Generalized Disappointment Aversion, Learning, and Asset Prices

Mykola Babiak *
CERGE-EI **

January 2018

Abstract

This paper provides a generalized disappointment aversion (GDA) interpretation of the variance and skew risk premia in equity returns and the volatility skew in index option prices. The key ingredients are Bayesian learning about the consumption growth rate and the investor’s tail aversion induced by GDA preferences which amplify the impact of consumption shocks. This model with disappointment risk reproduces salient properties of the variance and skew risk premia and generates a realistic volatility skew implied by index options, while simultaneously matching the mean and volatility of risk-free rate and equity returns, and the level of the price-dividend ratio.

Keywords: Equity Premium, Variance and Skew Risk Premia, Volatility Skew, Generalized Disappointment Aversion, Learning, Markov Switching

JEL: D81, E32, E44, G12

I am grateful to Roman Kozhan for numerous comments that significantly improved the paper. For helpful comments, I also thank Jaroslav Borovicka, Aleš Černý, Michael Johannes, Keneth L. Judd, Marek Kapicka, Michal Kejak, Michal Pakoš, Petar Sabtchevsky (discussant), Veronika Selezneva, Ctirad Slavik, Sergey Slobodyan and Ansgar Walther, as well as seminar participants at the 2016 EEA-ESEM Meeting, the 2016 Meeting of the Society for Computational Economics, the 2016 Zurich Initiative for Computational Economics, the 2016 Annual Conference of the Swiss Society for Financial Market Research, Columbia Business School, and Warwick Business School. Financial support from the Charles University Grant Agency, GAUK 151016, and the Czech Science Foundation project No. P402/12/G097 (DYME Dynamic Models in Economics) is gratefully acknowledged.

*Email: mykola.babiak@cerge-ei.cz

**CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic.
1. Introduction

U.S. financial data have provided many stylized facts that are difficult to reconcile in the standard rational framework. The equity market, in particular, is a source of several puzzling features of asset returns, including a high equity premium (Mehra and Prescott, 1985), a low risk-free rate (Weil, 1990), and high excess stock market volatility (Shiller, 1981). Based on this evidence, recent theories have revived the consumption-based model by introducing rare disasters, habits, or long-run risks to consumption growth.\footnote{See, for example, Rietz (1988), Campbell and Cochrane (1999), Bansal and Yaron (2004), Menzly, Santos and Veronesi (2004), Barro (2006), Bansal, Kiku and Yaron (2010).} Although these leading theories have proved to be successful in explaining the salient features of equity returns, a number of important facts from the derivatives market remain unexplained. For example, a growing body of the empirical literature documents the puzzling variance and skew risk premia in equity returns (Bollerslev, Tauchen and Zhou, 2009; Kozhan, Neuberger and Schneider, 2013). Additionally, the volatility skew, the pattern in the volatility curves implied by equity index options, has been a well-known feature since the 1987 stock market crash (Rubinstein, 1994; Jackwerth and Rubinstein, 1996). The risk premia associated with option prices and higher moments of equity returns are difficult to reproduce in the traditional consumption-based models.\footnote{Backus, Chernov and Martin (2011) note that the distribution of rare disasters in macroeconomic data is inconsistent with option prices. Du (2010) finds that the habit formation model cannot fit the observed volatility smirk. Drechsler and Yaron (2011); Drechsler (2013) and Shaliastovich (2015) show that the long-run risk model cannot reproduce both the variance premium and implied volatility skew in option prices.} Motivated by these shortcomings, the main goal of this paper is to explain the derivatives-related puzzles while simultaneously capturing salient features of equity returns.

In this paper, I construct an equilibrium representative agent model with generalized disappointment aversion (GDA) preferences and time-varying economic uncertainty induced by learning about an unobservable state of the consumption growth process. In the model, consumption growth follows a hidden two-state Markov chain and hence the agent filters a posterior belief about the hidden regime based on past consumption growth realizations. The negative consumption growth innovations suggest that, accord-
ing to Bayes’ rule, the posterior belief partially falls and so does the equity price. The combination of lower posterior beliefs and negative consumption growth rates raises the marginal utility. Crucially, the GDA preferences penalize continuation utilities below a scaled certainty equivalent and hence amplify the countercyclical dynamics of the pricing kernel, which make the agent particularly dislike stock market declines. This generates large premiums embedded in the equity returns and index option prices. Specifically, I show that this mechanism can reproduce a number of stylized facts observed in the equity and derivatives markets, including (i) a high equity premium and volatility of equity returns (ii) a low and stable risk-free rate (iii) a large and volatile variance premium (iv) predictability of excess returns by the variance premium (v) a large and volatile skew premium (vi) high prices of at-the-money (ATM) and out-of-the-money (OTM) European put options. Consequently, the results provide the evidence that the investor’s asymmetric preferences and incomplete information about fundamentals are important sources of equity return and option premiums. These theoretical results are consistent with three empirical observations:

The high prices of OTM put options can be attributed to the investor’s desire to hedge market declines (Bates, 2003). The investors are willing to pay a high premium for deep OTM put options, as they will provide the owner with a high payoff conditional on the occurrence of a market crash.

The increased realized variance volatility is associated with large stock market declines (French, Schwert and Stambaugh, 1987; Glosten, Jagannathan and Runkle, 1993). Therefore, the price of a variance swap contract, which pays the holder the realized return variance, can be thought of as an insurance premium for the market downside moves, similarly to the put option prices.

Kozhan, Neuberger and Schneider (2013) show that the skew and variance risk premia are manifestations of and compensate for the same risk. Consequently, the profits of a skew swap similarly embed an insurance premium for the downside risk.
In line with the empirical evidence above, the representative agent in the economy is willing to pay high prices for options, variance, and skew swaps in order to hedge against disappointment consumption shocks. In the calibrated model, a disappointment aversion coefficient of 2.33 implies that the investor penalizes disappointment events approximately 3.33 times more than other outcomes. Consequently, the representative agent’s desire to compensate for disappointment risks generates a steep implied volatility surface, sizable variance, and skew risk premia. In contrast to existing frameworks employing a number of risks, I propose a parsimonious framework with a single state variable, a time-varying posterior state belief, and investor’s generalized disappointment aversion to capture numerous salient features of the equity and derivatives markets. This framework is the first, to my knowledge, to jointly explain the equity returns and risk premiums of options, variance, and skew swaps.

Despite a rather simple structure, the asset pricing implications of the benchmark with GDA preferences are striking. Generalized disappointment aversion brings the benchmark model in line with stylized facts of the risk-free rate, excess equity returns, and the price-dividend ratio. The baseline calibration also reconciles salient moments of the variance and skew risk premia. Learning about a hidden state in the consumption growth dynamics produces the predictive power of the variance premium for the log excess returns with 1-, 3-, and 6-month horizons. Additionally, the model-based implied volatility curves closely match the empirical counterparts. The key role of generalized disappointment aversion can be seen when one looks at alternative preferences specifications: disappointment aversion preferences and Epstein-Zin utility.

Specifically, the framework with disappointment aversion predicts basically flat implied volatilities that are roughly equal to the realized equity return volatility of around 16%. The variance and skew risk premia statistics are also inconsistent with the data. Increasing risk aversion in the Epstein-Zin economy improves the results compared to a pure disappointment aversion model, though the mean variance and skew risk premia appear in about half of the baseline results. In addition, the high risk aversion alone cannot match the observed volatility skew in option prices. In a thorough comparative analysis,
I further show that these results are robust for different calibrations of key parameter values in three model specifications.

One of this article’s main contributions is to provide evidence that agent’s generalized disappointment aversion risk attitude is a source of considerable premiums embedded in equity returns and index option prices. This mechanism is a key differentiator from extant approaches, which mainly introduce volatility, jump, and/or persistence risks in the endowments dynamics in order to generate large equity and option premiums. Whilst in this paper I also introduce a rare bad state, the consumption decline in this regime is calibrated to the observed average consumption drop in the US historical data during the Great Depression and is certainly less severe than in Barro (2006). Furthermore, the derivative-related results are not driven by physical realizations of the ”depression” state, as I calculate derivative prices based on simulations without rare recessions consistent with the historical data over last three decades. Consequently, the model-based implied volatility skew, variance and skew risk premia are manifestations of endogenously fluctuating economic uncertainty induced by Bayesian learning and further amplified by investor’s disappointment aversion.

The other major contribution of this paper is that it jointly reconciles salient features of the implied volatility skew and higher-moment risk premia in equity returns. Drechsler (2013) is another good example in this respect as his model can explain the variance premium and option prices, though the skew premium is not analyzed in his paper. Drechsler (2013) also introduces a large number of risks in endowment dynamics (these include persistence shocks, stochastic volatility, as well as infrequent jumps in the mean growth and volatility of endowments) in order to match a wide range of salient pricing moments. In contrast, these features of the data endogenously arise in my model with a single state variable, a time-varying posterior state belief. Although the model operates through this single channel, the time-varying posterior belief generates rich conditional dynamics of the economy and hence the model is able to simultaneously reproduce numerous asset pricing phenomena. Finally, this paper points out the importance of the investor’s generalized disappointment aversion for asset prices, while Drechsler (2013)
This paper contributes to several strands of the literature. It is closely related to the growing literature studying the asset pricing implications of asymmetric preferences. A number of studies incorporate asymmetric preferences into the standard asset pricing frameworks (Campbell and Cochrane, 1999; Barberis and Huang, 2001; Barberis et al., 2001; Routledge and Zin, 2010) and find that these models better explain the asset returns compared to those with symmetric preferences. In the context of generalized disappointment aversion, Bonomo, Garcia, Meddahi and Tedongap (2011) construct a consumption-based asset pricing model with GDA preferences and long-run volatility risks to explain the equity premium. Bonomo, Garcia, Meddahi and Tedongap (2015) recalibrate their model at the daily frequency to reproduce the risk-return trade-off at high- and low-frequency. Additionally, Liu and Miao (2014) focus on the production-based asset pricing with GDA preferences. Augustin and Tedongap (2016) further shed light on the role of GDA preferences in explaining sovereign credit spreads. Recently, Delikouras (2017) employ disappointment aversion to explain the cross-section of expected returns. This paper contributes to the existing literature by exploring additional asset pricing implications of generalized disappointment aversion for option prices and higher-moment risk premia in equity returns, which are not analyzed by the aforementioned studies.

Furthermore, this paper is related to leading asset pricing theories advocating habit formation, rare disasters, and long-run risks in consumption for explaining the equity returns and option prices. In the context of habits, Du (2010) shows that the extension of the model with habit formation (Campbell and Cochrane, 1999) to include rare disasters can explain the observed implied-volatility skew and further reproduce state-dependent smirk patterns in the data. Under the rare disasters umbrella, the implied volatility surface can be explained with extensions to model uncertainty about rare events (Liu, Pan and Wang, 2005), rare jumps in persistence (Benzoni, Collin-Dufresne and Goldstein, 2011), or stochastic volatility of disasters (Seo and Wachter, 2017). The long-run risks literature generalizes the model of Bansal and Yaron (2004) by introducing jump risks.
(Eraker and Shaliastovich, 2008; Shaliastovich, 2015) to explain the high premium embedded in option prices. Additionally, a few papers can explain the variance premium in equilibrium. These include the long-run risks models with transient non-Gaussian shocks to fundamentals (Bollerslev, Tauchen and Zhou, 2009; Drechsler and Yaron, 2011) and multiple volatility risks (Zhou and Zhu, 2014). The mechanism in this paper is distinct from these papers since it provides a generalized disappointment aversion interpretation of risk premiums in options, variance, and skew swaps. Furthermore, the approach here jointly analyzes index option prices and higher-moment risk premia in equity returns.

The remainder of the paper is organized as follows: Section 2. describes the economy. Section 3. derives asset prices inside the model. Section 4. discusses the data. Section 5. provides asset pricing results of the benchmark model and two other specifications. Section 6. concludes. Appendix A. contains technical details of the representative agent’s maximization problem, Appendix B. outlines the application of the projection method.

2. Model Setup

2.1 Generalized Disappointment Aversion Risk Preferences

The environment is an infinite-horizon, discrete-time exchange economy with a representative agent extracting utility from a consumption stream. Following the recursive utility framework of Epstein and Zin (1989, 1991), the agent’s utility $V_t$ in period $t$ is defined as:

$$V_t = \left[ (1 - \beta)C_t^\rho + \beta \mu_t \right]^{1/\rho},$$

where $C_t$ is consumption at time $t$, $0 < \beta < 1$ is the subjective discount factor, $\frac{1}{1-\rho} > 0$ is the intertemporal elasticity of substitution (IES), and $\mu_t = \mu_t(V_{t+1})$ is the certainty equivalent of random future utility using the $t$-period conditional probability distribution.

The certainty equivalent captures the generalized disappointment aversion (GDA) risk attitude as defined by Routledge and Zin (2010). These risk preferences allocate more weight on the tail events compared to the expected utility. In the Routledge and Zin (2010) model, the representative agent perceives some outcomes as "disappointing" similarly to the disappointment aversion preferences of Gul (1991). For the Gul (1991)
model, an outcome is viewed as disappointing when it is below the certainty equivalent, whereas for the Routledge and Zin (2010) generalized disappointment aversion specification this outcome should be below some fraction of the implicit certainty equivalent. Formally, the certainty equivalent $\mu_t(V_{t+1})$ of GDA risk preferences is defined as:

$$\left[\frac{\mu_t(V_{t+1})}{\alpha}\right] = \mathbb{E}_t\left[\frac{V_{t+1}^\alpha}{\alpha}\right] - \theta \mathbb{E}_t\left[\mathbb{I}\left(\frac{V_{t+1}}{\mu_t(V_{t+1})} \leq \delta\right)\left(\frac{[\delta\mu_t(V_{t+1})]^{\alpha}}{\alpha} - \frac{V_{t+1}^\alpha}{\alpha}\right)\right]$$  (2)

or equivalently:

$$\mu_t(V_{t+1}) = \left(\mathbb{E}_t\left[\frac{V_{t+1}^\alpha}{\alpha} \cdot \frac{1 + \theta\mathbb{I}(V_{t+1} \leq \delta\mu_t(V_{t+1}))}{1 + \theta\delta\alpha\mathbb{E}_t[\mathbb{I}(V_{t+1} \leq \delta\mu_t(V_{t+1}))]}\right]\right)^{1/\alpha},$$

where $\mathbb{I}(\cdot)$ denotes the indicator function, $1 - \alpha > 0$ is the relative risk aversion, $\delta \in (0, 1]$ and $\theta \geq 0$ represent a disappointment threshold and a disappointment penalty, respectively. The GDA risk preferences enable to control for a disappointment threshold by changing $\delta$. In this case, the outcome $V_{t+1}$ is considered to be disappointing only when it is below the scaled certainty equivalent $\delta\mu_t(V_{t+1})$.

The Routledge and Zin (2010) preferences defined by (1) and (2) nest two preference specifications. The expected utility of Epstein and Zin (1989, 1991) can be obtained by setting $\theta = 0$, in which case the certainty equivalent $\mu_t(V_{t+1})$ simplifies to $\left(\mathbb{E}_t[V_{t+1}^{\alpha}]\right)^{1/\alpha}$.

Assuming $\theta \neq 0$ and $\delta = 1$, GDA preferences reduce to the Gul (1991) disappointment aversion utility.

2.2 Endowments and Inference Problem

A popular paradigm in the asset pricing literature is the application of a regime switching framework for modeling aggregate consumption growth. Since Hamilton (1989) and Mehra and Prescott (1985), researchers have used these models to embed business cycle fluctuations in the mean growth rates and volatility of consumption growth (Cecchetti, Lam and Mark, 1990; Veronesi, 1999; Ju and Miao, 2012; Johannes, Lochstoer and Mou, 2016; Collin-Dufresne, Johannes and Lochstoer, 2016). By changing the number of states and parameters controlling the persistence and conditional distribution of regimes, these models can also embed ‘peso problem’ in the growth rate (Rietz, 1988; Barro, 2006; Backus, Chernov and Martin, 2011; Gabaix, 2012) or persistence (Gillman, Kejak and Pakos, 2015) of consumption. Additionally, a proper calibration of a regime switching model can match the dynamics of long-run risks in consumption and dividend growth as studied in Bonomo, Garcia, Meddahi and Tedongap (2011, 2015).
in the asset pricing literature and subject log consumption growth to hidden regime switches:

\[ \Delta c_{t+1} = \mu_{s_{t+1}} + \sigma \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, 1). \]

The consumption volatility \( \sigma \) is assumed to be constant, whereas the mean growth rate \( \mu_{s_{t+1}} \) is driven by a two-state Markov-switching process \( s_{t+1} \) with the state space

\[ S = \{1 = \text{expansion}, 2 = \text{recession}\}, \]

and a transition matrix

\[
P = \begin{pmatrix}
\pi_{11} & 1 - \pi_{11} \\
1 - \pi_{22} & \pi_{22}
\end{pmatrix},
\]

where \( \pi_{ii} \in (0, 1) \) are transition probabilities. I further assume \( \mu_1 > \mu_2 \) in order to identify states \( s_{t+1} = 1 \) and \( s_{t+1} = 2 \) as expansion and recession, respectively.

I now specify a dividend stream of the equity. There are several approaches in the literature to model dividends. A standard Lucas-type model (Lucas Jr., 1973) implies dividends and consumption are the same in the equilibrium. However, dividends are more volatile than consumption in the data. I follow Bansal and Yaron (2004) and model consumption and equity dividends separately. In the endowment economy, it is commonly assumed that aggregate consumption is generated by several stochastic endowments, including dividends as one of them, and is the sum of all these endowments in the equilibrium. Similarly to Campbell (1996), one can interpret other endowments as labor income.

I seek to price the equity (a levered consumption claim) with monthly log dividend growth defined as:

\[ \Delta d_{t+1} = g_d + \lambda \Delta c_{t+1} + \sigma_d e_{t+1}, \quad (3) \]

where \( e_{t+1} \sim N(0, 1) \) is the idiosyncratic shock of dividend growth, \( \lambda > 0 \) is the leverage ratio on expected consumption growth. I use a growth rate of dividends \( g_d \) to match the long-run consumption growth, and the volatility of dividends \( \sigma_d \) to match the annual 11.04% dividend growth volatility observed in the data. In addition, the chosen value
of the leverage parameter enables me to match the observed correlation between annual consumption and dividend growth.

The investor knows the true parameters of the model (e.g., \( \pi_{11}, \pi_{22}, \mu_1, \mu_2, \sigma, g_d, \sigma_d, \lambda \)), but does not observe the state \( s_{t+1} \) of the economy. Consequently, he forms a posterior belief about the hidden state \( s_{t+1} \), conditional on the observable history of consumption and dividend growth rates at time \( t \):

\[
F_t = \left\{ (\Delta c_{\tau}, \Delta d_{\tau}) : 0 \leq \tau \leq t \right\}.
\]

The inference problem is to derive the evolution of \( \pi_t = \mathbb{P}(s_{t+1} = 1|F_t) \). The agent holds the initial belief \( \pi_0 \) (the stationary prior) and uses Bayes’ rule to update his belief \( \pi_{t+1} \) as follows:

\[
\pi_{t+1} = \frac{\pi_{11}f(\Delta c_{t+1}|s_{t+1} = 1)\pi_t + (1 - \pi_{22})f(\Delta c_{t+1}|s_{t+1} = 2)(1 - \pi_t)}{f(\Delta c_{t+1}|s_{t+1} = 1)\pi_t + f(\Delta c_{t+1}|s_{t+1} = 1)(1 - \pi_t)},
\]

where

\[
f(\Delta c_{t+1}|s_{t+1} = i) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(\Delta c_{t+1} - \mu_{st+1})^2}{2\sigma^2}}.
\]

I assume that the agent takes into account only consumption for belief updating, while dividends do not provide any additional information.

### 3. Asset Prices

The representative agent consumes \( C_t \) in period \( t \) and invests the remaining wealth in multiple assets. The agent maximizes his utility subject to a budget constraint:

\[
W_{t+1} = (W_t - C_t)R_{\omega,t+1},
\]

where \( R_{\omega,t+1} \) is the return on the total (unobservable) wealth \( W_t \). Additionally, the return \( R_{\omega,t+1} \) satisfies:

\[
R_{\omega,t+1} = \sum_{i=1}^{N} \omega_{i,t}R_{i,t+1} \land \sum_{i=1}^{N} \omega_{i,t} = 1,
\]

where \( \omega_{i,t} \) is the fraction of the \( t \)-period wealth invested in the \( i \)-th asset with gross real return \( R_{i,t+1} \). Whilst Appendix A. provides the solution of the agent’s consumption and
portfolio choice problem for an arbitrary number of assets, this paper is focused on three
asset classes: one period risk-free bonds, equities, and European put options. Bonds pay
zero coupons and act as purely discount real bonds with the realized gross rate of return
$R_{f,t+1}$. Equities entitle the owner to a stochastic amount of dividends in each period with
the realized gross rate of return $R_{e,t+1}$. The $t$-time price of a European put option with
the given maturity time $\tau$ and the strike price $K$ is denoted by $P^o_t(\tau, K)$.

3.1 The Impact of GDA and Learning

In equilibrium, the representative investor makes his consumption and portfolio de-
cisions subject to the endogenous determination of asset prices and markets clearing
conditions. Following Routledge and Zin (2010), it can be shown (see Appendix A.) that
the gross return $R_{i,t+1}$ on the $i$-th traded asset satisfies:

$$\mathbb{E}_t [M_{t+1} R_{i,t+1}] = 1,$$

where the pricing kernel of the economy is defined as:

$$M_{t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{\rho-1} \left( \frac{V_{t+1}}{\mu_t(V_{t+1})} \right)^{\alpha-\rho} \cdot \left( \frac{1 + \theta \mathbb{I}(V_{t+1} \leq \delta \mu_t(V_{t+1}))}{1 + \theta \delta \mathbb{E}_t \mathbb{I}(V_{t+1} \leq \delta \mu_t(V_{t+1}))} \right).$$

To better understand the mechanism that enables the model to match moments of the
data and to generate the dynamic properties of conditional asset prices, it is important to
note the role of generalized disappointment aversion and learning about the unobservable
state of the economy.

First, for the impact of GDA preferences, consider different components of the pricing
kernel. The first part $M_{t+1}^{CRA}$ is the pricing kernel of the time-separable power utility.
The second multiplier $M_{t+1}^{EZ}$ is the adjustment of Epstein-Zin preferences, which allow a
separation between the coefficient of risk aversion and elasticity of intertemporal substi-
tution. The third component $M_{t+1}^{GDA}$ represents the generalized disappointment aversion
adjustment. When the agent’s utility is below a certain fraction of the certainty equiva-
 lent, more weight is attached to the pricing kernel. The generalized disappointment
aversion thus magnifies the countercyclical dynamics of the pricing kernel.
Second, the presence of the disappointment threshold $\delta$, the major difference between DA and GDA frameworks, will allow to control which events will be disappointing. As noted in Backus, Routledge and Zin (2004), ”disappointment aversion implies proportionately greater aversion to small risks than large ones”. This feature of DA preferences is detrimental for generating steep implied volatility curves, large skew and variance risk premia. In contrast, GDA preferences fix this drawback: one can properly choose the disappointment threshold and hence generate endogenous countercyclical risk aversion, a desirable property of the dynamic asset pricing model.

Third, learning is crucial for reproducing conditional properties of asset prices. In the complete information setting, asset prices would fluctuate only upon the realization of a rare recession. Hence, this model would predict flat implied volatility curves, constant equity returns, constant variance and skew risk premia based on the post-war US history when the US economy did not experience a depression event. Learning about the unobservable state of the economy induces an endogenously time-varying belief. Consequently, the variation in the posterior belief will lead to endogenous fluctuations in equity returns, variance and skew risk premia, and option prices.

Finally, it is important to note the interaction between the agent’s generalized disappointment aversion and the posterior state belief fluctuations. I can express the continuation utility using the wealth-consumption ratio (see Appendix B.) and further rewrite the disappointment condition $\frac{V_{t+1}}{\mu(V_{t+1})} \leq \delta$ as:

$$\beta e^{\rho \Delta c_{t+1}} \cdot \frac{W_{t+1}}{C_{t+1}} \leq \delta \cdot \left( \frac{W_t}{C_t} - 1 \right).$$

As the equilibrium wealth-consumption ratio $\frac{W_t}{C_t}$ is a function of the posterior state belief $\pi_t$, the probability of disappointment events will be time-varying and depend on both current belief $\pi_t$ and future belief $\pi_{t+1}$, as well as actual consumption growth realizations $\Delta c_{t+1}$.

3.2 Risk-free Rate and Equity Returns

I solve the model numerically due to the lack of an analytical solution for equilibrium returns. I first need to solve for the return on the wealth portfolio $R_{t+1}^\omega$ (the return on
the aggregate consumption claim) and then the equity return \( R_{e,t+1} \) (the return on the aggregate dividend claim), which are implicitly defined by the equation (5). Denoting the equity price by \( P^e_t \), the returns on the wealth portfolio and equity can be rewritten as:

\[
R^{\omega}_{t+1} = \frac{W_{t+1}}{W_t - C_t} = \frac{\frac{W_{t+1}}{C_{t+1}}}{\frac{W_t}{C_t} - 1} \cdot e^{\Delta c_{t+1}} \quad \land \quad R^e_{t+1} = \frac{P^e_{t+1} + D_{t+1}}{P^e_t} = \frac{P^e_{t+1}}{D_t} + 1 \cdot e^{\Delta d_{t+1}}.
\]

I conjecture that the wealth-consumption ratio \( \frac{W_t}{C_t} = G(\pi_t) \) and the price-dividend ratio \( \frac{P^e_t}{D_t} = H(\pi_t) \) are functions of the state belief \( \pi_t \). Substituting \( R^{\omega}_{t+1} \) and \( R^e_{t+1} \) into (5), I apply the projection method (Judd, 1992) to approximate \( G(\pi_t) \) and \( H(\pi_t) \) by a basis of complete Chebyshev polynomials. The details of the numerical solution algorithm are provided in Appendix B. Having solved for wealth-consumption and price-dividend ratios, I can simulate asset pricing moments associated with the one-period risk-free rate, equity returns and price-dividend ratio. Furthermore, I can numerically calculate the pricing kernel and all asset prices, including implied volatilities and quantities in the skew and variance risk premia.

### 3.3 The Variance and Skew Risk Premia

In this paper, I focus on the monthly variance and skew risk premia associated with equity returns. The \( t \)-time monthly variance premium \( v p_t \) is defined as the difference between risk-neutral and physical expectations of the total return variance between time \( t \) and \( t + 1 \). As in Drechsler and Yaron (2011), the variance premium equals:

\[
v p_t = \mathbb{E}_Q^t(\text{var}_{t+1}^Q(r_{e,t+2})) - \mathbb{E}_P^t(\text{var}_{t+1}^P(r_{e,t+2})),
\]  

where \( \text{var}_{t+1}^Q(r_{e,t+2}) \) and \( \text{var}_{t+1}^P(r_{e,t+2}) \) are \((t + 1)\)-period conditional variances of the log return \( r_{e,t+2} = \ln(R_{e,t+2}) \) under the risk-neutral \( Q \) and physical \( P \) probability measures, respectively. As noted in Drechsler and Yaron (2011), the variance premium is decomposed of two components called the level difference and drift difference. The level difference, defined as

\[
\text{var}_{t+1}^Q(r_{e,t+1}) - \text{var}_{t+1}^P(r_{e,t+1}),
\]
reflects the difference in the conditional return variance under the risk-neutral and physical measures. The drift difference, defined as

$$\left[\mathbb{E}^Q_t(var^Q_t(r_{e,t+1})) - var^Q_t(r_{e,t+1})\right] - \left[\left(\mathbb{E}^P_t(var^P_t(r_{e,t+1})) - var^P_t(r_{e,t+1})\right)\right],$$

incorporates the difference in the expected change of $var_{t+1}(r_{e,t+2})$ under the measures $Q$ and $P$.

The $t$-time monthly skew premium is defined as the expected payoff of the skew swap, a contract paying the difference between the implied skew and the realized skew of the index return between time $t$ and $t+1$ (Kozhan, Neuberger and Schneider, 2013). The monthly implied and realized skews simply equal the risk-neutral and physical expectations of the index return skewness denoted by $\mathbb{E}^Q_t(skew^Q_t(r_{e,t+2}))$ and $\mathbb{E}^P_t(skew^P_t(r_{e,t+2}))$, respectively. As in Kozhan, Neuberger and Schneider (2013), I further express the skew risk premium as a percentage of the implied skew that results in the definition of the skew risk premium as stated below:

$$sk_t = \frac{\mathbb{E}^P_t(skew^P_t(r_{e,t+2}))}{\mathbb{E}^Q_t(skew^Q_t(r_{e,t+2}))} - 1.$$

The quantity $sk_t$ reflects the dollar amount of profit per $1 investment in the implied skew.

### 3.4 Implied Volatilities

I now describe how I compute model-based option prices and solve for their Black-Scholes implied volatilities. Consider a European put option written on the price of the equity that is traded in the economy. Note that the equity price should not include dividend payments; that is, options are written on the ex-dividend stock price index. Using the Euler condition (5), the relative price $\mathcal{O}_t(\pi_t, \tau, K) = \frac{P^P_t(\pi_t, \tau, K)}{P^P_t(\pi_t)}$ of the $\tau$-period European put option with the strike price $K$, expressed as a ratio to the initial price of the equity $P^P_t$, should satisfy:

$$\mathcal{O}_t(\pi_t, \tau, K) = \mathbb{E}_t \left[ \prod_{k=1}^{\tau} M_{t+k} \cdot \max \left( K - \frac{P^P_{t+\tau}}{P^P_t}, 0 \right) \right].$$  

(8)
It is worth noting that a put price $P_t^p$ depends on the equity price $P_t^e$, whereas the normalized price $O_t$ does not. One can express the ratio $\frac{P_{t+\tau}^e}{P_t^e}$ in terms of the dividend growth rates and price-dividend ratios on the equity and hence the state belief $\pi_t$ provides sufficient information for the calculation of the option prices. Specifically, I compute model-based European put prices $O_t = O_t(\pi_t, \tau, K)$ via Monte Carlo simulations. I convert them into Black-Scholes implied volatilities with properly annualized continuous interest rate $r_t = r_t(\pi_t)$ and dividend yield $q_t = q_t(\pi_t)$. Thus, given the time to maturity $\tau$, the strike price $K$, the risk-free rate $r_t$, and dividend yield $q_t$, the implied volatility $\sigma_t^{imp} = \sigma_t^{imp}(\pi_t, \tau, K)$ solves the equation:

$$O_t = e^{-r_t \tau} \cdot K \cdot N(-d_2) - e^{-q_t \tau} \cdot N(-d_1),$$

$$d_{1,2} = \ln\left(\frac{1}{K}\right) + \tau \left(r_t - q_t \pm \frac{(\sigma_t^{imp})^2}{2}\right) \sqrt{\tau}.$$  

4. Data

4.1 Consumption, Dividends, and Market Returns

I follow Bansal and Yaron (2004) and construct the real per-capita consumption growth series (annual due to the frequency restriction) for the longest sample available 1930-2016. In the literature, consumption is defined as the sum of personal consumption expenditures on nondurable goods and services. I download the data from the U.S. National Income and Product Accounts (NIPA) as provided by the Bureau of Economic Analysis. I first apply the seasonally adjusted annual quantity indexes from Table 2.3.3. (Real Personal Consumption Expenditures by Major Type of Product, Quantity Indexes, A:1929-2016) to the corresponding series from Table 2.3.6. (Real Personal Consumption Expenditures by Major Type of Product, Chained Dollars, A:1995-2016) to obtain real personal consumption expenditures on nondurable goods and services for the sample period 1929-2016. I further retrieve mid-month population data from NIPA Table 7.1. to convert both real consumption series to per capita terms.

I measure the total market return as the value-weighted return including dividends, and the dividends as the sum of total dividends, on all stocks traded on the NYSE, AMEX,
and NASDAQ. The dividends and value-weighted market return data are monthly and are retrieved from the Center for Research in Security Prices (CRSP). To construct the monthly nominal dividend series, I use the CRSP value-weighted returns including and excluding dividends of CRSP common stock market indexes (NYSE/AMEX/NASDAQ/ARCA), denoted by $RI_t$ and $RE_t$, respectively. Following Hodrick (1992), I construct the price series $P_t$ by initializing $P_0 = 1$ and iterating recursively $P_t = (1 + RI_t)P_{t-1}$. Next, I compute normalized nominal monthly dividends $D_t = (RI_t - RE_t)P_t$. The proxy of the risk-free return $R_{f,t+1}$ is the 1-month nominal Treasury bill. The nominal annualized dividends are constructed by summing the corresponding monthly dividends within the year. Finally, I retrieve the inflation index from CRSP to deflate all quantities to real values.

### 4.2 The Variance and Skew Risk Premia Data

I define risk premia associated with higher moments of equity returns consistent with the existing literature. For the variance risk premium, I closely follow Bollerslev, Tauchen and Zhou (2009), Bollerslev, Gibson and Zhou (2011), Drechsler and Yaron (2011) and Drechsler (2013), while the empirical strategy and key definitions of the skew risk premia are in line with Bakshi, Kapadia and Madan (2003) and Kozhan, Neuberger and Schneider (2013).

The variance premium is a phenomenon on the variance swap market that can be defined as a difference between expectations of return variance under the risk-neutral $Q$ and actual physical (i.e., true/statistical) $P$ probability measures for a given horizon. The focus of this paper is on the one-month variance premium defined as:

$$vp_t = E_Q^{t}[ \text{Return Variation}(t, t+1) ] - E_P^{t}[ \text{Return Variation}(t, t+1) ].$$

Under the no-arbitrage assumption, the risk-neutral conditional expectation of the return variance is equal to the price of a variance swap, a forward contract on the realized variance of the asset. Britten-Jones and Neuberger (2000) show that a continuous price process under the risk-neutral expectation of the underlying’s variance can be implied from prices of European calls on that asset. Their result follows from the fact that
the payoff of a variance swap can be replicated by a portfolio of options. Furthermore, Britten-Jones and Neuberger (2000) follow the so called "model-free" approach, meaning that their result is not based on any particular option pricing model (Jiang and Tian, 2005). Since the Chicago Board of Options Exchange (CBOE) calculates the VIX index as a measure of the 30-days ahead risk-neutral expectation of the variance of the S&P 500 index, I use the VIX index as a proxy for the risk-neutral expectation of the market’s return variation. The VIX is quoted in annualized standard deviation. Hence, I first take it to a second power to transform to variance units and then divide by 12 to obtain monthly frequency. Thus, I obtain a new series defined as $[\text{VIX}]^2_t = \frac{\text{VIX}_t^2}{12}$. I further use the last available observation of $[\text{VIX}]^2_t$ in a particular month as a measure of the risk-neutral expectation of return variance in that month.

For the objective expectation of return variance, a second component in the variance premium, I calculate a one-step-ahead forecast from a simple regression similar to Drechsler (2013). I first calculate the measure of the realized variance by summing the squared daily log returns on the S&P 500 futures and S&P 500 index obtained from the CBOE. The constructed series are denoted by FUT$_t^2$ and IND$_t^2$, respectively. Subsequently, I estimate the following regression:

$$FUT_{t+1}^2 = \beta_0 + \beta_1 \cdot IND_t^2 + \beta_2 \cdot [\text{VIX}]_t^2 + \epsilon_{t+1}. \quad (10)$$

The actual statistical expectation is measured by the one-period ahead forecast given by (10). I refer to the resulting series as the realized variance and denote it by RV$_t$. Theoretically, the variance premium should be non-negative in each period. Thus, I truncate the difference between the implied series of $[\text{VIX}]_t^2$ and RV$_t$ from below by 0.

For the empirical strategy above, I obtain the data series of the VIX index, S&P 500 index futures, and S&P 500 index from the CBOE. The main restriction on the length of the constructed monthly variance premium is the VIX index, reported by the CBOE only from January 1990. Using high-frequency data would provide a finer estimation precision of the realized variance, but the empirical statistics of the variance premium remain largely consistent with the existing estimates in the literature (Drechsler and
Yaron, 2011; Drechsler, 2013).

The skew risk premium can be considered as a payoff of a skew swap, which pays the holder the difference between the implied skew and realized skew of the index return. Kozhan, Neuberger and Schneider (2013) show that a skew swap contract can be replicated by a trading strategy, which involves long positions in OTM calls and short positions in OTM puts. As in Kozhan, Neuberger and Schneider (2013), I focus on one-month skew risk premium defined as an average profit from a skew swap. I further express this profit as a percentage of the implied skew. Formally, the definition of the skew risk premium reads as follows:

\[
sk_t = \frac{E_P^t \left[ \text{Return Skewness}(t, t+1) \right]}{E_Q^t \left[ \text{Return Skewness}(t, t+1) \right]} - 1.
\]

I use empirical estimates of the realized and implied skew from Kozhan, Neuberger and Schneider (2013)\(^4\). In their empirical strategy, the authors use European options written on the S&P 500 index and traded on the CSOB. The options dataset used to construct the skew risk premia is obtained from OptionMetrics and covers the period from January 1996 to January 2012. Further details about the empirical strategy of the skew risk premium can be found in Kozhan, Neuberger and Schneider (2013).

4.3 Option Prices

For the empirical implied volatility curves, I use European options written on the S&P 500 index and traded on the CBOE. The option data set covers the period from January 1996 to December 2016 and is from OptionMetrics. Option data elements include the type of options (call/put) along with the contract’s variables (strike price, time to expiration, Greeks, Black-Scholes implied volatilities, closing spot prices of the underlying) and trading statistics (volume, open interest, closing bid and ask quotes), among other details.

To construct the empirical implied volatility curves, I first compute the moneyness for each observed option using the daily S&P 500 index on a particular trading day. I filter out all data entries with non-standard settlements. I use the remaining observations

\(^4\)I would like to thank Roman Kozhan for providing the skew-risk-premium-related series.
to construct the implied volatility surface for a range of moneyness and maturities. In particular, I follow Christoffersen and Jacobs (2004) and perform polynomial extrapolation of volatilities in the maturity time and strike prices. This strategy makes use of all available options and not only those with a specific maturity time. The fitted values are further used to construct the implied volatility curves for 1-, 3-, and 6-month maturities.

5. Calibration and Quantitative Results

In this section, I first calibrate the cash-flow processes for consumption and dividend growth. The chosen parameters are consistent with the historical US data from January 1930 to December 2016. The consumption-based asset pricing economy in this paper has a key ingredient, generalized disappointment aversion. Therefore, I consider three specifications of preference parameters for the comparative statistics exercise: the benchmark model (GDA) with generalized disappointment preferences, a pure disappointment aversion economy (DA) with linear preferences and infinite elasticity of intertemporal substitution, and an Epstein-Zin framework (EZ). The comparison between GDA and DA isolates the contribution of disappointment aversion, while the comparison between GDA and EZ illustrates the impact of the representative agent’s preference for early resolution of uncertainty.

As the model does not admit an analytical solution, I solve for equilibrium pricing ratios using the projection method (Judd, 1992) with Chebyshev interpolation. Having solved the model, I generate 20,000 simulations of the economy and report annualized statistics of cash-flows and asset prices corresponding to their empirical counterparts. Specifically, for the returns data, consumption and dividend growth rates, I report the annualized moments based on the simulations with 1044 monthly observations, consistent with the data sample spanning the period from January 1930 to December 2016. The monthly variance premium statistics are obtained based on the simulations with 324 monthly observations, in line with the constructed variance premium series, in particular, the historical VIX index reported by CBOE from January 1990. The sample monthly skew risk premium statistics are calculated based on the simulations with 193 monthly
observations, consistent with the empirical series covering the period from January 1996 to January 2012. The model-based implied volatility curves for 1-, 3-, and 6-month maturities are unconditional averages of implied volatilities based on the simulations with a sample length of 252 months, corresponding to the empirical option prices data set from January 1996 to December 2016.

5.1 Calibrated Parameters

I begin with the parameters of a regime-switching process for aggregate consumption growth \( (\pi_{11}, \pi_{22}, \mu_1, \mu_2, \sigma) \). As in Bansal and Yaron (2004), I make the model’s time-averaged consumption statistics consistent with observed annual log consumption growth from 1930 to 2016. I calibrate a two-state regime-switching model of monthly consumption growth with the recession state mimicking large declines like the Great Depression and the expansion state reflecting normal business cycle fluctuations. Therefore, I set \( \pi_{11} = 1151/1152 \) and \( \pi_{22} = 47/48 \). These numbers imply the average duration of the high-growth state of about \( (1 - \pi_{11})^{-1} = 96 \) years and the low-growth state of about \( (1 - \pi_{22})^{-1} = 4 \) years. Furthermore, the unconditional probability of being in expansion \( \pi_{11} = (1 - \pi_{22})/(2 - \pi_{11} - \pi_{22}) \) results in \( \pi_{11} = 0.96 \) and hence the agent experiences one 4-year depression per century consistent with the historical data. For the mean growth rate, consumption tends to grow on average at the annualized rates of about \( \mu_1 \times 12 = 2.08\% \) and \( \mu_2 \times 12 = -4.6\% \) in the expansion and recession states, respectively. The depression state is consistent with an average annual decline in the real, per capita log consumption growth during the Great Depression and is less severe than rare disasters, defined as a drop in annual consumption growth larger than 10 percent (Rietz, 1988; Barro, 2006). I calibrate the consumption volatility \( \sigma \) to match the observed standard deviation 2.22%.

The reason for calibrating the model with two regimes only is twofold. First, I want to retain parsimony for the sake of convenient interpretation of results. Second, I do not introduce additional risks in consumption growth as considered in other papers (more states, non-gaussian shocks in consumption growth, alternative information settings, etc.) in order to isolate and emphasize the impact of learning and GDA risk preferences. Of course, the model with more regimes would lead to richer consumption dynamics. For example, it could introduce long-run risks in consumption as studied in Bonomo, Garcia, Meddahi and Tedongap (2011, 2015). Alternatively, the framework with a multidimensional learning problem (Collin-Dufresne, Johannes and Lochstoer, 2016; Johannes, Lochstoer and Mou, 2016) could
Table 1
Parameter values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta)</td>
<td>Discount factor</td>
<td>0.9989</td>
</tr>
<tr>
<td>(1 - \alpha)</td>
<td>Risk aversion</td>
<td>2.5</td>
</tr>
<tr>
<td>(1/(1 - \rho))</td>
<td>EIS</td>
<td>1.5</td>
</tr>
<tr>
<td>(\theta)</td>
<td>Disappointment aversion</td>
<td>2.33</td>
</tr>
<tr>
<td>(\delta)</td>
<td>Disappointment threshold</td>
<td>0.94</td>
</tr>
<tr>
<td>(\pi_{11})</td>
<td>Transition probability from expansion to expansion</td>
<td>0.9991</td>
</tr>
<tr>
<td>(\pi_{22})</td>
<td>Transition probability from recession to recession</td>
<td>0.9787</td>
</tr>
<tr>
<td>(\mu_1)</td>
<td>Consumption growth in expansion</td>
<td>0.17(3)</td>
</tr>
<tr>
<td>(\mu_2)</td>
<td>Consumption growth in recession</td>
<td>-0.38(3)</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>Consumption volatility</td>
<td>0.7217</td>
</tr>
<tr>
<td>(g_d)</td>
<td>Mean adjustment of dividend growth</td>
<td>-0.2417</td>
</tr>
<tr>
<td>(\sigma_d)</td>
<td>Std. deviation of dividend growth shock</td>
<td>3.28</td>
</tr>
<tr>
<td>(\lambda)</td>
<td>Leverage ratio</td>
<td>2.6</td>
</tr>
</tbody>
</table>

This table reports parameter values in the benchmark model (GDA). All parameters are calibrated at a monthly frequency.

I now turn to calibrating parameters in the dividend process. I regress the annual dividends on the annual consumption covering the period 1930-2016 and find the estimate of the leverage ratio of around 2.5, a conservative number within an interval of plausible values from 1.5 to 4. The leverage ratio is an important parameter for two reasons. First, it controls the volatility of dividends in normal times. Second, it determines the decline of dividends in the recession state. As a result, increasing the leverage would increase the payoffs of put options, conditional on the realization of recession. To compare model performance with the existing literature, particularly the literature on rare disasters with the asset pricing implication for option prices, I set the leverage ratio \(\lambda = 2.6\), corresponding to the value used in Seo and Wachter (2017). I further follow the empirical evidence and set \(g_d\) to equalize the long-run dividend and consumption growth. The standard deviation of the dividend process \(\sigma_d\) is used to match annual dividend volatility 11.04% observed in the data.

I can now choose parameter values for the preferences. In the benchmark model, additionally contribute to the dynamics of the model. Although taking into account all these risks channels would certainly improve the model’s performance, I show that a combination of learning about an unobservable state and the agent’s GDA risk attitude alone can reproduce a wide array of dynamic asset pricing phenomena observed in the equity and derivatives markets.
I would like to keep the GDA parameters in line with existing studies for comparison purposes. In particular, I fix the subjective discount factor $\beta$, the relative risk aversion $1 - \alpha$, the EIS $1/(1 - \rho)$, and the disappointment aversion $\theta$ at the values chosen by Bonomo, Garcia, Meddahi and Tedongap (2011, 2015). I only change the disappointment threshold $\delta$ to match the high equity premium observed in the data. Therefore, it appears that all other relevant moments of returns, variance and skew premia, observed patterns of the excess return predictability, and implied volatility surface are not directly targeted during the model calibration. Surprisingly, these complex features of the data compare well with the model-generated statistics, which endogenously arise in the model due to fluctuations in the posterior state belief. Table 1 summarizes the calibrated values of the benchmark model.

For the DA model, I shut off the generalized disappointment aversion channel by setting $\delta = 1$. Furthermore, the DA specification does not exhibit any curvature in the pricing kernel associated with the relative risk aversion, which is set $1 - \alpha = 0$, or elasticity of intertemporal substitution, which is equal $(1 - \rho)^{-1} = \infty$. I only adjust the disappointment aversion to match the observed equity premium. The remaining parameters are fixed at the benchmark values. For the EZ model, I turn off all (generalized) disappointment aversion by setting $\theta = 0$. In this case, the representative agent only exhibits the preference for early resolution of uncertainty, a popular workhorse in the asset pricing literature. By keeping the risk aversion at the benchmark value $1 - \alpha = 2.5$, the EZ model predicts (although not reported) too small equity and variance risk premium, in addition to too low implied volatility curves. Therefore, I increase the risk aversion to $1 - \alpha = 6$, where the model matches the equity premium observed in the data. Other parameters correspond to those in the benchmark model.

5.2 Asset Pricing Implications

Before discussing the asset pricing implications of GDA, DA, and EZ models, I look at the cash-flow dynamics predicted by a two-state regime switching model. Panel A in Table 2 compares the annualized consumption and dividends moments of the data with
those implied by the calibration in this paper. The model-based median estimates of the mean and volatility of consumption and dividends growth come out close to their empirical counterparts, although mean dividend growth is slightly higher in the simulations. The autocorrelation of cash-flows is also in line with the empirical estimates, although the number for consumption is somewhat smaller than in the data. The choice of the leverage parameter allows me to capture the observed correlation between consumption and dividends. Overall, one can see that a two-state Markov-switching model of consumption and dividend growth matches the key empirical statistics well.

5.2.1 Risk-free Rate and Equity Returns

Panel B in Table 2 reports the key annualized moments of the risk-free rate, equity returns, and price-dividend ratio for three model specifications: GDA, DA, and EZ. Overall, all three models do a good job of accounting for the salient features of the equity returns, as all predict the low risk-free rate, the large equity premium and volatility of excess returns. In addition, the volatility of the risk-free rate and level of the log price-dividend ratio correspond well to the empirical estimates under all specifications, with the benchmark model’s values closer to the data. The main shortcoming of the three models is the low volatility of the log price-dividend ratio.

For brevity, the bottom of Panel B also presents the higher moments of excess returns at the monthly (labeled [M]) and annual (labeled [A]) frequencies. It is well known that the distribution of log excess returns exhibits non-normality at the high frequency and, according to the central limit theorem, approach a normal distribution over longer horizons (because the \( n \)-period log return is the sum \( n \) one-period returns). In line with this evidence, the kurtosis of the monthly log excess returns in our sample is larger than the kurtosis of the corresponding annualized series. The annualized log excess returns are more negatively skewed compared to the monthly data. All three specifications qualitatively respect these features of the data, although quantitatively the GDA model better captures the excess kurtosis in monthly returns with its lower value at the annual frequency.
Table 2
Benchmark Calibration and Sensitivity Analysis: Cash Flows and Stock Market Returns

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>GDA</th>
<th>DA</th>
<th>EZ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>50% 95%</td>
<td>5% 50% 95%</td>
<td>5% 50% 95%</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.9989</td>
<td>0.9989</td>
<td>0.9989</td>
<td></td>
</tr>
<tr>
<td>1 – ( \alpha )</td>
<td>2.5</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>1/(1 – ( \rho ))</td>
<td>1.5</td>
<td>( \infty )</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>2.33</td>
<td>0.55</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.94</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel A: Cash Flows

<p>| | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( E(\Delta c) )</td>
<td>1.83</td>
<td>0.93</td>
<td>1.86</td>
<td>2.42</td>
<td>0.93</td>
<td>1.86</td>
<td>2.42</td>
<td>0.93</td>
<td>1.86</td>
<td>2.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma(\Delta c) )</td>
<td>2.22</td>
<td>1.86</td>
<td>2.24</td>
<td>3.17</td>
<td>1.86</td>
<td>2.24</td>
<td>3.17</td>
<td>1.86</td>
<td>2.24</td>
<td>3.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( acl(\Delta c) )</td>
<td>0.50</td>
<td>0.09</td>
<td>0.30</td>
<td>0.63</td>
<td>0.09</td>
<td>0.30</td>
<td>0.63</td>
<td>0.09</td>
<td>0.30</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E(\Delta d) )</td>
<td>1.44</td>
<td>-1.14</td>
<td>1.87</td>
<td>4.46</td>
<td>-1.14</td>
<td>1.87</td>
<td>4.46</td>
<td>-1.14</td>
<td>1.87</td>
<td>4.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma(\Delta d) )</td>
<td>11.04</td>
<td>9.51</td>
<td>11.06</td>
<td>13.01</td>
<td>9.51</td>
<td>11.06</td>
<td>13.01</td>
<td>9.51</td>
<td>11.06</td>
<td>13.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( acl(\Delta d) )</td>
<td>0.19</td>
<td>0.08</td>
<td>0.26</td>
<td>0.45</td>
<td>0.08</td>
<td>0.26</td>
<td>0.45</td>
<td>0.08</td>
<td>0.26</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( corr(\Delta c, \Delta d) )</td>
<td>0.55</td>
<td>0.37</td>
<td>0.54</td>
<td>0.70</td>
<td>0.37</td>
<td>0.54</td>
<td>0.70</td>
<td>0.37</td>
<td>0.54</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Returns

<p>| | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( E(r_f) )</td>
<td>0.81</td>
<td>0.49</td>
<td>1.34</td>
<td>1.82</td>
<td>1.32</td>
<td>1.32</td>
<td>1.32</td>
<td>0.65</td>
<td>1.44</td>
<td>1.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma(r_f) )</td>
<td>1.87</td>
<td>1.05</td>
<td>1.88</td>
<td>2.71</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.67</td>
<td>1.43</td>
<td>2.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E(r_e – r_f) )</td>
<td>5.22</td>
<td>3.29</td>
<td>5.74</td>
<td>8.06</td>
<td>2.59</td>
<td>5.40</td>
<td>7.93</td>
<td>3.00</td>
<td>5.45</td>
<td>7.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma(r_e – r_f) )</td>
<td>19.77</td>
<td>14.93</td>
<td>18.53</td>
<td>22.57</td>
<td>13.16</td>
<td>16.85</td>
<td>22.52</td>
<td>14.44</td>
<td>18.55</td>
<td>23.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E(pd) )</td>
<td>3.11</td>
<td>2.92</td>
<td>3.00</td>
<td>3.02</td>
<td>2.97</td>
<td>3.05</td>
<td>3.07</td>
<td>2.95</td>
<td>3.04</td>
<td>3.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma(pd) )</td>
<td>0.33</td>
<td>0.03</td>
<td>0.07</td>
<td>0.19</td>
<td>0.01</td>
<td>0.06</td>
<td>0.22</td>
<td>0.03</td>
<td>0.07</td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>skew(( r_e – r_f ))[M]</td>
<td>-0.47</td>
<td>-0.33</td>
<td>-0.08</td>
<td>0.18</td>
<td>-0.53</td>
<td>-0.13</td>
<td>0.23</td>
<td>-0.42</td>
<td>-0.08</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kurt(( r_e – r_f ))[M]</td>
<td>10.06</td>
<td>4.18</td>
<td>5.02</td>
<td>6.25</td>
<td>3.45</td>
<td>6.16</td>
<td>10.03</td>
<td>4.52</td>
<td>6.07</td>
<td>8.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>skew(( r_e – r_f ))[A]</td>
<td>-0.93</td>
<td>-0.73</td>
<td>-0.10</td>
<td>0.52</td>
<td>-1.28</td>
<td>-0.28</td>
<td>0.37</td>
<td>-0.90</td>
<td>-0.15</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kurt(( r_e – r_f ))[A]</td>
<td>4.24</td>
<td>2.75</td>
<td>4.00</td>
<td>6.59</td>
<td>2.60</td>
<td>4.45</td>
<td>9.97</td>
<td>2.78</td>
<td>4.58</td>
<td>8.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel A reports moments of consumption and dividend growth rates denoted by \( \Delta c \) and \( \Delta d \), respectively. Panel B reports the moments of the log risk-free rate \( r_f \), the excess log equity returns \( r_e – r_f \), and the log price-dividend ratio \( pd \). The statistics of macroeconomic quantities and asset returns are for the data and three models: the benchmark model with generalized disappointment aversion preferences GDA, a pure disappointment aversion specification with linear preferences and infinite elasticity of intertemporal substitution DA, and an Epstein-Zin economy EZ. The entries of the table are annualized statistics except for the [M] rows, which provide the higher-order moments of the excess log equity returns sampled at a monthly frequency. The empirical moments are for the U.S. data from January 1930 to December 2016. For each model, I simulate 20,000 economies sampled at a monthly frequency with a sample size equal to the empirical counterpart. The simulation results are percentiles of sample moments based on these 20,000 artificial series. I use the common notations for the average \( E \), standard deviation \( \sigma \), autocorrelation \( acl \), skewness \( skew \), and kurtosis \( kurt \) of the series.

5.2.2 The Variance and Skew Risk Premia

Benchmark Model: GDA. In Table 3, I collect moments of the variance premium and
Table 3
Benchmark Calibration and Sensitivity Analysis: Variance Premium and Predictability

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>GDA 5%</th>
<th>GDA 50%</th>
<th>GDA 95%</th>
<th>DA 5%</th>
<th>DA 50%</th>
<th>DA 95%</th>
<th>EZ 5%</th>
<th>EZ 50%</th>
<th>EZ 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.9989</td>
<td>0.9989</td>
<td>0.9989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1 - \alpha$</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>$1/(1 - \rho)$</td>
<td>1.5</td>
<td>$\infty$</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>$\theta$</td>
<td>2.33</td>
<td>0.55</td>
<td>0.55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.94</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Panel A: Variance Premium

$E(vp)$

$\sigma(vp)$

$skew(vp)$

$kurt(vp)$

$\sigma(var^\mathbb{P}(r_e))$

$acl(var^\mathbb{P}(r_e))$

$\sigma(var^\mathbb{Q}(r_e))$

$acl(var^\mathbb{Q}(r_e))$

$skew(var^\mathbb{Q}(r_e))$

$kurt(var^\mathbb{Q}(r_e))$

Panel B: Predictability

$\beta(1m)$

$R^2(1m)$

$\beta(3m)$

$R^2(3m)$

$\beta(6m)$

$R^2(6m)$

Panel A reports moments of the conditional variance premium $vp$, market return variances $var^\mathbb{P}(r_e)$ and $var^\mathbb{Q}(r_e)$ under the physical $\mathbb{P}$ and risk-neutral $\mathbb{Q}$ probability measures, respectively. The entries of Panel A are monthly statistics. Panel B reports results of the predictive regression of $h$-month future excess log equity returns constructed as $r^e_{t+1} = \sum_{i=1}^{h} (r_{e,t+i} - r_{f,t-1+i})$ on the lagged variance premium $vp_t$. Specifically, the slope estimates $\beta(h)$ and $R^2(h)$ are based on the linear projection:

$$100 \times r^e_{t+1} = \text{Intercept} + \beta(h) \times vp_t + \varepsilon_{t+h},$$

where $h = 1, 3$ and 6 months. The moments and regression outputs are for the data and three models: the benchmark model with generalized disappointment aversion preferences GDA, a pure disappointment aversion specification with linear preferences and infinite elasticity of intertemporal substitution DA, and an Epstein-Zin economy EZ. The empirical statistics are for the U.S. data from January 1990 to December 2016. For each model, I simulate 20,000 economies at a monthly frequency with a sample size equal to the empirical counterpart. I obtain moments, the slope coefficients $\beta(h)$ and $R^2(h)$ for each simulation and report percentiles of sample statistics over all 20,000 artificial series. I use the common notations for the average $E$, standard deviation $\sigma$, autocorrelation $acl$, skewness $skew$, and kurtosis $kurt$ of the series.
conditional variances of the market return under the actual and risk-neutral probability measures. Panel A in Table 3 shows that the GDA model is able to generate a large and volatile variance premium. Although the mean variance premium predicted by the benchmark is slightly lower than in the data, the empirical estimate easily falls into the 90% model-based confidence interval. It is well-known that the variance premium distribution is fat-tailed with positive skewness and large excess kurtosis. The GDA model qualitatively respects the non-normality of the distribution, although the benchmark statistics are smaller relative to the data.

Most importantly, the GDA model is able to account for the first and second moments of the variance premium with empirically consistent conditional return variances under both probability measures. Specifically, the total return variance is more volatile under the risk-neutral probability measure relative to the physical probability measure. Additionally, both volatilities are persistent as in the data. I also report the skewness and kurtosis of the VIX index. Interestingly, these statistics dramatically changed after the 2007-2008 Financial Crisis, which can be mainly attributed to large values of VIX during October and November 2008. Consequently, the skewness and kurtosis of the risk-neutral conditional variance increased from around 2 and 9 in the pre-crisis period to 3.45 and 20.72 in the full sample ranging until December 2016. The median statistics of the GDA model are closer to empirical estimates based on the pre-crisis period. This can be partially explained by the fact that I rule out realizations of rare recessions in consumption while generating model-based statistics.

Empirical literature further documents the predictability of excess returns by the variance premium. To study this predictive relationship, I regress the one-, three-, and six-month cumulative excess log returns, which are expressed in percentages, on the lagged monthly variance premium. Consistent with the existing literature, the ”Data” column of Panel B in Table 3 indicates a positive impact as measured by positive and slightly decreasing regression coefficients, in addition to an increasing predictive power as measured by increasing $R^2$ over longer horizons. It is striking that the GDA model replicates this empirical finding by closely matching the magnitude of coefficients and $R^2$. 

26
The literature on the risk premia in higher-order moments of equity returns also documents a puzzling skew premium. Kozhan, Neuberger and Schneider (2013) quantify its value by looking at the profit of a skew swap, a financial contract paying the difference between the implied skew and the realized skew of the equity return. Table 4 summarizes descriptive statistics of the skew premium and related variables. The benchmark model with GDA preferences produces a sizable negative skew premium that closely matches the historical data (-42%). Note that the possibility for a large jump in the conditional skew premium leads to wide simulation percentiles for the sample volatility, skewness, kurtosis, and first-order autocorrelation. Nevertheless, the calibration’s median statistics of these moments are reasonably close to the data and reflect the key features of the skew premium. Additionally, the conditional first and second moments of the return skewness under risk-neutral and physical probability measures are in line with their empirical counterparts, although the volatility of the realized skew in the model is too low relative to the data.

Pure Disappointment Aversion and Epstein-Zin Specifications. Although the model provided a good fit with equity returns in the data, comparable the performance of the GDA framework, the crucial difference between them can be observed in the light of the variance and skew risk premia statistics. Table 3 shows that disappointment aversion alone produces the mean and volatility of the variance premium that are approximately five times smaller than the benchmark values. Turning off the GDA channel also leads to a significant reduction in the volatility of return variance. As the variance premium decreases, its predictive power for the excess log returns also suffers. This is manifested in the lower median values of $R^2$ and empirically inconsistent regression coefficients. The asset pricing implications of the DA model are further augmented by Table 4. Specifically, Table 4 shows that a simple specification does not reproduce the skew premium statistics. The skew premium turns out to be positive with excessively median values of sample volatility, skewness, kurtosis and excessively high median autocorrelations. The bottom part of Table 4 shows that the DA model predicts 2-3 times smaller first and second
Table 4
Benchmark Calibration and Sensitivity Analysis: Skewness Premium

<table>
<thead>
<tr>
<th></th>
<th>Data 5%</th>
<th>Data 50%</th>
<th>Data 95%</th>
<th>GDA 5%</th>
<th>GDA 50%</th>
<th>GDA 95%</th>
<th>DA 5%</th>
<th>DA 50%</th>
<th>DA 95%</th>
<th>EZ 5%</th>
<th>EZ 50%</th>
<th>EZ 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>0.9989</td>
<td>0.9989</td>
<td>0.9989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( 1 - \alpha )</td>
<td>2.5</td>
<td>2.5</td>
<td></td>
<td>0</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( 1/(1 - \rho))</td>
<td>1.5</td>
<td></td>
<td>( \infty )</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>2.33</td>
<td>0.55</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.94</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table reports moments of the conditional skew premium \( sp \), market return skewness \( skew_r^2(r_e) \) and \( skew_Q^2(r_e) \) under the physical \( P \) and risk-neutral \( Q \) probability measures, respectively. The entries of the table are monthly statistics. The moments are for the data and three models: the benchmark model with generalized disappointment aversion preferences GDA, a pure disappointment aversion specification with linear preferences and infinite elasticity of intertemporal substitution DA, and an Epstein-Zin economy EZ. The empirical statistics are for the U.S. data from January 1996 to January 2012. For each model, I simulate 20,000 economies at a monthly frequency with a sample size equal to the empirical counterpart. The simulation results are percentiles of sample moments based on these 20,000 artificial series. I use the common notations for the average \( E \), standard deviation \( \sigma \), autocorrelation \( ac1 \), skewness \( skew \), and kurtosis \( kurt \) of the series.

Moments of return skewness.

Next, I turn off any source of (generalized) disappointment aversion and consider a representative agent with preferences for early resolution of uncertainty, a widely used utility specification in asset pricing studies. According to Table 3, the EZ model increases the mean and volatility of the variance premium by a factor 2.5 compared to the DA model. However, the mean variance premium is still only half of the benchmark value. In addition, the predictive regression of the excess log returns on the variance premium indicates a weaker predictability power. As an additional robustness exercise, I raise risk aversion to the value 10, an upper bound of reasonable values according to Mehra and Prescott (1985). In this case, the median value of the sample mean of the variance

28
premium increases marginally to 5.15, while the median volatility statistics are dampened at the level of around 4.11. Moreover, the regression of the excess log returns on the variance premium produces weaker predictive power with $R^2$ less than 2% at the 6-month horizon. Overall, these results indicate the important role of generalized disappointment aversion for the equity returns and variance premium, as DA and EZ models cannot account for salient features of the data.

5.2.3 The Term Structure of Implied Volatilities

In this section, I examine the asset pricing implications of all models for index options. The top graph of Figure 1 compares the 3-month volatility curves for the data and three models (GDA, DA, and EZ). The implied volatilities are expressed as a function of moneyness ranging from 0.9 to 1.05. The plot shows that the empirical implied volatilities are declining in moneyness, a pattern also known in the literature as the volatility skew. The top panel of Figure 1 shows that the DA implied volatilities for the 3-month maturity are very flat and approximately equal the realized stock market volatility. One apparent candidate to generate a steep volatility skew is high risk aversion. Although increased risk aversion in the EZ model improves the model performance, it cannot fully account for the level in implied volatilities. In contrast, the GDA framework can fit the option prices very closely.

The middle and bottom plots of Figure 1 additionally present the term structure of implied volatilities for ATM and 0.90 OTM options. In the data, ATM volatilities slightly increase over the horizon, while the downward trend can be observed for OTM volatilities. The model-based results clearly indicate that neither DA nor EZ models can match the empirical curves, although high risk aversion in EZ produces higher volatilities. The benchmark with generalized disappointment aversion can closely fit the empirical ATM volatilities for 1-, 3-, and 6-month maturities. Although the GDA model cannot fully account for the steep slope of the OTM volatilities, the model-predicted level of the term structure corresponds the one in the data. In general, the performance of the GDA framework clearly indicates the importance of generalized disappointment risk in option prices.
Figure 1: Benchmark Calibration and Sensitivity Analysis: Volatility Term Structure. The top panel plots the 3-month implied volatility curve as a function of moneyness (Strike/Spot Price) for the data and three models: the benchmark model with generalized disappointment aversion preferences GDA, a pure disappointment aversion specification with linear preferences and infinite elasticity of intertemporal substitution DA, and an Epstein-Zin economy EZ. The middle and bottom panels plot the empirical and model-based (GDA, DA, and EZ) implied volatility curves for ATM and OTM options as functions of the time to maturity expressed in months. The empirical statistics are for the U.S. data from January 1996 to December 2016. For each model, I simulate 20,000 economies at a monthly requency with a sample size equal to the empirical counterpart. The model-based curves are calculated for option prices using the annualized model-implied interest rate $r_t(\pi_t)$ and dividend-yield $q_t(\pi_t)$ in each period. The simulation results are medians of implied volatilities based on these 20,000 artificial series.
Figure 2: Sensitivity of Asset Prices to Preference Specifications and Parameter Values. The figure plots asset pricing implications of alternative preference calibrations in a benchmark GDA model, a pure disappointment aversion specification, and an Epstein-Zin economy. In each of these models, a single parameter is changed while others are fixed at original values. Specifically, I change a disappointment threshold, a disappopintment aversion and a relative risk aversion over a range of values in GDA, DA and EZ models, respectively. For each model specification, I simulate 20,000 economies at a monthly frequency. The entries of the figure are medians of sample statistics (annualized for the risk-free rate, the equity premium and the price-dividend ratio; monthly for the variance and skew risk premia) based on these 20,000 artificial series. I use the common notations for the average $E$ and standard deviation $\sigma$ of the series.
The (generalized) disappointment aversion and risk aversion both stand between the physical $\mathbb{P}$ and risk-neutral $\mathbb{Q}$ probability measures through the Radon-Nikodym derivative
\[ \frac{dQ}{dP} = \frac{M_{t+1}}{\mathbb{E}_t(M_{t+1})}. \]
However, the failure of DA and EZ models and the striking success of the GDA specification can be attributable to the difference in their impacts on a $\mathbb{Q}$-distribution. Notice that DA preferences penalize many more outcomes relative to GDA preferences due to higher disappointment threshold. Consequently, the risk-neutral measure implied by DA puts less weights on the adverse negative consumption growth rates and more weights on the milder consumption realizations relative to GDA. Similarly, a high risk aversion smoothly distorts a $\mathbb{Q}$-density towards a left tail and hence the impact is not strong enough to fully capture option prices. Since generalized disappointment aversion enables me to control the threshold of disappointment events, it becomes instrumental in generating a fatter left tail of a $\mathbb{Q}$-measure compared to alternative preferences. By penalizing the outcomes below a scaled certainty equivalent, the GDA model exhibits more countercyclical risk aversion that helps explain the volatility skew, variance and skew risk premia in the data.

### 5.3 Sensitivity Analysis

Figure 2 provides an extensive sensitivity analysis with respect to different preference specifications and parameter values. In the sensitivity exercise, I consider three model settings (GDA, DA, and EZ) and change a key parameter in each of them while holding remaining values as in original economies. In the GDA model, I vary a disappointment threshold $\delta$ between 0.92 and 0.955. In the DA model, a disappointment aversion $\theta$ changes between 0.35 and 0.7. In the EZ model, the results are provided for the relative risk aversion ranging from 4 to 7.5. I focus on the first and second moments of the risk-free rate, the equity premium, the price-dividend ratio, the variance and skew risk premia. The panels in Figure 2 present the model-based median statistics implied by the GDA (a dashed line), DA (a red line with dots), and EZ (a blue line) frameworks. The asset pricing moments are expressed as a function of a varying parameter.
Figure 2 shows that the risk-free rate decreases with a disappointment threshold $\delta$ in GDA and a risk aversion $1 - \alpha$ in EZ, while it is equal to a constant $-12 \ln \beta$ regardless of a disappointment aversion $\theta$ in DA. Furthermore, the equity premium increases and equity prices decline in $\delta$, $\theta$, and $1 - \alpha$, when the agent faces more disappointing outcomes or becomes more averse to negative consumption growth rates as induced by higher values of these parameters. The impact of $\delta$ and $1 - \alpha$ on the volatility of asset prices is similar in GDA and EZ: the higher disappointment threshold or the higher risk aversion leads to more volatile risk-free rate, while the volatility of equity returns and the price-dividend ratio exhibits a hump-shaped pattern with a maximum approximately in the middle of considered parameter intervals. In DA, increasing the disappointment aversion slightly increases the equity volatility, while the risk-free rate remains unaffected due to linear preferences and infinite elasticity of intertemporal substitution. Overall, the magnitude of changes in the risk-free rate, the equity returns, and the price-dividend ratio are quite comparable across three preference specifications, especially when looking at the performance of GDA and EZ. These findings suggest that all three preference specifications can reasonably explain first and second moments of equity returns by adjusting a key preference parameter. In contrast, the four bottom panels in Figure 2 indicate the crucial importance of generalized disappointment aversion for generating significant risk premiums in higher moments of equity returns.

It is evident from Figure 2 that a pure disappointment aversion model can produce a high mean and volatility of equity returns, however, the model-generated variance premium is too low and less volatile compared to the data. Moreover, disappointment aversion alone cannot reproduce salient moments of the skew premium at all. Increasing a disappointment aversion parameter $\theta$ in the DA setting does not improve the model performance, as the variance and skew risk premia moments are not very sensitive to changes in $\theta$. The Epstein-Zin economy provides a better fit with the data. In particular, when the risk aversion increases from 4 to 7.5, the mean variance premium increases from less than 2 to around 5, while the skew premium declines from around -10% to -20%. However, note that the mean and volatility of the variance premium actually start
declining at some time and, thus, the higher risk aversion will bring the model away from the data. The comparative analysis with respect to the disappointment threshold in GDA preferences provides the overall patterns in the variance and skew risk premia similar to those generated by different risk aversion parameters in Epstein-Zin preferences. However, with generalized disappointment aversion, the magnitude of the variance and skew risk premia is significantly amplified. The sensitivity analysis in Figure 2 confirms that the distribution of the stochastic discount factor, necessary to reconcile the empirical asset pricing moments, is attributable to the agent’s generalized disappointment aversion and cannot be supported by any parameter values in alternative preference specifications.

6. Conclusion

This paper builds a representative agent, consumption-based asset pricing framework in which the agent has generalized disappointment aversion risk preferences and consumption growth is modeled as a hidden Markov-switching process. I show that the combination of the investor’s aversion to tail events and fluctuating economic uncertainty due to Bayesian learning help to successfully explain a wide variety of asset pricing phenomena. The benchmark model is able to reproduce variance and skew risk premia and to generate a realistic volatility surface implied by equity index options, while simultaneously capturing the salient moments of equity returns.

The success of the model is attributable to the endogenously varying probability of disappointment events, which has a large impact on asset prices, particularly on the higher moments risk premia in equity returns and implied volatilities. The agent’s time-varying beliefs induce rich conditional dynamics that are reflected in the equity and option prices. To emphasize the importance of GDA preferences, I consider alternative models with disappointment aversion and preferences for early resolution of uncertainty. Although all three specifications can reasonably match moments of equity returns, the benchmark model outperforms the other two calibrations by additionally capturing the salient features of options prices, variance, and skew risk premia. These results suggest the important role of generalized disappointment aversion in asset pricing models.
References


Appendix

A. Representative Agent’s Maximization Problem

A representative agent starts with an initial wealth denoted by $W_0$. Each period $t$, the agent consumes $C_t$ consumption goods and invests in $N$ assets traded on the competitive market. Denote the fraction of the total $t$-period wealth $W_t$ invested in the $i$-th asset with gross real return $R_{i,t}^{\omega}$ by $\omega_{i,t}$. Then, the agent’s budget constraint in period $t$ takes the form:

$$W_{t+1} = (W_t - C_t)R_{t+1}^\omega$$  \hspace{1cm} (A1)$$

where

$$\sum_{i=1}^{N} \omega_{i,t} = 1 \quad \text{and} \quad R_{t+1}^\omega = \sum_{i=1}^{N} \omega_{i,t}R_{t+1}^{\omega_i}.$$  \hspace{1cm} (A2)$$

The agent chooses the allocation $\{C_t, \omega_{1,t}, ..., \omega_{N,t}\}$ in period $t$ in order to maximize (1) subject to (A1) and (A2).

The Bellman equation becomes:

$$J_t = \max_{C_t, \omega_{1,t}, ..., \omega_{N,t}} \left\{ (1 - \beta)C_t^\rho + \beta \left[ \mu_t(J_{t+1}) \right]^\rho \right\}^{1/\rho}$$

with the constraints (A1) and (A2). I guess optimal value function of the form $J_t = \phi_t W_t$. Using this conjecture of $J_t$ and the form of $\mu_t$ from (2), I rewrite the Bellman equation as:

$$\phi_t W_t = \max_{C_t, \omega_{1,t}, ..., \omega_{N,t}} \left\{ (1 - \beta)C_t^\rho + \beta \left[ \mathbb{E}_t \left[ (\phi_{t+1} W_{t+1})^\alpha K(\phi_{t+1} W_{t+1}) \right]^\rho \right]^{1/\rho} \right\},$$

where

$$K(x) = \frac{1 + \theta \mathbb{I}\{x \leq \delta \mu_t(x)\}}{1 + \theta \delta \mathbb{E}_t \mathbb{I}\{x \leq \delta \mu_t(x)\}}.$$  \hspace{1cm} (A3)$$

Note that the function $K$ defined above is homogeneous of degree zero.

The Return on the Aggregate Consumption Claim Asset. I further conjecture that the consumption $C_t$ is homogeneous of degree one in wealth at the optimum, that is $C_t = b_t W_t$. Then, I obtain the Bellman equation:

$$\phi_t^\rho = \left\{ (1 - \beta) \left( \frac{C_t}{W_t} \right)^\rho + \beta \left( 1 - \frac{C_t}{W_t} \right)^\rho \left[ \mathbb{E}_t \left[ (\phi_{t+1} R_{t+1}^\omega)^\alpha K(\phi_{t+1} R_{t+1}^\omega) \right]^\rho \right]^{1/\rho} \right\}^{\rho/\alpha}$$  \hspace{1cm} (A3)$$
or equivalently

$$\phi_t^\rho = \{(1 - \beta)b_t + \beta (1 - b_t)^\rho y_t^*\} \quad \text{(A4)}$$

where

$$y_t^* = \left[ \mathbb{E}_t \left[ (\phi_{t+1}R_{t+1}^\omega)^\alpha \mathcal{K}(\phi_{t+1}R_{t+1}^\omega) \right] \right]^{\rho/\alpha}.$$  

Taking the FOC of the right side of a simplified Bellman equation (A3) with respect to $C_t$, I find:

$$(1 - \beta) \left( \frac{C_t}{W_t} \right)^{\rho - 1} = \beta \left( 1 - \frac{C_t}{W_t} \right)^{\rho - 1} y_t^*,$$

or using the notations:

$$(1 - \beta)b_t^{\rho - 1} = \beta(1 - b_t)^{\rho - 1} y_t^*. \quad \text{(A5)}$$

Solving for $y_t^*$ from the last equation and substituting it into (A4), I deduce:

$$\phi_t = (1 - \beta)\left( \frac{1 - \beta}{b_t^{\rho - 1}} \right) = (1 - \beta)\left( \frac{1 - \beta}{b_t^{\rho - 1}} \right)^{\rho - 1}$$

Shifting one period ahead the formula for $\phi_t$ and substituting the resulting form of $\phi_{t+1}$ into (A5), I obtain:

$$(1 - \beta)C_t^{\rho - 1} = \beta(W_t - C_t)^{\rho - 1} \left[ \mathbb{E}_t \left[ (1 - \beta)^{\alpha/\rho} \left( \frac{C_{t+1}}{W_{t+1}} \right)^{\alpha - 1/\rho} \mathcal{K}(\phi_{t+1}R_{t+1}^\omega) \right] \right]^{\rho/\alpha}.$$

Then, I rewrite the equation above as:

$$C_t^{\rho - 1} = \beta \mathbb{E}_t \left[ \left( \frac{C_{t+1}}{W_{t+1}} \right)^{\alpha - 1/\rho} \mathcal{K} \left( \left( \frac{C_{t+1}}{W_{t+1}} \right)^{\rho - 1/\rho} \frac{R_{t+1}^\omega}{W_{t+1} - C_t} \right) \right]^{\rho/\alpha}.$$

and derive the asset pricing restriction for the return on the total wealth $R_{t+1}^\omega$:

$$\mathbb{E}_t \left[ \left( \beta \left( \frac{C_{t+1}}{C_t} \right)^{\rho - 1} \frac{R_{t+1}^\omega}{z_{t+1}} \right)^{1/\rho} \mathcal{K} \left( \left( \beta \left( \frac{C_{t+1}}{C_t} \right)^{\rho - 1} \frac{R_{t+1}^\omega}{z_{t+1}} \right)^{1/\rho} \right) \right]^{1/\alpha} = 1.$$

Define $R_{t+1}^c$ the return on the consumption endowment. In equilibrium, $R_{t+1}^c = R_{t+1}^\omega$ and, as in Routledge and Zin (2010), using the definition of the certainty equivalent (2) and the function $\mathcal{K}$, the return $R_{t+1}^c$ should satisfy the equation:

$$\mu_t(z_{t+1}) = 1 \quad \text{(A6)}$$
where
\[ z_{t+1} = \left( \beta \left( \frac{C_{t+1}}{C_t} \right)^{\rho-1} R_{t+1}^c \right)^{1/\rho} \].

Rewriting \( R_{t+1}^c \) in the form:
\[ R_{t+1}^c = \frac{W_{t+1}}{W_t - C_t} = \frac{W_{t+1}}{C_{t+1}} \cdot \frac{C_{t+1}}{C_t} = \frac{\xi_{t+1}}{\xi_t} \cdot \frac{C_{t+1}}{C_t}, \]
the wealth-consumption ratio \( \xi_t = \frac{W_t}{C_t} \) can be found from the functional equation:
\[ E_t \left[ \beta \alpha \left( \frac{C_{t+1}}{C_t} \right)^\alpha \cdot \left( \frac{\xi_{t+1}}{\xi_t} \right)^{\frac{\alpha}{\rho}} \cdot K(z_{t+1}) \right] = 1. \]

The Return on the Aggregate Dividend Asset. Following Routledge and Zin (2010), the portfolio problem for the obtained values \( \phi_{t+1} \) reads as follows:
\[
\max_{\omega_{1,t},...\omega_{N,t}} \mu_t(\phi_{t+1} R_{t+1}^\omega),
\]
subject to the constraints \( \sum_{i=1}^{N} \omega_{i,t} = 1 \) and \( R_{t+1}^\omega = \sum_{i=1}^{N} \omega_{i,t} R_{i,t+1} \). Taking the FOC with respect to the weight \( \omega_{i,t} \), I derive:
\[
E_t \left[ \phi_{t+1}^{\alpha} (R_{t+1}^\omega)^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R_{t+1}^\omega < \delta \mu_t)] R_{i,t+1} \right] = 0.
\]
Taking the difference between the \( i \)-th and \( j \)-th FOCs, I thus obtain:
\[
E_t \left[ \phi_{t+1}^{\alpha} (R_{t+1}^\omega)^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R_{t+1}^\omega < \delta \mu_t)] (R_{i,t+1} - R_{j,t+1}) \right] = 0.
\]
Multiplying the last equation by \( \omega_{j,t} \) and summing over all possible values of \( j \), I further obtain:
\[
E_t \left[ \phi_{t+1}^{\alpha} (R_{t+1}^\omega)^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R_{t+1}^\omega < \delta \mu_t)] \sum_{j=1}^{N} \omega_{j,t} \right] = \\
E_t \left[ \phi_{t+1}^{\alpha} (R_{t+1}^\omega)^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R_{t+1}^\omega < \delta \mu_t)] \sum_{j=1}^{N} R_{j,t+1} \omega_{j,t} \right] = R_{t+1}^\omega.
\]
or
\[
E_t \left[ \phi_{t+1}^\alpha (R_{t+1}^\alpha)^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R_{t+1}^\alpha < \delta \mu_t)] R_{t+1} \right] = \\
= E_t \left[ \phi_{t+1}^\alpha (R_{t+1}^\alpha)^\alpha [1 + \theta \mathbb{I}(\phi_{t+1} R_{t+1}^\alpha < \delta \mu_t)] \right].
\] (A7)

Following Epstein and Zin (1989, 1991), it is straightforward to show that \( \phi_{t+1} = \frac{z_{t+1}}{R_{t+1}^c} \) holds in equilibrium. Using these equilibrium conditions and the definition of \( \mu_t \), I have:
\[
E_t \left[ \phi_{t+1}^\alpha (R_{t+1}^\alpha)^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R_{t+1}^\alpha < \delta \mu_t)] R_{t+1} \right] = \\
= E_t \left[ z_{t+1}^\alpha [1 + \theta \mathbb{I}(z_{t+1} < \delta \mu_t)] \right].
\] (A8)

Combining (A7)-(A8) and using the equilibrium condition \( R_{t+1}^c = R_{t+1}^\omega \), I finally obtain the asset pricing restriction for the gross return \( R_{t+1} \):
\[
E_t \left[ z_{t+1}^\alpha (R_{t+1}^c)^{-1} (1 + \theta \mathbb{I}(z_{t+1} < \delta)) R_{t+1} \right] = 1,
\] (A9)

Moreover, the pricing kernel \( M_{t+1} \) is:
\[
M_{t+1} = z_{t+1}^\alpha (R_{t+1}^c)^{-1} (1 + \theta \mathbb{I}(z_{t+1} < \delta)) \frac{1 + \theta \delta^\alpha \mathbb{E}_t \mathbb{I}(z_{t+1} < \delta)}{1 + \theta \delta^\alpha \mathbb{E}_t \mathbb{I}(z_{t+1} < \delta)}.
\]

Rewriting \( R_{t+1} \) in the form:
\[
R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t},
\]

the price-dividend ratio of the \( i \)-th asset \( \lambda_t = \frac{P_{t+1}^i}{D_{t+1}^i} \) can be found from the functional equation:
\[
E_t \left[ \beta^\sigma \left( \frac{C_{t+1}}{C_t} \right)^{\alpha-1} \frac{D_{t+1}^i}{D_t^i} \cdot \left( \frac{\xi_{t+1}^i}{\xi_t^i} \right)^{\frac{\sigma-1}{\sigma}} \cdot \mathbb{K}(z_{t+1}) \cdot (\lambda_{t+1}^i + 1) \right] = \lambda_t.
\]

B. Numerical Algorithm

This technical appendix provides the description of the numerical method used to solve the model. Following the notation from the paper, aggregate consumption growth
\[
\Delta c_{t+1} = \ln \left( \frac{C_{t+1}}{C_t} \right)
\]
is given by:

$$\Delta c_{t+1} = \mu s_{t+1} + \sigma \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, 1).$$

Whereas the consumption volatility $\sigma$ is assumed to be constant, the mean growth rate $\mu s_{t+1}$ is driven by a two-state Markov-switching process $s_{t+1}$ with the state space:

$$S = \{1 = \text{expansion}, 2 = \text{recession}\},$$

a transition matrix

$$\mathbf{P} = \begin{pmatrix} \pi_{11} & 1 - \pi_{11} \\ 1 - \pi_{22} & \pi_{22} \end{pmatrix}$$

and transition probabilities $\pi_{ii} \in (0, 1)$. Let

$$\mathcal{X}(y_1, y_2, y_3) = \frac{1 + \theta \left\{ \beta e^{\rho y_1} \left( \frac{y_2}{y_3-1} \right) \leq \delta \rho \right\}}{1 + \theta \delta e^{\rho y_1} \left( \frac{y_2}{y_3-1} \right) \leq \delta \rho},$$

then, the wealth-consumption ratio $\xi_t = \frac{W_t}{C_t}$ satisfies the equation:

$$\mathbb{E}_t \left[ \beta \rho e^{\alpha \Delta c_{t+1}} \cdot \left( \frac{\xi_{t+1}}{\xi_t - 1} \right)^{\frac{\alpha}{\rho}} \cdot \mathcal{X}(\Delta c_{t+1}, \xi_{t+1}, \xi_t) \right] = 1, \quad \text{(B1)}$$

and the price-dividend ratio $\lambda_t = \frac{P_t}{D_t}$ of the asset with gross return $R_{t+1}$ (I skip the subscript $i$ for convenience) is given by:

$$\mathbb{E}_t \left[ \beta \rho e^{(\alpha-1) \Delta c_{t+1} + \Delta d_{t+1}} \cdot \left( \frac{\xi_{t+1}}{\xi_t - 1} \right)^{\frac{\alpha-1}{\rho}} \cdot \mathcal{X}(\Delta c_{t+1}, \xi_{t+1}, \xi_t) \cdot \frac{\lambda_{t+1} + 1}{\lambda_t} \right] = 1. \quad \text{(B2)}$$

I apply the projection method of Judd (1992) to solve for the equilibrium pricing functions when the state of the economy is unobservable.

**The Return on the Aggregate Consumption Claim Asset.** I conjecture the wealth-consumption ratio of the form $\xi_t = G(\pi_t)$, where $\pi_t$ is the posterior belief defined by (4). I seek to approximate the functional form of $G(\pi_t)$, which solves (B1). I approximate $G(\pi_t)$ by a basis of complete Chebyshev polynomials $\Psi = \{\Psi_k(\pi_t)\}_{k=0}^n$ of order $n$ with coefficients $\psi = \{\psi_k\}_{k=0}^n$:

$$G(\pi_t) = \sum_{k=0}^n \psi_k \Psi_k(\pi_t) \quad \forall \pi_t \in [0, 1]. \quad \text{(B3)}$$

\(^6\)I adjust the domain of the Chebyshev polynomials from $[-1, 1]$ to the domain of the state variable $\pi_t$ which is $[0, 1]$. 

44
I further define the function:

\[
\Gamma(\pi_t; j) = \mathbb{E}_{t,j} \left[ \beta \alpha e^{\alpha \Delta c_{t+1}} \cdot \left( \frac{\xi_{t+1}}{\xi_t - 1} \right)^{\frac{\alpha}{\rho}} \cdot X(\Delta c_{t+1}, \xi_{t+1}, \xi_t) \right] = \beta \alpha \int e^{\alpha y} \left( \frac{G(B(y, \pi_t))}{G(\pi_t)} - 1 \right)^{\frac{\alpha}{\rho}} \cdot X(y, G(B(y, \pi_t)), G(\pi_t)) f(y, j) dy \quad (B4)
\]

where \( B(y, \pi_t) \) is given by:

\[
B(y, \pi_t) = (1 - q) f(y, 1)(1 - \pi_t) + pf(y, 2)\pi_t
\]

and \( f(y, j) \) is the probability density function of a normal distribution \( N(\mu_{S_t}, \sigma^2) \) conditional on \( S_t = j \). Substituting \( G(\pi_t) \) from (B3) and \( \Gamma(\pi_t; j) \) from (B4) into (B1), I obtain the residual function:

\[
R^c(\pi_t; \psi) = (1 - \pi_t)\Gamma(\pi_t, 1) + \pi_t\Gamma(\pi_t, 2) - 1.
\]

The objective is to choose the unknown coefficients \( \psi \) in order to make the residual function \( R^c(\pi_t; \psi) \) close to zero \( \forall \pi_t \in [0, 1] \). I apply the orthogonal collocation method. Formally, I evaluate the residual function in the collocation points \( \{r_k\}_{k=1}^{n+1} \) given by the roots of the \( n+1 \) order Chebyshev polynomial \(^7\) and then solve the system of \( n+1 \) equations:

\[
R^c(r_k; \psi) = 0 \quad \forall k = 1, \ldots, n+1
\]

for \( n+1 \) unknowns \( \psi = \{\psi_k\}_{k=0}^{n} \). Let \( \tilde{\xi}_t = \tilde{G}(\pi_t) = \sum_{k=0}^{n} \tilde{\psi}_k \Psi_k(\pi_t) \) denote an approximation of the wealth-consumption ratio.

The Return on the Aggregate Dividend Asset. I conjecture the price-dividend ratio of the form \( \lambda_t = H(\pi_t) \), where \( \pi_t \) is the posterior belief defined by (4). Now, I seek to approximate the functional form of \( H(\pi_t) \), which solves the equation (B2). I approximate \( H(\pi_t) \) by a basis of complete Chebyshev polynomials \( \Upsilon = \{\Upsilon_k(\pi_t)\}_{k=0}^{n} \) of order \( n \) with coefficients \( \psi = \{\psi_k\}_{k=0}^{n} :\)

\[
H(\pi_t) = \sum_{k=0}^{n} \psi_k \Upsilon_k(\pi_t) \quad \forall \pi_t \in [0, 1]. \quad (B5)
\]

\(^7\)Again, I adjust the domain of the Chebyshev polynomials from \([-1, 1]\) to the domain of the state variable \( \pi_t \) which is \([0, 1]\).
Define the function:

\[
\Lambda(\pi_t; j) = \mathbb{E}_{t,j} \left[ \beta_{\pi} e^{(a-1) \Delta c_{t+1} + \Delta d_{t+1}} \left( \frac{\tilde{\xi}_{t+1}}{\xi_t - 1} \right)^{\frac{\alpha}{\rho} - 1} \cdot \mathcal{X}(\Delta c_{t+1}, \tilde{\xi}_{t+1}, \tilde{\xi}_t) \cdot \frac{\lambda_{t+1} + 1}{\lambda_t} \right] =
\]

\[= \beta_{\pi} \int \int e^{(a+\lambda-1)y+g+\xi} \left( \frac{\tilde{G}(B(y, \pi_t))}{\tilde{G}(\pi_t) - 1} \right)^{\frac{\alpha}{\rho} - 1} \cdot \mathcal{X}(y, \tilde{G}(B(y, \pi_t)), \tilde{G}(\pi_t)) \cdot \frac{H(B(y, \pi_t))}{H(\pi_t) - 1} f(y, j) g(z, j) dy dz , \quad (B6)
\]

where \(f(y, j)\) and \(g(z, j)\) are probability density functions of normal distributions \(N(\mu_{S_{t+1}}, \sigma)\) and \(N(g_{d}, \sigma_d)\), respectively, conditional on \(S_{t+1} = j\). Substituting \(H(\pi_t)\) from \((B5)\) and \(\Lambda(\pi_t; j)\) from \((B6)\) into \((B2)\), I obtain the residual function:

\[R^d(\pi_t; v) = (1 - \pi_t) \Lambda(\pi_t, 1) + \pi_t \Lambda(\pi_t, 2) - 1.\]

Next, I evaluate the residual function \(R^d(\pi_t; \psi)\) in the collocation points \(\{s_k\}_{k=1}^{n+1}\) given by the roots of the \(n + 1\) order Chebyshev polynomial and solve the system of \(n + 1\) equations

\[R^d(s_k; v) = 0 \quad \forall k = 1, ..., n + 1 \]

for \(n + 1\) unknowns \(v = \{v_k\}_{k=0}^{n}.\)