year horizon exceeds its data counterpart by a factor of six.

An interesting aspect of the results in Table 5 is that the variance premium and the price-dividend ratio differ in terms of their ability to predict returns over different horizons. The variance premium is a successful predictor over short horizons (it produces a R^2 of 6.4% at the *quarterly* horizon in the GDA-MSM model), whereas the price-dividend ratio works better at longer horizon (it produces a R^2 of 6.3% at the *annual* horizon). The model successfully captures this challenging dimension of the data because it incorporates variance shocks with different persistence levels. As discussed above, an increase in endowment variance leads to an increase in the probability of disappointments and higher expected returns. However, the nature of this effect differs substantially across variance components. This is illustrated in Figure 8, which shows expected 1-month returns at different horizons conditional on one of the multipliers being in its high state. The panels differ in terms of the multipliers being considered. Conditional on the most persistent component being in its high state, returns are expected to be high over a long horizon (top-left panel). On the other hand, expected returns only increase over a short horizon conditional on the least persistent component being in its high state (bottom-right panel).

How do these differences help in reconciling the differences in the predictive ability of the variance premium and the price-dividend ratio? Because the price-dividend ratio reflects the riskiness of cash flows over the long-run, it is strongly affected by persistent shocks and much less strongly affected by transient shocks. Specifically, the log price-dividend ratio equals 3.11 conditional on component 1 being high and 3.33 conditional on component 1 being low. On the contrary, it only changes from 3.22 to 3.23 when conditioning on the most transient component (component 6) being high rather than low. Because changes in the price dividend ratio are mostly associated with changes in persistent variance components, it is a better predictor over long horizons than over short horizons.

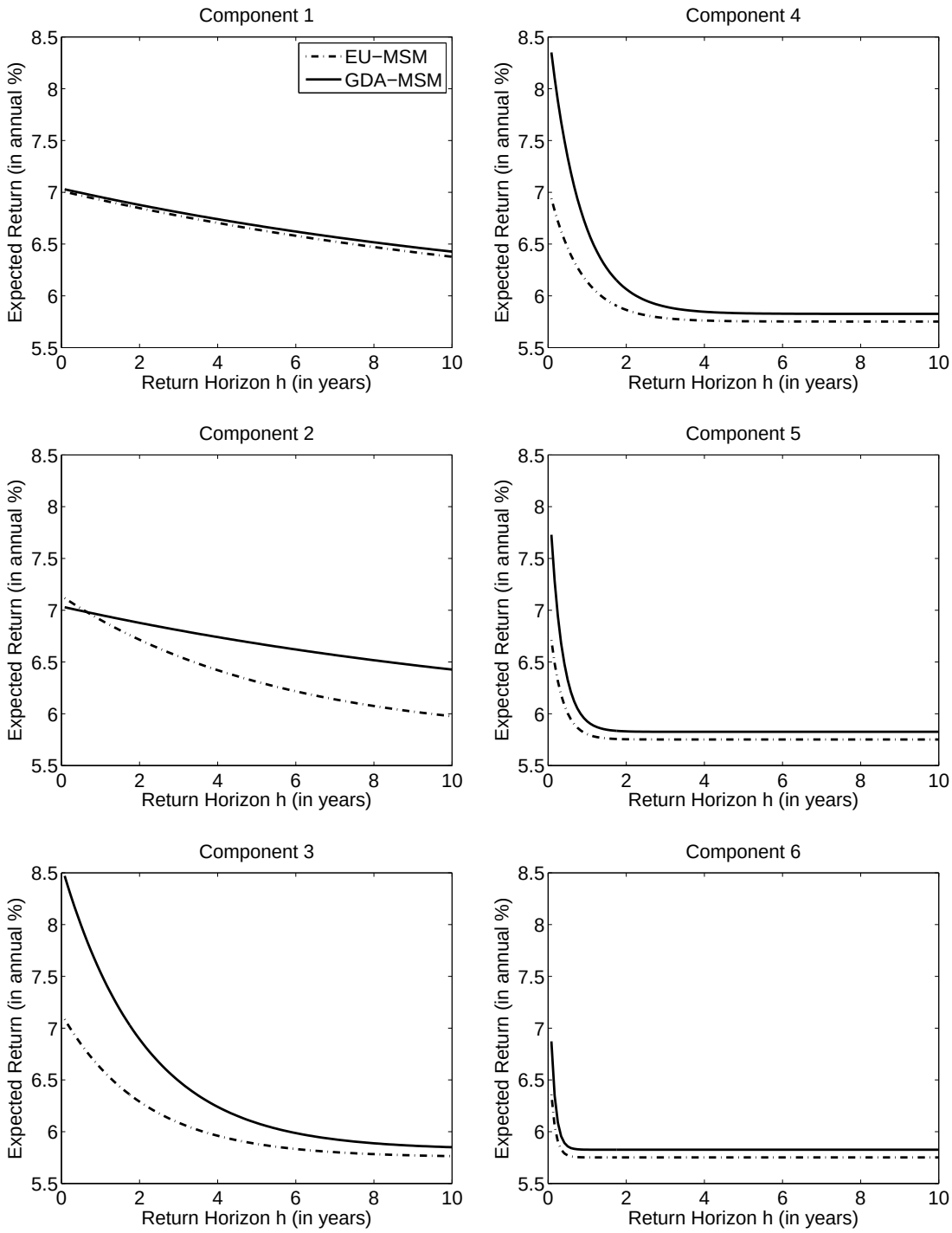

Figure 8: Shock Frequencies and Expected Returns

Figure 8 shows the expectation of the (annualized) 1-month log cum dividend return at time $t+h$ conditional on the k^{th} variance component being in the high state at time t (not conditioning on values of the other components). Component 1 has the highest persistence and component 6 has the lowest persistence. Dash-dotted lines refer to the EU model whereas solid lines refer to the GDA model. The unconditional mean return equals about 5.8% in both models. Note that returns are not cumulative, i.e. the figure shows single period (1-month) returns.

As the (one-month) variance premium equals the conditional covariance between realized return volatility and the pricing kernel over the next month (see Equation 3), its value depends more strongly on transient shocks than on persistent shocks. Specifically, the variance premium equals 11.99 conditional on component 1 being high and 11.53 conditional on component 1 being low. In contrast, it changes from 15.70 to 7.82 when conditioning on the most transient component (component 6) being high rather than low. The variance premium is therefore a better predictor over short horizons than over long horizons.

Lastly, Figure 8 shows that GDA risk preferences increase the importance of transient shocks relative to EU. For example, the intercept in the bottom-right panel (the expected 1-month return conditional on component 6 being in its high state) equals 6.87% the GDA model but only 6.36% in the EU model. Relative to the unconditional expected return of about 5.8%, transient shocks therefore increase the expected return by twice as much in the GDA model than in the EU model. On the other hand, the most persistent variance component affects expected returns equally for both preference specifications (top-left panel). This feature is interesting because a recent paper by Dew-Becker, Giglio, Le, and Rodriquez (2013) shows that recursive utility with EU risk preferences imply risk prices for low persistence shocks that are too low relative to what is implied by variance swap rates. The higher price for such shocks under GDA preferences represents a potential solution for this problem, and it likely contributes to the present model's ability to capture the dynamics of swap rates.

5 Additional Results

This section illustrates additional details about the model mechanism. First, I contrast the option pricing implications of the GDA model with those of the two nested cases, i.e. DA and EU risk preferences. Next, I quantify the risk aversion implied by different preference specifications and show explicitly that risk premia in the GDA-MSM model arise predominantly from aversion against tail risk. The analysis further shows that GDA preferences imply less risk aversion than both nested preference specifications when calibrated to match the equity premium.

5.1 The degree of tail sensitivity and option prices

The main mechanism for matching option prices in the present model is the high aversion toward tail risk implied by GDA preferences. In particular, risk aversion in the bench-

mark calibration is almost entirely determined by disappointment aversion because the period utility function, $u(x) = \log(x)$ has very little curvature. Furthermore, the disappointment threshold is set to a relatively low value of $\delta = 0.9625$, which implies an unconditional disappointment probability of 0.075% or about once per century.[26] This section investigates the importance of these choices for quantitatively matching option prices.

I begin by considering the role of the low disappointment *threshold*. Relative to the nested case of (pure) disappointment aversion (DA) risk preference, the GDA model shifts the disappointment threshold further into the left tail of the distribution, thereby lowering the probability of disappointment.[27] In the first experiment, I show how the variance premium and the implied volatility curve change as one gradually moves from a GDA calibration with a low threshold parameter ($\delta = 0.95$) to pure DA ($\delta = 1$). I keep the curvature parameter α fixed at the benchmark value of 0. To make the comparisons meaningful, I simultaneously adjust the disappointment magnitude θ so that the model-implied equity premium remains unchanged. In other words, I consider a set of economies that differ in terms of their risk preference calibrations, but not in their ability to match the historically observed equity premium. The left panel of Figure 9 shows the results. In all LHS panels, the horizontal axis is identical and scaled to be in units of δ. The MSM endowment calibration as well as the parameters controlling time preference are kept unchanged.

Panel A of Figure 9 shows that the average variance premium equals zero with DA preferences ($\delta = 1$). As δ is lowered, the variance premium increases and reaches the data value of 11.29 at the benchmark disappointment threshold of 0.9625. The reason for this effect is that the return variance is a convex function of the return itself, which implies that it is predominantly determined by extreme values. As the disappointment threshold is lowered, the utility function puts increasingly more weight on a smaller set of left-tail events, which are associated with large negative returns (see Section 4.4). Because variance swaps have high payoffs in these states, lowering δ increases the risk premium associated with them. Similar to the effect on the variance premium, both

[26] It should be noted that, while disappointing outcomes are overweighted in the utility computation ex-ante, the fact that an outcome is disappointing has no effect ex-post. In other words, events that fall just below and just above the disappointment threshold are not associated with systematically different changes in the price-dividend ratio or systematically different returns. In contrast to rare consumption disasters, infrequent disappointments therefore do not rely on the assumption of a Peso problem.

[27] DA preferences imply an unconditional disappointment probability of 39.7%, or about once per quarter when calibrated to match the equity premium.

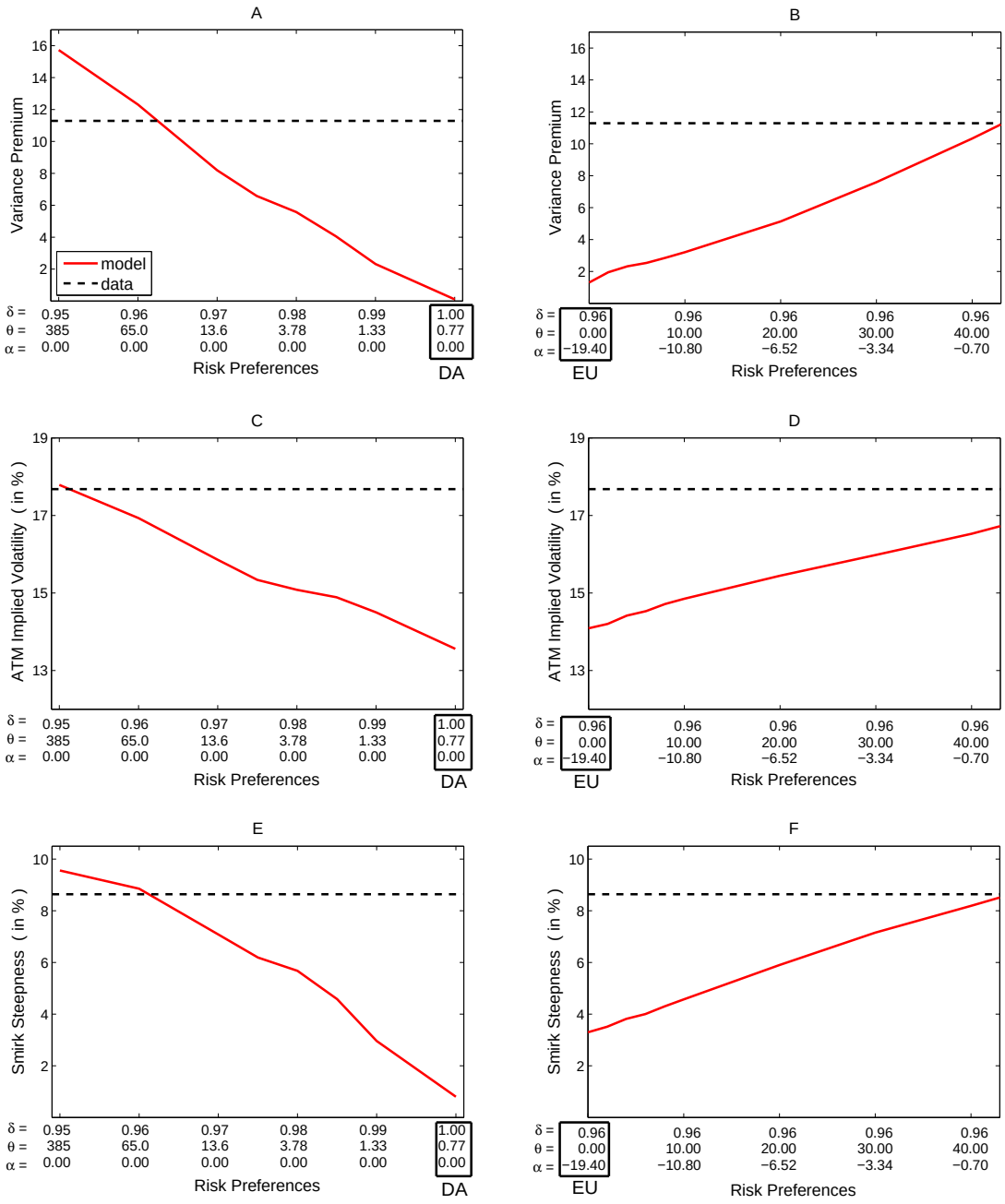

Figure 9: Tail Sensitivity of Preferences and Option Prices

Figure 9 shows how option-related moments change for different preference calibrations. All considered parameter combinations (shown on the horizontal axis) imply a mean annual equity premium equal to the historical average from 1930-2012. The graphs on the left consider changes in the disappointment threshold δ and include pure disappointment aversion ($\delta = 1$) as a special case. The graphs on the right consider changes in the disappointment magnitude θ and include expected utility risk preferences ($\theta = 0$) as a special case. Solid (red) lines show model-implied quantities, while dashed (black) lines show data equivalents. Panels A and B show the mean variance premium, Panels C and D show the at-the-money implied volatility, and Panels E and F show the steepness of the IV smirk, defined as the mean IV for options with a relative moneyness of -2 minus the mean IV for at-the-money options. Model-implied statistics equal small sample medians, computed as described in Section 3.3.

the level of the 1-month implied volatility curve (Panel C) and its slope (Panel E) are much too low for the DA model, increase as δ is lowered, and are close to their data counterparts for the benchmark calibration.[28] The level shift can be explained by the higher variance premium as well as the fact that lower δ values lead to more volatility feedback and therefore a higher return volatility. For example, the standard deviation of annual returns equals 14.7% for $\delta = 1$ and it increases to 18.2% for $\delta = 0.95$. The slope of the IV curve changes with the disappointment threshold for a similar reason as the variance premium, i.e. as δ is lowered, the agent's focus shifts toward more extreme returns, which makes put options with low strike prices particularly valuable.

In the second experiment, I illustrate the effect of changing the disappointment magnitude θ, and I adjust the curvature parameter α to hold the equity premium constant. The disappointment threshold δ is held fixed at the benchmark value of 0.9625. The results are shown in the right column of Figure 9, where the horizontal axis is scaled to be in units of θ. For $\alpha < 0$, the period utility function equals $u(x) = x^\alpha/\alpha$, wheras it equals $u(x) = \log(x)$ in my benchmark calibration ($\alpha = 0$). The value $\theta = 0$, shown at the left end of the plots, corresponds to expected utility (EU) risk preferences, i.e. the most popular version of Epstein-Zin preferences.[29] EU preferences imply a very low variance premium (Panel B), as well as an implied volatility curve that is both too low (Panel D) and too flat (Panel F) relative to the data. These results agree with those of previous studies that investigate option prices in the long run risks framework. As the disappointment magnitude is increased (and the curvature parameter lowered), risk aversion is increasingly determined by aversion to tail outcomes, which moves all three statistics closer to their data counterparts. Taken together, the results in this section illustrate that matching average option prices clearly requires GDA preferences ($\delta < 1$ and $\theta > 0$).

5.2 The degree and nature of risk aversion of a GDA agent

What is the degree of effective risk aversion implied by the benchmark GDA calibration? How does the answer compare to the one for the nested risk preference specifications of pure disappointment aversion (DA) and expected utility (EU)? To answer these questions, I conduct a welfare analysis in the spirit of Lucas (1987). Consider the following thought experiment. You are facing the consumption process described in

[28] I define the slope as the IV with a relative moneyness of -2 minus the at-the-money IV.

[29] The calibration is slightly different from the EU model in Section 4 because ρ is held fixed at the benchmark GDA calibration rather than that of the EU model. This is done to ensure that the benchmark GDA calibration appears among the considered cases.

Section 3 for $t, t+1, ...$ Parameter values are calibrated as shown in Table 2. How much would you pay to eliminate the risk inherent in the consumption process? More precisely, suppose you were offered to trade your current endowment for a consumption stream with the same current level of consumption but no future shocks (so that $\Delta c_s = \mu$ for $s = t+1, t+2, ...$). What is the maximum fraction of the mean consumption level that you would give up for this trade? Denote this fraction by Δ_t, and denote the value function evaluated at the alternative endowment by $\overline{V}_t$. Then Δ_t is defined by

$$\Delta_t = 1 - \frac{V_t}{\overline{V}_t} = 1 - \frac{\lambda_t^V}{\overline{\lambda}_t^V}, \tag{13}$$

where $\lambda_t^V = V_t/C_t$ and $\overline{\lambda}_t^V = \overline{V}_t/C_t$. For the deterministic endowment, there is no time-variation in the state, $\overline{V}_t$ and $\overline{\mu}_t$ grow at a constant rate of e^{μ}, and nothing is disappointing. Using these facts, the value function can be computed in closed form,[30] which allows me to write Equation 13 as

$$\Delta_t = 1 - \lambda_t^V \left(\frac{1-\beta}{1-\beta e^{\mu\rho}} \right)^{-\frac{1}{\rho}}. \tag{14}$$

Note that Δ_t depends on the state via λ_t^V. Similarly, Δ_t depends on risk preferences via λ_t^V. This latter fact allows me to use Δ_t to compare the degree of risk aversion across different risk preference calibrations. The results of this exercise are illustrated in Figure 10, where solid lines show the unconditional mean of Δ_t (the dashed lines will be discussed shortly). As in Section 5.1, all considered risk preference calibrations imply a mean equity premium equal to the historical mean.

The left panel considers changes in the disappointment threshold (δ) in order to contrast GDA with DA. The horizontal axis, which is identical to the one for the comparative statics results in Figure 9, is scaled to be in units of δ. The equity premium is held constant by adjusting the disappointment magnitude (θ) for each value of δ (the curvature parameter α is held constant at the benchmark value of $\alpha = 0$). The figure shows that the degree of risk aversion, as measured by $E[\Delta_t]$, is monotonically and nearly linearly increasing in δ. GDA thus implies less risk aversion than DA. The right panel considers changes in the disappointment magnitude (θ) in order to contrast GDA with EU risk preferences. The horizontal axis is scaled to be in units of θ. The equity premium is held constant by adjusting the curvature parameter, α, for each value of θ (the disappointment threshold δ is held constant at the benchmark value of $\delta = 0.9625$). In this

[30] Dividing the certainty equivalent (Equation 5) by C_t implies that $\overline{\lambda}^{\mu} = \overline{\lambda}^V e^{\mu}$, which can be substituted into the value function (Equation 4) to yield $\overline{\lambda}^V = \left(\frac{1-\beta}{1-\beta e^{\mu\rho}} \right)^{\frac{1}{\rho}}$. Note that the value function is time-invariant in this case because there is no time-variation in the state.

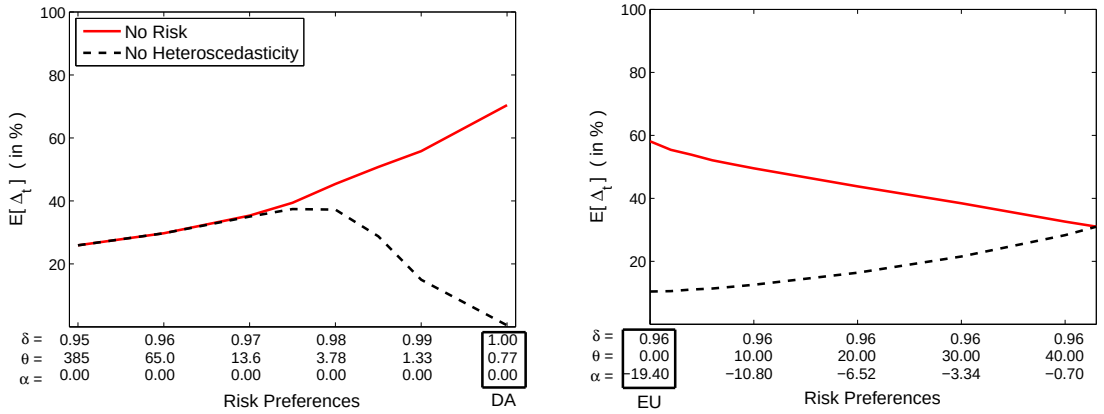

Figure 10: Welfare measures

Figure 10 shows welfare measures for different risk preference calibrations. All considered parameter combinations (shown on the horizontal axis) imply a mean annual equity premium equal to the historical average from 1930-2012. The endowment and time-preference parameters equal those of the benchmark GDA model calibration.

case risk aversion is linearly decreasing in θ, which implies that a GDA agent is less risk averse than an agent with EU preferences. To the extend that a low level of risk aversion is preferred because it corresponds more closely to estimates commonly found in Microeconomic studies, GDA risk preferences appear preferable compared to either DA or EU preferences.

To gain a clearer understanding about the interaction between risk preferences and stochastic endowment volatility in determining effective risk aversion, I compute a second, related welfare measure. Specifically, I consider an alternative endowment that sets volatility constant, so that $\Delta c_{t+1} = \mu + \overline{\sigma}\varepsilon^c_{t+1}$. As above, I ask how much the agent would give up in order to trade the benchmark endowment for this alternative endowment.[31] The results are shown with dashed lines in Figure 10. The following three features of the plot are helpful for understanding the mechanism of the GDA model:

1. (Left panel) The DA agent ($\delta = 1$) would give up close to nothing to eliminate the heteroscedasticity of his endowment (the dashed line is close to zero). The reason is that the disappointment threshold for DA equals the certainty equivalent, which implies that even outcomes close to the center of the distribution of V_{t+1} are considered disappointing. Shifting probability mass from the tails of the endowment distribution toward the center by removing stochastic volatility leads

[31] Different from the previous case, no closed-form solution exists for the value function for this alternative endowment, so that $\overline{\lambda}^V$ has to be computed numerically. The computation mimics the one for the benchmark model discussed in Section 3.2.

to no reduction in risk because these outcomes continue to be disappointing.

2. (Left panel) For sufficiently low disappointment thresholds (δ less than about 0.975), the GDA agent would give up just as much for eliminating stochastic volatility as he would give up for eliminating endowment risk all together (the two lines overlap for low δ's). This is due to the fact that without stochastic volatility, the probability of extreme tail outcomes becomes very small, so that nearly no outcomes are disappointing. Specifically, the i.i.d. endowment implies an unconditional disappointment probability of less than once per $100,000$ years and risk premia that are very close to zero. Eliminating stochastic volatility is therefore just as valuable as eliminating all risk.

3. (Right panel) Starting from EU ($\theta = 0$), increasing the disappointment magnitude increases the amount the GDA agent would give up for eliminating stochastic volatility. This holds true despite the fact that his overall degree of risk aversion (solid line) is decreasing in θ, which shows that tail risks account for a larger fraction of overall risk premia under GDA.

Overall, the results in this section illustrate that GDA implies somewhat less risk aversion than both DA and EU risk preferences. At the same time, the GDA agent is more averse to tail outcomes, resulting in a high price of tail risk.

6 Conclusion

The model presented in this paper provides a parsimonious explanation for a broad set of stylized facts of equity and equity index option markets. The model has two key components, a recursive utility function with GDA risk preferences and a multifractal process for endowment volatility. In equilibrium, regime switches in the conditional cash flow volatility result in substantial volatility feedback and strongly leptokurtic short-horizon returns. Because the most negative returns tend to coincide with particularly high marginal utility caused by disappointments, the representative agent is willing to pay large premia for assets that insure against them, such as variance swaps and out-of-the-money put options. This results in a large variance premium and a steep implied volatility smirk. In line with the empirical evidence in Bollerslev and Todorov (2011), risk premia in the model arise primarily from the representative agent's aversion to tail risk. In particular, the high price of tail risk allows the model to increase the importance of tail risks without relying on a Peso problem. This is a clear advantage relative to disaster models, which can only produce a realistic implied volatility smirk by assuming

no defaults on put options in times of macroeconomic disasters.

Return predictability arises in the model from the interaction between time-varying macroeconomic risk and GDA preferences. Specifically, times of high uncertainty are associated with higher disappointment probabilities, so that the representative agent behaves more risk-aversely. As a result, volatility shocks with low persistence cause strong co-movement between short horizon risk premia, giving the variance premium the ability to predict short-horizon excess returns. On the other hand, shocks with high persistence give the price-dividend ratio the ability to predict excess returns at longer horizons. Different shock frequencies also allow the model to generate realistic dynamics for the conditional return volatility, which in turn results in realistic time-series properties for option prices.

A number of avenues suggest themselves for future research. It would be interesting to extend the current framework to other asset classes such as credit default swaps (CDS) or currencies. Because payoffs to CDS and out-of-the-money put options occur in similar states, it is natural to conjecture that GDA can rationalize features of CDS that are hard to explain with standard preferences. Similarly, recent empirical evidence in Gavazzoni, Sambalaibat, and Telmer (2013) suggests that exchange rate and interest rate data imply properties for the pricing kernel that directly contradict those implied by log normal models. The authors conclude that future work should develop models in which higher moments play a more prominent role in the pricing kernel. GDA preferences represent a promising approach to do so. Lastly, it would be interesting to further analyze the effect that GDA preferences have on the pricing of shocks with different frequencies. Existing empirical evidence suggests that recursive utility models with EU risk preferences imply counterfactually low risk prices for high frequency shocks. The evidence presented in Section 4.7 shows that this property improves in the presence of GDA risk preferences. Future work could investigate the theoretical reasons for this finding as well as the general endowment properties necessary for it to result in a quantitatively important effect.

Appendix A: Data

Filters applied to option data

Observations with the following characteristics were removed.

1. Non-standard expiration dates
2. PM-settled options
3. Options with less than a week to expiration
4. Observations with error codes 999 (for ask) and 998 (for bid)
5. Observations whose bid-ask spread exceeds 10 times the bid
6. Singles (a call without matching put or a put without matching call)
7. The option's mid quote violates simple no arbitrage bounds. I allow for a small margin of error to reflect the fact that the assumed risk-free rate may not exactly reflect the prevailing lending rate for option brokers.

I further manually remove a few observations with obvious data errors.

Computing synthetic variance swap rates

With two exceptions, I use the same procedure that the CBOE uses to compute the VIX in order to compute synthetic variance swap rates. While the VIX is based on maturities bracketing 30 calendar days, I apply the same approach to compute swap rates with maturities of $30, 60, ..., 360$ days. The two exceptions are:

- The CBOE's replication formula is based on in-the-money options only because these options are most liquid. The at-the-money price level equals the forward price of the S&P 500, which is inferred from option data. Andersen, Bondarenko, and Ganzalez-Perez (2012) have documented that the CBOE's procedure for doing so lacks a certain robustness. Because I find that this problem becomes more severe at longer maturities, I use the approach suggested by Andersen et al. for inferring the forward price.
- I follow Jiang and Tian (2005) in approximating the integral over strike prices beyond the observed strike range. Once again, I find this to be particularly important at longer maturities.

Steps for estimating the variance premium

- To estimate a time-series of daily realized variance estimates, I implement the realized kernel estimator of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008). The basic idea of the estimator is akin to that of a HAC covariance estimator such as the well-known example by Newey and West (1987), i.e. it corrects for non-zero autocorrelations at higher lags. Based on the recommendations in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) as well as subsequent papers by the

same authors, I use a Parzen kernel and sample in business time (using every X^{th} transaction) in such a way that observations lie 60 seconds apart on average.

- I then produce one-month-ahead variance forecasts from a time-series model. Specifically, I follow Drechsler and Yaron (2011) in first computing a time series of monthly variance estimates by aggregating the daily estimates,[32] and then projecting the monthly realized variance onto its first lag and the variance swap rate at the end of the previous month. The OLS estimates for the model are as follows

$$RV_{t:t+22} = \underset{[0.16]}{0.00} + \underset{[4.73]}{0.34} RV_{t-22:t} + \underset{[7.63]}{0.47} \mathcal{V}_t(22), \quad R^2_{adj} = 50.4\%$$

  T-statistics are Newey-West (HAC) corrected using three lags.

- The one-step-ahead forecast from the time series model serves an an estimate of the conditional expectation of next-month realized variance. The difference between this estimate and the one-month variance swap rate proxies for the conditional variance premium.

Steps for computing implied volatilities on a fixed grid

- On any day in the sample, I first compute the (Black-Scholes) IV and the standardized moneyness for all available options. To standardize moneyness, I use the previously computed variance swap rate with the same maturity as the option.

- I then use a Gaussian kernel to interpolate IV to a fixed grid of relative moneyness from -2 to 1 and maturities from 1 to 12 months. This is the same procedure used by OptionMetrics to compute its well-known implied volatility surface.[33] I use a kernel weight of 0.05 in the maturity dimension (measured as the log of days to maturity) and a weight of 0.005 in the moneyness dimension, both of which are similar to the values used by OptionMetrics.

Data for annual cash flows and basic asset prices

Data for annual cash flows and asset prices are equivalent to those in Bansal, Kiku, and Yaron (2012) and Beeler and Campbell (2012) – extended to the end of 2012 – with one exception: I add the four quarterly ex-ante estimates of the risk-free rate within each year to compute the annual risk-free rate instead of multiplying the first-quarter estimate by four. The alternative corresponds more closely to the model-equivalent, which is based on the sum of the risk-free rates within the year, both in my paper and in the above-cited papers.

[32] The aggregation approach of Drechsler and Yaron (2011) is to sum up the daily estimates within each month. This has the disadvantage that the monthly estimates vary not only because the variance changes from month to month, but also because the number of trading days differs across months. Because the second source of variation is not what I intend to capture, I instead estimate the monthly variance as 22 times the average daily estimate within each month.

[33] OptionMetrics expresses IV as a function of maturity and the option's delta, which serves as a standardized moneyness measure.

Appendix B: Model solution details

The model solution is characterized by the $N \times 1$ vectors

$$\lambda^V, \lambda^\mu, \mathcal{D}, \mathcal{B}(\tau), \mathcal{C}(\tau, K), \mathcal{V}(\tau),$$

for $\tau \in \mathbb{N}^+$ and $K \in \mathbb{R}^+$, which are defined in Section 3.2. This appendix contains algebraic details of the solutions. In deriving these solutions, I consider the more general endowment

$$\begin{aligned}\Delta c_{t+1} =& \mu_t^c + \sigma_t^c \varepsilon_{t+1}^c \\ \Delta d_{t+1} =& \mu_t^d + \sigma_t^d \varepsilon_{t+1}^d,\end{aligned}$$

which allows for arbitrary Markov-switching processes for both the mean and the volatility of $(\Delta c, \Delta d)$. The model in the main text is a special case given by $\mu_t^c = \mu_t^d = \mu$, $\sigma_t^c = \sigma_t$, and $\sigma_t^d = \varphi\sigma_t$.

Notation. Throughout the appendix, denote the cumulative density function (cdf) of a standard normal by $\Phi(\cdot)$, and let the number of arguments indicate whether the cdf is univariate or bivariate. Let $\mathbf{1}\{\cdot\}$ denote the indicator function. Denote the Hadamard (element-wise) matrix product by $\odot$. Lastly, define

$$\begin{aligned}\mu_t^{cd} \equiv& (\alpha - 1)\mu_t^c + \mu_t^d \\ \sigma_t^{cd} \equiv& \frac{1}{2}(\alpha - 1)^2 (\sigma_t^c)^2 + \frac{1}{2}(\sigma_t^d)^2 + (\alpha - 1)\sigma_t^c \sigma_t^d\end{aligned}$$

The derivations presented below makes extensive use of the the following Lemma, the proof of which is contained in the online appendix.

Lemma 1. Let x, y be standard normal with correlation ρ, and let a, b, r, s be constants. Then

(a)

$$E[e^{rx+sy}\mathbf{1}\{x \le a\}] = e^{\frac{1}{2}(r^2+2\rho rs+s^2)}\Phi(a - r - \rho s)$$

(b)

$$E[e^{rx+sy}\mathbf{1}\{a \le y\}] = e^{\frac{1}{2}(r^2+2\rho rs+s^2)}\left[1 - \Phi(a - \rho r - s)\right]$$

(c)

$$E[e^{rx+sy}\mathbf{1}\{x \le a\}\mathbf{1}\{b \le y\}] = e^{\frac{1}{2}(r^2+2\rho rs+s^2)}\left[\Phi(a - r - \rho s) - \Phi(a - r - \rho s, b - \rho r - s)\right]$$

(d)

$$E\left[e^{rx}y^2\mathbf{1}\{x \le a\}\right] = e^{\frac{r^2}{2}}(1 + r^2\rho^2)\Phi(a - r) - e^{\frac{2ra-a^2}{2}}\frac{(r + a)\rho^2}{\sqrt{2\pi}} \equiv \Omega(r, a)$$

(e)

$$E\left[e^{rx}y\mathbf{1}\{x \le a\}\right] = \rho e^{\frac{r^2}{2}}\left(r\Phi(a - r) - \frac{1}{\sqrt{2\pi}}e^{-\frac{(a-r)^2}{2}}\right) \equiv \Gamma(r, a)$$

Solving for the Pricing Kernel

Define the utility-consumption and certainty equivalent-consumption ratios $\lambda_t^V \equiv V_t/C_t$ and $\lambda_t^\mu \equiv \mu_t/C_t$. To simplify notation, re-write the disappointment event as

$$V_{t+1} \leq \delta\mu_t \quad \Leftrightarrow \quad \frac{\lambda_{t+1}^V}{\lambda_t^\mu} e^{\Delta c_{t+1}} \leq \delta \quad \Leftrightarrow \quad \varepsilon_{t+1}^c \leq \frac{\log\left(\frac{\delta\lambda_t^\mu}{\lambda_{t+1}^V}\right) - \mu}{\sigma_t^c} \equiv \phi_{t+1}^c, \tag{15}$$

where I have defined the disappointment threshold ϕ_{t+1}^c. After some simple algebra, the pricing kernel (Equation 7) can be expressed as

$$M_{t+1} = \beta e^{(\alpha-1)\Delta c_{t+1}} \left(\frac{\lambda_{t+1}^V}{\lambda_t^\mu}\right)^{\alpha-\rho} \left(\frac{1+\theta\mathbf{1}\{\varepsilon_{t+1}^c \leq \phi_{t+1}^c\}}{1+\delta^\alpha\theta E_t[\Phi\left(\phi_{t+1}^c\right)]}\right), \tag{16}$$

where I have used the law of iterated expectations to write $E_t\left[\mathbf{1}\left(V_{t+1} \leq \delta\mu_t\right)\right]$ as $E_t[\Phi\left(\phi_{t+1}^c\right)]$. The next step is to solve for the ratios λ_t^V and λ_t^μ. Dividing the value function (Equation 4) by C_t gives

$$\lambda_t^V = \left[(1-\beta) + \beta(\lambda_t^\mu)^\rho\right]^{\frac{1}{\rho}}. \tag{17}$$

Dividing the certainty equivalent (Equation 5) by C_t and re-arranging terms gives

$$\lambda_t^\mu = \begin{cases} E_t\left[e^{\alpha\Delta c_{t+1}}\left(\lambda_{t+1}^V\right)^\alpha \frac{1+\theta\mathbf{1}\{\varepsilon_{t+1}^c \leq \phi_{t+1}^c\}}{1+\theta\delta^\alpha E_t[\Phi\left(\phi_{t+1}^c\right)]}\right]^{\frac{1}{\alpha}} & \text{, for } \alpha \leq 1,\ \alpha \neq 0 \\ \exp\left(E_t\left[\frac{\left(\log(\lambda_{t+1}^V)+\Delta c_{t+1}\right)(1+\theta\mathbf{1}\{\varepsilon_{t+1}^c \leq \phi_{t+1}^c\}) - \theta\log(\delta)\mathbf{1}\{\varepsilon_{t+1}^c \leq \phi_{t+1}^c\}}{1+\theta E_t[\Phi\left(\phi_{t+1}^c\right)]}\right]\right) & \text{, for } \alpha = 0 \end{cases} \tag{18}$$

The expectation over ε_{t+1}^c can be evaluated by using Lemma 1a. In the case $\alpha \leq 1$, $\alpha \neq 0$, the objects in the lemma are given by $x = \varepsilon_{t+1}^c$, $r = \alpha\sigma_t^c$, $s = 0$, and $a = \phi_{t+1}^c$. To solve the case $\alpha = 0$, note that for $\varepsilon \sim N(0,1)$, $E[\varepsilon\mathbf{1}\{\varepsilon \leq a\}] = \int_{-\infty}^a \varepsilon f(\varepsilon)d\varepsilon = -\frac{1}{\sqrt{2\pi}} e^{-\frac{\varepsilon^2}{2}}\Big|_{-\infty}^a = -\Phi'(a)$. It follows that

$$\lambda_t^\mu = \begin{cases} E_t\left[\frac{\left(\lambda_{t+1}^V\right)^\alpha}{1+\theta\delta^\alpha E_t[\Phi\left(\phi_{t+1}^c\right)]} e^{\alpha\mu_t^c + \frac{1}{2}\alpha^2(\sigma_t^c)^2}\left(1+\theta\Phi(\phi_{t+1}^c - \alpha\sigma_t^c)\right)\right]^{\frac{1}{\alpha}} & \text{, for } \alpha \leq 1,\ \alpha \neq 0 \\ \exp\left(E_t\left[\frac{\left(\log(\lambda_{t+1}^V)+\mu_t^c\right)(1+\theta\Phi(\phi_{t+1}^c)) - \sigma_t^c\theta\Phi'(\phi_{t+1}^c) - \theta\log(\delta)\Phi(\phi_{t+1}^c)}{1+\theta E_t[\Phi(\phi_{t+1}^c)]}\right]\right) & \text{, for } \alpha = 0 \end{cases} \tag{19}$$

Given a guess for the $N \times 1$ vectors λ^V and λ^μ, the remaining expectation in equation 19 can be evaluated as a matrix product. The system of equations in 17 and 19 has to be solved numerically.

Proof of Theorem 1

A. Risk-free Bonds. The price of a one-period bond equals

$$\begin{aligned}\mathcal{B}_t(1) =& E_t\left[M_{t+1}\right] \\ =& E_t\left[\beta e^{(\alpha-1)\Delta c_{t+1}}\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\left(\frac{1+\theta\mathbf{1}\{\varepsilon^c_{t+1}\leq\phi^c_{t+1}\}}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right)\right] \\ =& E_t\left[\beta e^{(\alpha-1)\mu^c_t+\frac{1}{2}(\alpha-1)^2(\sigma^c_t)^2}\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\left(\frac{1+\theta\Phi(\phi^c_{t+1}-(\alpha-1)\sigma^c_t)}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right)\right],\end{aligned} \tag{20}$$

where the last equality uses the law of iterated expectations and Lemma 1a with $x = \varepsilon^c_{t+1}$, $r = (\alpha-1)\sigma^c_t$, $s = 0$, and $a = \phi^c_{t+1}$. Prices of multi-period bonds can be expressed recursively as

$$\mathcal{B}_t(\tau) \equiv E_t[M_{t:t+\tau}] = E_t\left[\prod_{h=1}^{\tau} M_{t+h}\right] = E_t\left[\left(\prod_{h=1}^{\tau-1} M_{t+h}\right)E_{t+\tau-1}[M_{t+\tau}]\right]$$

To evaluate the remaining expectations, denote the term inside the expectation of Equation 20 by a^b_{ij} when the Markov chain is in state i at time t in in state j at time $t+1$. Collect the terms a^b_{ij} in a matrix A^b. Defining $\mathcal{B}(0) \equiv \iota_N$, bond prices can be written recursively in matrix form as

$$\mathcal{B}(\tau) = (P \odot A^b)\cdot\mathcal{B}(\tau-1)$$

B. Equity. After dividing by D_t, the Euler equation for the dividend claim is given by

$$\begin{aligned}\mathcal{D}_t \equiv \frac{S_t}{D_t} =& E_t\left[M_{t+1}\frac{S_{t+1}+D_{t+1}}{D_t}\right] \\ =& E_t\left[\beta\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\left(\frac{1+\theta\mathbf{1}\{\varepsilon^c_{t+1}\leq\phi^c_{t+1}\}}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right)e^{(\alpha-1)\Delta c_{t+1}+\Delta d_{t+1}}\left(\mathcal{D}_{t+1}+1\right)\right]\end{aligned}$$

One can integrate out $(\varepsilon^c_{t+1}, \varepsilon^d_{t+1})$ by using the law of iterated expectations. Applying Lemma 1a with $x = \varepsilon^c_{t+1}$, $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$, $s = \sigma^d_t$, and $a = \phi^c_{t+1}$ yields

$$\mathcal{D}_t = E_t\Bigg[\underbrace{\beta\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\left(\frac{1+\theta\Phi(\phi^c_{t+1}-(\alpha-1)\sigma^c_t-\varrho\sigma^d_t)}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right)e^{\mu^{cd}_t+\sigma^{cd}_t}}_{\equiv a^d_{ij}}\left(\mathcal{D}_{t+1}+1\right)\Bigg],$$

where μ^{cd}_t and σ^{cd}_t were defined under 'Notation' on page 48. To evaluate the remaining expectation and to solve for $\mathcal{D}$, denote the term pre-multiplying $(\mathcal{D}_{t+1}+1)$ inside the expectation by a^d_{ij} when the Markov chain is in state i at time t in in state j at time $t+1$. Collect the terms a^d_{ij} in a matrix A^d. Then the expectation can be evaluated in matrix form as

$$\begin{aligned}\mathcal{D} =& [P \odot A^d \odot (\mathbf{1}_N + \iota_N(\mathcal{D})')]\iota_N \\ =& [P \odot A^d]\iota_N + [(P \odot A^d) \odot (\iota_N(\mathcal{D})')]\iota_N \\ =& [P \odot A^d]\iota_N + (P \odot A^d)\mathcal{D} \\ \Leftrightarrow \mathcal{D} =& (I_N - P \odot A^d)^{-1}[P \odot A^d]\iota_N\end{aligned}$$

C. Call Options. The (relative) price of a 1-period call with moneyness K equals

$$\mathcal{C}_t(1,K) \equiv E_t\left[M_{t+1}\max\left(0, \frac{S_{t+1}}{S_t} - K\right)\right],$$

The call payoff is triggered if

$$\frac{S_{t+1}}{S_t} > K \Leftrightarrow \frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}e^{\Delta d_{t+1}} > K \Leftrightarrow \varepsilon^d_{t+1} > \frac{\log\left(K\frac{\mathcal{D}_t}{\mathcal{D}_{t+1}}\right) - \mu^d_t}{\sigma^d_t} \equiv \phi^d_{t+1}(K),$$

where I have defined the payoff threshold $\phi^d_{t+1}(K)$. The price can now be written as

$$\begin{aligned}\mathcal{C}_t(1,K) =& E_t\left[M_{t+1}\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}e^{\Delta d_{t+1}} - K\right)\mathbf{1}\left\{\phi^d_{t+1}(K) < \varepsilon^d_{t+1}\right\}\right] \\ =& E_t\left[\beta e^{(\alpha-1)\Delta c_{t+1}}\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\left(\frac{1+\theta\mathbf{1}\{\varepsilon^c_{t+1}\le\phi^c_{t+1}\}}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right)\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}e^{\Delta d_{t+1}} - K\right)\mathbf{1}\left\{\phi^d_{t+1}(K) < \varepsilon^d_{t+1}\right\}\right]\end{aligned}$$

To integrate out $(\varepsilon^c_{t+1}, \varepsilon^d_{t+1})$, one has to use the law of iterated expectations in combination with Lemma 1 several times. Specifically, the above expression contains four additive parts containing normal innovations. To keep notation manageable, I first integrate over the normal terms in each of these four expressions separately. Using Lemma 1b with $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$, $s = \sigma^d_t$, and $a = \phi^d_{t+1}(K)$,

$$E_t\left[e^{(\alpha-1)\Delta c_{t+1}+\Delta d_{t+1}}\mathbf{1}\left\{\phi^d_{t+1}(K) < \varepsilon^d_{t+1}\right\}\right] = E_t\Big[\underbrace{e^{\mu^{cd}_t+\sigma^{cd}_t}(1-\Phi(\phi^d_{t+1}(K) - (\alpha-1)\varrho\sigma^c_t - \sigma^d_t))}_{\equiv\chi^1_{t+1}}\Big]$$

Using Lemma 1b with $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$, $s = 0$, and $a = \phi^d_{t+1}(K)$,

$$E_t\left[e^{(\alpha-1)\Delta c_{t+1}}\mathbf{1}\left\{\phi^d_{t+1}(K) < \varepsilon^d_{t+1}\right\}\right] = E_t\Big[\underbrace{e^{(\alpha-1)\mu^c_t+\frac{1}{2}(\alpha-1)^2(\sigma^c_t)^2}(1-\Phi(\phi^d_{t+1}(K) - (\alpha-1)\varrho\sigma^c_t))}_{\equiv\chi^2_{t+1}}\Big]$$

Using Lemma 1c with $x = \varepsilon^c_{t+1}$, $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$, $s = \sigma^d_t$, $a = \phi^c_{t+1}$, and $b = \phi^d_{t+1}(K)$,

$$\begin{aligned}&E_t\left[e^{(\alpha-1)\Delta c_{t+1}+\Delta d_{t+1}}\mathbf{1}\{\varepsilon^c_{t+1}\le\phi^c_{t+1}\}\mathbf{1}\left\{\phi^d_{t+1}(K) < \varepsilon^d_{t+1}\right\}\right] \\ &= E_t\Big[\underbrace{e^{\mu^{cd}_t+\sigma^{cd}_t}\left(\Phi(\phi^c_{t+1} - (\alpha-1)\sigma^c_t - \varrho\sigma^d_t) - \Phi(\phi^c_{t+1} - (\alpha-1)\sigma^c_t - \varrho\sigma^d_t, \phi^d_{t+1}(K) + \gamma\rho\sigma^c_t - \sigma^d_t)\right)}_{\equiv\chi^3_{t+1}}\Big]\end{aligned}$$

Using Lemma 1c with $x = \varepsilon^c_{t+1}$, $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$, $s = 0$, $a = \phi^c_{t+1}$, and $b = \phi^d_{t+1}(K)$,

$$\begin{aligned}&E_t\left[e^{(\alpha-1)\Delta c_{t+1}}\mathbf{1}\{\varepsilon^c_{t+1}\le\phi^c_{t+1}\}\mathbf{1}\left(\phi^d_{t+1}(K) < \varepsilon^d_{t+1}\right)\right] \\ &= E_t\Big[\underbrace{e^{(\alpha-1)\mu^c_t+\frac{1}{2}(\alpha-1)^2(\sigma^c_t)^2}\left(\Phi(\phi^c_{t+1} - (\alpha-1)\sigma^c_t) - \Phi(\phi^c_{t+1} - (\alpha-1)\sigma^c_t, \phi^d_{t+1}(K) - (\alpha-1)\varrho\sigma^c_t)\right)}_{\equiv\chi^4_{t+1}}\Big]\end{aligned}$$

Recombining these terms, the call price can be written as

$$\mathcal{C}_t(1,K) = E_t\Bigg[\underbrace{\frac{\beta}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\times\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}(\chi^1_{t+1}+\theta\chi^3_{t+1})-K(\chi^2_{t+1}+\theta\chi^4_{t+1})\right)}_{\equiv a^c_{ij}(K)}\Bigg]$$

To evaluate the remaining expectations, denote the term inside the expectation by $a^c_{ij}(K)$ when the Markov chain is in state i at time t in in state j at time $t+1$. Collect the terms $a^c_{ij}(K)$ in a matrix $A^c(K)$. Then the expectation can be evaluated in matrix form for each K as

$$\mathcal{C}(1,K) = (P\odot A^c(K))\cdot\iota_N$$

D. Variance Swap Rates. The τ-period variance swap rate is given by

$$\begin{aligned}\mathcal{V}_t(\tau) = E^Q_t\left[\sum_{h=1}^{\tau} r^2_{t+h}\right] =& E_t\left[M_{t:t+\tau}\right]^{-1}\times E_t\left[M_{t:t+\tau}\sum_{h=1}^{\tau} r^2_{t+h}\right]\\ =&\mathcal{B}_t(\tau)^{-1}\times\sum_{h=1}^{\tau}E_t\left[M_{t:t+h}M_{t+h:t+\tau}r^2_{t+h}\right]\\ =&\mathcal{B}_t(\tau)^{-1}\times\sum_{h=1}^{\tau}E_t\left[M_{t:t+h}r^2_{t+h}E_{t+h}\left[M_{t+h:t+\tau}\right]\right]\\ =&\mathcal{B}_t(\tau)^{-1}\times\sum_{h=1}^{\tau}\underbrace{E_t\left[M_{t:t+h}r^2_{t+h}\mathcal{B}_{t+h}(\tau-h)\right]}_{\equiv\mathcal{V}_t(h,\tau)}\end{aligned} \tag{21}$$

The terms $\mathcal{V}_t(h,\tau)$, for $h\leq\tau$ and $h,\tau\geq 2$, can be computed recursively as

$$\mathcal{V}_t(h,\tau) = E_t\left[M_{t+1}\mathcal{V}_{t+1}(h-1,\tau-1)\right],$$

and evaluated in matrix form as

$$\mathcal{V}(h,\tau) = (P\odot A^b)\cdot\mathcal{V}(h-1,\tau-1) \tag{22}$$

where the matrix A^b was defined under "Risk-free Bonds" above. The recursion begins with the terms $\mathcal{V}_t(1,\kappa)$ for $\kappa=1,...,\tau$, which can be computed as

$$\begin{aligned}\mathcal{V}_t(1,\kappa) =& E_t\left[M_{t+1}r^2_{t+1}\mathcal{B}_{t+1}(\kappa-1)\right]\\ =&E_t\Bigg[\beta e^{(\alpha-1)\Delta c_{t+1}}\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\left(\frac{1+\theta\mathbf{1}\{\varepsilon^c_{t+1}\leq\phi^c_{t+1}\}}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right)\times\mathcal{B}_{t+1}(\kappa-1)\\ &\times\left(\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)^2+(\mu^d_t)^2+(\sigma^d_t)^2(\varepsilon^d_{t+1})^2+2\mu^d_t\sigma^d_t\varepsilon^d_{t+1}+2\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)(\mu^d_t+\sigma^d_t\varepsilon^d_{t+1})\right)\Bigg],\end{aligned}$$

where $\mathcal{B}_{t+1}(0)=1$. To integrate out $(\varepsilon^c_{t+1},\varepsilon^d_{t+1})$, one has to use the law of iterated expectations in combination with Lemma 1 several times. For notational convenience, I first integrate over the normal terms in parts of the above expression individually,

and I subsequently combine terms. First, using Lemma 1d with $x = \varepsilon^c_{t+1}$, $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$ and $a = \infty$,

$$E_t\left[e^{(\alpha-1)\Delta c_{t+1}}(\varepsilon^d_{t+1})^2\right] = \underbrace{e^{(\alpha-1)\mu_t}\Omega\left((\alpha-1)\sigma^c_t, \infty\right)}_{\equiv \xi_{1,t+1}},$$

where $\Omega(\cdot,\cdot)$ was defined in Lemma 1d. Note that $\Omega(r,\infty) = e^{\frac{r^2}{2}}(1+r^2\rho^2)$.[34] Second, using Lemma 1d with $x = \varepsilon^c_{t+1}$, $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$ and $a = \phi^c_{t+1}$,

$$E_t\left[e^{(\alpha-1)\Delta c_{t+1}}\mathbf{1}\{\varepsilon^c_{t+1} \le \phi^c_{t+1}\}(\varepsilon^d_{t+1})^2\right] = E_t\left[\underbrace{e^{(\alpha-1)\mu^c_t}\Omega\left((\alpha-1)\sigma^c_t, \phi^c_{t+1}\right)}_{\equiv \xi_{2,t+1}}\right].$$

Third, using Lemma 1e with $x = \varepsilon^c_{t+1}$, $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$ and $a = \infty$,

$$E_t\left[e^{(\alpha-1)\Delta c_{t+1}}\varepsilon^d_{t+1}\right] = \underbrace{e^{(\alpha-1)\mu_t}\Gamma\left((\alpha-1)\sigma^c_t, \infty\right)}_{\equiv \xi_{3,t+1}}.$$

Note that $\Gamma(r,\infty) = r\rho e^{\frac{r^2}{2}}$. Forth, using Lemma 1e with $x = \varepsilon^c_{t+1}$, $y = \varepsilon^d_{t+1}$, $r = (\alpha-1)\sigma^c_t$ and $a = \phi^c_{t+1}$,

$$E_t\left[e^{(\alpha-1)\Delta c_{t+1}}\mathbf{1}\{\varepsilon^c_{t+1} \le \phi^c_{t+1}\}\varepsilon^d_{t+1}\right] = E_t\left[\underbrace{e^{(\alpha-1)\mu_t}\Gamma\left((\alpha-1)\sigma^c_t, \phi^c_{t+1}\right)}_{\equiv \xi_{4,t+1}}\right],$$

Lastly, using Lemma 1a with $x = \varepsilon^c_{t+1}$, $r = (\alpha-1)\sigma^c_t$, $s = 0$, and $a = \phi^c_{t+1}$,

$$E_t\left[e^{(\alpha-1)\Delta c_{t+1}}\mathbf{1}\{\varepsilon^c_{t+1} \le \phi^c_{t+1}\}\right] = E_t\left[e^{(\alpha-1)\mu^c_t + \frac{1}{2}(\alpha-1)^2(\sigma^c_t)^2}\Phi(\phi^c_{t+1} - (\alpha-1)\sigma^c_t)\right].$$

Using this last result as well as the ξ-terms defined above, $\mathcal{V}_t(1,\kappa)$ can be written as

$$\begin{aligned}
\mathcal{V}_t(1,\kappa) = E_t\Bigg[\mathcal{B}_{t+1}(\kappa-1) \times \Bigg\{ & \beta e^{(\alpha-1)\mu^c_t + \frac{1}{2}(\alpha-1)^2(\sigma^c_t)^2}\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho}\left(\frac{1+\theta\Phi(\phi^c_{t+1} - (\alpha-1)\sigma^c_t)}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right) \\
& \times \left(\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)^2 + (\mu^d_t)^2 + 2\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)\mu^d_t\right)\Bigg] \\
& + \left(\frac{\beta}{1+\delta^\alpha\theta E_t[\Phi\left(\phi^c_{t+1}\right)]}\right)\left(\frac{\lambda^V_{t+1}}{\lambda^\mu_t}\right)^{\alpha-\rho} \\
& \times \left((\sigma^d_t)^2(\xi_{1,t+1} + \theta\xi_{2,t+1}) + \left[2\mu^d_t\sigma^d_t + 2\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)\sigma^d_t\right](\xi_{3,t+1} + \theta\xi_{4,t+1})\right)\Bigg\}\Bigg]
\end{aligned}$$

To evaluate the remaining expectation, denote the term in the curly brackets by a^ν_{ij} when the Markov chain is in state i at time t in in state j at time $t+1$. Collect the terms a^ν_{ij} in a matrix A^ν. Then the expectation can be evaluated in matrix form as

$$\mathcal{V}(1,\kappa) = (P \odot A^\nu)\cdot\mathcal{B}(\kappa-1)$$

[34] This holds because in the second term of Ω, the exponential factor goes to zero faster that the latter factor goes to infinity.

At this point one can apply the recursion in Equation 22 to compute $\mathcal{V}(h,\kappa)$ for $h>1$. Lastly, variance swap rates can be computed by summing appropriate terms via Equation 21.

E. Variance Premium. The 1-period variance premium is given by

$$\begin{aligned}
VP_t =& E_t^Q\left[\sum_{h=1}^{\tau} r_{t+h}^2\right] - E_t^P\left[\sum_{h=1}^{\tau} r_{t+h}^2\right] \\
=& \mathcal{V}_t(1) - E_t\left[\left(\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)^2 + (\mu_t^d)^2 + (\sigma_t^d)^2(\varepsilon_{t+1}^d)^2 + 2\mu_t^d\sigma_t^d\varepsilon_{t+1}^d + 2\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)(\mu_t^d+\sigma_t^d)\right)\right] \\
=& \mathcal{V}_t(1) - E_t\underbrace{\left[\left(\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)^2 + (\mu_t^d)^2 + (\sigma_t^d)^2 + 2\log\left(\frac{\mathcal{D}_{t+1}}{\mathcal{D}_t}\right)\mu_t^d\right)\right]}_{\equiv a_{ij}^p},
\end{aligned}$$

where the last equality used $E_t[\varepsilon_{t+1}^d]=0$ and $E_t[(\varepsilon_{t+1}^d)^2]=1$. Denote the term inside of the expectation by a_{ij}^p when the Markov chain is in state i at time t in in state j at time $t+1$. Collect the terms a_{ij}^p in a matrix A^p. Then the variance premium can be evaluated in matrix form as

$$VP = \mathcal{V}(1) - (P \odot A^p)\cdot \iota_N$$

.

References

Allais, M. (1979): "The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school," in *Expected Utility Hypothesis and the Paradox*, ed. by M. Allais, and O. Hagon. D. Reidel Publishing Co., Dordrecht, Holland.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001): "The distribution of realized stock return volatility," *Journal of Financial Economics*, 61, 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71(2), 579–625.

Andersen, T. G., O. Bondarenko, and M. T. Ganzalez-Perez (2012): "Uncovering Novel Features of Equity-Index Return Dynamics via Corridor Implied Volatility," Working Paper.

Backus, D., M. Chernov, and I. Martin (2011): "Disasters Implied by Equity Index Options," *Journal of Finance*, 66(6), 1969–2012.

Bansal, R., D. Kiku, and A. Yaron (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1(1), 183–221.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 57(4), 1481–1509.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008): "Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76(6), 1481–1536.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *Quarterly Journal of Economics*, 121(3), 823–866.

Barro, R. J., and J. F. Ursua (2008): "Macroeconomic crises since 1870," *Brookings Papers on Economic Activity*, p. 255335.

Beeler, J., and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1(1), 141–182.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2011): "Explaining asset pricin gpuzzles associated with the 1987 market crash," *Journal of Financial Economics*, 101, 552–573.

Boguth, O., and L.-A. Kuehn (2013): "Consumption Volatility Risk," *Journal of Finance*, forthcoming.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, (11), 4464–4492.

Bollerslev, T., and V. Todorov (2011): "Tails, fears, and risk premia," *Journal of Finance*, 66(6), 2165–2211.

Bonomo, M., R. Garcia, N. Meddahi, and R. Tedongap (2011): "Generalized Disappointment Aversion, Long-run Volatility Risk, and Asset Prices," *Review of Financial Studies*, 24(1), 82–122.

Breeden, D., and R. H. Litzenberger (1978): "State Contingent Prices Implicit in option Prices," *Journal of Business*, 51(4), 3–24.

Britten-Jones, M., and A. Neuberger (2000): "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 55(2), 839–866.

Calvet, L. E., and A. J. Fisher (2001): "Forecasting Multifractal Volatility," *Journal of Financial Econometrics*, 105, 27–58.

——— (2004): "How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes," *Journal of Financial Econometrics*, 2(1), 49–83.

——— (2007): "Multifrequency News and Stock Returns," *Journal of Financial Economics*, 86, 178–212.

Campanale, C., R. Castro, and G. L. Clementi (2010): "Asset Pricing in a Production Economy with Chew-Dekel Preferences," *Review of Economic Dynamics*, 13, 379–402.

Carr, P., and L. Wu (2003): "The Finite Moment Log Stable Process and Option Pricing," *Journal of Finance*, 58(2), 753–777.

——— (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22, 1311–1341.

Dew-Becker, I., S. Giglio, A. Le, and M. Rodriquez (2013): "The Term Structure of the Variance Risk Premium and Investor Preferences," Working Paper.

Ding, Z., C. W. J. Granger, and F. Engle, Robert (1993): "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1.

Drechsler, I. (2013): "Uncertainty, Time-Varying Fear, and Asset Prices," *Journal of Finance*, forthcoming.

Drechsler, I., and A. Yaron (2011): "Whats Vol Got to Do with It," *Review of Financial Studies*, 24.

Du, D. (2011): "General Equilibrium Pricng of Options with Habit Formation and Event Risks," *Journal of Financial Economics*, 99, 400–426.

Epstein, L. G., and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4), 937–969.

——— (2001): "The Independence Axiom and Asset Returns," *Journal of Empirical Finance*, 8, 537–572.

Foresi, S., and L. Wu (2005): "Crash-O-Phobia: A Domestric Fear or a Worldwide Concern?," *Journal of Derivatives*, pp. 8–21.

Gavazzoni, F., B. Sambalaibat, and C. Telmer (2013): "Currency Risk and Pricing Kernel Volatility," Working Paper.

Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59(3), 667–686.

Jiang, G., and Y. Tian (2005): "The Model-Free Implied Volatility and Its Information Content," *Review of Financial Studies*, 18, 1305–1342.

Lucas, R. E. J. (1987): *Models of Business Cycles*. Basil Blackwell, New York.

Martin, I. (2013): "Simple Variance Swaps," Working Paper.

Nakamura, E., D. Sergeyev, and J. Steinsson (2012): "Growth-rate and Uncertainty Shocks in Consumption: Cross-Country Evidence," Working Paper.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

Rietz, T. A. (1988): "The equity risk premium: a solution," *Journal of Monetary Economics*, 22, 117–131.

Routledge, B. R., and S. E. Zin (2010): "Generalized Disappointment Aversion and Asset Prices," *Journal of Finance*, 65(4), 1303–1332.

Rouwenhorst, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, pp. 294–330. Princeton University Press, Princeton, NJ.

Rubinstein, M. (1994): "Implied Binomial Trees," *Journal of Finance*, (3), 771–818.

Seo, S. B., and J. A. Wachter (2013): "Option Prices in a Model with Stochastic Disaster Risk," Working Paper.

Tamoni, A. (2011): "The Multi-horizon Dynamics of Risk and Return," Working Paper.

Wachter, J. (2002): "Comment on: Are behavioral asset-pricing models structural?," *Journal of Monetary Economics*, 49, 229–233.

Working, H. (1960): "Note on the Correlation of First Differences of Aggregates in a Random Chain," *Econometrica*, 28, 916–918.

Online Appendix

Proof of Lemma 1

Parts a-c. All three parts are special cases of the following more general result

$$E[e^{rx+sy}\mathbf{1}\{a \leq x \leq b\}\mathbf{1}\{c \leq y \leq d\}] = e^{\frac{1}{2}(r^2+2\rho rs+s^2)}\left[\Phi(b^*,d^*)+\Phi(a^*,c^*)-\Phi(a^*,d^*)-\Phi(c^*,b^*)\right],$$

where $a^* = a-r-\rho s$, $b^* = b-r-\rho s$, $c^* = c-\rho r-s$, and $d^* = d-\rho r-s$. In what follows, I prove this more general case.

$$E[e^{rx+sy}\mathbf{1}\{a \leq x \leq b\}\mathbf{1}\{c \leq y \leq d\}] = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_a^b\int_c^d e^{rx+sy-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}dydx$$

I next re-write the exponent. This is easier in matrix notation. Define $t = [r \quad s]'$, $z = [x \quad y]'$, $\Sigma \equiv \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, and $\varphi \equiv \Sigma t = \begin{bmatrix} r+s\rho \\ s+r\rho \end{bmatrix}$. Then

$$\begin{aligned}
rx+sy-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)} =& t'z-\frac{1}{2}z'\Sigma^{-1}z \\
=& \frac{1}{2}\left(z't+t'z-z'\Sigma^{-1}z\right) \\
=& \frac{1}{2}\left(z'\Sigma^{-1}\varphi+\varphi'\Sigma^{-1}z-z'\Sigma^{-1}z\right) \\
=& \frac{1}{2}\left(z'\Sigma^{-1}\varphi-(z-\varphi)'\Sigma^{-1}z\right) \\
=& \frac{1}{2}\left(\varphi'\Sigma^{-1}\varphi+(z-\varphi)'\Sigma^{-1}\varphi-(z-\varphi)'\Sigma^{-1}z\right) \\
=& \frac{1}{2}\left(\varphi'\Sigma^{-1}\varphi-(z-\varphi)'\Sigma^{-1}(z-\varphi)\right) \\
=& \frac{1}{2}t'\Sigma t-\frac{1}{2}(z-\varphi)'\Sigma^{-1}(z-\varphi)
\end{aligned}$$

Note that $t'\Sigma t = r^2+2\rho rs+s^2$. Let $v = z_1-\varphi_1 = x-r-s\rho$ and $w = z_2-\varphi_2 = y-s-r\rho$ so that $dv = dx$ and $dw = dy$. Plugging back the re-written exponent along with this change of variables yields the result:

$$E[e^{rx+sy}\mathbf{1}\{a \leq x \leq b\}\mathbf{1}\{c \leq y \leq d\}] = e^{\frac{1}{2}(r^2+2\rho rs+s^2)}\frac{1}{2\pi|\Sigma|^{1/2}}\int_{a-r-s\rho}^{b-r-s\rho}\int_{c-s-r\rho}^{d-s-r\rho}e^{\frac{v^2-2\rho vw+w^2}{2(1-\rho^2)}}dwdv$$

The special cases are obtained by noting that

1. $\Phi(-\infty,x) = \Phi(x,-\infty) = 0$
2. $\Phi(\infty,x) = \Phi(x,\infty) = \Phi(x)$

■

Part d.

$$E\left[e^{rx}y^2\mathbf{1}\{x\le a\}\right]=\int_{-\infty}^{a}\int_{-\infty}^{\infty}e^{rx}y^2\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}dydx$$

The exponent can be written as

$$\begin{aligned}rx-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}=&rx-\frac{x^2(1-\rho^2)+\rho^2x^2-2\rho xy+y^2}{2(1-\rho^2)}\\=&-\frac{x^2-2rx}{2}-\frac{y^2-2(\rho x)y+(\rho x)^2}{2(1-\rho^2)}\\=&\frac{r^2}{2}-\frac{(x-r)^2}{2}-\frac{(y-\rho x)^2}{2(1-\rho^2)}\end{aligned}$$

Let $z=\frac{y-\rho x}{\sqrt{1-\rho^2}}$, which implies $dz=\frac{dy}{\sqrt{1-\rho^2}}$ and $y^2=z^2(1-\rho^2)+2\rho\sqrt{1-\rho^2}xz+\rho^2x^2$. Substituting the re-written exponent along with the change of variables gives

$$\begin{aligned}E\left[e^{rx}y^2\mathbf{1}\{x\le a\}\right]=&\int_{-\infty}^{a}\int_{-\infty}^{\infty}\left(z^2(1-\rho^2)+2\rho\sqrt{1-\rho^2}xz+\rho^2x^2\right)\frac{1}{2\pi}e^{\frac{r^2}{2}-\frac{(x-r)^2}{2}-\frac{z^2}{2}}dzdx\\=&e^{\frac{r^2}{2}}\int_{-\infty}^{a}e^{-\frac{(x-r)^2}{2}}\frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty}\left(z^2(1-\rho^2)+2\rho\sqrt{1-\rho^2}xz+\rho^2x^2\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz\right)dx\\=&e^{\frac{r^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}e^{-\frac{(x-r)^2}{2}}\left(1-\rho^2+\rho^2x^2\right)dx,\end{aligned}$$

where the last equality used the fact that $z\sim N(0,1)$. Now let $w=x-r$, which implies $dw=dx$ and $x^2=w^2+2wr+r^2$, so that

$$\begin{aligned}E\left[e^{rx}y^2\mathbf{1}\{x\le a\}\right]=&e^{\frac{r^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a-r}e^{-\frac{w^2}{2}}\left(1-\rho^2+\rho^2(w^2+2wr+r^2)\right)dw\\=&e^{\frac{r^2}{2}}\left((1-\rho^2+\rho^2r^2)\Phi(a-r)+\rho^2\int_{-\infty}^{a-r}w^2\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw+2\rho^2r\int_{-\infty}^{a-r}w\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw\right)\\=&e^{\frac{r^2}{2}}\left((1-\rho^2+\rho^2r^2)\Phi(a-r)+\rho^2\int_{-\infty}^{a-r}w^2\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw-2\rho^2re^{-\frac{(a-r)^2}{2}}\right)\end{aligned}$$

The remaining integral can be evaluated using integration by parts.

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a-r}\underbrace{(-w)}_{\equiv u}\underbrace{(-we^{-\frac{w^2}{2}})}_{\equiv dv}dw=-\frac{1}{\sqrt{2\pi}}we^{-\frac{w^2}{2}}\Big|_{-\infty}^{a-r}+\int_{-\infty}^{a-r}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw=\frac{a-r}{\sqrt{2\pi}}e^{-\frac{(a-r)^2}{2}}+\Phi(a-r)$$

Substituting this back in and combining terms gives the result.

■

Part e.

$$E\left[e^{rx}y\mathbf{1}\{x\leq a\}\right]=\int_{-\infty}^{a}\int_{-\infty}^{\infty}e^{rx}y\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}dydx$$

The exponent can be written as

$$\begin{aligned}rx-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}=&rx-\frac{x^2(1-\rho^2)+\rho^2x^2-2\rho xy+y^2}{2(1-\rho^2)}\\=&-\frac{x^2-2rx}{2}-\frac{y^2-2(\rho x)y+(\rho x)^2}{2(1-\rho^2)}\\=&\frac{r^2}{2}-\frac{(x-r)^2}{2}-\frac{(y-\rho x)^2}{2(1-\rho^2)}\end{aligned}$$

Let $z=\frac{y-\rho x}{\sqrt{1-\rho^2}}$, which implies $dz=\frac{dy}{\sqrt{1-\rho^2}}$ and $y=z\sqrt{1-\rho^2}+\rho x$. Substituting the re-written exponent along with the change of variables gives

$$\begin{aligned}E\left[e^{rx}y\mathbf{1}\{x\leq a\}\right]=&\int_{-\infty}^{a}\int_{-\infty}^{\infty}\left(z\sqrt{1-\rho^2}+\rho x\right)\frac{1}{2\pi}e^{\frac{r^2}{2}-\frac{(x-r)^2}{2}-\frac{z^2}{2}}dzdx\\=&e^{\frac{r^2}{2}}\int_{-\infty}^{a}e^{-\frac{(x-r)^2}{2}}\frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty}\left(z\sqrt{1-\rho^2}+\rho x\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz\right)dx\\=&\rho e^{\frac{r^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}xe^{-\frac{(x-r)^2}{2}}dx,\end{aligned}$$

where the last equality used the fact that $z\sim N(0,1)$. Now let $w=x-r$, which implies $dw=dx$, so that

$$\begin{aligned}E\left[e^{rx}y\mathbf{1}\{x\leq a\}\right]=&\rho e^{\frac{r^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a-r}(r+w)e^{-\frac{w^2}{2}}dw\\=&\rho e^{\frac{r^2}{2}}\left(r\Phi(a-r)+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a-r}we^{-\frac{w^2}{2}}dw\right)\\=&\rho e^{\frac{r^2}{2}}\left(r\Phi(a-r)-\frac{1}{\sqrt{2\pi}}e^{-\frac{(a-r)^2}{2}}\right)\end{aligned}$$

■