Asset Prices in Turbulent Markets with Rare Disasters

Soo hun Kim*

January 12, 2013

JOB MARKET PAPER

Abstract

I propose a parsimonious econometric model for the stochastic process governing the evolution of per capita consumption and stock market dividend over time. The model features stochastic volatility of consumption and dividend growth rates, and time-varying likelihood of rare disasters. I embed this time-variation of risk in an endowment economy with a representative agent and estimate the parameters from U.S. stock market data using Maximum Likelihood. Allowing for time-varying likelihood of rare disasters improves the model’s performance. My model successfully explains a number of empirical puzzles: the high equity risk premium, excessive volatility of equity return, predictability of market returns through the price-to-dividend ratio, and the cyclical patterns observed in the term structure of the yield on dividend strips. In addition, the model-implied correlations between equity premium, variance risk premium, and the implied volatility of deep OTM put options are consistent with empirical findings in the literature.

*Kellogg School of Management, Northwestern University, 2001 Sheridan Rd, Evanston, IL, 60208, soohun-kim@kellogg.northwestern.edu: I am deeply indebted to my advisors, Ravi Jagannathan and Dimitris Papanikolaou, for their constant encouragement, support and constructive criticisms. I wish to thank Torben Andersen, Akash Bandyopadhyay, Jules van Binsbergen, Anna Cieslak, Kent Daniel, Ralph Koijen, Robert Korajczyk, Edwin Mills, Andreas Neuhierl, Seongkyu Gilbert Park, Costis Skiadas, Viktor Todorov for helpful comments on the earlier version of the paper. I thank Nicola Fusari, Maria T. Gonzalez-Perez, Yan Li, David T. Ng, and Bhaskaran Swaminathan for data. All errors are mine.
1 Introduction

The recent events of the financial crisis have led to a resurgence of interest in models with rare disasters affecting the economy. These models have been fairly successful in generating features of financial markets. Notable among these models are Barro and Ursua (2011), Gabaix (2012) and Wachter (2012). A key feature in these models is that the likelihood of rare disasters is time-varying. However, rare disasters, by their very nature, are rarely observed. The lack of sufficient observations makes the econometric estimation of time-varying likelihood difficult. I address this issue by proposing a tractable and parsimonious econometric model, that nests rare disasters with stochastic volatility in the endowment process.

I build a model with time-varying risk using multiple regimes. Risk is captured by the variance of small, frequent shocks and the likelihood of large, infrequent disasters. Both risks co-vary over time as a function of the state of the economy. The state of the economy is characterized by a vector of binary basis regimes; a model with $k$ basis regimes implies a total of $2^k$ possible states. Each binary regime evolves independently. Their dynamics are identical except for their persistence. Specifically, the regime $j$ has a persistence that is a power function of the persistence of regime $j - 1$. Hence, the persistency cascades geometrically in the ordering of regimes. As a result, the transition of economy with arbitrarily many states can be parsimoniously specified with only three parameters; the steady state distribution, the persistency of the most persistent regime, and the geometric decay rate of the persistency.

The advantage of using multiple basis regimes is that it gives a simple framework that allows for flexible dynamics for time-variation in systematic risk at multiple frequencies. Financial markets are potentially subject to shocks which operate at different frequencies. My econometric model allows me to parse high-frequency shocks – for instance, temporary demand imbalances – from medium-run fluctuations due to macroeconomic shocks. Despite its flexibility, my model is highly tractable and can be easily estimated using Maximum Likelihood.

I embed the specification of endowments into an economy with a representative agent. After imposing equilibrium pricing restrictions, I estimate the joint dynamics of stochastic volatility and the time-varying likelihood of rare disasters in the endowment process from U.S. stock market returns over the years 1927-2011 using Maximum Likelihood. The return crashes in the Great Depression, the Second World War, the 1987 Crash, the Russian Crisis of 1998, and the recent Financial Crisis, are statistically identified as the occurrence of rare disasters. The Likelihood-Ratio test confirms that the model with a time-varying likelihood

2
of rare disasters is a closer fit to the data than, either, a model without disasters, or a model with a constant likelihood of disasters. Using the estimated parameters of the endowment process, I study the implied properties of asset returns.

My model successfully explain a number of empirical puzzles. First, it generates a high and time-varying equity premium. The agent requires a high compensation for bearing time-varying risk due to the variance of frequent shocks and the likelihood of rare disasters; the compensation for this risk varies with the state of the economy. Further, the price-to-dividend ratio varies along with the level of systematic risk, generating excess volatility of stock returns relative to dividends and return predictability through the price-to-dividend ratio.

Second, the model economy reproduces time-variation in the term structure of equity yields. In the worst regime, the term structure of the yield on dividend strips is downward sloping, consistent with the empirical observation by van Binsbergen, Hueskes, Koijen, and Vrugt (2011) during the recent financial crisis. I show that stochastic volatility is an important component of this relation. In the high volatility state, the agent requires high yield for short duration dividends due to high covariance between consumption and dividends. However, since the volatility would converge down in the long run, the agent does not discount dividend strips with longer maturities as heavily as short-term dividends, inducing the term structure to be inverted in the high risk regime.

Third, the model generates high variance risk premium and the smirk pattern in the implied volatility from option prices. In the model economy, the payoffs of VIX or deep out-of-the-money (OTM) put options at the maturity are closely related to the regime-shifts or the occurrence of rare disasters over the life of the assets. These derivatives can protect the investors from the structural change of the economy or the severe damage on the endowments induced by rare disasters. Market participants are willing to pay high prices for VIX or deep OTM puts to hedge these risks. Hence, VIX contains a sizeable variance risk premium, and the implied volatility of deep OTM put options are significantly higher than that of at-the-money put option.

I use the econometric model to extract the time-varying premiums of equity and VIX along with the time-varying implied volatility of deep OTM puts from the data. The standard filter by Hamilton (1989) enables the econometrician to update her prior over hidden states in real time and to extract those time series. Those series exhibit a strong correlation; hence the model replicates the strong comovement in risk premia across asset classes. In the risky state of the model economy, the agent requires high premium on dividend assets due to the high variance and the significant likelihood of rare disasters, and VIX and deep OTM put
options provide more valuable hedge against the risk of endowments.

A distinguishing feature of my econometric model is its flexibility in handling data at mixed frequencies. I demonstrate this by providing an alternative estimation using quarterly consumption data and monthly returns. When using the aggregate consumption series, the model is a mixed success. I obtain qualitatively similar results regarding the time-varying risk in endowments and the slope of the term structure of risk premia. However, the maximum likelihood estimator detects no evidence of rare disasters in the consumption process, due to the absence of large drops in the post-war quarterly consumption series. I interpret this result as suggestive of the need to move beyond the representative agent paradigm and highlighting the need for models with limited risk sharing.

The rest of the paper is organized as follows: Section 2 briefly reviews related literature. Section 3 presents the model economy where the time-varying risk evolves with multiple binary regimes, and drives equilibrium prices. In Section 4, I explain the data and show how to estimate the parameters for the endowment process with the Maximum Likelihood method. Section 5 applies parameters estimated in Section 4 to the asset prices derived in Section 3 and relates the implication of the model economy to the empirically observed stylized facts. Section 6 provides an alternative estimation of the endowment process as a robustness test. Section 7 discusses the possible extensions of the model economy. Section 8 is the conclusion.

2 Related Literature

Rare Disaster

Rare disasters are economic events that occur rarely but are disastrous in terms of magnitude. Rietz (1988) suggested that infrequent, but catastrophic shocks to consumption, can explain the equity premium puzzle proposed by Mehra and Prescott (1985). Even though Mehra and Prescott (1988) criticize that the assumption on extreme events is extreme by itself, Barro (2006) supports the hypothesis as a solution for many asset-pricing puzzles with empirical evidence of GDP over the 20th century and across countries. Barro and Ursua (2008) strengthen the scope of observed disaster events to international consumption data. Barro and Ursua (2009) investigate the relation between stock market crashes and depressions.\footnote{However, the explanatory power of rare disasters, or how to estimate disasters, is still under debate (Ghosh and Julliard, 2012).}
Figure 1 shows the scatter plot of consumption disasters over the years 1870-2006 across 21 OECD countries and 13 non-OECD countries.\textsuperscript{2} Based on the realized consumption disasters in Figure 1, Gabaix (2012) introduces the time-varying severity of disasters and Wachter (2012) suggests the time-varying likelihood of disaster, generating important features of financial markets. Since we observe disasters only 3-4 times over a century, due to the rarity of rare disasters, samples are not sufficient to estimate the dynamic risk of rare disasters.\textsuperscript{3} As a consequence, in most studies, the time-varying risk of rare disasters is calibrated to explain the target moments of our interest.

The main contribution of this paper is providing a highly tractable econometric model which can estimate and test the time-varying risk of rare disasters. Built on Markov-Switching Multifrequency\textsuperscript{4}(MSM), the econometric specification can be easily estimated using Maximum Likelihood(ML). Also, exploiting ML estimation, I test competing specifications on the risk of rare disasters with Likelihood-Ratio criteria.\textsuperscript{5} I get a statistically significant result in favor of the time-varying likelihood of rare disasters.

The model economy has some limitations as a realistic model for rare disaster risk. First, I do not consider the possibility of recoveries following disasters. Gourio (2008) measures the recoveries and examines the effect of recoveries on equity premium and return predictability. However, the speed and magnitude of recoveries are not uniform across realized disasters. Incorporating recoveries, I should lose the tractability of the model, making statistical estimation complicated. Hence, in this project, I do not investigate the possibility of recoveries. However, since the probability of rare disasters is linked to the variance of Gaussian shocks, the model economy features the turbulent economy around rare disasters.

Second, I infer rare disasters in endowment from monthly return data. Thus, estimated rare disasters may not be reflected in the endowment process. Because it is hard to find sufficient rare disasters in the U.S. endowment data, this limitation is inevitable in estimating a statistical model using U.S. data. In Section 6, I estimate parameters with quarterly consumption data and show some of the results are qualitatively consistent.

\textbf{Regime-Switching Model}

Since the ground-breaking work by Hamilton (1989), the regime-switching model has been

\textsuperscript{2}Data on consumption disasters are from Table 6 of Barro and Ursua (2008). Consumption disasters are defined as more than 10\% drops in the level of consumption.

\textsuperscript{3}Berkman, Jacobsen, and Lee (2011) provides empirical support for time-varying risk of rare disaster; They construct time-varying crisis risk focusing on potential disasters and document the empirical effects of crisis risk.

\textsuperscript{4}Or, Markov-Switching Multifractal

\textsuperscript{5}Specifically, I use the test by Vuong (1989).
utilized in economics and finance to describe the statistical properties of interest in an
economy. In terms of volatility modeling, regime-switching was incorporated into shifts
in parameters of generalized autoregressive conditional heteroscedastic (GARCH) processes
so that parameter shifts at low frequencies (Cai (1994), Hamilton and Susmel (1994), So,
Lam, and Li (1998), Gray (1996), Smith (2002), Klaassen (2002)).

For most models with regime shifts, the number of regimes is usually assumed to be
minimal (from two to four). Even though a small number of regimes hardly identifies every
possible state of the economy, the assumption is inevitable not to encounter the curse of
dimensionality: The number of parameters increases quadratically with the dimension of
regimes. In practical applications, researchers put restrictions on the switching probability
such as the independence of a couple of Markov regimes (Bollen, Gray, and Whaley, 2000),
dependence on durations (Durland and McCurdy, 1994) or dependence on state variables (Gray,
1996). Along with these strands of restrictions on a regime transition matrix, MSM provides
a tool to generate switching probability among an arbitrarily large number of regimes only
with a minimal number of parameters.

A multifractal model was advanced into asset returns by Mandelbrot, Fisher, and Calvet
(1997) and extended to be bridged with regime-switching by Calvet and Fisher (2001, 2002,
2004). MSM manifests how regime-switching models can collaborate with multifractality
in a tractable way, enabling econometricians to estimate parameters with the conventional
estimation method used in Hamilton (1989).

The methodology offered in this paper extends MSM into modeling the stochastic inten-
sity of the jump process. Calvet and Fisher (2008) show multifrequency jumps in returns
can be endogenously generated in the price process even with continuous endowment pro-
cess. To incorporate disastrous events in the endowment process, I directly put jumps in
the endowment process with time-varying intensity moving along with stochastic volatility.
In an econometric perspective, this paper suggests a simple tool for the joint estimation of
time-varying volatility and intensity of jumps within MSM framework. Since I use monthly
data in this paper, multiple jump-sizes are hard to estimate due to the limitation of data
and I focus on the time-varying intensity of fixed-size jumps. More refined modeling on the
nature of jumps will be done in future research using daily or intra-day data.

6MSM is the application of multifractality into binary regimes. For the application to continuous regimes,
see Ajello, Benzoni, and Chyruk (2012) and Calvet, Fisher, and Wu (2010). They restrict the dynamics of
multiple continuous factors in a similar way so that a small number parameters can specify the dynamics of
large number of factors.
Asset Pricing with a Regime-Switching Model

Regime-switching models have been widely used in the asset pricing literature. Even in the seminal work on the equity premium puzzle by Mehra and Prescott (1985) or in the rare disaster model by Rietz (1988), they used the regime-switching model to design the non-i.i.d. endowment process. The recent survey by Ang and Timmermann (2011) summarizes the application of the basic regime-switching model in various assets: equity, bonds, and currencies.

Various forms of the regime-switching endowment process have been applied to the equilibrium model in an attempt to resolve anomalies/puzzles. While the theoretical formulation of the regime-switching model is very general, most of those studies should have employed only a small number of possible regimes for tractability in statistical estimation. However, empirical characteristics of financial data are hard to capture by such a restricted model. Shocks to the underlying economy are differentiated across their persistency. Temporary shocks due to market frictions evaporate suddenly. Business cycles reflect the transition of the economy over the frequency of a few years (Hamilton and Gang, 1996). Over the last decade, many studies have documented the importance of long-sustained movements of state variables, which may appear as technology innovations or demographic changes. MSM provides a framework unifying multifrequency dynamics in a regime-switching formulation. The contribution of this paper is that I cast the time-varying risk of rare disasters into a pure regime-switching formulation, which is tractable in equilibrium pricing as well as statistical estimation.

The econometric specification of this article can be easily extended to pricing bonds or derivatives. Regime-shifts in interest rates have been identified in many empirical works. Hence, it is natural to develop bond pricing models with regime-switching attributes. Augmenting popular affine pricing models such as Duffie and Kan (1996) with a regime-switching model, researchers introduce non-linear models of the term structure of interest rates. Other than through a regime-switching formulation, nonlinearity can be brought about by generalizing the affine silicification to a quadratic form as in Ahn, Dittmar, and Gallant (2002), or by incorporating a jump component as in Ahn and Thompson (1988) and Das (2002).
especially, Bansal, Tauchen, and Zhou (2004) document that term structure with regime-shifts can account for long-term bond return predictability with a tent-shaped combination of multiple forward rates (Cochrane and Piazzesi, 2005), a serious challenge to affine factor models. With the hyperinflation in the event of rare disasters to the model economy, preliminary results show the potential in addressing upward sloped nominal yield curve, high credit spread, and bond return predictability.

For pricing derivatives, a regime-switching formulation has rarely been put to use. The reason is obvious. While the regime-switching model is used primarily for infrequent structural breaks, the maturity of derivatives is relatively short. However, since MSM can handle multifrequency risks in a single framework, a high frequency movement of state variables can be described within a regime-switching formulation. Also, this paper incorporates the time-varying intensity of jump into a regime-switching framework. The importance of jumps in derivatives pricing has been documented in many studies. Naik and Lee (1990) derive option prices where underlying assets exhibit discontinuous returns. Liu, Pan, and Wang (2005) find that uncertainty aversion toward rare events bring the smirk into the implied volatility surface. Todorov (2010) shows that jumps play an important role in pricing variance swaps. Drechsler and Yaron (2011) add jumps into the Long-run risk model, augmenting the size of variance risk premium.

In this paper, I show the capacity of MSM in explaining the prices of derivatives. The model economy validates a high level of variance risk premium and reproduces the smirk in the implied volatility surface. Further, the filtered time series show strong correlation between variance risk premium and equity risk premium. Sophisticated derivative pricing models will be designed with more frequently observed data in future research.

Long-run Risk and Term Structure of Equity Yields

Risks are spread out over the investment horizon. Researchers have tried to find the answer to the high premium of equity returns in the compensation for the risk over a long horizon. Daniel and Marshall (1997) suggest that measuring consumption data over a long horizon can resolve the issues of high premium and a low risk free rate. Parker and Julliard (2005) resolve size and value premium with a similar approach. Malloy, Moskowitz, and Vissing-Jorgensen (2009) take a deeper step into the micro-level data to measure the consumption risk of long-run stock holders, explaining cross sectional differences in returns. Since the advent of Long-run Risk (LRR) model by Bansal and Yaron (2004), many studies explain the importance of the risk over long horizons.
Still, due to the difficulties in statistical identification of Long-run risk,\textsuperscript{11} there is a controversy over whether the Long-run risk is actually priced. As a way to explain the value premium, Lettau and Wachter (2007) suggested a reduced form economy where short term cash flow risk is heavily priced, supported by the empirical findings of van Binsbergen, Brandt, and Koijen (2011) using prices of synthetic dividend strips.\textsuperscript{12}

It is natural to visualize the relation between risk and return over a long horizon through the term structure of yields. Term structure of equity yields is an up-to-date empirical observation. Using a new set of data from OTC and dividend futures, van Binsbergen, Hueskes, Koijen, and Vrugt (2011) compute equity yields, which is analogous to zero coupon bond yields for fixed assets.\textsuperscript{13} van Binsbergen, Hueskes, Koijen, and Vrugt (2011) document stylized facts for the term structure of equity yields: \textit{counter cyclical level} of equity premium, \textit{high variance} of short term premium, and \textit{inverted} term structure in the recent financial crisis in 2008 and 2009. The contribution of this article is to show that empirically observed features for the term structure of equity yields can be explained by the equilibrium of an endowment economy. Especially, I verify that the high volatility in the risky state implies the yield curve to be downward sloped.\textsuperscript{14}

Hansen, Heaton, and Li (2008) derive term structure of equity yield in a economy with Gaussian state variables, in contrast with the discrete changes in regimes of this paper. Lettau and Wachter (2011) extend the model of Lettau and Wachter (2007) and construct the term structure of equity, showing the inverted term structure can be obtained in average with some parameter values. However, this paper focuses on the dynamics for the term structure of equity yields. Building on He and Krishnamurthy (2012), Muir (2012) uses intermediary capital to model financial crises and reproduces the dynamic features of slope for equity yields: positive during normal time and negative during crises. The approach of this paper is different in that I utilize a consumption based asset pricing model to explain

\begin{itemize}
  \item \textsuperscript{11}See Hansen and Sargent (2010).
  \item \textsuperscript{13}Equity yields by van Binsbergen, Hueskes, Koijen, and Vrugt (2011) is different from cross-sectional expected returns. Many papers use the phrase of \textit{Term Structure of Equity} to describe the cross-sectional differences in the expected returns of dividend assets with different durations even though it is not related to the investment term or investment horizon.
  \item \textsuperscript{14}Fusari and Gonzalez-Perez (2012) show that the high volatility derives the term structure of variance risk premium to be downward sloped.
\end{itemize}
the features of the dynamic term structure of equity yields.

3 Asset Pricing Model

A regime-dependent bivariate process of consumption and dividend is specified. Given the recursive preference of a representative agent, the risk of regime switching is directly incorporated into the stochastic discount factor. With a Markovian property of the state variables of the model economy, prices of assets across different maturities are computed through recursion.

Notation

For simplicity, I define the following notation: $M_{i,j}$ is defined as the $(i,j)$ -th element of a matrix $M$. In the regime switching economy, many quantities depend solely on the regime. I use a bold character to denote a vector notation of the corresponding function of regime. That is, in the case that $f(\cdot)$ is a function of a random regime which takes a discrete value from 1 to $K$, a $K \times 1$ vector of $\left[ f(1) \cdots f(K) \right]'$ is denoted as $f$. I use $f(k)$ to denote $k$ -th element of $f$. That is, $f(k) = f(k)$. And, $1_K$ represents a $K$ - dimensional column vector of ones. To express operations in a vector space, I use the symbols of $f^x$, $\frac{f}{g}$, $\exp(f)$, $fg$, and $\text{diag}(f)$, defined as follows:

$$f^x = \left[ f(1)^x \cdots f(K)^x \right]'$$

$$\frac{f}{g} = \left[ \frac{f(1)}{g(1)} \cdots \frac{f(K)}{g(K)} \right]'$$

$$\exp(f) = \left[ \exp(f(1)) \cdots \exp(f(K)) \right]$$

$$fg = \left[ f(1)g(1) \cdots f(K)g(K) \right]'$$

$$\text{diag}(f) = \begin{bmatrix} f(1) & 0 & \cdots & 0 \\ 0 & f(2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(K) \end{bmatrix}.$$
3.1 Bivariate Process of Consumption and Dividend

Let $C_t$ and $D_t$ denote the consumption and the dividend at time $t$ and $c_t$ and $d_t$ the logarithms of $C_t$ and $D_t$, respectively. The logarithms of consumption and dividend follow a bivariate process, the parameters of which change according to a discrete Markov process $S_t \in \{1, \cdots, K\}$. That is, the number of possible values for $S_t$ is $K$. The realization of $S_t$ in the sequence of $\{S_\tau\}_{\tau=0}^\infty$ evolves according to the following transition probability matrix:

$$
\Pi = \begin{bmatrix}
\pi_{1,1} & \pi_{1,2} & \cdots & \pi_{1,K} \\
\pi_{2,1} & \pi_{2,2} & \cdots & \pi_{2,K} \\
\vdots & \vdots & \ddots & \vdots \\
\pi_{K,1} & \pi_{K,2} & \cdots & \pi_{K,K}
\end{bmatrix}
$$

(1)

where $\pi_{i,j} = \text{Pr} (S_t = j | S_{t-1} = i)$. Conditional on the realization of $S_t$, the bivariate process for the logarithms of consumption and dividend is written as:

$$
\begin{bmatrix}
c_t - c_{t-1} \\
d_t - d_{t-1}
\end{bmatrix} =
\begin{bmatrix}
\mu_c \\
\mu_d
\end{bmatrix} +
\begin{bmatrix}
\Sigma^{1/2}(S_t) \mathbf{Z} \\
J(S_t) \begin{bmatrix} B_c \\ B_d \end{bmatrix}
\end{bmatrix}.
$$

(2)

The logarithms of endowment growth are decomposed into three parts. The first part is constant.\textsuperscript{15} The second part is Gaussian shock. $\mathbf{Z}$ is a (2 by 1) standard Gaussian random vector:

$$
\mathbf{Z} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).
$$

(3)

Hence, $\Sigma(S_t)$ is the regime-dependent covariance matrix of Gaussian disturbance of the endowment process. The third part illustrates the risk of a rare disaster. $J(S_t)$ is a Bernoulli trial which has a value of 1 with the probability of $p(S_t)$ and a value of 0 with the probability of $1 - p(S_t)$. $p(S_t)$ is the time-varying likelihood of rare disasters, and $B_c$ and $B_d$ are the fixed size of disaster to logarithms of consumption and dividends, respectively.

The endowment process described above is quite general but difficult to be identified\textsuperscript{15} I explicitly put the expected growth to be constant across regimes since the dynamics of expected growth is hardly identified. See Hansen and Sargent (2010) for more discussion on this issue. Also, through this restriction, I can turn aside the predictability of consumption and dividend through price-to-dividend ratio observed in Bansal and Yaron (2004) but not in real world, as pointed out by Beeler and Campbell (2009). However, Jagannathan and Marakani (2011) shows that aggregate consumption and dividends are predictable with the extracted factors from price-to-dividend ratios of portfolio, supporting Long-run risk models.

\textsuperscript{15}
statistically. In what follows, I explain the structure I impose on the regime-switching economy.

### 3.1.1 Regime Transition Matrix

I adopt the MSM on the regime transition matrix of (1) to handle a large number of regimes effectively.\(^{16}\) The regime of \(S_t\) is driven by \(k\) basis binary regimes, \(\{S_{k,t}\}_{k=1}^{k}\). Each binary regime evolves independently, and their dynamics are identical except for time-scaling. Specifically, the regime \(k\) has a persistency that is a power function of the persistence of regime \(k-1\). Hence, the persistency cascades geometrically in the ordering of regimes.

Consider a Bernoulli trial of \(S\), which has a value of 0 with the probability of \(\lambda\) and a value of 1 with the probability of \(1 - \lambda\). The transition of \(S_{k,t}\) is depicted as follows:

\[
\begin{align*}
S_{k,t} & \text{ is drawn from } S \quad \text{with probability } \alpha_k \\
S_{k,t} &= S_{k,t-1} \quad \text{with probability } 1 - \alpha_k
\end{align*}
\]

That is, the transition matrix of \(S_{k,t}\) is written as follows:

\[
\Pi_k \equiv \begin{bmatrix} \pi_{k:0,0} & \pi_{k:0,1} \\ \pi_{k:1,0} & \pi_{k:1,1} \end{bmatrix} = \alpha_k \begin{bmatrix} \lambda & 1 - \lambda \\ \lambda & 1 - \lambda \end{bmatrix} + (1 - \alpha_k) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},
\]

where \(\Pr (S_{k,t} = j | S_{k,t-1} = i) = \pi_{k:i:j}\). The steady state distribution of \(S_{k,t}\) over \(\{0, 1\}\) is \(\{\lambda, 1 - \lambda\}\) for all \(k\). The only difference of \(S_{k,t}\) across \(k\) is the persistency parameterized by \(\alpha_k\).

Once a value of \(S_{k,t}\) is determined, it is expected that for \(1/\alpha_k\) periods the value would not be drawn from the Bernoulli trial of \(S\). The persistency of \(\{\alpha_k\}_{k=1}^{k}\) is modeled with the geometric cascade:

\[
1 - \alpha_k = (1 - \alpha_1)^{b_{k-1}}
\]

where \(\alpha_1 \in (0, 1)\) and \(b \in (1, \infty)\). The restriction of (5) shows that the persistency of regime, \(1 - \alpha_k\), cascades geometrically by the rate of \(b\) as \(k\) grows. Since \(\alpha_1 < \cdots < \alpha_F\), \(S_{k,t}\) is more persistent than \(S_{k',t}\) for \(k < k'\).

With \(k\) binary basis regimes, \(2^k\) possible states can be generated. The regime at time \(t\),

---

$S_t$, can be expressed as $k$-dimensional vector as follows:

$$S_t = (S_{1,t}, S_{2,t}, \cdots, S_{k-1,t}, S_{k,t}).$$

However, for the sake of computational convenience, I define a regime of $S_t$ as follows:

$$S_t = 1 + \sum_{k=1}^{\bar{k}} 2^{k-1} S_{k,t}. \quad (6)$$

Then, $S_t \in \{1, \cdots, K\}$ where $K = 2^{\bar{k}}$. Consistent with (6), the complete transition matrix of (1) is computed as follows:

$$\Pi = (\cdots ((\Pi_1 \otimes \Pi_2) \otimes \Pi_3) \cdots \otimes \Pi_{\bar{k}}),$$

where $\Pi_k$ is defined in (4) and $\otimes$ represents Kronecker-product. The complete transition matrix is pinned down by only three parameters of $\{\lambda, b, \alpha_1\}$.

In the empirical section, I use five frequencies, $\bar{k} = 5$, which generates $32 (= 2^5)$ regimes. A general transition matrix for 32 regimes has $992 (= 32^2 - 32)$ parameters, the estimation of which is not practically feasible. With the restriction of MSM, we can construct a transition matrix for arbitrarily many regimes with only three parameters.

Regime-dependent variance and the likelihood of rare disaster will be formalized in later subsections. Here, I define binary processes, the product of which implies the time-varying risk of the economy. $A(S_{k,t})$ is expressed as follows:

$$A(S_{k,t}) = A_0(1 - S_{k,t}) + A_1 S_{k,t} \quad (7)$$

such that

$$1 = \lambda A_0 + (1 - \lambda) A_1
A_0 < A_1
0 < A_0, A_1.$$ 

From the properties of $\{S_{k,t}\}_{k=1}^{\bar{k}}$, it follows that $A(S_{k,t})$ is a positive binary process, the unconditional expectation of $A(S_{k,t})$ is one, and $A(S_{k,t})$ is independent of $A(S_{k',t})$ when $k \neq k'$.

Lastly, I define a positive unit-mean process that governs the time-varying risk of the
overall economy:

\[ A(S_t) \equiv \prod_{k=1}^{\bar{k}} A(S_{k,t}) \]  

(8)

where \( A(S_{k,t}) \) is defined in (7).

### 3.1.2 Covariance of Gaussian Shocks

Now, I go for the stochastic variance of Gaussian shock. The regime-dependent covariance matrix for Gaussian shocks is specified as follows:

\[ \Sigma(S_t) = A(S_t)\bar{\Sigma} \]  

(9)

where

\[ \Sigma(S_t) = \begin{bmatrix} \sigma^2_c(S_t) & \sigma_{cd}(S_t) \\ \sigma_{cd}(S_t) & \sigma^2_d(S_t) \end{bmatrix} \]

and

\[ \bar{\Sigma} = \begin{bmatrix} \bar{\sigma}^2_c & \bar{\sigma}_{cd} \\ \bar{\sigma}_{cd} & \bar{\sigma}^2_d \end{bmatrix}. \]

The above specification implies that the variance of consumption moves with that of dividend by the same relative size. Also, the correlation does not change across regimes. Since \( A(S_t) \) is a unit mean process, the unconditional expectation of \( \Sigma(S_t) \) is \( \bar{\Sigma} \).\(^{17}\)

### 3.1.3 Likelihood of Rare Disasters

Recall the specification of the endowment process in (2). Conditional on that the current regime is \( S_t \), the rare disaster, which can wipe out logarithms of consumption and dividends by \( B_c \) and \( B_d \), respectively, can occur with probability \( p(S_t) \).

In an analogous way of modeling stochastic variance of Gaussian shocks, the dynamics of \( p(S_t) \) is specified as:

\[ p(S_t) = A(S_t)\bar{p}. \]  

(10)

Since \( A(S_t) \) is a unit mean process, the unconditional probability for the event of rare disaster

\(^{17}\)I put the sample covariance matrix of logarithms of consumption and dividend growth for \( \Sigma \) in the ML estimation.
is $\bar{p}$.\footnote{I use $\bar{p} = 1 - (1 - 3.63\%)^{1/12}$ as an unconditional monthly probability for the rare disaster to occur in the ML estimation. Barro and Ursua (2008) estimates the probability for consumption disasters as 3.63 times over a century. Gabaix (2012) also used this value for the calibration of parameters.} For $p(S_t) \not\geq 1$ for all $S_t$, we need to put a restriction on the relative size between $p$ and $A_1$ such that $\bar{p}(A_1)^{\bar{p}} < 1$. This condition is not restrictive since rare disasters happen rarely, implying $\bar{p}$ is small.

Up to this point, endowment processes of the economy are enumerated. In what follows, I illustrate the preference of the representative agent and derive the equilibrium with the specified endowment process.

### 3.2 Preference

The information of a representative agent is expressed as a filtration of $\mathcal{F} = \{\mathcal{F}_t | t = 0, 1, \cdots\}$ where $\mathcal{F}_t$ constitutes the information of the realization of consumption, dividends, and regimes up to time $t$, $\{C_r, D_r, S_r\}_{r=0}^t$.\footnote{I assume that the representative agent knows the regime. Many papers consider the case that an agent has incomplete information about the underlying regimes. See David (1997), Veronesi(1999, 2000, 2004), Lettau, Ludvigson, and Wachter (2008) and Ghosh and Constantinides (2011). Most of these models assume a small number of regimes where numerical integration is feasible. However, since a model with multiple binary regimes requires multidimensional numerical integration, the computational burden is extremely heavy. While Calvet and Fisher (2007) use SMM(simulated method of moments) as a detour, due to the rarity of rare disasters, simulation is not a legitimate substitute of integration for the learning model of the suggested economy. The learning model will be investigated in future research. In Section 7, I derive the stochastic discount factor with the extension of the model economy into the learning environment.} A representative agent evaluates a consumption process for infinite horizon with the recursive preference of Epstein and Zin (1989) and Weil (1990), based on Kreps and Porteus (1978). Let $\{C_r\}_{r=0}^\infty$ denote the optimal consumption process. Then, $V_t$, the continuation value at time $t$ for the consumption process $\{C_r\}_{r=t}$, is solved with a CES(constant elasticity of substitution) recursion:

$$V_t = \left[(1 - \delta) C_t^{1-\psi^{-1}} + \delta R_t^{1-\psi^{-1}}\right]^{\frac{1}{1-\psi}} \text{ when } \psi \neq 1$$

$$V_t = C_t^{1-\delta} R_t^\delta \text{ when } \psi = 1$$

where $R_t$ represents the certainty equivalent of $V_{t+1}$, the continuation value of the consumption plan at time $t + 1$, and the agent evaluates the certainty equivalent $R_t$ with the following risk adjustments:

$$R_t = \mathbb{E}_t\left[V_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}} \text{ when } \gamma \neq 1$$

$$R_t = \exp \mathbb{E}_t[\log V_{t+1}] \text{ when } \gamma = 1$$

\[18\]
where $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_t]$.

An agent with the above preference can be characterized by the three parameters of $\delta, \gamma, \psi$: where $\delta \in (0, 1)$ is the time discount factor, $\gamma \in (0, \infty)$ is CRRA (constant relative risk aversion) parameter, and $\psi \in (0, \infty)$ is EIS (elasticity of intertemporal substitution) parameter. When consumption plans are deterministic, the value of $\psi$ determines the elasticity of intertemporal substitution between today’s consumption and tomorrow’s consumption. In the case of time-additive preference, the intertemporal substitution is conflated with the risk aversion through the restriction that $\gamma = \psi^{-1}$. The proposed preference provides a convenient separation between risk aversion and the elasticity of intertemporal substitution, allowing us to investigate the effects of risk-appetite with a modest level for risk-free rate.

Then, $V_t$, the continuation value of the consumption process of $\{C_\tau\}_{\tau=t}^\infty$, can be solved as follows:

$$V_t = C_t v(S_t) \tag{13}$$

Given parameters for the endowment process and preference, the numerical value of $v(\cdot)$ can be found as a fixed point which satisfies (11), (12), and (13) simultaneously.\footnote{Marinacci and Montrucchio (2010) cover the solution of value process under general recursive preferences. Skiadas (2009) provides an excellent and rigorous explanation on the properties of recursive utility.} A detailed explanation on how to find numerical values is presented in Appendix A.

To derive the stochastic discount factor, I use the shadow valuation of a given consumption process as in Hansen, Heaton, Lee, and Roussanov (2007) and Hansen, Heaton, and Li (2008). Since the current continuation value is homogenous of degree one in the current consumption and the future continuation value, the continuation value of $V_t$ should satisfy the following condition due to Euler’s Theorem:

$$V_t = (MC_t) C_t + \mathbb{E}_t [(MV_{t+1}) V_{t+1}] \tag{14}$$

where

$$MC_t = (1 - \delta) V_t^{\psi^{-1}} C_t^{-\psi^{-1}}$$
$$MV_{t+1} = \delta V_t^{\psi^{-1}} \mathcal{R}_t^{\gamma^{-\psi^{-1}}} V_{t+1}^{-\gamma}.$$

Letting consumption be the numeraire, $W_t$, the wealth process, can be derived by dividing

the expression of (14) by $MC_t$:

$$W_t = \frac{V_t}{MC_t} = C_t + E_t \left[ \frac{MV_{t+1}MC_{t+1}}{MC_t} \frac{V_{t+1}}{MC_{t+1}} \right] = C_t + E_t \left[ \frac{MV_{t+1}MC_{t+1}}{MC_t} W_{t+1} \right].$$

Thus, $M_{t+1}$, the stochastic discount factor is expressed as:

$$M_{t+1} = \frac{MV_{t+1}MC_{t+1}}{MC_t} = \delta V_t^{\psi-1} \frac{R_t^{\gamma-\psi} V_t^{-\gamma}}{(1 - \delta) V_t^{\psi-1} C_t^{-\psi-1}} V_{t+1} \frac{C_t+1}{C_t} = \delta \frac{(C_{t+1})^{-\psi-1} (V_{t+1})^{\psi-1-\gamma}}{(R_t)^{\psi-1-\gamma}}\frac{C_t}{C_t^{\psi-1-\gamma}}.$$

(15)

Applying the regime switching structure and exploiting that $R_t$ is homogenous of degree one, I rewrite $M_{t+1}$ as:

$$M_{t+1} = \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\psi-1} \left( \frac{C_t+1 v_S(S_{t+1})}{R_t (C_{t+1} v_S(S_{t+1})))} \right)^{\psi-1-\gamma} \left( \frac{C_{t+1} v(S_{t+1})}{C_t \frac{C_t+1 v_S(S_{t+1})}{\kappa(S_t)}} \right)^{\psi-1-\gamma}.$$

(16)

where $\kappa (S_t) \equiv R_t \frac{C_{t+1} v_S(S_{t+1})}{C_t}$ and $\kappa (S_t)$ is evaluated as follows:

$$\kappa (S_t) = \begin{cases} E_t \left( \left( \frac{C_{t+1}}{C_t} v_S(S_{t+1}) \right)^{1-\gamma} \right)^{\frac{1}{1-\gamma}} & \text{when } \gamma \neq 1 \\ \exp E_t \left( \log \left( \frac{C_{t+1}}{C_t} v_S(S_{t+1}) \right) \right) & \text{when } \gamma = 1. \end{cases}$$

(17)

In expression (16), I can separate two sources of risk for the stochastic discount factor. The first part is directly related to one-period consumption growth, familiar from consumption based asset pricing literatures. The second part is related to the risk for shifts of the underlying regime.

Usually, the intertemporal marginal rate of substitution of Epstein-Zin preference is ex-
pressed with the return of aggregate consumption portfolio, which is a function of consumption growth and the wealth-to-consumption ratio. Actually, the components of the second part can be written in terms of the wealth-to-consumption ratio as follows:

\[ v(S_t) = \left( \frac{1}{1 - \delta} \left( \frac{W_t}{C_t} \right) \right)^\frac{\psi}{\psi - 1} \]

\[ \kappa(S_t) = \left( \frac{1}{\delta(1 - \delta)} \left( \frac{W_t}{C_t} - 1 \right) \right)^\frac{\psi}{\psi - 1} . \]

3.3 Price and Yield

With the stochastic discount factor derived in the previous section, the price of future consumption flow can be derived. For most assets, I can find the price through matrix multiplication.

3.3.1 Dividend Claims

Prices of Dividend Strips

The price of dividend strips for each maturity is derived recursively. Let \( (P_D)_n (S_t) \) denote the ratio of the price of dividend strips which pays \( D_{t+n} \) after \( n \) periods to the current dividend of \( D_t \) when the current regime is \( S_t \). That is, the current price of dividend strips with maturity \( n \) is \( D_t \left( \frac{P_D}{D} \right)_n (S_t) \). \( (P_D)_n (\cdot) \) can be computed with the following recursion:

\[
D_t \left( \frac{P}{D} \right)_n (S_t) = \mathbb{E}_t \left[ M_{t,t+1} D_{t+1} \left( \frac{P}{D} \right)_{n-1} (S_{t+1}) \right]
\]

\[
\left( \frac{P}{D} \right)_n (S_t) = \mathbb{E}_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{v(S_{t+1})}{\kappa(S_t)} \right)^{\psi - 1 - \gamma} \left( \frac{D_{t+1}}{D_t} \right) \left( \frac{P}{D} \right)_{n-1} (S_{t+1}) \right],
\]

with the boundary condition of \( (P_D)_0 (\cdot) = 1. \)

The recursion (18) can be simplified to matrix multiplication as follows:

\[
\left( \frac{P}{D} \right)_n = M_p \left( \frac{P}{D} \right)_{n-1},
\]
with the boundary condition of \((\frac{P}{D})_0 = 1_K\). \((\frac{P}{D})_n\) is the vectorized expression of \((\frac{P}{D})_n(\cdot)\) and the matrix multiplier \(M_p\) is defined as

\[
M_p = \delta \kappa^{\psi^{-1}} \Pi [p \exp (\gamma B_c - B_d + a) + (1_K - p) \exp(a)] u^{\psi^{-1} - \gamma},
\]

(20)

where

\[
a = -\gamma \mu_c + \mu_d + \frac{\gamma^2}{2} \sigma_c^2 + \sigma_d^2 - \gamma \sigma_{cd}.
\]

Return of Aggregate Dividend Claims

The return of aggregate dividend claims is derived from the regime-dependent price of dividend strips. \((\frac{P}{D})(\cdot)\), the ratio of the price of aggregate dividend claims to the current dividend, is simply the summation of the ratio of the price to current dividend for all dividend strips.

\[
\left(\frac{P}{D}\right)(S_t) = \sum_{n=1}^{\infty} \left(\frac{P}{D}\right)_n(S_t).
\]

In a vector notation,

\[
\begin{align*}
\left(\frac{P}{D}\right) &= \sum_{n=1}^{\infty} (\frac{P}{D})_n = \sum_{n=1}^{\infty} (M_p)^n 1_K = (I_K - M_p)^{-1} M_p 1_K
\end{align*}
\]

where \(M_p\) is defined as in (20).

With the price-to-dividend ratio of aggregate dividend claims, one-period return of the aggregate dividend claims is expressed as follows:

\[
R_{m,t+1} = \frac{D_{t+1} 1 + \frac{P}{D}(S_{t+1})}{D_t \frac{P}{D}(S_t)}
\]

\[
r_{m,t+1} \equiv \log R_{m,t+1} = \Delta d_{t+1} + \log \left( \frac{1 + \frac{P}{D}(S_{t+1})}{\frac{P}{D}(S_t)} \right)
\]

(21)

where the subscript of \(m\) represents the market portfolio. The equilibrium return (21) will be used in estimating parameters using ML.

Expected Dividend Growth

To derive the yield curve of dividend strips, we need the expected growth of dividends. Let \(G_n^d(S_t)\) denote the expected growth of dividend for the next \(n\) periods when the current
regime is $S_t$. $G^d_n(\cdot)$ can be computed with the following recursion:

$$ G^d_n(S_t) = \mathbb{E}_t \left[ \left( \frac{D_{t+1}}{D_t} \right) G^d_{n-1}(S_{t+1}) \right], \quad (22) $$

with the boundary condition of $G^d_0(\cdot) = 1$.

The recursion (22) can be simplified to matrix multiplication as follows:

$$ G^d_n = M_g G^d_{n-1}, $$

with the boundary condition of $G^d_0 = 1_K$. $G^d_n$ is the vectorized expression of $G^d_n(\cdot)$ and the matrix multiplier $M_g$ is defined as

$$ M_g = \Pi \left[ p \exp (-B_d + b) + (1_K - p) \exp(b) \right] \psi^{\psi-1} - \gamma, $$

and

$$ b = \mu_d + \frac{1}{2} \sigma_d^2. $$

**Yield Curve of Dividend Strips**

We complete the computation of $(P_D)_n(\cdot)$ and $G^d_n(\cdot)$. Then, $y^d_n(S_t)$, the yield of a dividend strip which matures after $n$ periods, is defined as the discount rate to equate the discounted expected value to the current price:

$$ D_t \left( \frac{P}{D} \right)_n(S_t) = \mathbb{E}_t [D_{t+n}] \exp(-ny^d_n(S_t)) $$

$$ \left( \frac{P}{D} \right)_n(S_t) = \mathbb{E}_t \left[ \frac{D_{t+n}}{D_t} \right] \exp(-ny^d_n(S_t)) $$

$$ \left( \frac{P}{D} \right)_n(S_t) = G^d_n(p_t) \exp(-ny^d_n(S_t)) $$

$$ y^d_n(S_t) = \frac{g^d_n(S_t) - p^d_n(S_t)}{n} $$

where $g^d_n(S_t) = \log(G^d_n(S_t))$ and $p^d_n(S_t) = \log\left(\left( \frac{P}{D} \right)_n(S_t)\right)$. Then, it is natural to define $y^{f,d}_n(\cdot)$, the forward yield of the dividend strips from $n-1$ to $n$, as follows:

$$ y^{f,d}_n(\cdot) = ny^d_n(\cdot) - (n-1)y^d_{n-1}(\cdot) $$

$$ = - \left( p^d_n(\cdot) - p^d_{n-1}(\cdot) \right) + \left( g^d_n(\cdot) - g^d_{n-1}(\cdot) \right) \quad (23) $$
n-Asymptotic Forward Yields of Dividend Strips

Following the expression (23), I vectorize the expression \( y_{n}^{f,d}() \) as follows:

\[
y_{n}^{f,d} = -\log\left( \frac{P_{Dn}}{P_{Dn-1}} \right) + \log\left( \frac{G_{n}^{d}}{G_{n-1}^{d}} \right) = -\log\left( \frac{(M_{p})^{n}1_{K}}{(M_{p})^{n-1}1_{K}} \right) + \log\left( \frac{(M_{g})^{n}1_{K}}{(M_{g})^{n-1}1_{K}} \right),
\]

the asymptotic limit of which is determined by the principal eigenvalues of \( M_{p} \) and \( M_{g} \), which shall be verified in what follows.

Let \( M \) be a general \( K \times K \) full rank matrix with positive elements. Then, I find \( K \) pairs of eigenvalues and eigenvectors of \( M \). Let \( \{ (\lambda_{k}, v_{k}) \}_{k=1}^{K} \) denote the set of pairs of eigenvalues and eigenvectors. Without loss of generality, \(|\lambda_{1}| > |\lambda_{2}| > \cdots > |\lambda_{K}|\).\(^{21}\) Since every element of \( M \) is positive, I apply the Perron-Frobenius theorem, implying that \( \lambda_{1} > 0 \) and \( v_{1} > 0_{K} \).

\[ 1_{K} \] is rewritten as a linear combination of eigenvectors:

\[
1_{K} = \sum_{k=1}^{K} c_{k} v_{k}
\]

where \( c_{k} \) is a coefficient. Then, the recursive multiplication of \( M \) on a vector of ones can be written as:

\[
M^{n}1_{K} = M^{n}\left( \sum_{k=1}^{K} c_{k} v_{k} \right) = \sum_{k=1}^{K} c_{k}(M^{n})v_{k} = \sum_{k=1}^{K} c_{k} \lambda_{k}^{n} v_{k} = \lambda_{1}^{n} \sum_{k=1}^{K} c_{k} \left( \frac{\lambda_{k}}{\lambda_{1}} \right)^{n} v_{k}.
\]

As \( n \to \infty \), the asymptotic limit of the marginal decaying rate through recursive multiplication:

\[
\frac{M^{n}1_{K}}{M^{n-1}1_{K}} = \frac{\lambda_{1}^{n} \sum_{k=1}^{K} c_{k} \left( \frac{\lambda_{k}}{\lambda_{1}} \right)^{n} v_{k}}{\lambda_{1}^{n-1} \sum_{k=1}^{K} c_{k} \left( \frac{\lambda_{k}}{\lambda_{1}} \right)^{n-1} v_{k}} = \lambda_{1} \sum_{k=1}^{K} c_{k} \left( \frac{\lambda_{k}}{\lambda_{1}} \right)^{n} v_{k} \rightarrow \lambda_{1} 1_{K}
\]

and, in logarithms,

\[
\log\left( \frac{M^{n}1_{K}}{M^{n-1}1_{K}} \right) \rightarrow \log(\lambda_{1}) 1_{K}.
\]

That is, the decaying rates converge uniformly across all elements.\(^{22}\)

---

\(^{21}\)I neglect boundary cases where inequalities are replaced with equalities.

\(^{22}\)This result does not depend on whether the initial vector is a vector of ones. That is, for any \( v \in \mathbb{R}^{K}, \)
I apply the result (25) to the forward yield of dividend strips. Let \( \{(\lambda^p_k, v^p_k)\}_{k=1}^K \) and \( \{(\lambda^g_k, v^g_k)\}_{k=1}^K \) denote the set of pairs of eigenvalues and eigenvectors of \( M_p \) and \( M_g \), respectively. Without loss of generality, I assume that the eigenvalues are ordered in decreasing absolute values, \(|\lambda^p_1| > |\lambda^p_2| > \cdots > |\lambda^p_K|\) and \(|\lambda^g_1| > |\lambda^g_2| > \cdots > |\lambda^g_K|\). Since every element of \( M_p \) and \( M_g \) is positive, the result (25) is applicable. Recalling the expression (24), I find the asymptotic limit of forward yields as follows:

\[
y_n^{f,d} \to (-\log(\lambda^p_1) + \log(\lambda^g_1)) \mathbf{1}_K.
\]

Hence, the asymptotic forward yield of dividend strips is constant and does not change across regimes. When forward yield converges, the yield should converge to the limit of forward yield. That is, long forward and dividend strips yield can never fall.\(^{23}\)

### 3.3.2 Bonds

I derive the price of zero coupon bonds across different maturities, which would be computed recursively as in the case of dividend strips. Let \( P^z_n(S_t) \) denote the price of real zero coupon bond which pays 1 unit of consumption good after \( n \) periods when the current regime is \( S_t \). The superscript of \( z \) represents a zero coupon bond. \( P^z_n(\cdot) \) can be computed with the following recursion:

\[
P^z_n(S_t) = \mathbb{E}_t \left[ M_{t,t+1} P^z_{n-1}(S_{t+1}) \right]
\]

\[
P^z_n(S_t) = \mathbb{E}_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{v(S_{t+1})}{\kappa(S_t)} \right)^{\psi-1-\gamma} \right] P^z_{n-1}(S_{t+1})
\]

(27)

with the boundary condition of \( P^z_0(\cdot) = 1 \).

The recursion (27) can be simplified to matrix multiplication as follows:

\[
P^z_n = M_d P^z_{n-1}
\]

(28)

with the boundary condition of \( P^z_0 = \mathbf{1}_K \). \( P^z_n \) is the vectorized expression of \( P^z_n(\cdot) \) and the matrix multiplier \( M_d \) is defined as

\[
M_d = \delta \kappa^{-\gamma} \Pi \left[ \exp (\gamma B_c + \mathbf{1}_K - p) \exp(c) \right] \psi^{\psi-1-\gamma},
\]

\( M_d \to \lambda_1 \mathbf{1}_K \) as \( n \to \infty. \)

\(^{23}\)This sentence is the equity version of *Long Forward and Zero-Coupon Rates Can Never Fall* by Dybvig, Ingersoll, and Ross (1996). Similar results were shown by Hansen, Heaton, and Li (2008) in an economy where state variables are evolved with Gaussian shocks.
and
\[ c = -\gamma \mu_c + \gamma^2 \frac{\sigma_c^2}{2}. \]

The logarithms of one-period risk free return is defined as:
\[ r_{f,t} = \log \left( \frac{1}{P_{t}^2(S_{t-1})} \right), \tag{29} \]
which will be used to estimating parameters using ML.

### 3.3.3 Variance Risk Premium

For the recent decade, the market size of variance trading has increased substantially. Commonly traded assets are instruments using VIX, a ticker symbol of Chicago Board Options Exchange Market Volatility Index.\(^{24}\) Often referred to as the fear index or the fear gauge, VIX represents the market’s valuation of stock market volatility over the next one month period. The gap between VIX and realized volatility has been noticed and named as the variance risk premium.\(^{25}\) I study the implications of the model economy on the variance risk premium – specifically, the difference in physical expectation and risk-neutral expectation of the return variations – where the equilibrium market return is generated as in (21).

Since I will set the decision interval of the model economy to be monthly, I use a discrete analogue of quadratic variation in the continuous time set-up as variation of returns over a month. I decompose the return equation (21) as follows:

\[ r_{m,t+1} = \underbrace{\mu_d} + \underbrace{\sigma_d (S_{t+1}) \varepsilon_t} - \underbrace{J(S_{t+1}) B_d} + \underbrace{\log \left( \frac{1 + P_{t} (S_{t+1})}{P_{t} (S_{t})} \right)}, \]

and define the quadratic variation as follows:

\[ QV_{t+1} = \underbrace{\sigma_d^2 (S_{t+1})} + \underbrace{J(S_{t+1}) B_d^2} + \underbrace{\left( \log \left( \frac{1 + \left( \frac{P_{t}}{P_{D}} \right) (S_{t+1})}{\left( \frac{P_{t}}{P_{D}} \right) (S_{t})} \right) \right)^2}. \tag{30} \]

\(^{24}\)The VIX is interpreted as the square root of the par variance swap rate for a 30 day term initiated today. A variation of returns can be statically replicated through blending puts and calls with continuous strike prices and maturities. The CBOE VIX is quoted as the value of portfolios of puts and calls, which approximates the replication with discrete strike prices and maturities.

I define the variance risk premium as the difference between physical expectation and risk-neutral expectation of quadratic variation:\(^{26}\)

\[
VRP(S_t) = E_t^Q [QV_{t+1}] - E_t [QV_{t+1}]
= \frac{E_t [M_{t+1} QV_{t+1}]}{E_t [M_{t+1}]} - E_t [QV_{t+1}],
\]

the first part of which is equivalent to \(VIX^2/12\) and the second part of which is the expected return variation, not directly observable. The expression of quadratic variation in (30) manifests good reasons for the sizable amount of variance risk premium. Consider a VIX instrument the final payoff of VIX at maturity is the quadratic variation. The payoff is positively correlated with the high variance of Gaussian shock, the shifts of regimes, and the realization of rare disasters. Hence, such a VIX instrument can provide an investor with the protection against those tail risks, leading to a sizable amount of variance risk premium.

### 3.3.4 Put Options

Modeling volatility and jumps is a crucial part for pricing put or call options due to the nature of the skewed payoff. Especially, for a deep OTM put option, the strike price of which is far below the current price of the underlying, the positive cashflow at the maturity would be highly probable conditional on that the underlying regime is switched into a riskier state or the rare disaster is realized over the life of the put. Thus, the payoff of deep OTM puts provides a hedge against those tail risks. Naturally, we can expect the high prices of deep OTM puts, compared with those computed through the pricing formula by Black and Scholes (1973).

Since the payoff of put option depends on the level of the underlying, the regime, \(S_t\), is not sufficient to track the prices of put options. We need to increase the state space by adding \(D_t\).\(^{27}\)

Let \(P_{n,\text{put}}(S_t, D_t, K)\) denote the price of the put option which matures in \(n\) periods and pays \(max\{P_{t+n} - K, 0\}\) where \(P_{t+n} = D_{t+n} \left( \frac{P}{D} \right) (S_{t+n})\). The price of put options can be

\(^{26}\)The risk premium is usually computed as the physical expectation minus risk-neutral expectation. I do not follow this convention not to carry a negative sign in front of the variance risk premium.

\(^{27}\)This is equivalent to adding the price of aggregate dividend claims because the price-to-dividend ratio of the aggregate dividend claims is a function of \(S_t\).
computed recursively as follows:\(^{28}\)

\[
P_{n}^{\text{put}}(S_t, D_t, K) = \mathbb{E}_t \left[ M_{t+1} P_{n-1}^{\text{put}}(S_{t+1}, D_{t+1}, K) \right] \\
= \mathbb{E}_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{\nu(S_{t+1})}{\kappa(S_t)} \right)^{\psi^{-1}-\gamma} P_{n-1}^{\text{put}}(S_{t+1}, D_{t+1}, K) \right],
\]

with the boundary condition \(P_{0}^{\text{put}}(S_t, D_t, K) = \max\{P_t - K, 0\}\) where \(P_t = D_t \left( \frac{P}{D} \right) (S_t)\).

4 Estimation

The estimation of parameters for the endowment process is performed with matching moments and maximizing the likelihood. Annual endowment processes are used to identify the unconditional moments of endowment processes. The evolution of time-varying risk is identified through maximizing the likelihood of monthly market excess returns. Parameters for the preference and the leverage in the event of rare disasters are calibrated.

4.1 Data and Estimation Process

I use the time series of annual endowments to identify the unconditional moments of model economy. Details on how to construct annual consumption and dividend growth are explained in Appendix B. Table 1 summarizes the sample moments of growth of consumption and dividends.\(^{29}\) Reported numbers in Table 1 are similar to those used in other studies.\(^{30}\)

The parameters related to unconditional first and second moment of endowment process, conditional on no disasters, are calibrated to monthly counterparts of the sample moments of annual endowment growth. That is, the constant part of (2), \(\begin{bmatrix} \mu_c \\ \mu_d \end{bmatrix}\), and the unconditional covariance in (9), \(\Sigma\), are matched to the monthly counterparts of the sample moments of the annual endowment process.

For the unconditional probability of rare disaster in (10), \(\overline{p}\), I use the data in Barro and Ursua (2008). Barro and Ursua (2008) identify the consumption disasters, more than 10% drops in the level of aggregate consumption over large panel data. Figure 1 plots consumption

\(^{28}\)Computation will be done through a numerical integration.

\(^{29}\)I use a simple summation of monthly cash dividends as annual dividends, not considering market reinvestment strategy in other studies(van Binsbergen and Koijen, 2010). With the market reinvestment strategy, the variation of dividends is boosted up, but the correlation between consumption and dividend growth becomes practically zero, which is not proper for pricing dividend assets with consumption based asset pricing.

\(^{30}\)For example, see Bansal and Yaron (2004) or Calvet and Fisher (2007).
disasters over 35 countries from 1870-2006 and the unconditional likelihood of rare disasters is estimated as 3.63 times over a century. I convert this value for the unconditional likelihood of rare disasters, $\bar{p}$.\textsuperscript{31}

Specifically, I match the unconditional moments with the sample moments as follows:

\[
\begin{bmatrix}
\mu_c \\
\mu_d
\end{bmatrix} = \begin{bmatrix}
\frac{1.85\%}{12} \\
\frac{0.93\%}{12}
\end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix}
\frac{(2.18)^2\%^2}{12} & 0.66 \cdot \frac{2.18 \cdot 11.23\%^2}{12} \\
0.66 \cdot \frac{2.18 \cdot 11.23\%^2}{12} & \frac{(11.23)^2\%^2}{12}
\end{bmatrix}
\]

\[
\bar{p} = 1 - (1 - 3.63\%)^{\frac{1}{12}}
\]

The evolution of time-varying risk is estimated with monthly returns. I use the value weighted return of CRSP Stock Market Indexes\textsuperscript{32} as aggregate market return and one-month Treasury bill rate from Ibbotson Associates as risk-free return. The monthly return data cover 1020 months from Jan/1927 to Dec/2011. Table 2 depicts the summary stats of the return data, demonstrating negative skewness and thick-tails of excess returns.

The parameters governing the time-varying risk of $A(S_t)$, $\{A_0, b, \alpha_1, \lambda\}$, and the size of disaster for dividends, $B_d$, are estimated with maximizing the likelihood of the market excess returns. Since most studies identify disasters in terms of consumption or GDP, I do not use the estimated size of those disasters for the size of dividend disasters. Rather, I let the return data to reveal the size of dividend disasters. The set of parameters estimated with ML is $\{A_0, b, \alpha_1, \lambda, B_d\}$, and the estimation results are reported in the following subsection.

The remaining parameters are calibrated: the preference of the representative agent, $\delta, \gamma, \psi$, and the leverage of dividends disasters relative to consumption disasters, $\phi$. I set CRRA and EIS parameters to be identical to those in Bansal and Yaron (2004). Also, as in Abel (1999), I interpret dividends as levered assets, more susceptible to rare disasters than consumption, $B_d > B_c$. The leverage of $\phi = \frac{B_d}{B_c}$ is calibrated so that model-implied moments are matched with the sample moments in Table 2.\textsuperscript{33} Calibrated values for these parameters are as follows:

\textsuperscript{31}Gabaix (2012) also uses this value to calibrate the model economy.

\textsuperscript{32}This includes NYSE, AMEX, NASDAQ, and ARCA.

\textsuperscript{33}The interpretation of leverage differs across many studies. In Abel (1999), the dividend growth is shocked by the leverage multiple of the innovation on the consumption growth. In Bansal and Yaron (2004), the expected dividend growth is shocked by the leverage multiple of the shock on the expected consumption growth. In this article, the dividend is wiped out by the leverage multiple of the innovation on the consumption in the event of rare disasters.
• Preference: $\gamma = 7.5$, $\psi = 1.5$, $\delta = (0.98)^{1/2}$.

• Leverage: $\phi = 3.0$.

4.2 ML Estimation

I find parameters of $\{A_0, b, \alpha_1, \lambda, B_d\}$ using monthly excess returns of a U.S. equity index over one-month T-bill rate from Jan/1927 - Dec/2011. In the equilibrium of the model economy, the excess return has the following closed form:

$$r_{m,t} - r_{f,t} = \Delta d_t + \log \left( \frac{1 + \left( \frac{P}{D} \right) (S_t)}{\left( \frac{P}{D} \right) (S_{t-1})} \right) + \log (P^c_t(S_{t-1})). \quad (32)$$

Given $\{S_t, S_{t-1}\}$, the last two terms of (32) are constant and we know the distribution of the first term, $\Delta d_t$. Thus, once we solve out the equilibrium with given parameters, the conditional likelihood for the realized excess return of $r_{m,t} - r_{f,t}$ is available. It is straightforward to extend the standard filter by Hamilton (1989) to get the closed form of the log likelihood for the whole sample. Details on how to compute the log likelihood are explained in Appendix C.

A key parameter of MSM is the number of frequencies, $k$. Table 3 reports the ML estimation results for each possible value of $k$. The effects of $k$ on $\alpha_1$, the inverse of which is the duration of the most persistent regime, and $b$, the inverse of which is the geometric decay rates of the persistency, are similar as in Calvet and Fisher (2007). As $k$ increases, $\alpha_1$ and $b$ tend to decrease. That is, with more frequency, the duration of the most persistent regime is inclined to be lengthened and the persistency of basis regimes cascades at a slower rate. The size of disaster is quite stable across different number of frequencies.

The optimal level of $k$ is naturally chosen by listening to what data tell us. Within the range of $k \in \{1, \cdots, 6\}$, the log likelihood is maximized at $k = 5$.\(^{34}\) Table 4 reports the results of Likelihood Ratio test by Vuong (1989) on the null that the model with $k = 5$ does not explain the data better than other models with $k' = 1, 2, 3, 4, 6$. We can reject the null for $k' = 1, 2$ with a reasonable level of significance. However, we cannot reject the null for $k' = 3, 4, 6$.

However, I choose $k = 5$ using AIC and BIC.\(^{35}\) For the remainder of the paper, I will use

---

\(^{34}\)I also estimate the observations with $k = 7, 8$ and obtain a lower level of log likelihood compared with that with $k = 5$.

\(^{35}\)Since the number of parameters is constant for $k = 2, 3, 4, 5, 6$, the ordering of models in the log-likelihood is identical to that in AIC or BIC.
ML estimates with \( \bar{k} = 5 \) as benchmark parameters.

Implication of estimated parameters with \( \bar{k} = 5 \) on the time-varying risk is as follows. The expected durations of five basis regimes are 87, 60, 42, 29, and 20 months in the increasing order of \( k \). The regime-shifts of each basis regime will increase the risk by 2.5 times. Hence, the size of risk in the most risky regime is almost 100 times of that in the most stable regime.

### 4.3 Inference on Time-Varying Risk and Occurrence of Rare Disasters

In the model economy, the time variation of risk is derived by the dynamics of the product of multiple binary processes, \( A(S_t) \). Since the underlying regime is hidden to the econometrician, s/he does not observe the time series of the time-varying risk, \( A(S_t) \). However, making Bayesian inferences over the hidden regimes through observing the time series of excess returns, \( I_t = \{r_{m,\tau} - r_{f,\tau}\}_{\tau=0}^t \), the econometrician can extract the information on the realization on the time-varying risk, \( A(S_t) \).

The econometrician uses either the previous returns, \( I_{t-1} \), or the whole returns, \( I_T \), in making Bayesian inferences over hidden regimes, usually referred to prediction, or smoothing, respectively. I use the filter by Hamilton (1989) and the smoother by Kim (1994).

I extract the time-varying risk as follows:

\[
A_{t,\text{prediction}} = \sum_{S_t} (A(S_t) \Pr(S_t|I_{t-1}))
\]

and

\[
A_{t,\text{smoothing}} = \sum_{S_t} (A(S_t) \Pr(S_t|I_T))
\]

Figure 2 shows the time series of predicted time-varying risk and smoothed time-varying risk. We observe the spikes of risk in the Great Depression, the Second World War, Oil Shocks, and the recent Financial Crisis.

Figure 3 shows the Ex-Ante and Ex-Post probability for a rare disaster to have occurred, \( \mathbb{E} [J_t = 1|I_{t-1}] \) and \( \mathbb{E} [J_t = 1|I_T] \), respectively. The Ex-Ante probability is \( \bar{p} \) times of \( A_{t,\text{prediction}} \). Since I use the unconditional probability for a rare disaster to occur 3.63 times over a hundred years, the econometric specification would identify 3 or 4 extreme outliers as the occurrence of rare disasters. We can observe a couple of peaks of the Ex-Post probability, \( \mathbb{E} [J_t = 1|I_T] \), corresponding to the Great Depression, the Second World War, the 1987 Crash, the Russian Crisis, and the recent Financial Crisis. However, as I noted as a...
limitation in Section 2, identified disasters are not perfectly aligned with the sudden drops in the endowment process. This mismatch should somewhat be expected from the popular criticism of rare disaster models; We have not observed a sufficient number of endowment disasters in the US data. Especially after the Second World War, endowment processes are quite smooth so far.

Even though the model economy has the features of rare disasters, the simulated endowment series is not far from the observed consumption and dividend data. Table 5 compares the summary stats of simulated annual series of consumption and dividends with those of observed data. Since I impose sample moments to be the first two moments of endowments conditional on no disaster, model-implied first moment is a little lower than the sample mean and model-implied second moment is a little higher than sample variance. Nevertheless, those differences are minor. The sample moments of real data are within the 80% confidence interval of the distribution of simulated sample moments with 80 years of data. In Section 6, I provide an alternative estimation for a regime-switching economy using the quarterly consumption process and show some of the results are robust to the different set of data.

4.4 Test on the Time-Varying Likelihood of Rare Disasters

Distinguished features of the model are the inclusion of rare disasters and the comovement of volatility and the likelihood of rare disasters. I test whether a model with time-varying likelihood of rare disasters performs better than other competing models. Specifically, I put the following restrictions in the estimation and perform the Likelihood Ratio test by Vuong (1989).

\[
p(S_t) = \begin{cases} 
  A(S_t)p : & \text{SI, stochastic intensity} \\
  \bar{p} : & \text{CI, constant intensity} \\
  0 : & \text{NI, no intensity}
\end{cases}
\]

Table 6 reports ML estimates with a different specification on the risk of rare disasters and the results of the Likelihood Ratio test by Vuong (1989). I restrict the number of frequencies as \( \tilde{k} = 5 \) to compare ML estimates across SI, CI, and NI. The sizes of dividend disaster do not vary much on whether the intensity is constant or stochastic. The remarkable result is that the likelihood is maximized with the original specification of stochastic intensity. Furthermore, the Likelihood Ratio tests by Vuong (1989) show statistically significant results in favor of the time-varying likelihood of rare disasters.
The difference is not only statistically significant but also economically meaningful. Table 7 reports the model implied moments with a different specification on the risk of rare disasters. Only with stochastic intensity of rare disasters, the economy can reproduce the high premium and excessive volatility of market returns. Without the time-varying likelihood of rare disasters, the agent is not concerned about the effect of time-varying risk as much. Consequently, the price of market portfolio becomes stable and does not need to be heavily discounted.

5 Asset Pricing Implications

5.1 High Premium and Excessive Volatility of Returns for Aggregate Dividend Claims

Since Mehra and Prescott (1985) proposed the equity premium puzzle – the inability of standard asset pricing models to rationalize high premium of market return –, generating high equity premium has been a long-lasting challenge in the consumption based asset pricing literature. Furthermore, as Shiller (1981) pointed out, the difference between the volatility of dividends and that of returns for aggregate dividend claims is hard to reconcile with a rational behavior of investors.

Extending the static rare disaster model by Barro (2006), recent studies show that a model with time-varying risk of rare disasters can reproduce these features of financial market. In this paper, I compound the time-varying risk of rare disasters with stochastic volatility. This joint specification intensifies the effect of time-varying risk, boosting the premium and volatility of market returns.

Since it is already well-known that the Rare Disaster model can explain equity premium with CRRA level less than 5, readers may wonder why the economy needs the stochastic volatility in addition to the time-varying likelihood of rare disasters. However, the size of consumption disasters used in this paper is relatively small, compared with the size used in other literature. In this paper, the size of disaster for dividend, $B_d$, is -0.26 and that for consumption, $B_c$, is -0.09 (=-0.26/3). Barro and Ursua (2008) identify consumption disasters as more than 10% drops in the level of consumption and report the average of consumption disasters as -0.24 (=-log(0.78)), 2.5 times $B_c$. Both of Gabaix (2012) and Wachter (2012) use

\[36\] Drechsler and Yaron (2011) also model the joint dynamics of volatility and jump. However, in this paper, I restrict jumps as rare disasters by setting the unconditional likelihood of the arrival of jumps to be very small.
the distribution of consumption disasters by Barro and Ursua (2008) for their calibration. Hence, in the model economy of this paper, the time-varying likelihood of rare disasters alone is not enough. In addition to the stochastic volatility, we also need a relatively high level of risk aversion. With the CRRA of $\gamma = 7.5$, it is necessary to separate risk aversion from intertemporal rate substitution for a reasonable size of risk free rate.

To show that it is necessary to jointly specify stochastic volatility and stochastic intensity and to use the recursive preference by Epstein and Zin (1989), I report the implied moments of models with different specifications on endowments or preference in Table 8. If I shut off either stochastic volatility or stochastic intensity by using the following specifications:

\[
\Sigma(S_t) = \Sigma, \text{ CV: constant volatility} \tag{35}
\]

\[
p(S_t) = \bar{p}, \text{ CI: constant intensity} \tag{36}
\]

the economy does not reproduce the high premium and volatility of the market return. As the stochastic discount factor of (16) implies, the regime switching risk is not directly priced with time-additive preference, resulting in low premium and volatility of market returns.

With the joint effect of stochastic volatility and time-varying likelihood of rare disasters, the representative agent is concerned severely about the regime shifts of the economy. Since dividend assets are highly exposed to the risk of regime shifts, high returns are required for the aggregate dividend claims. Also, the price of aggregate dividend claims becomes vulnerable to the evolution of regimes, leading to excessive volatility in the return of aggregate dividend claims.

### 5.2 Return Predictability

Since the inspiring work by Campbell and Shiller (1988), researchers have striven to explain the predictability of long-horizon return through the price-to-dividend ratios.\(^{37}\) The following regression equation is commonly used to show the predictability of returns:

\[
\sum_{\tau=t+1}^{t+h_p} r_\tau = a + b \log \left( \frac{P_t}{D_t} \right) + \varepsilon, \tag{37}
\]

where $t$ is measured annually. $P_t$ is the price of market portfolio at the end of year $t$ and $D_t$ is computed as explained in Appendix B. $r_t$ is the summation of logarithms of monthly market return for the year $t$.

Table 9 reports regression results of (37) using real return data and simulated return data. The population column shows the regression results using all of the simulated annual series. The distribution of $b$ and $R^2$ are obtained by performing a regression for non-overlapped simulated 80 years annual series repeatedly. The model economy replicates the empirically observed return predictability fairly well; both the size of coefficients and the magnitude of $R^2$.

Where does this predictability come from? The regime-switching formulation provides a simple and clear explanation on the return predictability: The shift of each basis regime affects the expected return of aggregate dividend claims and the price-to-dividend ratio in the opposite direction. As a result, the low level of price-to-dividend ratio is associated with the high expected return in the future. As a confirmation of this explanation, I test the following regression equation:

$$
\log \left( \left( \frac{P}{D} \right) (S_t) \right) = a + \sum_{k=1}^{K} b_k S_{k,t} + \varepsilon \tag{38}
$$

$$
\mathbb{E}[r_{m,t+1}|S_t] = a + \sum_{k=1}^{K} b_k S_{k,t} + \varepsilon \tag{39}
$$

Table 10 reports the regression estimates of (38) and (39). Negative coefficients of (38) show that high risk regimes lower the price-to-dividend ratio. Positive coefficients of (39) illustrate that high returns are required in high risk regimes. That is, the shifts of regime into a risker state push down the price-to-dividend ratio and push up the expected return, yielding the predictability of return through the price-to-dividend ratio.

Table 11 provides the summary statistics for the logarithms of annual price-to-dividend ratio. Compared with real data, the simulated price-to-dividend series shows the features of lower level and less volatility. However, these disparities are not far off those of other competing models.

$^{38}$Since the decision interval is monthly in our model, I convert the simulated monthly data to annual time series of return and price-to-dividend ratios. I simulate $10^6$ monthly data to convert those into annual returns and annual valuation ratios.
5.3 Yield Curve of Dividend Strips

The model economy captures interesting features of the dynamic yield curve of dividend strips. As I examine in section 3.3.1, *long forward and dividend strips yield can never fall*. Since forward yields converge faster than yields, for the sake of convenience, I present forward yields in most figures. Figure 4 shows the average forward yield curve with 2SD bounds. On average, the forward yield curve is upward sloping, which is a general feature for the yield curve of dividend strips over the 2000s as shown in van Binsbergen, Hueskes, Koijen, and Vrugt (2011).

It would be more interesting to test the cyclical features of the yield curve. Figure 5 shows the forward yield curve for the low risk regime, $S_{k,t} = 0$ for $k = 1, \ldots, 5$, and for the high risk regime, $S_{k,t} = 1$ for $k = 1, \ldots, 5$. The yield curve is upward sloping in the low risk regime and downward sloping in the high risk regime. Hence, in the model economy, the slope of yield curve of dividend strips shows procyclical feature. Empirically observed yield curve is shown in Figure 6. The short-term yield is lower than the long-term yield for most of the time. However, the relation is flipped during the financial crisis from 2008-2009.

The model economy can reproduce this procyclical movement of the slope.

Among two key ingredients of time-varying risk, stochastic volatility and stochastic intensity, which derives the procyclical slope of equity yields? To answer this question, I shut off either stochastic volatility or stochastic intensity as specified in (35) or (36). Figure 7 shows the yield curve of dividend strips of high and low risk regimes with the restriction of in (35) or (36). High and low risk regimes are identically defined as specified before. The upper plot shows yield curves only with stochastic intensity. The lower plot depicts yield curves only with stochastic volatility. The yield curve is higher in the high intensity regime in the upper plot. However, for short-maturity, downward slope does not appear to be pronounced due to the lower growth rates from high probability for a rare disaster to occur. In contrast, short-term yield is very high in the high volatility regime in the lower plot. A sharp decline of yield is driven from the high temporary volatility shocks. With an alternative estimation of the endowment process in Section 6, this explanation will be reconfirmed. The underlying intuition on the cyclical feature is clear. The short duration assets are heavily discounted.

---

39During the recent crisis of 2008-2009, the yield curve of dividend strips had been negatively sloped as documented by van Binsbergen, Hueskes, Koijen, and Vrugt (2011).

40The values in Figure 6 is not directly comparable with those in Figure 5. Nominal risk free rates and expected dividend growths should be incorporated to convert values in Figure 6 to those in Figure 5. However, the effect of those are relatively small for short-term and the flipped feature is robust to those adjustments. Thus, I just show two figures for qualitative comparison.
when the current volatility is high – due to the increased covariance between consumption and dividends. With the mean-reverting property of volatility process, the covariance over a longer horizon would decrease, generating the downward sloped term structure of equity yields.\footnote{Nonetheless, I do not assert that every model with stochastic volatility can generate the downward sloped yield curve in the high volatility state. Even in the high volatility state, the variance of endowments can be upward sloped when the persistent movement of expected growth plays an important role as in Bansal and Yaron (2004), resulting in the upward sloped equity yields.}

### 5.4 Variance Risk Premium

The variance risk premium is defined as the difference between the statistical and risk-neutral expectation of forward variations of return. In the model economy, the variance risk premium is defined as in (31), repeated as follows:

\[
VRP(S_t) = \frac{E_t[M_{t+1}QV_{t+1}]}{E_t[M_{t+1}]} - E_t[QV_{t+1}],
\]

where

\[
QV_{t+1} = \frac{\sigma^2_d(S_{t+1})}{\text{Stochastic Volatility}} + \frac{J(S_t)B^2_p}{\text{Rare Disaster Jump}} + \left(\log \left(\frac{1 + (\frac{P}{D}) (S_{t+1})}{(\frac{P}{D}) (S_t)}\right)\right)^2.
\]

The payoff of VIX, \(QV_{t+1}\), is high with high variance of Gaussian shocks, the realization of rare disasters, and the regime-shifts to the low valuation regimes. Hence, VIX provides the agent with a hedge against those tail risks, leading to a sizable amount of variance risk premium.\footnote{Within a general equilibrium framework, the agent can require the variance risk premium due to the compensation for volatile fundamentals(Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011), and Zhou and Zhu (2012)) or due to the time-varying risk aversion(Bakshi and Madan (2006), and Wu (2012)).}

Table 12 exhibits the summary stats of variance risk premium of real data and various asset pricing models. Even though the price part(risk-neutral expectation) is directly observable, we need assumptions to find the statistical expectation. Panel A of Table 12 shows the variance risk premium with different estimates for the statistical expectation of return variation. Panel B of Table 12 reports the model-implied moments for the variance risk premium. My model generates variance risk premium equivalent to an empirically observable level, strictly higher than the level reproduced by Drechsler and Yaron (2011) and Zhou and Zhu (2012).
Further, since quadratic variation is decomposed into stochastic volatility, rare disaster jumps, and regime-switching jumps, I can decompose VRP into three corresponding parts as follows:

\[
VRP(S_t) = E^Q_t [QV_{t+1}] - E^P_t [QV_{t+1}]
\]

= \[
E^Q_t \left[ \sigma^2_d(S_{t+1}) \right] - E^P_t \left[ \sigma^2_d(S_{t+1}) \right] + E^Q_t \left[ J(S_t)B^2_d \right] - E^P_t \left[ J(S_t)B^2_d \right]
\]

\[
+ \ E^Q_t \left[ \left( \log \left( \frac{1 + \left( \frac{P}{D} \right) (S_{t+1})}{(S_t)} \right) \right)^2 \right] - E^P_t \left[ \left( \log \left( \frac{1 + \left( \frac{P}{D} \right) (S_{t+1})}{(S_t)} \right) \right)^2 \right].
\]

Table 13 shows that most of the variance risk premium is from jumps, due to regime-shifts.

Bollerslev, Tauchen, and Zhou (2009) document that short-term returns are predictable through variance risk premium and the predictability is maximal for quarterly returns. Fusari and Gonzalez-Perez (2012) compute the variance risk premium with dual frequency models and show that variance risk premium can predict returns even with a longer horizon. Specifically, they test the following regression equation:

\[
\frac{1}{h_p} \sum_{\tau=t+1}^{t+h_p} r_\tau = a + b \cdot VRP_t + \varepsilon,
\]

Table 14 reports regression results using real data along with those of the simulated data. In the model economy, variance risk premium can predict returns as in the real data.

5.5 Implied Volatility

Implied volatility of an option is the level of volatility that is implied by the market price of the option based on the option pricing formula by Black and Scholes (1973). That is, it is the volatility with which the Black-Scholes pricing model produces a theoretical value for the option equal to the market price of that option.

Since the Crash of 1987, the implied volatility of at-the-money option is mostly lower than that of in-the-money or out-of-the-money options, commonly referred to as volatility smile. Furthermore, the implied volatility of out-of-the-money puts tends to be higher than that of out-of-the-money calls, known as volatility smirk.43 In addition, Carr and Wu (2003)

\footnote{See Rubinstein (1994), Jackwerth and Rubinstein (1996), and Ait-Sahalia and Lo (1998) among others.}
document an empirical regularity of implied volatility in the maturity direction; the volatility smirk does not flatten out as maturity increases.\footnote{This empirical observation is hard to account for by most models using the log return as a Lévy process. See Heston (1993), Carr and Wu (2004), and many others.}

In Section 3.3.4, I showed how to compute the price of put options on aggregate dividends. Figure 8 shows the average implied volatility surface of put prices of the model economy where moneyness is defined as

\[
d = \frac{\log(K/P)}{\sigma \sqrt{\tau}},
\]  

where \(K\) is the strike price, \(P\) is the current price of aggregate dividend claims and \(\tau\) is the time to maturity expressed in years. I use \(\sigma = 27.4\%\) as in Carr and Wu (2003). Despite that there are no option specific parameters in the model, the model implies that the implied volatility surface shows the smirked pattern and the persistency of the smirk over maturities.

The volatility surface of Figure 8 is qualitatively similar to what is observed in the financial market. Figure 9 shows the estimated implied volatility surface by Carr and Wu (2003). The underlying intuition for the smirked pattern is clear. Due to the volatility dynamics modeled with regime-switching, the cashflow in the tails probably is involved with regime-shifts to higher volatility regimes or the realization of rare disasters. That is, deep OTM puts provide investors with valuable insurance. Hence, the investors are willing to pay high prices for deep OTM puts, resulting in the high implied volatility. Furthermore, persistent regimes appear as the persistency of smirk over maturities. Even though this result is a preliminary application to puts, I expect that joint modeling of variance and jump with MSM is a promising venue for option pricing in future research.

5.6 Comovement of Premiums

The proposed econometric model can filter out the prior over hidden regimes with Bayesian updating process. It would be interesting to extract the time series of premiums across different asset classes. Specifically, I am interested in the comovements of three objects; equity premium, variance risk premium, and the implied volatility.

Making inferences on the time series of those processes, I use the information of market excess return. The time series of equity premium is computed as:

\[
\mathbb{E}[r_{m,t+1} - r_{f,t+1}|\mathcal{I}_t] = \sum_{S_t} \mathbb{E}[r_{m,t+1} - r_{f,t+1}|S_t] \Pr(S_t|\mathcal{I}_t). \tag{42}
\]
Similarly, the time series of variance of risk premium is derived as:

$$
\mathbb{E}[VRP_t|\mathcal{I}_t] = \sum_{S_t} VRP(S_t) \Pr(S_t|\mathcal{I}_t).
$$

(43)

I denote $IV_t$ as the implied volatility of a deep OTM put option with the maturity of 1 month and the moneyness of $d = -2$. The time series of the implied volatility is evaluated as:

$$
\mathbb{E}[IV_t|\mathcal{I}_t] = \sum_{S_t} IV(S_t) \Pr(S_t|\mathcal{I}_t).
$$

(44)

Figure 10 shows the time series of three computed processes. Since the risks of three assets are commonly driven by the underlying regime, premiums across those assets comove. The following table shows the sample correlation of those extracted time series.

<table>
<thead>
<tr>
<th></th>
<th>EP</th>
<th>VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>VP</td>
<td></td>
<td>0.60</td>
</tr>
<tr>
<td>IV</td>
<td>0.96</td>
<td>0.60</td>
</tr>
</tbody>
</table>

The shift of a basis regime into the risky state elevates the equity premium, variance risk premium and implied volatility simultaneously, generating the positive correlations among those series. This is consistent with the findings by Bollerslev and Todorov (2011). They investigate the relation between equity premium and variance premium and verify that the rare but large tails account for a large fraction of the premiums of both assets.

I compare the extracted time series with empirical counterparts. The extracted time series of equity premium is annualized 1-month premium. Since it is hard to find time series of monthly premium over a long horizon, I refer to the Implicit Risk Premium found by Li, Ng, and Swaminathan (2012). Following the approach of Li, Ng, and Swaminathan (2012), I find the Implicit Risk Premium as ICC minus one-month risk-free return where ICC solves the following:

$$
P_t = \sum_{n=1}^{\infty} \frac{\mathbb{E}_t[D_{t+n}]}{(1 + ICC_t)^n},
$$

(45)

where $P_t$ is the value of future aggregate dividends and $D_t$ is the aggregate dividend at time $t$.\footnote{Li, Ng, and Swaminathan (2012) compute ICC for individual stocks in the first step and find the market-wide ICC as a value-weighted average of individual ICCs.}

Figure 11 shows the time series of Implicit Risk Premium over Jan/1977-Dec/2011.
Divergence for two series is pronounced about the early 2000s, dotcom bubble period. In those years, many internet business firms tended to be highly valued even with negative accounting profits. Thus, through the present-value approach by Li, Ng, and Swaminathan (2012), the required yield of equity market is supposedly low. However, my model detects high risk – probably due to the high variance – during those periods, implying high yield of equity. However, the movement around the recent financial crisis is well matched.

Further, I plot the filtered variance risk premium of my model economy against the variance risk premium computed by Fusari and Gonzalez-Perez (2012) using high frequency option data. Figure 12 shows monthly variance risk premiums over Feb/1996-Apr/2010. The estimates by Fusari and Gonzalez-Perez (2012) are rather volatile. Both models feature the spike of variance risk premium in the recent crisis.

Lastly, I overlay the implied volatility obtained from Option Metrics on the extracted implied volatility of deep OTM puts.\textsuperscript{46} I interpolate implied volatilities of option prices with similar moneyness and maturity to get the comparable implied volatility. Figure 13 reports those series. The model-implied implied volatility is relatively higher than the implied volatility obtained from Option Metrics. Still, the long-term patterns are matched.

6 Robustness

In Section 4, I use solely monthly return process for the estimation of the time-varying risk. In this section, as robustness tests, I estimate the regime-switching endowment process with different sets of data. In the first subsection, I estimate the process with quarterly consumption data in addition to the monthly market returns. In the second subsection, only quarterly endowments are used in the assessment of parameters.\textsuperscript{47}

6.1 Monthly Returns and Quarterly Consumption

The tractability of the econometric specification comes out in handling data at mixed frequencies. In this subsection, I use the mixed-frequency data – quarterly consumption process from Q1/1947-Q4/2011 and monthly market excess returns from Jan/1927-Dec/2011 – to estimate the time-varying risk of the model economy. The estimation methodology is identical as in Section 4 but I update the likelihood quarterly by extending the state space to a

\textsuperscript{46}With the moneyness of -2 and the maturity of 1 month

\textsuperscript{47}In addition to these tests, I also try annual returns in ML estimation. The parameters of the model economy with multifrequency risk are not properly estimated with ML procedure. Multifrequency risks are faded out in the annual data. There is a significant advantage in the use of high frequency data.
sequence of regimes over a quarter.

I plug in the following monthly sample moments as in Section 4:

\[
\begin{bmatrix}
\mu_c \\
\mu_d 
\end{bmatrix} = \begin{bmatrix}
\frac{1.85\%}{12} \\
\frac{0.93\%}{12}
\end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix}
\frac{(2.18)^2\%^2}{12} & 0.66 \cdot \frac{2.18-11.23\%^2}{12} \\
0.66 \cdot \frac{2.18-11.23\%^2}{12} & \frac{(11.23)^2\%^2}{12}
\end{bmatrix}
\]

\[
\bar{p} = 1 - (1 - 3.63\%)^{\frac{1}{12}}.
\]

I use the same preference parameters:

- \(\gamma = 7.5\), \(\psi = 1.5\), \(\delta = (0.98)^{\frac{1}{12}}\).

In Section 4, I estimate the parameters governing the time-varying risk of \(A(S_t), \{A_0, b, \alpha_1, \lambda\}\), and the size of disaster for dividends, \(B_d\), using ML. Here, since I observe quarterly consumption, I add \(B_c\) as a parameter to be estimated. Table 15 reports parameters estimated using ML along with the estimates in Section 4 for a reference. I estimate parameters only with \(\bar{k} = 5\) to compare parameters between two models. In general, individual parameters are similar across two data sets.

It is interesting to test the differences in parameters. Using quarterly consumption makes the data richer. That is, we can consider the estimation in this subsection as Full Information Maximum Likelihood (FIML) and that in Section 4 as Limited Information Maximum Likelihood (LIML). Specifically, I test the following joint hypothesis:

\[
\begin{align*}
\hat{A}_0,FIML &= \hat{A}_0,LIML \\
\hat{b}_{FIML} &= \hat{b}_{LIML} \\
\hat{\alpha}_{1,FIML} &= \hat{\alpha}_{1,LIML} \\
\hat{\lambda}_{FIML} &= \hat{\lambda}_{LIML} \\
\hat{B}_{d,FIML} &= \hat{B}_{d,LIML} \\
\hat{B}_{c,FIML} &= \hat{B}_{c,LIML} = \psi \hat{B}_{d,LIML}.
\end{align*}
\]

Table 16 reports results of the Wald test\textsuperscript{48} on the null hypothesis. The result of Test 1 confirms that the differences in first five parameters are not statistically significant. However,\hfill

\textsuperscript{48}I use the time series of quarterly likelihood for the estimation of joint covariance matrix.
in Test 2, the difference in the size of consumption disaster is fairly large and rejects the null hypothesis. This should be expected since we cannot observe disasters in the post-war U.S. consumption data yet.

The differences are economically meaningful. Table 17 shows the model-implied moments of interest. Since the estimated $B_d$, the size of consumption disasters, is quite small, the dividend disasters are not directly priced and the premium and volatility of market returns are not as high as the empirically observed level.

Still, the model economy has the feature of stochastic volatility and reproduces the inverted term structure of dividend strips yield. Figure 14 plots the forward yield curve of dividend strips. Consistent with the results using baseline parameters, the yield curve is downward sloped in the risky regime.

### 6.2 Quarterly Endowments

As shown in Figure 3, disasters in returns do not seem to be well-aligned with disasters in endowments. I estimate the regime-switching endowment process with quarterly consumption and dividend process from Q1/1947 to Q4/2011.

The quarterly endowment process is constructed similarly to the annual endowment process described in Appendix B. As in Section 4, I impose quarterly sample moments using annual endowments to the likelihood function so that the sample moments are matched with the model-implied moments conditional on no rare disaster. Specifically, I match the following unconditional moments:

\[
\begin{align*}
\begin{bmatrix} \mu_c \\ \mu_d \end{bmatrix} &= \begin{bmatrix} 1.85 \% \\ 0.93 \% \end{bmatrix} \\
\Sigma &= \begin{bmatrix} \frac{(2.18)^2}{4} \%^2 & 0.66 \cdot \frac{2.18 \cdot 11.23}{4} \%^2 \\ 0.66 \cdot \frac{2.18}{4} \cdot \frac{11.23}{4} \%^2 & \frac{(11.23)^2}{4} \%^2 \end{bmatrix} \\
p &= 1 - (1 - 3.63\%)^{\frac{1}{4}}
\end{align*}
\]

In the ML estimation in Section 4, I estimate parameters of \{ $A_0, b, \alpha_1, \lambda, B_d$ \}. Since I have a separate process of consumption growth, I add the size of consumption disaster, $B_c$, as a parameter to be estimated. Also, since dividends are concentrated in the fourth quarter, I adjust the seasonality of the dividend process by adding quarterly dummies, \{ $\delta_1, \delta_2, \delta_3$ \}, so that the logarithms of $i$-th quarter’s observed dividend is smaller than $i$-th quarter’s
supposed-to-be dividend by $\delta_i$ due to the seasonality.

Table 18 shows ML estimates and log-likelihood of the sample with a different number of frequencies. Based on BIC, I use estimated parameters with $\bar{k} = 2$ as baseline parameters. As we have expected, since jumps in aggregate consumption or aggregate dividends are hardly observable in post-war U.S. data, the size of jumps is almost zero. Thus, estimated parameters would generate a process only with stochastic volatility. As shown in subsection 5.1, stochastic volatility alone was not enough with $\gamma = 7.5$ to get the high-risk premium with the endowment process estimated in Section 4. We should use the high level of CRRA parameter to achieve the equivalent level of risk premium. Table 19 shows the implied moments with CRRA, $\gamma$, of 18 and EIS, $\psi$, of 1.5. Even though the level of risk premium is matched to the empirically observed level, the model implies strictly smaller standard deviation of market return compared with sample moments.

As discussed in subsection of 5.3, stochastic volatility plays an important role in determining the sign of the slope for the yield curve of dividend strips. Figure 15 shows the regime dependent term structure of forward yields of dividend strips. Whether the long-run risk regime, $S_{1,t}$, is high or low determines the slope of the yield curve in long maturity; whereas the short-run risk regime, $S_{2,t}$, governs the slope of yield curve for short maturity.

In sum, rare disasters are not identified with the post-war quarterly endowment process. We need a high level of risk aversion to match the model-implied risk premium with the sample counterparts. However, the high volatility of returns is not reproduced. Still, the slope of the yield curve of dividend strips can be explained with the stochastic volatility of endowment.

7 Model Extensions

In this section, I suggest possible extensions of the model economy. In the first subsection, I derive the stochastic discount factor where the regime is not observed directly to a representative agent, who infers about the underlying regime through realized consumption and dividend process. In the second subsection, I consider the case that the dividend and consumption are cointegrated and show how to price dividend strips.

7.1 Learning about Regimes

In the model economy of this paper, I assume that the representative agent knows about the underlying regime which governs the evolution of volatility and time-varying risk of rare
disaster. In this subsection, I show how to extend the baseline case to the environment where the representative agent learns about the underlying regime through observing consumption and dividends. The learning model with MSM is investigated in Calvet and Fisher (2007). To by-pass the computational burden of multi-dimensional integration, they use a simulation-based approach. However, with the inclusion of rare disasters, simulation is not a proper substitute for the integration. At this point, I cannot afford the luxury of multi-dimensional numerical integration. Hence, I suggest the equilibrium pricing with learning about hidden regimes for future research. Especially with high frequency data and multiple-sizes of jumps, learning might explain why small jumps, like a swing by a butterfly, can end up as a crash of stock prices, much like a disastrous storm.

The key state variable under the learning environment is the prior of the representative agent over the $k$ regimes. I denote $\Xi_t \in [0, 1]^k$ as a prior of the representative agent at time $t$. For a compact exposition, I consider the case that $\gamma \neq 1$ and $\psi^{-1} \neq 1$. With the same approach as in the full information model, the utility process is expressed as $V_t = C_t v(\Xi_t)$, where $v(\cdot)$ solves the following equation as a fixed point:

$$v(\Xi_t) = \left[ (1 - \delta) + \delta \left( \mathbb{E} \left[ \exp \left( (1 - \gamma) (c_{t+1} - c_t) \right) v(\Xi_{t+1}) \right] \right)^{1 - \psi^{-1}} \right]^{1/(1 - \gamma)}.$$  \hfill (46)

And, $\Xi_t$, a $K$-dimensional vector representation of $\Xi_t$, is updated with the Bayes’ rule as follows:

$$\Xi_{t+1} = \frac{(\Xi, \Pi) \odot f(\Delta c_{t+1}, \Delta d_{t+1})}{1_K \left[ (\Xi, \Pi) \odot f(\Delta c_{t+1}, \Delta d_{t+1}) \right]},$$  \hfill (47)

where $f(\Delta c_{t+1}, \Delta d_{t+1})$ is $K$-dimensional regime dependent joint density of logarithms of consumption and dividend growth, and $\odot$ represents Hadamard product.

Then, it is straight forward to derive the stochastic discount factor, $M_{t+1}$, as follows:

$$M_{t+1} = \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \times \left( \frac{v(\Xi_{t+1})}{\kappa(\Xi_t)} \right)^{\psi^{-1} - \gamma},$$  \hfill (48)

which is analogous to (16) and $\kappa(\cdot)$ is solved as follows:

$$\kappa(\Xi_t) = \mathbb{E} \left[ \left( \frac{C_{t+1}}{C_t} v(\Xi_{t+1}) \right)^{1 - \gamma} | \Xi_t \right]^{\frac{1}{1 - \gamma}}.$$

42
7.2 Cointegration

In the model economy of this paper, I assume that the level of consumption can be diverted away from the level of dividends by an arbitrarily large distance. Since the portion of stock market in the aggregate economy does not need to be stationary in the long run, the assumed joint process of consumption and dividend is not necessarily unrealistic. However, the recent studies show that modeling the cointegration between consumption and dividend can explain the cross-sectional differences in expected returns. Since I use a univariate return series to estimate the parameters of the model economy, I do not consider the cointegrated model in the current project. The suggested model will be pursued in future research using bivariate series of endowments and return.

Assume that the consumption process is identical to the model economy and the dividend process is cointegrated with the consumption as follows:

\[ c_t - \phi d_t = u_t, \]

where \( u_t \) is a stationary AR(1) process:

\[ u_t = a + b_u u_{t-1} + \varepsilon_t. \]

Then, the cointegrated process implies that the logarithm of dividend growth is evolved as follows:

\[ \Delta c_t - \phi \Delta d_t = \Delta u_t \]
\[ \Delta d_t = \phi^{-1} \Delta c_t - \phi^{-1} \Delta u_t, \]

Since the consumption process is identical, the stochastic discount factor is the same as (16). As an application of the cointegrated model, I show how to compute the dividend strips. Let \( (\frac{p}{D})_n (S_t, u_t) \) denote the ratio of price of dividend strips with maturity \( n \) to the current level of dividend. It is obvious that \( (\frac{p}{D})_0 (S_t, u_t) = 1 \). Then, I derive \( (\frac{p}{D})_n (S_t, u_t) \)

\[ 49 \text{See Bansal, Dittmar, and Kiku (2009) among others.} \]
for each $n > 1$ recursively as follows:

\[
\left( \frac{P}{D} \right)_n (S_t, u_t) = E \left[ M_{t+1} \left( \frac{D_{t+1}}{D_t} \right) \left( \frac{P}{D} \right)_{n-1} (S_{t+1}, u_{t+1}) | S_t, u_t \right] \\
= E \left[ \delta \left( \frac{v(S_{t+1})}{\kappa(S_t)} \right) \psi^{1-\gamma} \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{D_{t+1}}{D_t} \right) \left( \frac{P}{D} \right)_{n-1} (S_{t+1}, u_{t+1}) | S_t, u_t \right] \\
= E \left[ \delta \left( \frac{v(S_{t+1})}{\kappa(S_t)} \right) \psi^{1-\gamma} \exp \left( -\gamma \Delta c_{t+1} + \Delta d_{t+1} \right) \left( \frac{P}{D} \right)_{n-1} (S_{t+1}, u_{t+1}) | S_t, u_t \right] \\
= E \left[ \delta \left( \frac{v(S_{t+1})}{\kappa(S_t)} \right) \psi^{1-\gamma} \exp \left( -\gamma + \phi^{-1} \Delta c_{t+1} - \phi^{-1} (u_{t+1} - u_t) \right) \left( \frac{P}{D} \right)_{n-1} (S_{t+1}, u_{t+1}) | S_t, u_t \right].
\]

Since we know the distribution of $\Delta c_{t+1}$ and $u_{t+1}$ given $S_t, u_t$, we can compute the above object recursively through numerical integration.

8 Conclusion

I propose a parsimonious econometric model which can estimate and test the time-varying risk of rare disasters. In my model, the time-varying risk of rare disasters was cast into a pure regime-switching formulation, which is convenient in statistical estimation as well as in equilibrium pricing. Based on ML estimation, I can utilize the full range of diagnostics on the model-selection or hypothesis-test. The Likelihood-Ratio test confirms that the model with a time-varying likelihood of rare disasters is a closer fit to the data than either a model without disasters or a model with a constant likelihood of disasters. The Wald test shows that ML estimates from monthly market returns are not statistically distinguished from those estimated with mixed-frequency data of monthly market returns and quarterly consumption series.

The suggested econometric specification can be applied to any time series where the non-linear dynamics of stochastic volatility of Gaussian shocks and the stochastic intensity of jumps play important roles. Especially, the application of my model into derivatives market is promising. The stylized behavior of VIX or deep OTM puts in financial markets are reproduced in the model economy – estimated even without data on those derivatives. The application of my model to derivatives data will be pursued in the future research.

Further, the application of my model to the nominal bond market is straightforward.
Preliminary results on bond pricing show that my model can reproduce the qualitative features of predictive regressions in bond market (Fama and Bliss (1987), Cochrane and Piazzesi (2005), and Campbell and Shiller (1991)). With the higher probability of hyperinflation in the event of rare disasters, the nominal yield curve becomes upward-sloped and even tiny probabilities of default can generate a substantial size of credit spread.
References


A Computation of the Utility Process

In this section, I show how to find the numerical values of $u(\cdot)$ in the recursive utility process of (13), which satisfies (11) and (12) simultaneously. In this section, we use $\rho$ as the inverse of $\phi$, EIS. The numerical values of $u(\cdot)$ is found as a fixed point of a function mapping through iteration.\textsuperscript{50} In what follows, I show the form of function mapping, of which $u(\cdot)$ is the fixed point.

The functional space we consider is a space of $K$-dimensional vectors, $\mathbb{R}^K$. Let $v$ be an element of $V \equiv \mathbb{R}^K$. Let $T : V \rightarrow V$ be a function mapping that we are interested in. The form of $V$ is specified for each of following four cases:

**When $\gamma = 1$ and $\rho = 1$**

\[
C_t u(S_t) = V_t = C_t^{1-\delta} (\exp \mathbb{E}_t [\log V_{t+1}])^\delta \\
= C_t^{1-\delta} (\exp \mathbb{E}_t [\log (C_{t+1} u(S_{t+1}))])^\delta \\
= C_t (\exp \mathbb{E}_t [c_{t+1} - c_t + \log u(S_{t+1})])^\delta \\
= C_t \left( \exp \sum_{S_{t+1}} \Pr (S_{t+1}|S_t) (\mu_c (S_{t+1}) + \log u(S_{t+1})) \right)^\delta
\]

Thus, the form of function mapping of $T : V \rightarrow V$ is specified as follows:

\[
T(v) = (\exp (\prod (\mu_c - B_c \mathbf{p} + \log(v))))^\delta
\]

**When $\gamma = 1$ and $\rho \neq 1$**

\textsuperscript{50}Except the case that $\rho = 1$, the function mapping is not a contraction mapping. However, within a reasonable range of parameters, the solution is always obtained through iteration.
\[
C_t v(S_t)
= V_t = [(1 - \delta) C_t^{1-\rho} + \delta R_t^{1-\rho}]^{\frac{1}{1-\rho}}
= [(1 - \delta) C_t^{1-\rho} + \delta (\exp \mathbb{E}_t [\log (C_{t+1} v(S_{t+1}))])^{1-\rho}]^{\frac{1}{1-\rho}}
= C_t [(1 - \delta) + \delta (\exp \mathbb{E}_t [c_{t+1} - c_t + \log v(S_{t+1})])^{1-\rho}]^{\frac{1}{1-\rho}}
= C_t \left[ (1 - \delta) + \delta \left( \exp \sum_{S_{t+1}} \Pr(S_{t+1}|S_t) (\mu_c(S_{t+1}) - p(S_{t+1}) B_c + \log v(S_{t+1})) \right)^{1-\rho} \right]^{\frac{1}{1-\rho}}
\]

Thus, the form of function mapping of \( T : V \to V \) is specified as follows:

\[
T(v) = \left( (1 - \delta) \mathbf{1}_K + \delta (\exp (\Pi (\mu_c - B_c p + \log(v)))) \right)^{1-\rho}. \tag{50}
\]

When \( \gamma \neq 1 \) and \( \rho = 1 \)

\[
C_t v(S_t)
= V_t = C_t^{1-\delta} \mathbb{E}_t [V_{t+1}^{1-\gamma}]^{\frac{1}{1-\gamma}}
= C_t^{1-\delta} \mathbb{E}_t \left[ (C_{t+1})^{1-\gamma} v(S_{t+1})^{1-\gamma} \right]^{\frac{1}{1-\gamma}}
= C_t \mathbb{E}_t \left[ \left( \frac{C_{t+1}}{C_t} \right) \left( \frac{1}{1-\gamma} \right) v(S_{t+1})^{1-\gamma} \right]^{\frac{1}{1-\gamma}}
= C_t \mathbb{E}_t \left[ \exp \left( (1 - \gamma) (c_{t+1} - c_t) v(S_{t+1})^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma}}
= C_t \mathbb{E}_t \left[ \exp \left( (1 - \gamma) \left( \mu_c(S_{t+1}) + \sqrt{\sigma_c^2 (S_{t+1}) \varepsilon_{t+1} - J(S_{t+1}) B_c} \right) \right) v(S_{t+1})^{1-\gamma} \right]^{\frac{1}{1-\gamma}}.
\]

Thus, the form of function mapping of \( T : V \to V \) is specified as follows:

\[
T(v) = \left( \Pi [p \exp ((\gamma - 1) B_c + a) + (1_K - p) \exp(a)] v^{1-\gamma} \right)^{\frac{1}{1-\gamma}}, \tag{51}
\]

where

\[
a = (1 - \gamma) \mu_c + \frac{(1 - \gamma)^2}{2} \sigma_c^2
\]
When $\gamma \neq 1$ and $\rho \neq 1$

\[
C_t v(S_t)
= V_t = [(1 - \delta) C_t^{1-\rho} + \delta R_t^{1-\rho}]^{\frac{1}{1-\rho}}
= \left[(1 - \delta) C_t^{1-\rho} + \delta \left(\mathbb{E}_t \left[(C_{t+1})^{1-\gamma} v(S_{t+1})^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}\right]^{\frac{1}{1-\rho}}
= C_t \left[(1 - \delta) + \delta \left(\mathbb{E}_t \left[\exp\left(\left((1 - \gamma) (c_{t+1} - c_t) + (S_{t+1})^{1-\gamma}\right)\right)^{\frac{1}{1-\gamma}}\right]\right]\right]^{\frac{1}{1-\rho}}
= C_t \left[(1 - \delta) + \delta \left(\mathbb{E}_t \left[\exp\left(1 - \gamma\right) \left(\mu_c (S_{t+1}) + \sqrt{\sigma_c^2 (S_{t+1} \varepsilon_{t+1} - J(S_{t+1}) B_c)}\right) v(S_{t+1})^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}\right]^{\frac{1}{1-\rho}}.
\]

Thus, the form of function mapping of $T : V \rightarrow V$ is specified as follows:

\[
T(v) = \left((1 - \delta) 1_K + \delta \left(\Pi [\mathbf{p} \exp((1 - \gamma) B_c + \mathbf{b}) + (1_K - \mathbf{p}) \exp(\mathbf{b})] v^{1-\gamma}\right)^{\frac{1}{1-\gamma}}\right)^{\frac{1}{1-\rho}}, \quad (52)
\]

where

\[
b = (1 - \gamma) \mu_c + \frac{(1 - \gamma)^2}{2} \sigma_c^2
\]

For each of the four cases, the solution of $v(\cdot)$ would be found as the fixed point of the corresponding mapping of $T$. That is, $T(v) = v$.

**B Data**

**B.1 Consumption Data**

- NIPA Table 2.4.4. Price Indexes for Personal Consumption Expenditures by Type of Product (Annual), Bureau of Economic Analysis

  - This table contains the price index for consumption expenditures on each of non-durables and services where 2005 expenditures are normalized to be 100.
NIPA Table 2.4.5. Personal Consumption Expenditures by Type of Product (Annual), Bureau of Economic Analysis

- This table contains the nominal dollar expenditure on each of nondurables and services.

I find the annual consumption process from 1929 to 2011 as follows:

\[ C_y = \frac{1}{100} \left( \frac{C_y^{ND}}{i_y^{ND}} + \frac{C_y^{S}}{i_y^{S}} \right), \]

where \( C_y \) is the nominal dollar expenditure on nondurable goods for the year \( y \), \( C_y^{S} \) is the nominal dollar expenditure on services for the year \( y \), \( i_y^{ND} \) is the price index for non-durable goods expenditure for the year \( y \), and \( i_y^{S} \) is the price index for services for the year \( y \). The data of \( C_y^{ND} \) and \( C_y^{S} \) are available from NIPA table 2.4.5 and the data of \( i_y^{ND} \) and \( i_y^{S} \) are available from NIPA table 2.4.4.

Then, per capita real consumption level is calculated as \( \frac{C_y}{p_y} \) where \( p_y \) is the total aggregate population for the year \( y \). And, the difference in the log per-capita consumption, \( \Delta c_y \) is defined as:

\[ \Delta c_y \equiv \log \left( \frac{C_y/p_y}{C_{y-1}/p_{y-1}} \right) \]

### B.2 Dividend Data

- From CRSP, we obtain value-weighted monthly returns (including distributions and excluding distributions) for all AMEX/NYSE/NASDAQ stocks from Jan/1929 to Dec/2011. Let \( R_m \) denote the return including distribution and \( R_m^e \) denote the return excluding distribution for the month \( m \).


Let \( P_m \) and \( D_m \) denote the price of all the stocks at the end of month \( m \) and the (cash and stock) dividends for the month \( m \). Then,

\[ R_m = \frac{P_m + D_m}{P_{m-1}} \]
and

$$R_m^e = \frac{P_m}{P_{m-1}}.$$  \hfill (54)

Thus,

$$R_m - R_m^e = \frac{D_m}{P_{m-1}}.$$  \hfill (55)

Aggregating the expression of (54) and (55), we do the following manipulation:

$$\frac{R_m - R_m^e}{R_{m-1} - R_{m-1}^e} = \frac{D_m P_{m-2}}{P_{m-1} D_{m-1}} = \frac{D_m P_{m-2}}{D_{m-1} P_{m-1}} = \frac{D_m}{D_{m-1}} \frac{1}{R_m^e - R_{m-1}^e},$$

eyielding

$$\frac{D_m}{D_{m-1}} = R_{m-1}^e \frac{R_m - R_m^e}{R_{m-1} - R_{m-1}^e}. \hfill (56)$$

Thus, we find the monthly growth rate of dividend, $\frac{D_m}{D_{m-1}}$, from Feb/1929 and Dec/2011. Since we are interested in the growth rate, we can normalize the dividend for the first month of Jan/1929 to be 1. With this normalization, we can recover $D_m$ for the whole months from Jan/1929 and Dec/2011.

Then, $D_y$, the annual dividend from 1929 to 2011, can be calculated as a summation of monthly dividends for the corresponding year:

$$D_y = \sum_{m \in y} D_m.$$

Lastly, we convert the level growth of the nominal dividend into the difference in the log of the real dividend, $\Delta d_y$ from 1930 to 2011:

$$\Delta d_y \equiv \log \frac{D_y}{\pi_y} \frac{\pi_{y-1}}{D_{y-1}}, \hfill (57)$$

where $\pi_y$ is the consumer price index for year $y$.  

59
B.3 Market Return and Risk-free Return Data

- From CRSP, I obtain value-weighted monthly returns (including distributions) for all AMEX/NYSE/NASDAQ stocks from Jan/1927 to Dec/2011.
- From Ibbotson Associates, I get the monthly Treasury Bill rate from Jan/1927 to Dec/2011.

C Likelihood Function

- Parameters Calibrated or Matched to sample counterparts:
  \[ \gamma, \delta, \rho = \psi^{-1}, \bar{p}_c, \bar{p}_d, \sigma^2_c, \sigma^2_d, \rho_{cd}, \phi = \frac{B_d}{B_c} \]

- Parameters Estimated:
  \[ A_0, b, \alpha_1, \lambda, B_d \]

In equilibrium, the log of market return is expressed as:

\[
\log(R_{m,t}) = \log\left(\frac{D_t + P_t}{P_{t-1}}\right) = \log\left(\frac{D_t}{D_{t-1}} \frac{1 + \left(\frac{P_t}{D_t}\right) (S_t)}{\left(\frac{P_{t-1}}{D_{t-1}}\right) (S_{t-1})}\right) = \Delta d_t + \log\left(\frac{1 + \left(\frac{P_t}{D_t}\right) (S_t)}{\left(\frac{P_{t-1}}{D_{t-1}}\right) (S_{t-1})}\right) = \mu_d(S_t) + \sqrt{\sigma^2_d(S_t)} \epsilon_t - J(S_t) B_d + \log\left(\frac{1 + \left(\frac{P_t}{D_t}\right) (S_t)}{\left(\frac{P_{t-1}}{D_{t-1}}\right) (S_{t-1})}\right)
\]

where \( \epsilon_t \) is an i.i.d draw from standard normal distribution, and the log of risk free return is expressed as:

\[
\log\left(\frac{1}{P^z_1(S_{t-1})}\right)
\]

Thus, the conditional likelihood of the observed excess return \( y_t \equiv r_{m,t} - r_{f,t} \) is expressed as
follows:

\[
\mathcal{L}(y_t|S_t, S_{t-1}) = \frac{1 - p(S_t)}{\sqrt{2\pi\sigma_d^2(S_t)}} \exp \left( -\frac{(m_t)^2}{2\sigma_d^2(S_t)} \right) + \frac{p(S_t)}{\sqrt{2\pi\sigma_d^2(S_t)}} \exp \left( -\frac{(m_t + B_d)^2}{2\sigma_d^2(S_t)} \right),
\]

(58)

where

\[
m_t = r_{m,t} - r_{f,t} - \mu_d(S_t) + B_d - \log \left( \frac{1 + (\frac{P}{D}(S_t))}{\frac{P}{D}(S_{t-1})} \right) + \log \left( \frac{1}{F^z(S_{t-1})} \right)
\]

(59)

Our objective is to compute the likelihood of sample \( \{y_{\tau}\}_{\tau=1}^T \). The expression of (58) is the building block for the computation. To distinguish the information set \( \mathcal{F}_t \) of the representative agent from the information set of the econometricians, let \( \mathcal{I}_t \) denote the information of the excess returns up to time \( t \), \( \{y_{\tau}\}_{\tau=1}^t \). Then, the log likelihood of the \( \{y_{\tau}\}_{\tau=1}^T \) is computed as follows:

\[
\log \left( \mathcal{L} \left( \{y_{\tau}\}_{\tau=1}^T \right) \right) = \sum_{t=1}^{T} \log \mathcal{L}(y_t|\mathcal{I}_{t-1})
\]

\[
= \sum_{t=1}^{T} \sum_{S_t, S_{t-1}} \log \mathcal{L}(y_t|S_t, S_{t-1}, \mathcal{I}_{t-1}) \Pr(S_t, S_{t-1}|\mathcal{I}_{t-1})
\]

\[
= \sum_{t=1}^{T} \sum_{S_t, S_{t-1}} \log \mathcal{L}(y_t|S_t, S_{t-1}) \Pr(S_{t-1}|\mathcal{I}_{t-1}) \Pr(S_t|S_{t-1})
\]

(60)

Since we already compute \( \mathcal{L}(y_t|S_t, S_{t-1}) \) in (58), it suffices to show how to calculate \( \Pr(S_{t-1}|\mathcal{I}_{t-1}) \) for \( t = 1, \cdots, T \).

### C.1 Step1: Initializing Probability Distribution over Hidden Regimes

To start a calculation of likelihood function, I set the initial probability over the hidden regimes; that is, I need \( \Pr(S_0|y_0) \). I set \( \Pr(S_0|y_0) \) as a steady state distribution over regimes,
\( \theta \) such that
\[
\begin{align*}
\theta' \Pi &= \theta' \\
\theta' 1_K &= 1.
\end{align*}
\]

Then, \( \theta \) is computed as:
\[
\begin{align*}
\theta' [\Pi 1_K] &= [0'_K 1] \\
\theta' &= [0'_K 1] [\Pi 1_K]' ([\Pi 1_K]' [\Pi 1_K]')^{-1}
\end{align*}
\]

We set \( \theta \) as the initial value of \( \Pr (S_0|y_0) \):
\[
\begin{bmatrix}
\Pr (S_0 = 1|y_0) \\
\vdots \\
\Pr (S_0 = K|y_0)
\end{bmatrix} = \theta.
\]

C.2 Step 2: Updating the Belief over the Hidden Regimes

In this step, I show how to get \( \Pr (S_t|\mathcal{I}_t) \) from a new information of \( y_t \) and \( \Pr (S_{t-1}|\mathcal{I}_{t-1}) \). Applying Bayes’ rule, we can express \( \Pr (S_t|\mathcal{I}_t) \) as follows:
\[
\Pr (S_t|\mathcal{I}_t) = \frac{\Pr (S_t, \mathcal{I}_t)}{\Pr (\mathcal{I}_t)} = \frac{\Pr (S_t, \mathcal{I}_t)}{\Pr (\mathcal{I}_t|\mathcal{I}_{t-1})} = \Pr (S_t|\mathcal{I}_{t-1}) \Pr (\mathcal{I}_t|\mathcal{I}_{t-1}),
\]

the numerator of which is calculated as:
\[
\Pr (S_t, \mathcal{I}_t|\mathcal{I}_{t-1}) = \sum_{S_{t-1}} \Pr (S_t, S_{t-1}, \mathcal{I}_y|\mathcal{I}_{y-1}) = \sum_{S_{t-1}} \log (L (y_t|S_t, S_{t-1})) \Pr (S_t|S_{t-1}) \Pr (S_{t-1}|\mathcal{I}_{t-1})
\]

and the denominator of which is calculated as:
\[
\Pr (\mathcal{I}_t|\mathcal{I}_{t-1}) = \sum_{S_t} \sum_{S_{t-1}} \Pr (S_t, S_{t-1}, \mathcal{I}_y|\mathcal{I}_{y-1}) = \sum_{S_t} \sum_{S_{t-1}} \log (L (y_t|S_t, S_{t-1})) \Pr (S_t|S_{t-1}) \Pr (S_{t-1}|\mathcal{I}_{t-1}).
\]

Then, we can apply Step 2 recursively to get the likelihood of all sample in (60).
Figure 1: **Realized Consumption Disasters from 1870-2006**: X-axis corresponds to calendar years. Y-axis represents the size of realized consumption disasters. Consumptions disasters are defined as more than 10% drops in the level of consumption. Data are from table 6 of Barro and Ursua (2008). Red dots represent disasters in 21 OECD countries, Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Italy, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, U.K., and U.S. Blue dots represent disasters in 13 non OECD countries, Argentina, Brazil, Chile, Colombia, India, Malaysia, Mexico, Singapore, South Korea, Taiwan, Turkey, Uruguay, and Venezuela. The total number of disasters is 135, 67 of which occurred in OECD countries, and 68 of which occurred in non OECD countries. The unconditional likelihood of consumption disasters is 3.63 times over a century.
Figure 2: **Inference on the Time-Varying Risk**: Bayesian inference over the hidden regimes at time $t$ is made with the previous returns, $I_{t-1}$, or the whole sample, $I_T$. I extract the inferred time-varying risk of $A_{t,prediction}$ and $A_{t,smoothing}$ as in (33) and (34), respectively. Figure 2 shows the time series of $A_{t,prediction}$ and $A_{t,smoothing}$. Four distinguished peaks correspond to Great Depression, the Second World War, Oil Shocks in mid-1970s, and the recent Financial Crisis.
Figure 3: **Ex-Ante and Ex-Post Probability for the event of a Rare Disaster**: The first plot shows the *Ex-Ante* probability for a rare disaster to occur, $\mathbb{E}[J_t = 1|\mathcal{F}_{t-1}]$. The second plot depicts the *Ex-Post* probability for a rare disaster to have occurred, $\mathbb{E}[J_t = 1|\mathcal{F}_t]$. Spikes in the Ex-Post probability are matched to the realized crashes of excess returns. Stock market crashes tend to be aligned with the sudden drops in the endowment process before the Second World War. However, in post-war data, the relation does not seem salient.
Figure 4: **Average Forward Yield Curve of Dividend Strips**: The solid line represents the average forward yield. In the average, the forward yield curve is upward sloping. The dotted lines represent 2SD bounds around the average forward yield.
Figure 5: **Forward Yield Curve of Dividend Strips in High and Low Risk Regimes:**
The top line corresponds to the forward yield curve of the high risk regime, $S_{k,t} = 1$ for all $k = 1, \ldots, 5$. The bottom line corresponds to the forward yield curve of the low risk regime, $S_{k,t} = 0$ for all $k = 1, \ldots, 5$. The yield curve becomes downward sloped in the high risk regime.

Figure 6: **Forward Equity Yields: S&P 500**: This figure is from Figure 1 of van Binsbergen, Hueskes, Koijen, and Vrugt (2011). Each line displays the time series of yields of dividend strips for maturities of 1, 2, 5, and 7 years between Oct/07/2002 and Apr/08/2011. The yield curve becomes downward sloped during 2008-2009 financial crisis.
Figure 7: Driver of Downward Sloping Forward Yield Curve of Dividend Strips:
The upper plot presents the forward yield curve of dividend strips only with the time-varying likelihood of rare disaster. For the high intensity regime, the forward yield exhibits a hump-shaped curve. The lower plot gives the forward yield curve of dividend strips only with stochastic volatility. The sharp decline of the forward yield is highly pronounced in the high volatility regime. Stochastic volatility is the driver of downward sloped yield curve in the high risk regime.
Figure 8: Volatility Smirk in the Model Economy: Figure 8 shows the average implied volatility surface of put option prices of the model economy where moneyness is defined as in (41). Implied volatility of deep OTM put options is high across maturities.

Figure 9: Volatility Smirk in S&P 500 Index Options: This figure is from Figure 1 by Carr and Wu (2003). The implied volatility surface is computed through nonparametric smoothing of daily implied volatility quotes on S&P 500 index options from Apr/04/1999 to May/31/2000. Moneyness is defined as in (41).
Figure 10: **Comovement of Premiums**: Figure 10 plots the time series of equity premium, variance risk premium, and the implied volatility of deep OTM puts with the maturity of 1 month and the moneyness of $d = -2$. See the specifications of (42), (43), and (44).
Figure 11: **Time Series of Equity Premium**: Figure 10 plots the time series of Implicit Risk Premium by Li, Ng, and Swaminathan (2012) along with my model-implied Implicit Risk Premium over Jan/1977-Dec/2011. Implicit Risk Premium is the risk premium computed with the present-value approach to equate the discounted expected future dividends with the current price.

Figure 12: **Time Series of Variance Risk Premium**: Figure 10 plots the time series of filtered variance risk premium of the model economy along with that computed by Fusari and Gonzalez-Perez (2012) over Feb/1996-Apr/2010. Fusari and Gonzalez-Perez (2012) compute daily VRP using high frequency option data. I plot the average of daily estimates of VRP over a month.
Figure 13: **Time Series of Implied Volatility of Deep OTM Puts**: Figure 13 plots the time series of filtered implied volatility of deep OTM puts with the moneyness of -2 and the maturity of 1 month along with the comparable implied volatility from Option Metrics.

Figure 14: **Forward Yield Curve using Quarterly Consumption and Monthly Returns**: Figure 14 shows the regime dependent term structure of forward yields of dividend strips of the highest risk regime and the lowest risk regime. Yield curve is downward sloped in the highest risk regime as in the case using baseline parameters.
Figure 15: **Forward Yield Curve using Quarterly Endowments**: Figure 15 shows the regime dependent term structure of forward yields of dividend strips. Whether long-run risk regime, $S_{1,t}$, is high or low determines the slope of the yield curve for long maturity, whereas the short-run risk regime, $S_{2,t}$, governs the slope of the yield curve for short maturity.
E  Tables

Table 1: **Summary Stats of Annual Endowment and Inflation Process**

<table>
<thead>
<tr>
<th></th>
<th>Mean(%)</th>
<th>Std(%)</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta c_t$</td>
<td>1.85</td>
<td>2.18</td>
<td>-1.49</td>
<td>7.85</td>
<td>NA</td>
</tr>
<tr>
<td>$\Delta d_t$</td>
<td>0.93</td>
<td>11.23</td>
<td>-0.85</td>
<td>8.46</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 1 reports sample moments of the logarithms of annual growth of consumption and dividends. These values (after converted to monthly values) are plugged into the likelihood function of monthly returns for ML estimation.

Table 2: **Summary Stats of Monthly Return Data**

<table>
<thead>
<tr>
<th></th>
<th>Monthly</th>
<th>Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean(%)</td>
<td>Std(%)</td>
</tr>
<tr>
<td>$\log R_{m,t}$</td>
<td>0.76</td>
<td>5.46</td>
</tr>
<tr>
<td>$\log R_{f,t}$</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td>$\log R_{m,t} - \log R_{f,t}$</td>
<td>0.46</td>
<td>5.47</td>
</tr>
</tbody>
</table>

Table 2 reports sample moments of the logarithms of market returns and risk free returns. The annualized mean is 12 times that of the monthly mean and the annualized Std is $\sqrt{12}$ times that of monthly Std. Note that the mean level of $\log R_{m,t}$ and $\log R_{f,t}$ are nominal.
Table 3: ML estimates for the Parameters of Stochastic Volatility and time-varying likelihood of Rare Disasters

<table>
<thead>
<tr>
<th>( \bar{k} )</th>
<th>ML Estimates (SE)</th>
<th>( \log \mathcal{L} ) (×10³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₀</td>
<td>( b )</td>
<td>( \alpha_1(%) )</td>
</tr>
<tr>
<td>1</td>
<td>0.91</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>2</td>
<td>0.62</td>
<td>2.32</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.67)</td>
</tr>
<tr>
<td>3</td>
<td>0.77</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>4</td>
<td>0.82</td>
<td>1.67</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>5</td>
<td>0.86</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>6</td>
<td>0.88</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.20)</td>
</tr>
</tbody>
</table>

Table 3 reports ML estimates. The likelihood function is defined in Appendix C. The decaying rates of persistency, \( b \), decreases as the number of frequencies increases. The duration of the basis regime with the longest frequency is about 5 to 10 years for \( \bar{k} > 2 \). The size of a rare disaster is stable across different values of \( \bar{k} \).
Table 4: **Vuong’s Likelihood Ratio Test**

<table>
<thead>
<tr>
<th>$\overline{k}'$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z$-stat</td>
<td>-4.63**</td>
<td>-1.86*</td>
<td>-0.91</td>
<td>-0.96</td>
<td>-0.42</td>
</tr>
<tr>
<td>Result</td>
<td>Reject</td>
<td>Reject</td>
<td>Accept</td>
<td>Accept</td>
<td>Accept</td>
</tr>
</tbody>
</table>

Table 4 reports the results of Vuong’s Likelihood ratio test on the null that the model with $\overline{k} = 5$ does not explain the data better than other models with $\overline{k} = 1, 2, 3, 4, 6$. We can reject the null for $\overline{k} = 1, 2$ with a reasonable level of significance, where ** and * represent significance levels with 1% and 5% respectively. However, we cannot reject the null for $\overline{k} = 3, 4, 6$.

Table 5: **Model Implied Moments of Endowment Process**

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th></th>
<th></th>
<th>10th</th>
<th>50th</th>
<th>90th</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta c$</td>
<td>Data</td>
<td>Population</td>
<td>Mean</td>
<td>1.85</td>
<td>1.53</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>Std</td>
<td></td>
<td>Std</td>
<td>2.18</td>
<td>2.59</td>
<td>2.00</td>
</tr>
<tr>
<td>$\Delta d$</td>
<td>Data</td>
<td>Population</td>
<td>Mean</td>
<td>0.93</td>
<td>0.00</td>
<td>-1.60</td>
</tr>
<tr>
<td></td>
<td>Std</td>
<td></td>
<td>Std</td>
<td>11.23</td>
<td>12.27</td>
<td>10.66</td>
</tr>
</tbody>
</table>

Table 5 reports model-implied moments of endowment process. I simulate consumption and dividend process for $10^6$ months with baseline parameters, which are estimated with five frequencies, $\overline{k} = 5$. Then, monthly processes are aggregated to annual processes.
Table 6: **ML estimates for Different Specifications on Rare Disasters**

<table>
<thead>
<tr>
<th></th>
<th>ML Estimates (SE)</th>
<th>log $\mathcal{L}$ ($\times 10^3$)</th>
<th>Vuong’s z-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI</td>
<td>$A_0$ 0.86 (0.02)</td>
<td>1.7017 NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b$ 1.45 (0.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$ (%) 1.15 (0.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\lambda$ 0.89 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$B_d$ 0.26 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td>$A_0$ 0.88 (0.02)</td>
<td>1.6904 -3.14*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b$ 1.67 (0.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$ (%) 0.69 (0.45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\lambda$ 0.91 (0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$B_d$ 0.24 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NI</td>
<td>$A_0$ 0.87 (0.02)</td>
<td>1.6818 -3.29*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b$ 2.14 (0.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$ (%) 0.51 (0.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\lambda$ 0.90 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$B_d$ NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6 reports ML estimates with different specification on the risk of rare disasters. SI, CI, and NI represent stochastic intensity, constant intensity, and no intensity, respectively. All parameters are estimated with $k = 5$. The log likelihood is maximized with SI. I perform Vuong’s Likelihood-Ratio test. We can reject the null that the stochastic intensity of the model is not explaining the data better than the model without rare disasters or with constant intensity of rare disaster.

*: Significant with a confidence level of 1%

Table 7: **Economic Implication of Different Specifications on Rare Disasters**

<table>
<thead>
<tr>
<th></th>
<th>Annualized Moments(%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mathbb{E}[r_{m,t} - r_{f,t}]$</td>
<td>$\mathbb{E}[r_{f,t}]$</td>
</tr>
<tr>
<td>SI</td>
<td>6.56</td>
<td>1.87</td>
</tr>
<tr>
<td>CI</td>
<td>1.07</td>
<td>3.64</td>
</tr>
<tr>
<td>NI</td>
<td>0.50</td>
<td>2.97</td>
</tr>
<tr>
<td>Data</td>
<td>5.55</td>
<td>0.38$^i$</td>
</tr>
</tbody>
</table>

Table 7 reports implied annualized moments of our interests with different specifications on the risk of rare disasters. $r_{m,t}$ and $r_{f,t}$ represent the logarithms of market return and risk-free return, respectively. SI, CI, and NI represent stochastic intensity, constant intensity, and no intensity, respectively. Parameters for each specification is reported in Table 6. Annualized sample moments are computed with the logarithms of monthly returns from Jan/1927-Dec/2011.

$i$: The sample mean of risk free returns are adjusted with CPI inflation.
Table 8: **Model Implied Moments**

<table>
<thead>
<tr>
<th></th>
<th>( \mathbb{E}[r_{m,t} - r_{f,t}] )</th>
<th>( \mathbb{E}[r_{f,t}] )</th>
<th>STD ( r_{m,t} )</th>
<th>STD ( r_{f,t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>6.56</td>
<td>1.87</td>
<td>17.80</td>
<td>0.46</td>
</tr>
<tr>
<td>CV+SI</td>
<td>2.12</td>
<td>1.89</td>
<td>13.00</td>
<td>0.46</td>
</tr>
<tr>
<td>SV+CI</td>
<td>1.80</td>
<td>1.87</td>
<td>12.27</td>
<td>0.46</td>
</tr>
<tr>
<td>TA(^i)</td>
<td>0.94</td>
<td>11.22</td>
<td>13.57</td>
<td>1.39</td>
</tr>
<tr>
<td>Data</td>
<td>5.55</td>
<td>0.38(^{ii})</td>
<td>18.91</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 8 reports implied annualized moments of our interests. \( r_{m,t} \) and \( r_{f,t} \) represent the logarithms of market return and risk-free return, respectively. The numbers in the Model are computed with ML estimates for \( \bar{k} = 5 \) in Table 3. Annualized sample moments are computed with the logarithms of monthly returns from Jan/1927-Dec/2011. SV, CV, SI, and CI represents stochastic volatility, constant volatility, stochastic intensity, and constant intensity, respectively. TA represents the implied moments of the time-additive preference.

\(^{i}\): Time-additive preference with \( \gamma = \psi^{-1} = 7.5 \)

\(^{ii}\): The sample mean of risk free returns are adjusted with CPI inflation.
Table 9: \textbf{Return Predictability}

<table>
<thead>
<tr>
<th>$hp$ (year)</th>
<th>Data</th>
<th>Model Population</th>
<th>Distribution of $b$</th>
<th>Distribution of $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$</td>
<td>$R^2$</td>
<td>10th</td>
<td>50th</td>
</tr>
<tr>
<td>1</td>
<td>-0.33</td>
<td>0.02</td>
<td>-0.22</td>
<td>0.07</td>
</tr>
<tr>
<td>2</td>
<td>-0.41</td>
<td>0.07</td>
<td>-0.40</td>
<td>0.12</td>
</tr>
<tr>
<td>3</td>
<td>-0.49</td>
<td>0.11</td>
<td>-0.53</td>
<td>0.15</td>
</tr>
<tr>
<td>4</td>
<td>-0.59</td>
<td>0.18</td>
<td>-0.63</td>
<td>0.17</td>
</tr>
<tr>
<td>5</td>
<td>-0.61</td>
<td>0.22</td>
<td>-0.72</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 9 reports regression estimates of $\sum_{t=1}^{t+h_p} r_\tau = a + b \log \left( \frac{P_t}{D_t} \right) + \varepsilon$. The figures in the Data column is from regression of annual returns on price-to-dividend ratio from 1927 to 2011, constructed from CRSP monthly return series. See Appendix B for details. I simulate monthly data of $10^6$ months and convert monthly data into annual series to estimate the suggested regression. Distribution of $b$ and $R^2$ are obtained by repeating the regression with 80 years simulated data 1,040 times.

Table 10: \textbf{Dependency of the Price-to-Dividend ratio and Expected Return on State Variables}

<table>
<thead>
<tr>
<th>LHS</th>
<th>RHS</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\log \left( \frac{P_t}{D_t} (S_t) \right)$</td>
<td>$S_{1,t}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S_{2,t}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S_{3,t}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S_{4,t}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S_{5,t}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table 10 shows the dependency of price-to-dividend ratio and one-period expected return on the underlying basis regimes by estimating the regression $LHS = a + \sum_{k=1}^{\kappa} b_k S_{k,t} + \varepsilon$. The numbers in parenthesis are $t$-stats. The results show that the movement of each basis regime shifts the price-to-dividend ratio and the expected return into opposite directions from each other, leading to the predictability of return with valuation ratios.
Table 11: **Comparison of Implied Annualized Price-to-Dividend Ratio Dynamics to Other Models**

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>Other Models</th>
<th>MSM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BY\textsuperscript{i}</td>
<td>ZZ\textsuperscript{ii}</td>
</tr>
<tr>
<td>$\mathbb{E}[p - d]$</td>
<td>3.06</td>
<td>3.00</td>
<td>2.75</td>
</tr>
<tr>
<td>$\sigma(p - d)$</td>
<td>0.45</td>
<td>0.16</td>
<td>0.42</td>
</tr>
<tr>
<td>$AC1(p - d)$</td>
<td>0.88</td>
<td>0.77</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table 11 reports implied moments of logarithms of annual price-to-dividend ratio. Sample moments of real data are obtained from the annual data from 1927 to 2011, constructed from CRSP monthly return series. See Appendix B for details. Moments of MSM are the implied annual moments of the model economy in this paper. I simulate monthly data of $10^6$ months and aggregate monthly data into annual series to get annual price-to-dividend ratio.

i: See table 3 of Beeler and Campbell (2009)

ii: See table 2 of Zhou and Zhu (2012)
Table 12: **Variance Risk Premium**

<table>
<thead>
<tr>
<th>Panel A: Data</th>
<th>RV$_{index,t-1}$</th>
<th>̂RV$_{index,t}$</th>
<th>Daily$_{index,t}$</th>
<th>̂RV$_{fut,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>18.56</td>
<td>18.61</td>
<td>12.67</td>
<td>11.27</td>
</tr>
<tr>
<td>Std</td>
<td>15.34</td>
<td>15.06</td>
<td>14.38</td>
<td>7.61</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.50</td>
<td>0.69</td>
<td>0.54</td>
<td>0.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Model Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>BY$^{ii}$</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std</td>
</tr>
<tr>
<td>AR(1)</td>
</tr>
</tbody>
</table>

Panel A is from Drechsler and Yaron (2011). For each column, they use different statistical expectation for the forward return variation, $E_t[Q_{V_{t+1}}]$. $RV_{index,t-1}$ is the summation of the 5-minutes S&P 500 cash index return over the last month. $\hat{RV}_{index,t}$ is the estimated level of return variation over the next month through regressing $RV_{index,t}$ on $RV_{index,t-1}$ and $VIX^2_{t-1}$. Daily$$_{index,t}$ is the estimated level of the squared sum of the daily index return through regressing Daily$_{index,t}$ on Daily$_{index,t-1}$, realized summation of squared daily returns for the previous month, and GARCH(1,1) variance estimates. $RV_{fut,t}$ is the summation of the 5-minutes S&P 500 futures index return over the last month. $\hat{RV}_{fut,t}$ is the estimated realized variance using the futures index through regressing $RV_{fut,t}$ on $RV_{index,t-1}$ and $VIX^2_{t-1}$. Panel B is from Zhou and Zhu (2012). The model economy of this paper generates a higher level of variance risk premium compared with long-run risk based models of Bansal and Yaron (2004), Drechsler and Yaron (2011), and Zhou and Zhu (2012).

i: See Table 3 of Drechsler and Yaron (2011).

ii: See Table 5 of Zhou and Zhu (2012) except for the last column.

iii: Bansal and Yaron (2004)

iv: Drechsler and Yaron (2011)

v: Zhou and Zhu (2012)
Table 13: **Variance Risk Premium Decomposition**

<table>
<thead>
<tr>
<th></th>
<th>SV</th>
<th>RD-Jump</th>
<th>RS-Jump</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (Annual Std$^2$/12)</td>
<td>0.45</td>
<td>1.94</td>
<td>12.56</td>
<td>14.95</td>
</tr>
<tr>
<td>Ratio (%)</td>
<td>3</td>
<td>13</td>
<td>84</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 13 reports decomposed variance risk premium into three parts. SV, RD-Jump, and RS-Jump represent stochastic volatility, rare disaster jump, and regime switching jump, respectively. The most of the variance risk premium come from jumps in returns, especially due to jumps from regime-switching.

Table 14: **Return Predictability with Variance Risk Premium**

<table>
<thead>
<tr>
<th>$hp$</th>
<th>Data$^i$</th>
<th></th>
<th>Data$^ii$</th>
<th></th>
<th>Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$</td>
<td>Adj.$R^2$ (%)</td>
<td>$b$</td>
<td>Adj.$R^2$ (%)</td>
<td>$b$</td>
<td>Adj.$R^2$ (%)</td>
</tr>
<tr>
<td>1</td>
<td>0.39</td>
<td>1.07</td>
<td>0.20</td>
<td>0.69</td>
<td>0.19</td>
<td>0.52</td>
</tr>
<tr>
<td>3</td>
<td>0.47</td>
<td>6.87</td>
<td>0.33</td>
<td>2.16</td>
<td>0.19</td>
<td>1.47</td>
</tr>
<tr>
<td>6</td>
<td>0.30</td>
<td>5.42</td>
<td>0.42</td>
<td>9.13</td>
<td>0.18</td>
<td>2.72</td>
</tr>
<tr>
<td>9</td>
<td>0.17</td>
<td>2.30</td>
<td>0.39</td>
<td>11.28</td>
<td>0.17</td>
<td>3.72</td>
</tr>
<tr>
<td>12</td>
<td>0.12</td>
<td>1.23</td>
<td>0.35</td>
<td>11.46</td>
<td>0.16</td>
<td>4.57</td>
</tr>
</tbody>
</table>

Regression estimates using real data are from Bollerslev, Tauchen, and Zhou (2009) and Fusari and Gonzalez-Perez (2012). Regression results with Model are estimated with $10^6$ monthly simulated data using baseline parameters.

i: See Table 2 of Bollerslev, Tauchen, and Zhou (2009).

ii: Regression results are reported in Figure 16 of Fusari and Gonzalez-Perez (2012).
Table 15: **ML estimates using Quarterly Consumption and Monthly Returns**

<table>
<thead>
<tr>
<th>Data</th>
<th>ML Estimates (SE)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_0$</td>
<td>$b$</td>
<td>$\alpha_1$ (%)</td>
<td>$\lambda$</td>
<td>$B_d$</td>
<td>$B_c$</td>
</tr>
<tr>
<td>Quarterly Consumption &amp; Monthly Returns</td>
<td>0.85 (0.02)</td>
<td>1.52 (0.28)</td>
<td>0.80 (0.24)</td>
<td>0.89 (0.01)</td>
<td>0.28 (0.03)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Monthly Returns</td>
<td>0.86 (0.02)</td>
<td>1.45 (0.25)</td>
<td>1.15 (0.36)</td>
<td>0.89 (0.02)</td>
<td>0.26 (0.03)</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 15 reports the ML estimates using quarterly consumption and monthly returns. The estimates of individual parameters are similar across two data sets.

Table 16: **Wald Test on the Differences in Parameters**

<table>
<thead>
<tr>
<th>Test</th>
<th>$\chi^2$</th>
<th>p-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>0.036</td>
<td>1.00</td>
<td>Accept</td>
</tr>
<tr>
<td>Test 2</td>
<td>46.85</td>
<td>0.00</td>
<td>Reject</td>
</tr>
</tbody>
</table>

Table 16 reports results of Wald test on the null hypothesis. The result of Test 1 confirms that the differences in first five parameters are not statistically significant. However, in Test 2, the difference in the size of consumption disaster is fairly large and the null hypothesis is rejected. This should be expected since we cannot observe disasters in the post-war U.S. consumption data.
Table 17: Model Implied Moments using Quarterly Consumption and Monthly Returns

<table>
<thead>
<tr>
<th></th>
<th>Annualized Moments(%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mathbb{E}[r_{m,t} - r_{f,t}]$</td>
<td>$\mathbb{E}[r_{f,t}]$</td>
<td>$\text{STD}[r_{m,t}]$</td>
<td>$\text{STD}[r_{f,t}]$</td>
</tr>
<tr>
<td>Quarterly Consumption</td>
<td>0.53</td>
<td>2.94</td>
<td>13.01</td>
<td>0.10</td>
</tr>
<tr>
<td>&amp; Monthly Returns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly Returns</td>
<td>6.56</td>
<td>1.87</td>
<td>17.80</td>
<td>0.46</td>
</tr>
<tr>
<td>Data</td>
<td>5.55</td>
<td>0.38</td>
<td>18.91</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 17 reports implied moments with ML estimates using quarterly consumption and monthly returns. Sample moments are identical to those in Table 8.
### Table 18: ML Estimates using Quarterly Endowments

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ML Estimates (SE)</th>
<th>$\bar{k} = 1$</th>
<th>$\bar{k} = 2$</th>
<th>$\bar{k} = 3$</th>
<th>$\bar{k} = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_0$</td>
<td>(0.05)</td>
<td>0.83</td>
<td>0.64</td>
<td>0.75</td>
<td>0.81</td>
</tr>
<tr>
<td>$b$</td>
<td>NA (9.84)</td>
<td>13.71</td>
<td>4.27</td>
<td>2.71</td>
<td></td>
</tr>
<tr>
<td>$\alpha_1(%)$</td>
<td>(13.94)</td>
<td>27.64</td>
<td>4.88</td>
<td>4.76</td>
<td>4.57</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>(0.01)</td>
<td>0.98</td>
<td>0.89</td>
<td>0.93</td>
<td>0.95</td>
</tr>
<tr>
<td>$B_c$</td>
<td>(0.00)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>$B_d$</td>
<td>(0.00)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>$\delta_1(%)$</td>
<td>(0.42)</td>
<td>9.69</td>
<td>8.64</td>
<td>8.64</td>
<td></td>
</tr>
<tr>
<td>$\delta_2(%)$</td>
<td>(0.71)</td>
<td>6.11</td>
<td>6.13</td>
<td>6.14</td>
<td>6.13</td>
</tr>
<tr>
<td>$\delta_3(%)$</td>
<td>(0.88)</td>
<td>9.15</td>
<td>8.82</td>
<td>8.83</td>
<td>8.82</td>
</tr>
<tr>
<td>$\log L(\times 10^3)$</td>
<td>1.117</td>
<td>1.148</td>
<td>1.147</td>
<td>1.147</td>
<td></td>
</tr>
<tr>
<td>-BIC ($\times 10^3$)</td>
<td>2.120</td>
<td>2.234</td>
<td>2.232</td>
<td>2.231</td>
<td></td>
</tr>
</tbody>
</table>

Table 18 reports ML estimates using quarterly endowment process.

### Table 19: Model Implied Moments using Quarterly Endowments

<table>
<thead>
<tr>
<th>Annualized Moments(%)</th>
<th>$E[r_{m,t} - r_{f,t}]$</th>
<th>$E[r_{f,t}]$</th>
<th>$STD[r_{m,t}]$</th>
<th>$STD[r_{f,t}]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5.29</td>
<td>2.13</td>
<td>12.48</td>
<td>0.55</td>
</tr>
<tr>
<td>Data</td>
<td>5.55</td>
<td>0.38</td>
<td>18.91</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 19 reports implied moments with parameters described in Section 6. Sample moments are identical to those in Table 8.