Term structure of risk in expected returns [*]

Irina Zviadadze[†]

This Draft: March 8, 2018

Abstract

Return predictability reveals economic variables that drive expected returns. Alternative economic theories relate fluctuations in predictive variables to different sources of risk. I develop an empirical approach that exploits these observations and measures how economically interpretable shocks propagate in the term structure of expected buy-and-hold returns. Shock propagation patterns constitute term structure of risk in expected returns whose shape and level serve as informative moments to test competing equilibrium theories of return predictability. As an application, I examine sources of stock return predictability. I find that equilibrium shocks in the long-run mean of the variance of consumption growth can justify the level and the shape of the term structure of expected stock returns, in contrast to consumption disasters or long-run risk.

Keywords: term structure of risk, return predictability, incremental expected return

PRELIMINARY

---

1 Introduction

We have seen compelling evidence of return predictability in different asset classes: bonds, stocks, commodities, and currencies. This evidence conveys at least two economic messages. One is about designing profitable investment strategies to pocket risk premia. The other one is about understanding which economic variables drive expected returns and aggregate discount rates. While economic performance of investment strategies may be fragile at turbulent times, we can forcefully use the second message as a direct source of information about economic shocks reflected in time varying expected returns at alternative investment horizons.

The dynamics of expected returns reveals interactions between risk preferences of economic agents and macroeconomic shocks across alternative horizons. A vector autoregressive model (VAR) of a return and a predictive variable captures these interactions implicitly via multi-horizon return predictability, whereas economic theory describes these interactions explicitly. A natural question then is what multi-horizon return predictability implies for sources of the multi-period risk-return tradeoff in equilibrium models, and whether known equilibrium models are in line with these implications.

For any well-accepted predictive variable, there are multiple theories advocating competing risk channels behind the return predictability. In this paper, I propose an empirical methodology which identifies how alternative economic sources of return predictability propagate in expected buy-and-hold returns of different horizons. I label these shock propagation patterns the term structure of risk in expected returns. The level and the shape of the term structure of risk are informative metrics about the multi-period dynamics of prices and quantities of risk, and therefore, they are useful moments to target in calibrations of structural models. As a result, my empirical methodology delivers a new set of facts that can be used to discipline an equilibrium theory of time-varying risk premia.

To identify economic shocks in expected returns, I estimate a predictive system, in which I explicitly relate the predictive variable to economic states, suggested by theory. This system is a VAR of a one-period return and state variables augmented by an observation equation describing an unrestricted mapping between the predictive variable and the economic states. As an implication of return predictability, the vector autoregressive model necessarily features nonnormal shocks. These shocks are particularly interesting because they reflect a notion of extreme events and innovations in the macroeconomic variance, default probabilities, or disaster intensities, and they also represent central sources of risk compensated in financial markets.

To describe how the identified economic shocks propagate in the term structure of expected returns, I compute an incremental expected return ($\mathcal{IER}$). The $\mathcal{IER}$ measures how a future multi-period expected return changes if next period one of the state variables or return itself experiences an injection of an extra amount of risk; and it applies to both normal and nonnormal shocks. The collection of $\mathcal{IER}s$ of different holding periods constitute the term structure of risk in expected returns.

To the best of my knowledge, this paper is first to describe the term structure of risk for nonnormal shocks in asset pricing. The closest predecessor of the $\mathcal{IER}$ is a shock elasticity of Borovička and Hansen (2014), which was developed to characterize marginal sensitivities of multi-period expected returns or cash flows to normal sources of risk. I extend the notion of shock elasticity to the case of nonnormal shocks and show that the two metrics coincide for normal shocks but differ for nonnormal shocks. I explain the source of difference.

I develop three tractable cases of $\mathcal{IER}$ for nonnormal shocks: (i) Poisson mixture of gamma distributions with a scale parameter of one for volatility shocks, jump intensity shocks, or default probability shocks, (ii) Poisson mixture of exponentials for one-sided jumps such as jumps in variance, and (iii) Poisson mixture of normals for two-sided jumps, for example jumps in currency returns. The aforementioned sources of volatility and jump risk are dominant drivers of predictability and sources of risk premia in the state-of-the-art asset pricing literature.[1]

To illustrate how my approach works in practice, I examine the term structure of macroeconomic risk in equity returns. I rely on the predictive ability of the price-dividend ratio to forecast future expected returns and map fluctuations in the price-dividend ratio to alternative sources of equity risk premium advocated by several leading structural models. I choose tractable, parsimonious, and yet stylized macro-based asset pricing models that account for many salient data properties.

I study how alternative risk channels, such as (i) a long run risk with stochastic volatility of consumption growth (Bansal and Yaron (2004)), (ii) time-varying consumption disasters (Wachter (2013)), and (iii) stochastic volatility of consumption growth with a time-varying long-run mean and self-exciting jumps (Drechsler and Yaron (2011)), are reflected in the term structure of expected returns. These risk channels are interesting to explore on their own but also because they are prevalent in applications in which the entire term structure of risk must be a natural object of interest. Examples include but not limited to a climate change research (e.g., Bansal, Kiku, and Ochoa (2015) and Barro (2013)), impact of monetary policy and fiscal policy on asset prices (e.g., Gallmeyer, Hollifield, Palomino, and Zin (2007) and Liu (2016)), and impact of innovation risk on financial markets (e.g., Croce, Nguyen, Raymond, and Schmid (2017) and Kung and Schmid (2015)).

To accomplish my task, I estimate three systems containing a predictive model for stock returns and a law of motion for the state vector featured in one of the aforementioned structural models. I document that in every case the term structure of expected buy-and-hold returns is downward sloping. This property is a manifestation of the multivariate mean-reversion in returns (Cochrane,

---

[1]Augustin (2016), Le, Singleton, and Dai (2010), and Zviadadze (2017) use an autoregressive Gamma process to model the credit default process, the inverse consumption surplus ratio, and the stochastic volatility of macroeconomic fundamentals, respectively. Hassler and Marfe (2016) use the Poisson mixture of exponentials to model disasters and recoveries in the expected consumption growth and in the expected dividend growth. Drechsler and Yaron (2011) use the Poisson mixture of normals to model disasters in the long-run risk component of consumption growth.

2001), and it also serves as a reality check of my empirical procedure. As a consequence, every empirical model should have at least one shock associated with a downward sloping term structure of risk in expected returns. My empirical results identify such a shock: it is the long-run risk shock in the economic environment motivated by Bansal and Yaron (2004); it is a disaster intensity shock in the economic environment motivated by Wachter (2013), and it is a long-run volatility shock in the economic environment motivated by Drechsler and Yaron (2011).

Next, I analyze the term structures of risk implied by original calibrations of the structural models, that is theoretical term structures of risk. I find discrepancy between empirical and theoretical moments. In Bansal and Yaron (2004), the long-run risk shock has an upward sloping (not a downward sloping, like in the data) term structure of risk in expected returns. In Wachter (2013), the intensity shock features a downward sloping term structure of risk like in the data; yet the disaster shock has a flat negative (not a positive, like in the data) term structure of risk. Thus, in these models there is a tension between getting both the level and the shape of the term structure of risk right. The model of Drechsler and Yaron (2011) shows success along the term structure dimension: it matches both the shape and the level of the term structure of risk for all sources of risk.

Because the empirical term structures of risk in expected returns are new moments, perhaps, it is not surprising that the original calibrations of Bansal and Yaron (2004) and Wachter (2013) fall short to match them. Instead of re-calibrating the models, I ask what are the necessary conditions to match, at least qualitatively, the shape of the term structure of the long-run risk shock and the level of the term structure of the disaster risk. They turn out to be (i) a negative value for the parameter of the intertemporal elasticity of substitution in Bansal and Yaron (2004) and (ii) modeling an aggregate dividend as a levered consumption claim with a *negative* leverage parameter in Wachter (2013). Because these conditions are economically implausible, I conclude that alternative calibrations would not allow these models to match the evidence on the term structure of risk in expected returns.

This simple empirical exercise illustrates how an asset return predictability translates into a set of informative moments about the term structure of risk in expected returns. An interesting aspect of this analysis is that I exploit a joint likelihood of macroeconomic and asset pricing data for highlighting weak aspects of structural models. In the literature, there is a well-grounded scepticism regarding the use of asset prices when estimating macro-based structural models (e.g., Chen, Dou, and Kogan (2017)). The concern is that asset prices imply such properties of macroeconomic data that are impossible to test, and therefore impossible to reject. I turn exactly this feature of a tight dependence of the distribution of macroeconomic risk on asset prices, via predictability, into an informative set of moments about the properties of economic shocks in the data and in asset pricing models.

Related literature

This paper speaks to several strands of literature. First, it contributes to a rapidly growing empirical literature on the term structure of risk premia. Recent contributions include Bansal,

3

Miller, and Yaron (2017), Binsbergen, Brandt, and Koijen (2012), Binsbergen, Hueskes, Koijen, and Vrugt (2013), Dew-Becker, Giglio, Le, and Rodriguez (2015), Gormsen (2017), Binsbergen and Koijen (2017) and references therein. The aforementioned studies rely on novel datasets of zero-coupon yields of different maturities and compare the risk-return profile of short-term and long-term instruments. The challenge for general equilibrium theorists is to relate these empirical facts to specific sources of macroeconomic risk. In this paper, I propose a model-based methodology, which (i) overcomes the above mentioned challenge, (ii) provides a complementary evidence on the multi-period risk-return tradeoff in financial markets, and (iii) does not use data on zero-coupon yields of different maturities whose availability is still limited. I exploit empirical patterns of return predictability to describe how economically interpretable shocks propagate in the term structure of buy-and-hold returns on the same asset.

My methodology is motivated by macroeconomics literature on testing economic theories by means of examining shock propagation patterns in key macroeconomic indicators across alternative horizons, also known as, impulse response functions (e.g., Christiano, Eichenbaum, and Evans (2005), Fernandez-Villaverde, Rubio-Ramirez, and Schorfheide (2016), Giacomini (2013), Smets and Wouters (2007), among others).[2] I apply a similar idea to structural dynamic asset pricing models. I introduce a novel metric labeled an incremental expected return $\mathcal{IER}$ that applies to both normal and nonnormal shocks. The metric quantifies the impact of a change in the amount of risk on expected returns of alternative holding periods. The $\mathcal{IER}$ for normal shocks coincides with the shock elasticity of Borovička and Hansen (2014). I extend the methods of dynamic value decomposition (Borovička and Hansen, 2014; Borovička, Hansen, Hendricks, and Scheinkman, 2011; Borovička, Hansen, and Scheinkman, 2014; Hansen (2012)) to nonnormal shocks and delineate the difference between the $\mathcal{IER}$ and shock elasticity for nonnormal sources of risk. This extension is a separate methodological contribution of this paper. The $\mathcal{IER}$ and shock elasticity for nonnormal sources of risk may serve as a useful tool for diagnostics methods of discrete-time macro-based asset pricing models with time-varying risk premia.

Last but not least, it contributes to the methodological literature that relates properties of data to model-based metrics with the purpose of guiding progress in building equilibrium models. Examples include but not limited to Alvarez and Jermann (2005), Backus, Chernov, and Zin (2014), Backus, Boyarchenko, and Chernov (2015), Bakshi and Chabi-Yo (2012), Bakshi, Gao, and Panayotov (2017), Chabi-Yo and Colacito (2017), Cochrane and Hansen (1992), Hansen and Jagannathan (1991). Especially relevant studies are those that use properties of return predictability (e.g., Bekaert and Liu, 2004, Chabi-Yo, 2008, Favero, Ortu, Tamoni, and Yang, 2017, Gallant, Hansen, and Tauchen, 1990, Yang, 2014) for developing performance measures for structural models. A distinctive feature of my approach is that I use an empirical model for summarizing properties of the data, whereas the previous studies rely on model-free statistics. Extra assumptions that I make decode properties of specific economic shocks in the term structure of asset returns. Thus,

---

[2]There is a growing literature on estimating macroeconomic dynamic stochastic general equilibrium (DSGE) models by matching impulse response functions. For example, see Guerron-Quintana, Inoue, and Kilian (2017) and references therein.

I can analyze models not as a whole but on a shock by shock basis. As a result, my approach is more granular in nature.

Relatedly, this paper adds to an on-going discussion of alternative channels and mechanisms that may provide a realistic macroeconomic view on the term structure of risk premia in financial markets. Andries, Eisenbach, and Schmalz (2017), Belo, Collin-Dufresne, and Goldstein (2015), Croce, Lettau, and Ludvigson (2015), Hassler and Marfe (2016), and Marfe (2016) are some examples of prominent contributions to this debate.

The paper is organized as follows. Section I discusses how to examine propagation of various economic shocks (normal and nonnormal) in the term structure of expected returns. Section II develops an empirical framework for analyzing the term structure of macroeconomic risk in equity returns. Section III documents the estimation results and examines the term structures of alternative sources of risk premium in structural models and in the data. Section IV concludes. The Internet Appendix contains supplementary material.

2   Term structure of risk in expected returns

A joint model of an asset return and of an observable predictive variable is informative about the multi-period dynamics of expected returns. Trajectories of shocks that originate in the predictive variable and propagate in future expected returns reveal the presence or absence of such patterns in returns as mean reversion and momentum. In equilibrium, these patterns depend on the asset's exposures to risk and prices of risk across alternative horizons, and therefore they characterize the multi-period risk-return tradeoff.

What are the economic origins of the multi-period risk-return tradeoff? For every established return predictor, the existing literature proposes a number of alternative theories of (i) why the predictive variable varies over time and (ii) why it has a forecasting power for future returns. An economic interpretation of return predictability goes hand in hand with shock identification. In the language of econometric analysis, every structural model represents a number of shock identifying assumptions.

In this section, I describe a methodology for examining the term structure of economic shocks in expected buy-and-hold returns. This methodology has two important ingredients. One ingredient is an identification of economic sources of predictability by means of augmenting a reduced-form predictive system with a number of theoretically motivated shock identifying assumptions. Another ingredient is a characterization of the trajectories of alternative economic shocks in the term structure of expected returns. An empirical application of this methodology, which I discuss in the next section, delivers a collection of stylized facts about the multi-horizon properties of economic sources of stock return predictability.

To fix ideas, I start with a simple illustrative example. I denote a one-period log return $\log r_{t,t+1}$ and its predictive variable y_t and model their joint dynamics as the following system

$$\log r_{t,t+1} = f(y_t) + w_{rt+1}, \tag{1}$$
$$y_{t+1} = \kappa(y_t) + w_{yt+1}. \tag{2}$$

Fluctuations w_{rt+1} and w_{yt+1} can be correlated; functions f and κ can be nonlinear; y_t can be a vector of predictors (without loss of generality and for simplicity, y_t is a single predictor here); and the system may contain extra exogenous variables, which are without loss of generality omitted here.[3] The drift of the return equation does not contain the lagged return because returns are nearly iid.[4] The order of the predictive system does not have to be necessarily equal to 1. If lags of the predictive variable larger than 1 have forecasting power for future returns, the vector y_t can be extended by incorporating those lags as separate predictive variables.

I specify the predictive equation (1) for returns in logs rather than in levels because of convenience of linear time series analysis. The multi-period log returns are naturally a sum of one-period log returns, and therefore, the autoregressive structure of the model (1)-(2) provides direct evidence about how the predictive variable y_t affects future multi-period returns. As emphasized in Cochrane (2008), separate regressions for returns of alternative horizons do not convey any additional economic message regarding predictability over and above that reflected already by the system like one described by expressions (1)-(2). In addition, the long-run predictive evidence obtained through the lens of multivariate predictive systems exhibits greater statistical power (Campbell, 2001; Cochrane, 2008; Valkanov, 2003).[5]

Return predictor is an endogenous variable. Every equilibrium model maps it, through the first-order optimality conditions, to a model-specific state vector, and therefore, to a number of economic shocks. As a result, the system (1)-(2), augmented by a model-implied representation of the predictive variable y_t as a function of a state vector, encodes properties of economic sources of return predictability. This observation lies at the core of my analysis.

For example, consider a hypothetical structural model which describes the risk attitude of an economic agent to the distribution of underlying shocks. The model features a state vector s_t which summarizes time variation in the aggregate discount rate and follows the law of motion[6]

$$s_{t+1} = \tilde{\kappa}(s_t) + w_{st+1}. \tag{3}$$

---

[3]In the empirical section of this paper, I use tractable processes with linear functions f and κ.

[4]See Cochrane (2001) for a discussion of why returns can appear iid in the presence of return predictability.

[5]Recently Giglio and Kelly (2017) argue that the variability of prices on short-term and long-term claims on variety of different cash flows rejects the internal consistency conditions of autoregressive processes predominantly used in asset pricing modeling. In this paper, I stick to predictive systems modeled as autoregressive multivariate processes because I have to augment a reduced-form model with identifying assumptions stemming from structural models. To the best of my knowledge, there are no structural models reconciling evidence of Giglio and Kelly (2017) yet.

[6]Without loss of generality, s_t is a scalar random variable here.

The first-order optimality conditions imply that the predictive variable y_t is some function g of the state s_t

$$y_t = g(s_t), \tag{4}$$

and the parameters of the function g depend on the preference parameters and the parameters governing the distribution of risk in the economy. In addition, the underlying assumptions of the structural model relate reduced-form fluctuations in returns w_{rt+1} to structural shocks w_{st+1} via mapping η

$$w_{rt+1} = \eta(w_{st+1}), \tag{5}$$

so that the return dynamics is

$$\log r_{t,t+1} = \tilde{f}(s_t) + \eta(w_{st+1}), \tag{6}$$

where $\tilde{f}(s_t) = f(g(s_t))$.[7] As a result, the predictive system (1)-(2) augmented with economic restrictions (3)-(5) is equivalent to the state-space model represented by equations (3) and (6). The state-space formulation of the predictive system encodes multi-horizon properties of economic sources of return predictability w_{st+1}.

I examine the multi-period risk-return tradeoff in asset markets by focusing on how alternative economic shocks propagate through expected buy-and-hold returns on the same asset across investment horizons. I define these propagation patterns as a term structure of risk in expected returns. I analyze both the level and the shape of the term structure of risk for each individual shock, that is (i) whether a positive shock shifts returns up or down and (ii) whether the corresponding impact is horizon-dependent. These seemingly basic properties of economic shocks serve as informative and powerful metrics for evaluating economic mechanisms in macro-based asset pricing models.

Mathematically, I define the term structure of risk in expected returns as a set of revisions of expected multi-period returns, conditional on an information set $\mathcal{I}_t$, that coincides with one individual economic shock arriving at $t+1$. Effectively, I measure the impact of a shock on returns of alternative horizons, and therefore, I use terms "risk" and "term structure". I reserve term "incremental expected return" $\mathcal{IER}$ to describe one element of the term structure of risk corresponding to a horizon τ

$$\mathcal{IER}(r_{t,t+\tau}, \text{shock}_{t+1}, \mathcal{I}_t) = E(r_{t,t+\tau}|\mathcal{I}_t, \text{shock}_{t+1}) - E(r_{t,t+\tau}|\mathcal{I}_t). \tag{7}$$

The $\mathcal{IER}$ takes into account the presence of other sources of risk in the economic environment, that is it measures an *incremental* effect of a shock. It applies both to normal and nonnormal shocks, and therefore, it is a suitable metric for analyzing the term structure of risk for sources of time-varying risk premium, such as variance shocks and time-varying disaster shocks, among others.

---

[7]In the macroeconomic literature mapping η is usually represented by a linear combination of shocks.

In what follows, I illustrate how to characterize the term structure of economic shocks in two simple examples. I consider a predictive system (1)-(2) for stock returns $\log r_{t,t+1}$ and a price-dividend as a predictive variable y_t. I borrow shock identification restrictions from two competing theories of stock return predictability: (1) price-dividend ratio reflects variance risk and (2) price-dividend ratio reflects risk driving the time-varying probability of crashes in returns; and then characterize how the identified economic sources of return predictability propagate in returns across alternative horizons.[8] In *Example 1*, I examine the case of a normal variance shock, and in *Example 2*, I examine the case of a nonnormal disaster risk shock.

Example 1: Term structure of variance risk in expected returns

For simplicity and tractability, I consider the system (1)-(2) with linear mappings f and κ. The vector of the log return and of the price-dividend ratio follows a vector autoregressive process of order 1 with two restrictions: (i) the lagged return does not Granger-cause the future return and (ii) the lagged return does not predict the price-dividend ratio,

$$\log r_{t,t+1} \quad = \quad a + b \cdot \log pd_t + w_{rt+1}, \tag{8}$$

$$\log pd_{t+1} \quad = \quad (1 - \nu_{pd})v_{pd} + \nu_{pd} \cdot \log pd_t + \sigma_{pd}w_{pd,t+1}. \tag{9}$$

The first parameter restriction is standard because of the absence of serial correlation in stock returns. The second restriction stems from the fact that empirically the price-dividend ratio Granger-causes the stock returns, but the stock returns do not Granger-cause the price-dividend ratio.

As a persistent variable, the price-dividend ratio follows an autoregressive process. Strictly speaking the autoregressive process is not a suitable choice for modeling the price-dividend ratio, because it does not eliminate the possibility of a negative realization of the variable. However, in the context of my illustrative example it is not a concern. I take a proper care of modeling nonnegative variables in the empirical section by replacing an autoregressive process with an autoregressive gamma process (Gourieroux and Jasiak (2006), Le, Singleton, and Dai (2010)).

I identify economic sources of return predictability by applying the shock identification scheme summarized in *Hypothesis 1* to the system (8)-(9): (i) the variance of log returns is stochastic, $Var_t(w_{rt+1}) = v_t$ and perfectly (positively or negatively) correlated with the price-dividend ratio

$$\log pd_t = q_0 + q_v v_t, \tag{10}$$

and (ii) conditional on the current information set $\mathcal{I}_t$ the shocks to returns $w_{rt+1} = v_t^{1/2}\varepsilon_{rt+1}$ are orthogonal to the variance shocks ε_{vt+1}, $\mathrm{corr}_t(v_t^{1/2}\varepsilon_{rt+1}, \varepsilon_{vt+1}) = 0$.[9] As a result, the predictive

---

[8]In this section, I am agnostic about the exact nature of the sources of risk. For example, I do not take a stand on whether these are macroeconomic shocks.

[9]The restriction $\mathrm{corr}_t(v_t^{1/2}\varepsilon_{rt+1}, \varepsilon_{vt+1}) = 0$ is an over-identifying restriction. At this stage, I do not contemplate whether some of or all the identifying restrictions are empirically realistic.

system given in (8)-(9) can be represented as

$$\log r_{t,t+1} = a_v + b_v v_t + v_t^{1/2} \varepsilon_{rt+1}, \tag{11}$$
$$v_{t+1} = (1 - \nu_v)v_v + \nu_v v_t + \sigma_v \varepsilon_{vt+1}, \tag{12}$$

where

$$
\begin{aligned}
a_v &= a + bq_0, \\
b_v &= bq_v, \\
\nu_v &= \nu_{pd}, \\
v_v &= [(1 - \nu_{pd})v_{pd} - q_0 + \nu_{pd}q_0]/[q_v(1 - \nu_{pd})], \\
\sigma_v &= \sigma_{pd}/q_v.
\end{aligned}
$$

The state-space representation (11) and (12) describes the dynamics of stock returns $r_{t,t+\tau}$ as a function of the state variable v_t and *economic* shocks ε_{rt+1} and ε_{vt+1}, rather than as a function of the predictive variable pd_t and reduced-form shocks w_{rt+1} and w_{pdt+1}. The primer representation is useful for analyzing how economic sources of variation in stock returns propagate in the term structure of expected returns across alternative horizons.

The main challenge in computing the $\mathcal{IER}$ is to dissect the impact of alternative shocks at different horizons on expected returns. I achieve my goal by (i) representing the multi-period expected returns as a function of state variables at time t and shocks at time $t+1$ and (ii) comparing the trajectories of expected returns, if shocks at time $t + 1$ experience and do not experience a predetermined exogenous shift. The first step guarantees that the future propagation of the shocks arrived at $t+1$ are taken into account and their impact is isolated from that of shocks arriving later. The second step operationalizes computation of the revisions in the expected returns coinciding with a shock at time $t + 1$.

Below I illustrate step by step how to compute the $\mathcal{IER}$s associated with ε_{vt+1} and ε_{rt+1}. I start by characterizing the expected returns

$$\log(E_t r_{t,t+\tau}) = \mathcal{A}_0(\tau) + \mathcal{A}_v(\tau)v_t,$$

via a standard recursive system

$$
\begin{aligned}
\mathcal{A}_0(\tau) &= a_v + \mathcal{A}_0(\tau - 1) + \mathcal{A}_v(\tau - 1)(1 - \nu_v)v_v + \mathcal{A}_v^2(\tau - 1)\sigma_v^2/2, \\
\mathcal{A}_v(\tau) &= b_v + \mathcal{A}_v(\tau - 1)\nu_v + \mathcal{A}_v^2(\tau - 1)/2
\end{aligned}
$$

with initial conditions

$$
\begin{aligned}
\mathcal{A}_0(1) &= a_v, \\
\mathcal{A}_v(1) &= b_v + 1/2.
\end{aligned}
$$

Next I apply the law of iterated expectations to $E_t r_{t,t+\tau}$ in order to represent the expected multi-period returns in terms of the variables belonging to the information set $\mathcal{I}_t$ and shocks of interest arriving at $t+1$

$$
\begin{aligned}
\log E_t r_{t,t+\tau} &= \log(E_t r_{t,t+1} \cdot E_{t+1} r_{t+1,t+\tau}) \\
&= \log E_t(\exp(a_v + \mathcal{A}_0(\tau-1) + \mathcal{A}_v(\tau-1)(1-\nu) + (b_v + \mathcal{A}_v(\tau-1))v_t + v_t^{1/2}\varepsilon_{rt+1} \\
&+ \mathcal{A}_v(\tau-1)\sigma_v \varepsilon_{vt+1})).
\end{aligned}
$$

Finally, I consider the following thought experiment. First, I introduce an additional amount of variance risk, so that the stochastic variance of returns at time $t+1$ is equal to $\tilde{v}_{t+1} = v_{t+1} + \Delta_v$, where v_{t+1} is a random variable and Δ_v is known, or equivalently I shift the shock ε_{vt+1} by the amount Δ_v/σ_v, $\tilde{\varepsilon}_{vt+1} = \varepsilon_{vt+1} + \Delta_v/\sigma_v$

$$
\tilde{v}_{t+1} = v_{t+1} + \Delta_v = (1-\nu)v + \nu v_t + \sigma_v(\varepsilon_{vt+1} + \Delta_v/\sigma_v).
$$

Then, I measure how the expected multi-period returns change as a result of the extra amount of variance risk Δ_v/σ_v at $t+1$

$$
\begin{aligned}
\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{vt+1}, v_t) &= \log E(r_{t,t+\tau}|\mathcal{I}_t, \tilde{v}_{t+1} = v_{t+1} + \Delta_v) - \log E(r_{t,t+\tau}|\mathcal{I}_t) \\
&= \log E_t(\exp(a_v + b_v v_t + v_t^{1/2}\varepsilon_{rt+1} + \mathcal{A}_0(\tau-1) + \mathcal{A}_v(\tau-1)(v_{t+1} + \Delta_v))) \\
&- \log E_t(\exp(a_v + b_v v_t + v_t^{1/2}\varepsilon_{rt+1} + \mathcal{A}_0(\tau-1) + \mathcal{A}_v(\tau-1)v_{t+1})) \\
&= \mathcal{A}_v(\tau-1) \cdot \Delta_v. \tag{13}
\end{aligned}
$$

The choice of Δ_v is a matter of normalization. I set Δ_v to be equal to an unconditional standard deviation of the state variable shifted by the shock of interest, that is, in this case, at one standard deviation of the stochastic variance v_t, $\Delta_v = \sigma_v/(1-\nu_v^2)^{1/2}$.

Similarly I characterize the term structure of the direct return shock ε_{rt+1}. I introduce an additional amount of the direct return risk Δ_r into the economic environment at time $t+1$, $\log \tilde{r}_{t,t+1} = \log r_{t,t+1} + \Delta_r$, which is equivalent to shifting the shock ε_{rt+1} by Δ_r/v_t, $\tilde{\varepsilon}_{rt+1} = \varepsilon_{rt+1} + \Delta_r/v_t$, and compute the incremental expected returns for alternative τ[10]

$$
\begin{aligned}
\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{rt+1}, v_t) &= \log E(r_{t,t+\tau}|\mathcal{I}_t, v_t^{1/2}\tilde{\varepsilon}_{rt+1} = v_t^{1/2}\varepsilon_{rt+1} + \Delta_r) - \log E(r_{t,t+\tau}|\mathcal{I}_t) \\
&= \log E_t(\exp(a_v + b_v v_t + (v_t^{1/2}\varepsilon_{rt+1} + \Delta_r) + \mathcal{A}_0(\tau-1) + \mathcal{A}_v(\tau-1)v_{t+1})) \\
&- \log E_t(\exp(a_v + b_v v_t + v_t^{1/2}\varepsilon_{rt+1} + \mathcal{A}_0(\tau-1) + \mathcal{A}_v(\tau-1)v_{t+1})) \\
&= \Delta_r. \tag{14}
\end{aligned}
$$

I set Δ_r to be equal to one standard deviation of the log return (the only variable affected by the shock), $\Delta_r = Var^{1/2}(\log r_{t,t+1})$. The shape and the level of the term structure of risk associated

---

[10]The state-dependent shift in the shock ε_{rt+1} offsets the state-dependence in the stochastic volatility. There are alternative ways to treat state-dependence which I leave out of scope in this paper. Questions of state dependence are relevant for characterizing the conditional dynamics of shocks in asset prices.

with the shock ε_{vt+1} is determined by a recursion for $\mathcal{A}_v(\tau)$. The term structure of ε_{rt+1} is flat at the level of Δ_r, because ε_{rt+1} is a permanent shock in returns.

Example 2: Term structure of disaster risk in expected returns

The starting point is the same reduced-form predictive system given in expressions (8) and (9). I consider an alternative hypothesis of return predictability, and therefore, I identify alternative economic shocks that potentially lie at the core of the multi-period risk-return relationship. *Hypothesis 2* reflects the idea that the only source of variation in log returns is time-varying crashes, $-z_{t+1}$, and that their probability λ_t is a source of return predictability, that is

$$\log pd_t = q_0 + q_\lambda \lambda_t. \tag{15}$$

The negative of a crash, that is a random variable z_{t+1}, follows a distribution defined as a Poisson mixture of Gammas,

$$z_{t+1}|p_{t+1} \sim \Gamma(p_{t+1}, \theta), \quad \text{where} \quad p_{t+1} \sim \text{Poisson}(\lambda_t), \tag{16}$$

so that z_{t+1} is a nonnegative shock, and therefore a crash, if realized, is a negative disturbance $-z_{t+1}$. The crashes $-z_{t+1}$ are conditionally uncorrelated with the shocks driving the probability of crashes $\varepsilon_{\lambda t+1}$: $\text{corr}_t(-z_{t+1}, \varepsilon_{\lambda t+1}) = 0$.

The reduced-form predictive model (8) and (9) augmented with the identifying assumptions for shocks in returns and price-dividend ratio given in expressions (15) and (16) is equivalent to the system

$$\log r_{t,t+1} = a_\lambda + b_\lambda \lambda_t - z_{t+1}, \tag{17}$$
$$\lambda_{t+1} = (1 - \nu_\lambda)v_\lambda + \nu_\lambda \lambda_t + \sigma_\lambda \varepsilon_{\lambda,t+1}, \tag{18}$$

where

$$a_\lambda = a + bq_0,$$
$$b_\lambda = bq_\lambda,$$
$$\nu_\lambda = \nu_{pd},$$
$$v_\lambda = [(1 - \nu_{pd})v_{pd} - q_0 + \nu_{pd}q_\lambda]/[q_\lambda(1 - \nu_{pd})].$$
$$\sigma_\lambda = \sigma_{pd}/q_\lambda.$$

I use the state-space model specified in expressions (17)-(18) to characterize how the crash risk and intensity shock $\varepsilon_{\lambda t+1}$, that is an economic source of predictability, affect future returns at alternative horizons.

There is one distinct feature of nonnormal sources of risk, such as a crash risk. These risks are characterized by two random shocks: (i) one controling how many crashes occur per period of time and (ii) the other controling the size of a crash. Analyzing these shocks separately is challenging

11

because they work as a whole. See Backus (2014) for a discussion on this issue. For this reason, I bundle these shocks together into one demeaned random variable and examine its corresponding pricing implications.

Specifically, I use an insight from Gourieroux and Jasiak (2006) and represent the random shock z_{t+1} as a random process

$$z_{t+1} \quad = \quad \underbrace{\theta\lambda_t}_{E_t(z_{t+1})} + \underbrace{\underbrace{(2\theta^2\lambda_t)^{1/2}}_{Var_t^{1/2}(z_{t+1})}\varepsilon_{zt+1}}_{\text{martingale difference}} \quad .$$

The shock $\varepsilon_{zt+1} \sim \mathcal{D}(0,1)$ is nonnormal with known nonzero moments of order higher than 1. I characterize the term structure implications of the disaster risk z_{t+1} by studying how the shock ε_{zt+1} propagates in the term structure of expected returns. This is a legitimate and sensible exercise because what an investor cares about is an uncertain component of z_{t+1} represented here by ε_{zt+1}.

To characterize the term structures of ε_{zt+1} and $\varepsilon_{\lambda t+1}$, I follow the same procedure as in *Example 1*. I start by computing the term structure of expected returns

$$\log(E_t r_{t,t+\tau}) \quad = \quad \mathcal{A}_0(\tau) + \mathcal{A}_\lambda(\tau)\lambda_t,$$

where

$$
\begin{aligned}
\mathcal{A}_0(\tau) &= a_\lambda + \mathcal{A}_0(\tau-1) + \mathcal{A}_\lambda(\tau-1)(1-\nu_\lambda)v_\lambda + \mathcal{A}_\lambda^2(\tau-1)\sigma_\lambda^2/2, \\
\mathcal{A}_\lambda(\tau) &= b_\lambda + \mathcal{A}_\lambda(\tau-1)\nu_\lambda - \theta/(1-\theta), \\
\mathcal{A}_0(1) &= a_\lambda, \\
\mathcal{A}_\lambda(1) &= -\theta/(1-\theta).
\end{aligned}
$$

Next, I apply the law of iterated expectations to $E_t r_{t,t+\tau}$ and represent the expected multi-period returns as a function of variables belonging to the information set $\mathcal{I}_t$, and shocks of interest arriving at $t+1$

$$
\begin{aligned}
\log E_t r_{t,t+\tau} &= \log(E_t(r_{t,t+1} \cdot E_{t+1,t+\tau})) \\
&= \log E_t(\exp(a_\lambda + b_\lambda\lambda_t - z_{t+1} + \mathcal{A}_0(\tau-1) + \mathcal{A}_\lambda(\tau-1)\lambda_{t+1})) \\
&= \log E_t(\exp(a_\lambda - \theta\lambda + \mathcal{A}_0(\tau-1) + \mathcal{A}_\lambda(\tau-1)(1-\nu_\lambda)v_\lambda \\
&\quad + (b_\lambda + \mathcal{A}_\lambda(\tau-1)\nu_\lambda)\lambda_t - (2\theta^2\lambda_t)^{1/2}\varepsilon_{zt+1} + \mathcal{A}_\lambda(\tau-1)\sigma_\lambda\varepsilon_{\lambda t+1}). \quad (19)
\end{aligned}
$$

Finally, I inject an extra amount Δ_z of crash risk, that is equivalent to shifting the shock ε_{zt+1} by $\Delta_z/(2\theta^2\lambda_t)^{1/2}$, $\tilde{\varepsilon}_{zt+1} = \varepsilon_{zt+1} + \Delta_z/(2\theta^2\lambda_t)^{1/2}$,

$$\tilde{z}_{t+1} = z_{t+1} + \Delta_z = \theta\lambda_t + (2\theta^2\lambda_t)^{1/2}(\varepsilon_{zt+1} + \Delta_z/(2\theta^2\lambda_t)^{1/2}),$$

12

or an extra amount Δ_λ of risk driving the probability of a crash, that is equivalent to shifting the shock $\varepsilon_{\lambda t+1}$ by $\Delta_\lambda/\sigma_\lambda$, $\tilde{\varepsilon}_{\lambda t+1} = \varepsilon_{\lambda t+1} + \Delta_\lambda/\sigma_\lambda$,

$$\tilde{\lambda}_{t+1} = \lambda_{t+1} + \Delta_\lambda = (1 - \nu_\lambda)v_\lambda + \nu_\lambda \lambda_t + \sigma_\lambda(\varepsilon_{\lambda t+1} + \Delta_\lambda/\sigma_\lambda),$$

and measure the resulting revisions of expected multi-period returns

$$\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{zt+1}, \lambda_t) = \log E(r_{t,t+\tau}|\mathcal{I}_t, \tilde{z}_{t+1} = z_{t+1} + \Delta_z) - \log E(r_{t,t+\tau}|\mathcal{I}_t) = \Delta_z, \tag{20}$$
$$\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{\lambda t+1}, \lambda_t) = \log E(r_{t,t+\tau}|\mathcal{I}_t, \tilde{\lambda}_{t+1} = \lambda_{t+1} + \Delta_\lambda) - \log E(r_{t,t+\tau}|\mathcal{I}_t) = \mathcal{A}_\lambda(\tau - 1) \cdot \Delta_\lambda. \tag{21}$$

I set Δ_z to be equal to one standard deviation of z_{t+1}

$$\Delta_z = \left((b_\lambda - \theta)^2 \frac{\sigma_\lambda^2}{1 - \nu_\lambda^2} + 2\theta^2 v_\lambda \right)^{1/2}$$

for characterizing the term structure of ε_{zt+1}, and I set Δ_λ to be equal to one standard deviation of λ_t

$$\Delta_\lambda = \frac{\sigma_\lambda}{(1 - \nu_\lambda^2)^{1/2}}$$

for charactering the term structure of $\varepsilon_{\lambda t+1}$. The shape and the level of the term structure of risk associated with the shock $\varepsilon_{\lambda t+1}$ is determined by a recursion for $\mathcal{A}_\lambda(\tau)$. The term structure of ε_{zt+1} is flat at the level of Δ_z, because ε_{zt+1} is a permanent shock in log returns.

3   Empirical application. Term structure of risk in expected stock returns

In this section, I describe an empirical strategy for examining the term structure of risk in expected stock returns. I build my analysis on the examples discussed above but consider them in the context of explicit economic models. Alternative economic models explain why the price-dividend ratio has a forecasting power for future stock returns, but little is known about the empirical properties of competing economic channels of return predictability. There are at least two natural reasons for this lack of evidence. One is in the complexity of estimating fully-fledged structural models with latent states and nonlinearities on a short sample of available macroeconomic data.[11] Another reason is that any fully parametrized structural model, as a simplification of the real world, is misspecified.

---

[11]Bansal, Gallant, and Tauchen (2007), Bansal, Kiku, and Yaron (2014), Calvet and Czellar (2013), Constantinides and Ghosh (2011), Colacito and Croce (2011), Schorfheide, Song, and Yaron (2017) are among few prominent examples of thorough empirical evaluation of various macro-based asset pricing models.

My approach for examining the term structure of risk uses structural models partially and solely for the purpose of identifying sources of fluctuations in expected returns. I model the joint dynamics of returns and macroeconomic fundamentals without imposing cross-equation restrictions which are stemming from the interaction of risk preferences and fundamental risk. I identify shocks responsible for return predictability by augmenting the empirical model with an observation equation that maps the return predictor into a state vector from economic theory. I let the data show, how different economic shocks propagate in future returns across alternative holding periods.

There are two alternative ways to identify the competing economic sources of time-variation in multi-period equity returns. One is to develop and estimate a big model nesting alternative hypothesis of return predictability, the other one is to estimate the models separately on a case-by-case basis. I follow the latter approach because of the limited informational content of scarce macroeconomic data. For example, the available data are not sufficient to tell apart whether relatively big fluctuations in economic variables are a consequence of jumps or normal shocks with a high level of volatility.

I base my empirical analysis on three compelling hypotheses of time-varying equity risk premium: (i) long-run risk with stochastic variance of consumption growth (Bansal and Yaron, 2004), (ii) time-varying consumption disasters (Wachter, 2013), and (iii) a multi-factor volatility model of consumption growth with crashes in the stochastic variance of consumption growth (Drechsler and Yaron, 2011). These different risk channels generate realistic properties of excess stock return predictability by the price-dividend ratio in the structural models. I examine how these shocks propagate in expected gross returns in the data.[12]

I conduct my empirical analysis on a sample of quarterly U.S. consumption growth, stock returns, and price-dividend ratios from 1947 to 2015. I identify alternative sources of equity risk premium by estimating three empirical models of a joint evolution of consumption growth and stock returns with latent states and an observation equation that linearly maps the price-dividend ratio into the latent states. I use the Bayesian MCMC methods, which are particularly useful in the context of this exercise, as they allow me to identify latent states (e.g., stochastic variance of consumption growth or intensity of consumption disasters) and jumps (e.g., consumption disasters, jumps in the variance of consumption growth) in the data.

I do not pursue the goal of discriminating structural models based on the goodness of fit criterion, yet my empirical results have a bite for diagnostics of economic models. I can confront the empirical term structures of economic shocks in the expected buy-and-hold returns with those in structural models. The propagation trajectories of economic shocks in the future expected returns, as implied by *estimated empirical* models, that is models without cross-equation restrictions, are the term structures of risk in the data. An economic mechanism of return predictability has an empirical support, if the term structure of risk in the structural model is the same as in the data.

---

[12]The patterns of gross and excess stock return predictability are very similar. By looking at gross returns, I avoid the need of making assumptions about how to measure the term structure of real risk-free rates.

I characterize the term structures of risk in the structural models based on standard realistic calibrations, which account for salient properties of macroeconomic data and asset prices.

I use two metrics for comparing the term structures of risk in the data and in structural models: (i) the level and (ii) the shape. The level indicates how much a shock contributes to the time variation in returns (a criterion similar to a variance decomposition exercise). The shape indicates whether returns are more exposed to a specific risk over a short or long horizon. In structural models, cross-equation restrictions determine the level and the shape of the term structures of risk, and therefore, a comparison of theoretical and empirical term structures of risk serves as an implicit test of cross-equation restrictions.

The empirical term structures of risk is a set of stylized facts about the properties of economic shocks in the context of a specific state-space representation of a joint dynamics of macroeconomy and asset returns. There are many structural models sharing the same or similar state-space representations, and therefore, the empirical term structures of risk can be viewed as standard moments that those structural models should match. A GMM exercise minimizing the distance between theoretical and empirical moments, including the moments characterizing the term structure of risk, can deliver a calibration with realistic short-term and long-term asset pricing implications or indicate rejection of the structural model.

A natural question is whether it is economically interesting to match moments characterizing the term structure of risk over and above the multi-period covariances of returns and the predictive variable, that is return predictability. The answer is yes and the reason is simple – within one structural model there could be alternative ways to generate the desired properties of return predictability. At the same time, understanding how various sources of risk propagate in returns across alternative investment horizons delivers a lot more information about the multi-period dynamics of prices and quantities of risk. On a separate note, in general equilibrium models featuring monetary and/or fiscal policy or endogenous growth, a shock-by-shock analysis is a necessary tool for evaluating policy implications.

The methodology for analyzing the term structure of risk in the data and in structural asset pricing models is motivated by the way how macroeconomists test and estimate DSGE models. It is customary in the macroeconomics literature to test economic mechanisms of DSGE models by comparing how key economic shocks propagate across alternative horizons in the data and in different models. Alternatively, macroeconomists use empirical trajectories of structural shocks in economic indicators for estimating DSGE models by means of matching empirical and model-implied impulse response functions. Empirical impulse response functions are obtained after estimating a VAR implied by a linear or log-linear approximation of the DSGE model of interest with relaxed cross-equation restrictions.

Below I describe the empirical models that I use to identify alternative sources of return predictability in stock returns. I relegate the details of the estimation procedure to the online Appendix and describe the estimation output in Appendix A.

A   The model with the long-run risk and stochastic volatility of consumption growth

The model of Bansal and Yaron (2004) has become a standard reference in modern asset pricing as a successful framework rationalizing the magnitude of and the countercyclical time variation in the equity risk premium. The key modeling ingredients, such as the presence of a slow moving component in the expected consumption growth, stochastic volatility in consumption growth, and recursive preferences, have proved useful in accounting for the risk-return trade-off in other asset markets as well (e.g., Bansal and Shaliastovich, 2013, Colacito and Croce, 2011, among others). In this subsection, I develop an empirical model that is suitable for an empirical identification of the long-run risk in consumption growth and of the shock in the variance of consumption growth.

I group variables like consumption growth $\log g_{t-1,t}$, latent expected consumption growth x_t, stock returns $\log r_{t-1,t}$, and dividend growth $\log d_{t-1,t}$ into a vector $Y_t = (\log g_{t-1,t}, \; x_t, \; \log r_{t-1,t}, \; \log d_{t-1,t})'$ and posit that it follows a discrete-time vector autoregressive model with the state $X_t = (1, \; \log g_t, \; x_t)'$, stochastic variance v_t, and shocks $\varepsilon_t = (\varepsilon_{gt}, \; \varepsilon_{xt}, \; \varepsilon_{dt})'$

$$Y_{t+1} = \underset{4\times 3}{G}X_t + \underset{4\times 1}{H_v}(v_{t+1} - E_t v_{t+1}) + \underset{4\times 3}{H}v_t^{1/2}\varepsilon_{t+1}, \tag{22}$$

$$v_{t+1} = (1 - \varphi_v) + \varphi_v v_t + \sigma_v((1 - \varphi_v + 2\varphi_v v_t)/2)^{1/2}\varepsilon_{vt+1}. \tag{23}$$

The stochastic variance v_t follows a scalar autoregressive process of order one $v_t \sim \mathcal{ARG}(1)$ with the scale parameter $\sigma_v^2/2$, the degrees of freedom $2(1-\varphi_v)/\sigma_v^2$, and the serial correlation φ_v. The shocks $\varepsilon_{gt} \sim \mathcal{N}(0,1)$ and $\varepsilon_{xt} \sim \mathcal{N}(0,1)$ are normal, whereas the stochastic variance shock $\varepsilon_{vt} \sim \mathcal{D}(0,1)$ is nonnormal with a mean of zero and standard deviation of one. All the shocks are orthogonal to each other. While the original formulation of the model features an autoregressive process for the stochastic variance, I work with the autoregressive gamma process.[13] The $\mathcal{ARG}$ process guarantees positivity of the stochastic variance regardless of the realization of the shock ε_{vt}. The change of the process does not alter quantitative implications of the model.

The original model of Bansal and Yaron (2004) implies that the price-dividend ratio is (approximately) an affine function of the expected consumption growth x_t and the conditional variance of consumption growth v_t. I use this implication of the structural model as an identifying assumption for the shocks driving latent states, that is ε_{xt+1} and ε_{vt+1}. To this end, I add an observation equation

$$\log pd_t = q_0 + q_x x_t + q_v v_t + \sigma_{pd}\varepsilon_{pdt}, \tag{24}$$

to the system of equations (22)-(23), where the independent and identically distributed (iid) shock ε_{pdt} is an observation error orthogonal to the economic shocks, and σ_{pd} is very small and to be

---

[13]The autoregressive gamma process can be viewed as the discrete-time counterpart to the continuous-time square-root model. See, for example, Gourieroux and Jasiak (2006) and Le, Singleton, and Dai (2010).

estimated. At the estimation stage, I leave parameters q_0, q_x, and q_v, as well as the elements of the matrices F, G, H_v, and H, free, that is I do not impose the cross-equation restrictions implied by the original structural model. Appendix B presents the solution of the structural model and illustrates the cross-equation restrictions that I relax.

I exclude the dividend equation from the system given in (22)-(24), because of an approximate linear relationship between the log price-dividend ratio, log stock returns, and log dividend growth

$$\log r_{t,t+1} \quad \approx \quad \kappa_0 + \kappa_1 \log pd_{t+1} - \log pd_t + \log d_{t,t+1}, \tag{25}$$

that is in the current system with six economic variables there are only five independent shocks. I recover the implied dividend dynamics by using the log-linear approximation (25) and the estimated dynamics of returns and that of the price-dividend ratio. I exclude the dividend equation rather than the return equation or the price-dividend mapping (24), because (i) dividends are poorly measured (for example, seasonalities), (ii) dividends feature weaker predictability than returns, and (iii) dividends are not as informative about latent states of the consumption growth as the price-dividend ratio.

B   The model with time varying consumption disasters

Wachter (2013) argues that a risk exposure of dividends to time-varying consumption disasters is a leading risk channel that generates a sizeable and time varying equity risk premium. A representative investor fears big negative fundamental shocks that lead to dramatic drops in cash flows, hence a large magnitude of the risk premium, with a time-varying rate of arrival, hence time-variation in the equity risk premium. This idea has proved useful in rationalizing the risk-return tradeoff not only in the stock market, but also in the foreign exchange market (e.g., Du, 2013, Farhi and Gabaix, 2016), fixed-income market (e.g., Gabaix, 2012, Tsai, 2014), and credit market (e.g., Gourio, 2013).[14] In this subsection, I develop an empirical model suitable for an identification of the time-varying consumption disasters in the data.

I group variable like consumption growth $\log g_{t-1,t}$, log return $\log r_{t-1,t}$, and dividend growth $\log d_{t-1,t}$ into a vector $Y_t = (\log g_{t-1,t},\ \log r_{t-1,t},\ \log d_{t-1,t})'$, and posit that it follows a discrete-time autoregressive process with a state $X_t = (1,\ \log g_{t-1,t},\ \lambda_t)'$, shocks $\varepsilon_t = (\varepsilon_{gt},\ \varepsilon_{dt})'$ and $\varepsilon_{\lambda t}$, and time-varying disasters $-z_t$ with time-varying intensity $h_\lambda \lambda_{t-1}$

$$Y_{t+1} \quad = \quad \underset{3\times 3}{G}\, X_t + \underset{3\times 1}{H_{\varepsilon\lambda}}(\lambda_{t+1} - E_t\lambda_{t+1}) + \underset{3\times 2}{H}\,\varepsilon_{t+1} + \underset{3\times 1}{\Gamma}\, z_{t+1}, \tag{26}$$

$$\lambda_{t+1} \quad = \quad (1 - \varphi_\lambda) + \varphi_\lambda \lambda_t + \sigma_\lambda((1 - \varphi_\lambda + 2\varphi_\lambda \lambda_t)/2)^{1/2}\varepsilon_{\lambda t+1}. \tag{27}$$

The negative of a consumption disaster, z_{t+1}, is modeled as a Poisson mixture of gammas: $z_{t+1}|j_{gt+1} \sim \text{Gamma}(j_{gt+1}, \theta_g)$. Its central ingredient j_{gt+1} is a Poisson random variable, $j_{gt+1} \sim$

---

[14]Tsai and Wachter (2015) is an excellent survey on the role of the hypothesis of time-varying disaster risk in asset pricing.

Poisson($h_\lambda\lambda_t$), which controls how many jumps of average size θ_g arrive per period of time. The normalized jump intensity λ_t follows the scalar autoregressive process of order one $\lambda_t \sim \mathcal{ARG}(1)$ with the scale parameter $\sigma_\lambda^2/2$, the degrees of freedom $2(1-\varphi_\lambda)/\sigma_\lambda^2$, and the serial correlation φ_λ. The normal shocks $\varepsilon_{gt+1} \sim \mathcal{N}(0,1)$ and $\varepsilon_{dt+1} \sim \mathcal{N}(0,1)$ are orthogonal to each other. The nonnormal shock $\varepsilon_{\lambda t+1} \sim \mathcal{D}(0,1)$ has a mean of zero and standard deviation of one and is independent of ε_{gt+1} and ε_{dt+1}.

The original model of Wachter (2013) implies that the price-dividend ratio $\log pd_t$ is (approximately) an affine function of the disaster intensity λ_t. I use this implication of the structural model to identify the shock driving the probability of consumption disasters. To this end, I add the following observation equation

$$\log pd_t \quad = \quad q_0 + q_\lambda\lambda_t + \sigma_{pd}\varepsilon_{pdt} \tag{28}$$

to the system (26)-(27), where ε_{pdt} is an observation error orthogonal to the economic shocks ε_{gt}, ε_{dt}, and $\varepsilon_{\lambda t}$, and σ_{pd} is small and to be estimated. The parameters q_0 and q_λ, as well as the elements of the matrices H, H_λ, G, and Γ, are left free. Appendix C presents the solution of the structural model and illustrates the implied cross-equation restrictions which I relax at the estimation stage.

I exclude the dividend equation from the system given in (26)-(28) because of the approximate linear relationship between the log price-dividend ratio, log stock returns, and log dividend growth described in (25). I recover the implied dynamics of the dividend growth based on the estimation output. I exclude the dividend equation from the system (26)-(28) rather than the return equation or the price-dividend mapping because (i) dividends are poorly measured (e.g., seasonalities), (ii) dividends feature weaker predictability than returns and (iii) dividends are not as informative about the unobservable disaster intensity as the price-dividend ratio.

C   The two-volatility factor model with jumps in variance

The model of Drechsler and Yaron (2011) extends the structural model of Bansal and Yaron (2004) by adding time variation in the long-run mean of the variance of consumption growth and jumps in the conditional mean and in the conditional variance of consumption growth. The model generates realistic properties of stock return predictability, successfully matches standard moments of macro and asset pricing data, and also delivers realistic properties of the variance risk premium. Benzoni, Collin-Dufresne, and Goldstein (2011) build on this model to explain asset pricing puzzles associated with the 1987 crash; Drechsler (2013) extends the modeling environment of Drechsler and Yaron (2011) by adding model uncertainty for explaining the option-implied volatility surface. In this subsection, I develop an empirical model with multiple sources of variance risk as in Drechsler and Yaron (2011) and show how to identify them.[15] I exclude from consideration the long-run

---

[15]Multi-factor volatility models are known to capture more realistically the term structure of return volatilities (Bates, 1997, Bollerslev and Mikkelsen, 1996, Gallant, Hsu, and Tauchen, 1999, Duffie, Pan, and Singleton, 2000).

component of consumption growth as it is already explored in the model motivated by Bansal and Yaron (2004) in subsection A.

I group variables like consumption growth $\log g_{t-1,t}$, log stock returns $\log r_{t-1,t}$, and log dividend growth $\log d_{t-1,t}$ into a vector $Y_t = (\log g_{t-1,t},\ \log r_{t-1,t},\ \log d_{t-1,t})'$, and posit that it follows a discrete-time autoregressive model with a state $X_t = (1,\ \log g_{t-1,t},\ v_t,\ v_t^*)'$, shocks $\varepsilon_t = (\varepsilon_{gt},\ \varepsilon_{dt})'$, $\varepsilon_{vt},\ \varepsilon_{vt}^*$, and time-varying jumps in the stochastic variance of consumption growth z_{vt} with time-varying intensity $h_v v_{t-1}$

$$Y_{t+1} = \underset{3\times4}{G} X_t + \underset{3\times1}{H_v}(v_{t+1} - E_t v_{t+1} - z_{vt+1}) + \underset{3\times1}{H_v^*}(v_{t+1}^* - E_t v_{t+1}^*) + \underset{3\times2}{H}\varepsilon_{t+1} + \underset{3\times1}{\Gamma} z_{vt+1}, \quad (29)$$

$$v_{t+1} = (1 - \tilde{\varphi}_v)v_t^* + (1 - \varphi_v)v + \varphi_v v_t + \sigma_v(((1 - \varphi_v)v + 2\varphi_v v_t)/2)^{1/2}\varepsilon_{vt+1} + z_{vt+1}, \quad (30)$$

$$v_{t+1}^* = (1 - \varphi_v^*)v^* + \varphi_v^* v_t^* + \sigma_v^*(((1 - \varphi_v^*)v^* + 2\varphi_v^* v_t^*)/2)^{1/2}\varepsilon_{vt+1}^*. \quad (31)$$

The stochastic variance of consumption growth v_{t+1} follows an autoregressive gamma process of order one shifted by a positive process $(1 - \tilde{\varphi}_v)v_t^*$ which defines the time variation in the long-run mean of v_{t+1}, and includes self-exciting jumps z_{vt+1}. The time-varying mean of the stochastic variance v_{t+1}^* follows an autonomous autoregressive gamma process of order 1, $v_{t+1}^* \sim \mathcal{ARG}(1)$, with the scale parameter $\sigma_v^{*2}/2$, the degrees of freedom $2(1 - \varphi_v^*)v^*/\sigma_v^{*2}$, and the serial correlation φ_v^*. For the ease of referral, I label the shock ε_{vt+1}^* that feeds the stochastic trend of v_{t+1} a trending variance risk. The self-exciting jump z_{vt+1} is modeled as a Poisson mixture of Gammas: $z_{vt+1}|_{j_{vt+1}} \sim \Gamma(j_{vt+1}, \theta_v)$. Its central ingredient j_{vt+1} is a Poisson random variable, $j_{vt+1} \sim \text{Poisson}(h_v v_t)$, which controls how many jumps of average size θ_v arrive per period of time. When these jumps arrive, the stochastic variance spikes up. Despite the absence of a separate state variable controling the persistent component of consumption growth, the model allows for the time-varying expected growth via the lagged consumption growth and multiple variance factors.

The simplified version of the structural model of Drechsler and Yaron (2011) without the long-run risk in consumption growth, implies that the price-dividend ratio $\log pd_t$ is (appoximately) an affine function of the variance factors v_t and v_t^*. I use this implication of the structural model to identify the multiple sources of variance risk ε_{vt+1}, ε_{vt+1}^* and z_{vt+1}. To this end, I add the following observation equation

$$\log pd_t = q_0 + q_v v_t + q_v^* v_t^* + \sigma_{pd}\varepsilon_{pdt}, \quad (32)$$

to the system given in (29)-(31), where ε_{pdt} is an observation error orthogonal to the economic shocks ε_{gt}, ε_{dt}, ε_{vt}, and ε_{vt}^*, and σ_{pd} is small and to be estimated. The parameters q_0, q_v, and q_v^*, as well as the elements of the matrices G, H_λ, $H_{\varepsilon\lambda}$, H, and Γ, are left unrestricted. Appendix D presents the solution of the structural model and illustrates the implied cross-equation restrictions that are relaxed at the estimation stage.

I exclude the dividend equation from the system (29)-(31) because of the approximate linear relationship between the log price-dividend ratio, the log stock returns, and the log dividend growth given in (25) and recover the implied dividend dynamics based on the estimation output. I exclude

19

the dividend equation rather than the return equation or the price-dividend mapping because (i) dividends are poorly measured (e.g., seasonalities), (ii) dividends feature weaker predictability than returns and (iii) dividends are not as informative about the unobservable multiple stochastic variance factors as the price-dividend ratio.

4   Results

In this section, I describe the empirical term structures of risk for (i) the long run consumption growth shock and the shock in the stochastic variance of consumption growth, (ii) the consumption disaster risk and the shock in the intensity of consumption disasters, and (iii) the multiple shocks in the stochastic variance of consumption growth: the direct shock, the trending risk, and the jump risk. I focus on the two properties of the term structure: the level and the shape. I analyze whether upon arrival of a positive shock, which increases one of the state variables in the next period, the expected stock returns move up or down, and whether these effects are horizon-dependent.

Independently of the shock identification scheme, the term structure of expected buy-and-hold stock returns implied by the empirical model is downward sloping (Figure 1). The slope of the term structure of expected returns is defined in a standard way as the difference between the long-term and short-term expected returns. The negative slope is a manifestation of the multivariate mean-reversion in returns (Cochrane, 2001). The fact that my results are consistent with this property serves as a reality check for my empirical procedure. A necessary condition for observing a negative slope is the presence of at least one shock associated with a monotonically decreasing absolute value of the $\mathcal{IER}$ as the function of holding period. My empirical results indicate what this shock is, that is they reveal the source of stock return predictability. Appendix E describes how properties of individual shocks generate different shapes of the term structures of expected returns.

Under the shock identification scheme motivated by Bansal and Yaron (2004), I find that the long-run risk shock ε_{xt+1} has a downward sloping term structure of risk in expected returns (Figure 2). As a result, this shock is the main driver of predictability in *gross* stock returns. At short- and medium-term horizons, the expected returns exhibit a positive sensitivity to the shock: the returns go up upon arrival of a shock ε_{xt+1} that improves the future growth prospects of the macroeconomy. At horizons longer than 3 years, the expected returns are immune to the long-run risk shock.

Under the identification scheme motivated by Wachter (2013), the shock driving the time-varying probability of a consumption disaster $\varepsilon_{\lambda t+1}$ exhibits a downward sloping term structure of risk in expected returns (Figure 3): returns go down and more so in the short run upon arrival of a positive shock $\varepsilon_{\lambda t+1}$. As the shock $\varepsilon_{\lambda t+1}$ is tightly tied up with the consumption disaster risk ε_{zt+1}, I also analyze the term structure of ε_{zt+1} and find that the returns of any horizon go slightly up upon arrival of a consumption disaster.

The identification scheme motivated by the parsimonious version of Drechsler and Yaron (2011), reveals that the only source of predictability in gross stock returns is a trending variance

risk ε^*_{vt+1} (Figure 4). Upon arrival of a shock ε^*_{vt+1}, which increases the long-run mean of the stochastic variance of consumption growth, the stock returns decrease and more so in the short run. The direct variance shock ε_{vt+1} and the jump risk ε_{zt+1} have the term structures of risk in expected returns with a significant negative level but insignificant slope.

The documented above properties of the alternative sources of equity risk premium can be viewed as stylized facts. Structural models that have similar state-space representations to those behind the shock identification schemes, have to be consistent with these empirical facts. As a next step, I pursue an analysis of theoretical term structures of risk. I use original calibrations of the underlying structural models to infer theoretical (or model-implied) term structures of risk in expected returns. I compare theoretical and empirical term structures along the level and the shape of $\mathcal{IER}$s of different holding periods. These two simple metrics of the term structure of risk turn out to be informative about the plausibility of alternative predictability channels advocated by competing economic theories.

Figure 2 displays the term structure of risk in expected stock returns in the model of Bansal and Yaron (2004) and compares it to that in the data. In the model the shock ε_{xt+1} cumulates in expected returns as the investment horizon grows, whereas it gradually dies out in the data. The difference in the sensitivities of the long-term expected stock returns to the shock in the model and in the data is large. There are two questions of interest here: (i) does the difference between the empirical and theoretical term structures of risk matter economically and (ii) would an alternative calibration of the structural model imply a realistic term structure of the long-run risk shock in expected returns.

The answer to the first question is affirmative. An interaction of risk preferences and the distribution of risk produces the dynamics of the multi-period prices and quantities of risk, and therefore, implies how the shape and the level of the term structure of risk in expected returns look like. Whereas my methodology is silent about the individual dynamics of price and quantity of risk in the data, the empirical results set off alarm bells as it concerns the key assumptions in structural models. The answer to the second question is negative. A simple back-of-envelope computation implies that unless the long-run risk shock features the negative price of risk $q_x < 0$ or the elasticity of intertemporal substitution is negative $\rho > 1$, the theoretical term structure of the long-run risk is upward sloping (see Appendix E). The wrong slope of the term structure of ε_{xt+1} means that the structural model produces predictability of gross stock returns by a price-dividend ratio with an opposite sign to that in the data.

The disagreement between the data and the structural model in Bansal and Yaron (2004) is pervasive. The long-run risk shock ε_{xt+1}, which is an economic channel generating the sizeable magnitude of the one-period risk premium in the structural model, has unrealistic implications for the dynamics of asset returns. This result has important implications for a production-based asset pricing literature relying on models featuring an equilibrium consumption growth process with the long-run risk shock only (e.g., Croce, 2014; Hitzemann, 2016; Kung and Schmid, 2015).

Figure 3 displays the term structure of risk in the structural model of Wachter (2013) and compares it to that in the data. The theoretical term structure of the disaster intensity risk $\varepsilon_{\lambda t+1}$ has a realistic shape, yet its level at the short end is not as pronounced as in the data. The structural model posits that the consumption disaster z_{t+1}, a rare big negative shock in consumption growth, coincides with an even larger negative shock in dividends and stock returns, so that the term structure of risk ε_{zt+1} has a negative level. The data imply that (i) consumption disasters in the US post-world war data are not as severe as the structural model assumes and (ii) consumption disasters coincide with a slightly positive returns, so that the term structure of ε_{zt+1} is positive.

The discrepancy between the theoretical and empirical levels of the disaster risk in the term structure of expected returns represents a major challenge for the hypothesis of rare consumption disasters as it is formulated in Wachter (2013). If a consumption disaster that occurs in a bad state of the world and thus has a negative price of risk coincides with a positive shock in the stock market return, then the resulting risk premium is negative. It is not easy to overturn such a result by choosing an alternative calibration of the structural model. In the presence of a negative risk premium earned as a compensation for dividends' risk exposure to the shock ε_{zt+1}, there would be a lot of pressure on the stochastic intensity shock to produce a sizeable positive risk premium.

Given the presence of the negative skewness in consumption growth and stock returns, a model with consumption disaster risk appears to be a potentially successful working hypothesis for explaining the joint behavior of macroeconomy and asset markets. However, as the quarterly data suggest, the sources of the negative skewness in the two variables of interest are different. I check whether the annual data provide more support to the hypothesis of rare consumption disasters.

I analyze the sample of real annual consumption growth and stock returns from 1871 to 2017 and non-parametrically identify historical episodes corresponding to either consumption disasters or stock market crashes. I do not formally estimate the model on annual data given that the available sample is half of that for quarterly data. I classify the episodes, when the realized consumption growth falls two standard deviations (std dev is 3.55%) below its mean of 1.84%, as consumption disasters, and the episodes, when the realized stock return falls two standard deviations (std dev is 18.43%) below its mean of 8.20% as stock market crashes. Panel A of Figure 5 illustrates that only 1 episode that I interpret as a consumption disaster out of four possible ones coincides with a sizeable negative return; yet the return is within one standard deviation bound around the mean: the real stock return is -18.6% in 1930. Panel B of Figure 5 shows that only one out of six market crashes coincides with a sizeable negative consumption growth: consumption growth is -3.8% in 1931. As a result, I leave it for future research to find a more realistic configuration of the disaster hypothesis in the macroeconomy and stock market.

Figure 4 displays the term structure of equity risk in the parsimonious version of Drechsler and Yaron (2011) and compares it to that in the data. Because I consider the model without the long-run risk in consumption growth, I have to modify the original calibration so that the model still successfully reproduces key macroeconomic and asset pricing moments. Based on this calibration, I find that the theoretical terms structures of alternative sources of variance risk in

expected returns have realistic properties. Evidently, the main source of return predictability is the trending variance shock ε_{vt+1}^* that features a term structure of risk with a negative slope. The other sources of variance risk exhibit flat term structures with significantly negative levels. As a result, upon arrival of any type of variance shock, the multi-period returns go down. The effect of the trending variance risk is more pronounced in the short run.

The modified version of Drechsler and Yaron (2011) generates the realistic level and shape of the term structure of risk in expected returns. As a result, the model-implied term structure of expected returns coincides with that in the data (Figure 1). I dig deeper and analyze the pricing mechanism of the model. I find that in contrast to standard calibrations of various long-run risk models with stochastic variances, for example, Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012)), in which the variance shock does not account for a large portion of the risk premium, in the current calibration the variance shocks are the only source of equity premium. Thus, the variance shocks play two important roles here: (i) they generate a *sizeable* risk compensation, so that they are not a sideshow for the level of the equity risk premium and (ii) they produce a right amount and sign of return predictability.

To conclude, empirical properties of return predictability suggest that perhaps we need to change the way we think about the risk-return tradeoff in the stock market. The customary theoretical sources of the level of the risk premium fall short to account for empirical regularities. The theoretical term structure of the long-run risk shock cannot match the downward sloping term structure of risk in expected returns as it is observed in the data. The consumption disaster shock is associated with the negative incremental expected returns in the model but positive ones in the data. Finally, the variance shocks, which are almost predominantly used for generating time variation in the risk premium, may be leading sources of the level and time variation in the expected stock returns.

5   Conclusion

I propose an empirical methodology that unites an equilibrium theory of time-varying risk premia with an empirical evidence on multi-horizon return predictability. My approach identifies economic shocks that are key sources of risk premium in the leading macro-based models, and measures how they propagate in expected buy-and-hold returns. Shock propagation patterns constitute the term structure of risk in expected returns. The shape and the level of the term structure of risk serve as informative moments for testing competing theories of time-varying risk premia.

As an application, I examine three leading equilibrium models of stock return predictability: long-run risk with stochastic volatility in consumption growth (Bansal and Yaron, 2004), time-varying consumption disaster risk (Wachter, 2013), and consumption model with multiple sources of variance risk, such as the direct variance shock, the trending variance risk, and the jump in variance risk (in spirit of Drechsler and Yaron, 2011). I find that the models of (Bansal and Yaron,

2004) and (Wachter, 2013) fall short to account simultaneously for the level and the shape of the term structure of risk in expected returns. The long-run risk shock has a downward sloping term structure of risk in the data but an upward sloping term structure of risk in the model. The consumption disaster shock has a negative term structure of risk in the data, but a positive term structure of risk in the model. On a positive side, I find support for an equilibrium model with multiple variance shocks a-la Drechsler and Yaron (2011). The trending variance shock is at the core of return predictability both in the model and in the data.

References

Alvarez, Fernando, and Urban Jermann, 2005, Using asset prices to measure the persistence of the marginal utility of wealth, *Econometrica* 73, 1977–2016.

Andries, Marianne, Thomas M. Eisenbach, and Martin C. Schmalz, 2017, Horizon-dependent risk aversion and the timing and pricing of uncertainty, manuscript, January.

Augustin, Patrick, 2016, The term structure of cds spreads and sovereign credit risk, manuscript, April.

Backus, David K., 2014, Notes on risk and the price of risk, notes.

——— , Nina Boyarchenko, and Mikhail Chernov, 2015, Term structure of asset prices and returns, manuscript, August.

Backus, David K., Mikhail Chernov, and Stanley E. Zin, 2014, Sources of entropy in representative agent models, *Journal of Finance* 69, 51–99.

Bakshi, Gurdip, and Fousseni Chabi-Yo, 2012, Variance bounds on the permanent and transitory components of stochastic discount factors, *Journal of Financial Economics* 105, 191–208.

Bakshi, Gurdip, Xiaohui Gao, and George Panayotov, 2017, A theory of dissimilarity between stochastic discount factors, Working paper.

Bansal, Ravi, A. Ronald Gallant, and George Tauchen, 2007, Rational pessimism, rational exuberance, and asset pricing models, *Review of Economic Studies* 74, 1005–1033.

Bansal, Ravi, Dana Kiku, and Marcelo Ochoa, 2015, Climate change and growth risks, Working paper.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An empirical evaluation of the long-run risks model for asset prices, *Critical Finance Review* 1, 183–221.

——— , 2014, Risks for the long run: estimation with time aggregation, manuscript, July.

Bansal, Ravi, Shane Miller, and Amir Yaron, 2017, Is the term structure of equity risk premia upward sloping, Working paper.

Bansal, Ravi, and Ivan Shaliastovich, 2013, A long-run risks explanation of predictability puzzles in bond and currency markets, *Review of Financial Studies* 26, 1–33.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barro, Robert J., 2013, Environmental protection, rare disasters, and discount rates, Working paper, NBER.

Bates, David, 1997, Post-'87 crash fears in s&p 500 futures options, manuscript, NBER.

Bekaert, Geert, and Jun Liu, 2004, Conditioning information and variance bounds on pricing kernels, *Review of Financial Studies* 17, 339–378.

Belo, Frederico, Pierre Collin-Dufresne, and Robert Goldstein, 2015, Dividend dynamics and the term structure of dividend strips, *Journal of Finance* 70, 1115–1160.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2011, Explaining asset pricing puzzles associated with the 1987 market crash, *Journal of Financial Economics* 101, 552–573.

Binsbergen, Jules H. van, and Ralph S.J. Koijen, 2017, The term structure of returns: facts and theory, Journal of Financial Economics, forthcoming.

Binsbergen, Jules van, Michael Brandt, and Ralph Koijen, 2012, On the timing and pricing of dividends, *American Economic Review* 102, 1596–1618.

Binsbergen, Jules van, Wouter Hueskes, Ralph Koijen, and Evert Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110, 503–519.

Bollerslev, Tim, and Hans Ole Mikkelsen, 1996, Modeling and pricing long memory in stock market volatility, *Journal of Econometrics* 73, 151–184.

Borovička, Jaroslav, and Lars P. Hansen, 2014, Examining macroeconomic models through the lens of asset pricing, *Journal of Econometrics* 183, 67–90.

———— , Mark Hendricks, and José A. Scheinkman, 2011, Risk-price dynamics, *Journal of Financial Econometrics* 9, 3–65.

Borovička, Jaroslav, Lars P. Hansen, and José A. Scheinkman, 2014, Shock elasticities and impulse responses, *Mathematics and Financial Economics* 8, 333–354.

Calvet, Laurent, and Veronika Czellar, 2013, Through the looking glass: Indirect inference via simple equilibria, *Journal of Econometrics* 185, 343–358.

Campbell, John Y., 2001, Why long horizons? a study of power against persistent alternatives, *Journal of Empirical Finance* 8, 459–491.

Chabi-Yo, Fousseni, 2008, Conditioning information and variance bounds on pricing kernels with higher-order moments: theory and evidence, *Review of Financial Studies* 21, 181–231.

———— , and Riccardo Colacito, 2017, The term structure of co-entropy in international financial markets, manuscript, January.

Chen, Hui, Winston Dou, and Leonid Kogan, 2017, Measuring the Dark Matter in asset pricing models, March.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, 2005, Nominal rigidities and the dynamic effects of a shock to a monetary policy, *Journal of Political Economy* 113, 1–45.

Cochrane, John, 2001, *Asset Pricing* (Princeton University Press: Princeton).

———— , and Lars Hansen, 1992, Asset pricing explorations for macroeconomics, in Olivier J. Blanchard, and Stanley Fischer, ed.: *NBER Macroeconomics Manual* (MIT Press: Cambridge, MA).

Cochrane, John H., 2008, The dog that did not bark: A defence of return predictability, *Review of Financial Studies* 21, 1533–1575.

Colacito, Riccardo, and Mariano M. Croce, 2011, Risks for the long run and the real exchange rate, *Journal of Political Economy* 119, 153–183.

Constantinides, George M., and Anisha Ghosh, 2011, Asset pricing tests with long run risks in consumption growth, *Review of Asset Pricing Studies* 1, 96–136.

Croce, Max, 2014, Long-run productivity risk: a new hope for production-based asset pricing, *Journal of Monetary Economics* 66, 13–31.

———— , Thien Nguyen, Steve Raymond, and Lukas Schmid, 2017, Government debt and return to innovation, Working paper.

Croce, Mariano M, Martin Lettau, and Sydney C. Ludvigson, 2015, Investor information, long-run risk, and the term structure of equity, *Review of Financial Studies* 28, 706–742.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2015, The price of variance risk, Journal of Financial Economics, forthcoming.

Drechsler, Itamar, 2013, Uncertainty, time-varying fear, and asset prices, *Journal of Finance* 68, 1843–1889.

———— , and Amir Yaron, 2011, What's vol got to do with it, *Review of Financial Studies* 24, 1–45.

Du, Du, 2013, General equilibrium pricing of currency and currency options, *Journal of Financial Economics* 110, 730–751.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Farhi, Emmanuel, and Xavier Gabaix, 2016, Rare disasters and exchange rates, *Quarterly journal of economics* 131, 1–52.

Favero, Carlo A., Fulvio Ortu, Andrea Tamoni, and Haoxi Yang, 2017, Implications of return predictability across horizons for asset pricing models, Working paper.

Fernandez-Villaverde, Jesus, Juan Rubio-Ramirez, and Frank Schorfheide, 2016, Solution and estimation methods for DSGE models, in John B. Taylor, and Harald Uhlig, ed.: *Handbook of Macroeconomics, vol. 2B* (Elsevier: Amsterdam).

Gabaix, Xavier, 2012, Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.

Gallant, A. Ronald, Lars P. Hansen, and George Tauchen, 1990, Using conditional moments of asset payoffs to infer the volatility of intertemporal marginal rates of substitution, *Journal of Econometrics* 45, 141–179.

Gallant, A. Ronald, Chien-Te Hsu, and George Tauchen, 1999, Using daily range data to calibrate volatility diffusions and extract the forward integrated variance, *The review of economics and statistics* 81, 617–631.

Gallmeyer, Michael, Burton Hollifield, Francisco Palomino, and Stanley Zin, 2007, Arbitrage free bond pricing with dynamic macroeconomic models, *The Federal Reserve Bank of St. Louis Review* pp. 305–326.

Giacomini, Rafaela, 2013, The relationship between DSGE and VAR models, in Thomas B. Fomby, Lutz Kilian, and Anthony Murphy, ed.: *VAR models in Macroeconomics – New Developments and Applications: Essays in Honor of Christopher A. Sims (Advances in Econometrics), vol. 32* . pp. 1–25 (Emerald Group Publishing Limited).

Giglio, Stefano W., and Bryan T. Kelly, 2017, Excess volatility: beyond discount rates, Quarterly Journal of Economics, forthcoming.

Gormsen, Niels Joachim, 2017, Time variation of the equity term structure, manuscript, December.

Gourieroux, Christian, and Joann Jasiak, 2006, Autoregressive gamma processes, *Journal of Forecasting* 25, 129–152.

Gourio, Francois, 2013, Credit risk and disaster risk, *American Economic Journal: Macroeconomics* 5, 1–34.

Guerron-Quintana, Pablo, Atsushi Inoue, and Lutz Kilian, 2017, Impulse response matching estimators for dsge models, *Journal of Econometrics* 196, 144–155.

Hansen, Lars P., 2012, Dynamic valuation decomposition within stochastic economies, *Econometrica* 80, 911–967.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Hassler, Michael, and Roberto Marfe, 2016, Disaster recovery and the term structure of dividend strips, *Journal of Financial Economics* 122, 116–134.

Hitzemann, Steffen, 2016, Macroeconomic fluctuations, oil supply shocks, and equilibrium oil futures prices, September.

Kung, Howard, and Lukas Schmid, 2015, Innovation, growth, and asset prices, *Journal of Finance* 70, 1001–1037.

Le, Anh, Kenneth J. Singleton, and Qiang Dai, 2010, Discrete-time dynamic term structure models with generalized market prices of risk, *Review of Financial Studies* 23, 2184–2227.

Liu, Yang, 2016, Government debt and risk premia, Working paper.

Marfe, Roberto, 2016, Income insurance and the equilibrium term structure of equity, Journal of Finance, forthcoming.

Schorfheide, Frank, Dongho Song, and Amir Yaron, 2017, Identifying long-run risks: a bayesian mixed-frequency approach, Econometrica, forthcoming.

Smets, Frank, and Rafael Wouters, 2007, Shocks and frictions in us business cycles: A bayesian dsge approach, *American Economic Review* 97, 586–610.

Tsai, Jerry, 2014, Rare disasters and the term structure of interest rates, manuscript, December.

———— , and Jessica A. Wachter, 2015, Disaster risk and its implications for asset pricing, *Annual Review of Financial Economics* 7, 219–252.

Valkanov, Rossen, 2003, Long-horizon regressions: theoretical results and applications, *Journal of Financial Economics* 68, 201–232.

Wachter, Jessica, 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* LXVIII.

Yang, Haoxi, 2014, A robust variance bound on pricing kernels, Working paper.

Zviadadze, Irina, 2017, Term structure of consumption risk premia in the cross-section of currency returns, *Journal of FInance* 72, 1529–1566.
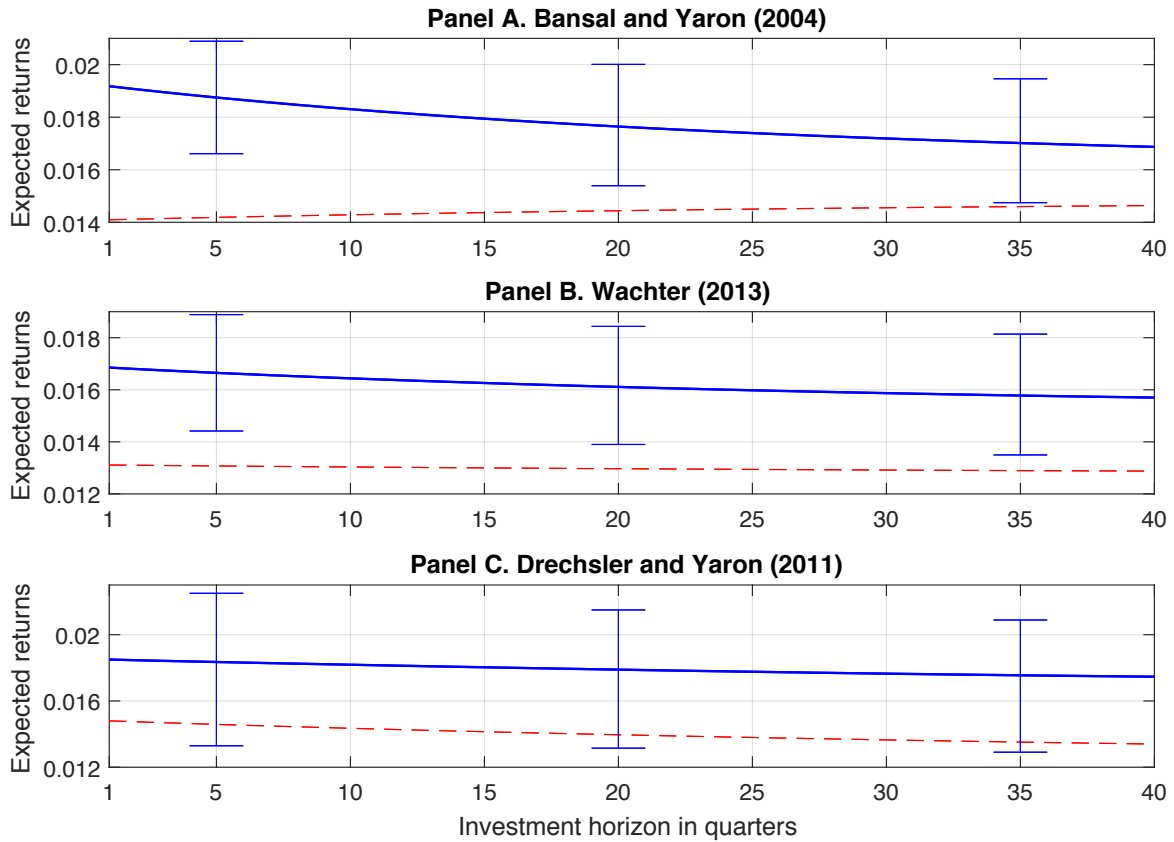
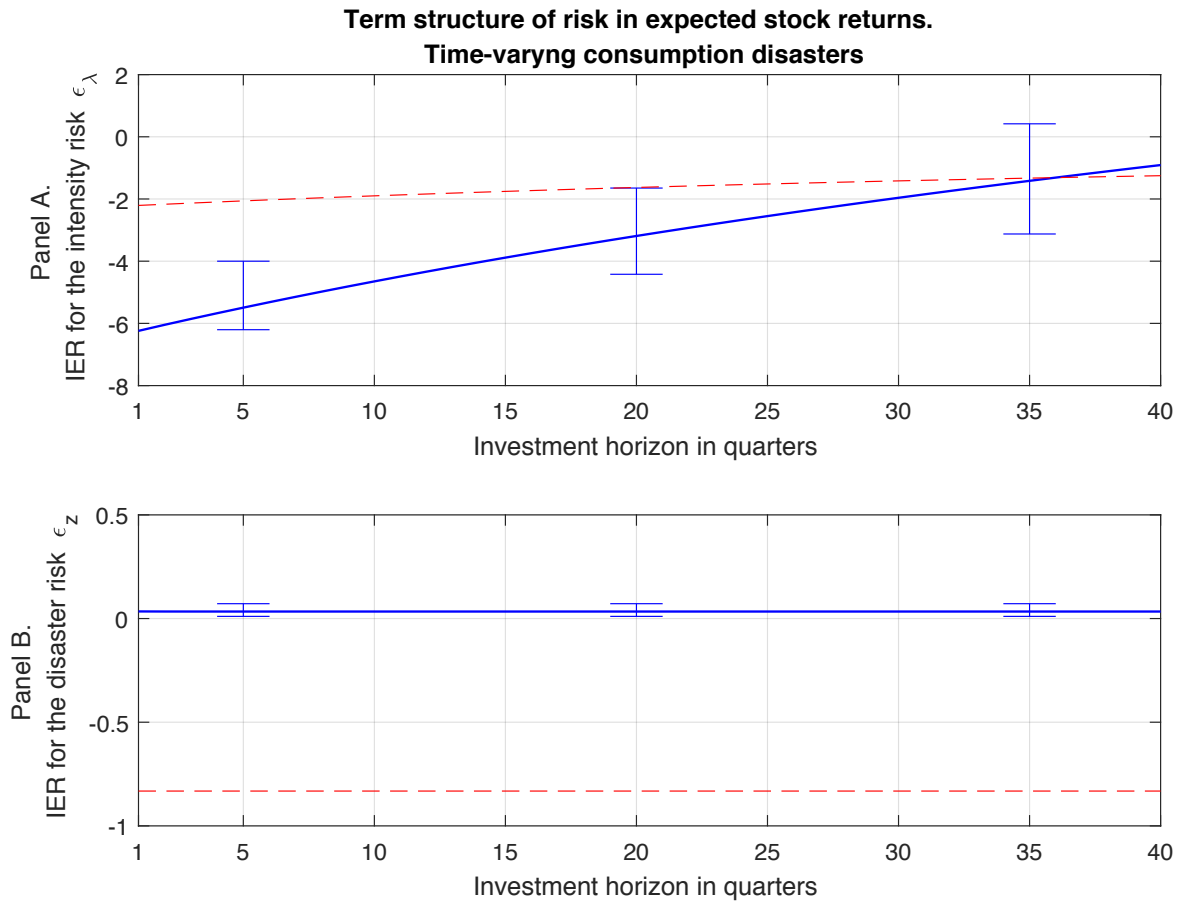**Figure 1**
Term structure of expected stock returns.
The red lines correspond to theoretical term structures of expected buy-and-hold returns, that is as implied by the original calibrations of Bansal and Yaron (2004), Wachter (2013), and Drechsler and Yaron (2011). The solid blue lines correspond to empirical term structures of expected buy-and-hold returns, that is as implied by the estimated empirical models. Vertical bars indicate 95% credible intervals for the estimated term structures of expected returns. Panel A represents the economic environment of Bansal and Yaron (2004). Panel B represents the economic environment of Wachter (2013). Panel C represents the economic environment of Drechsler and Yaron (2011). Quarterly.
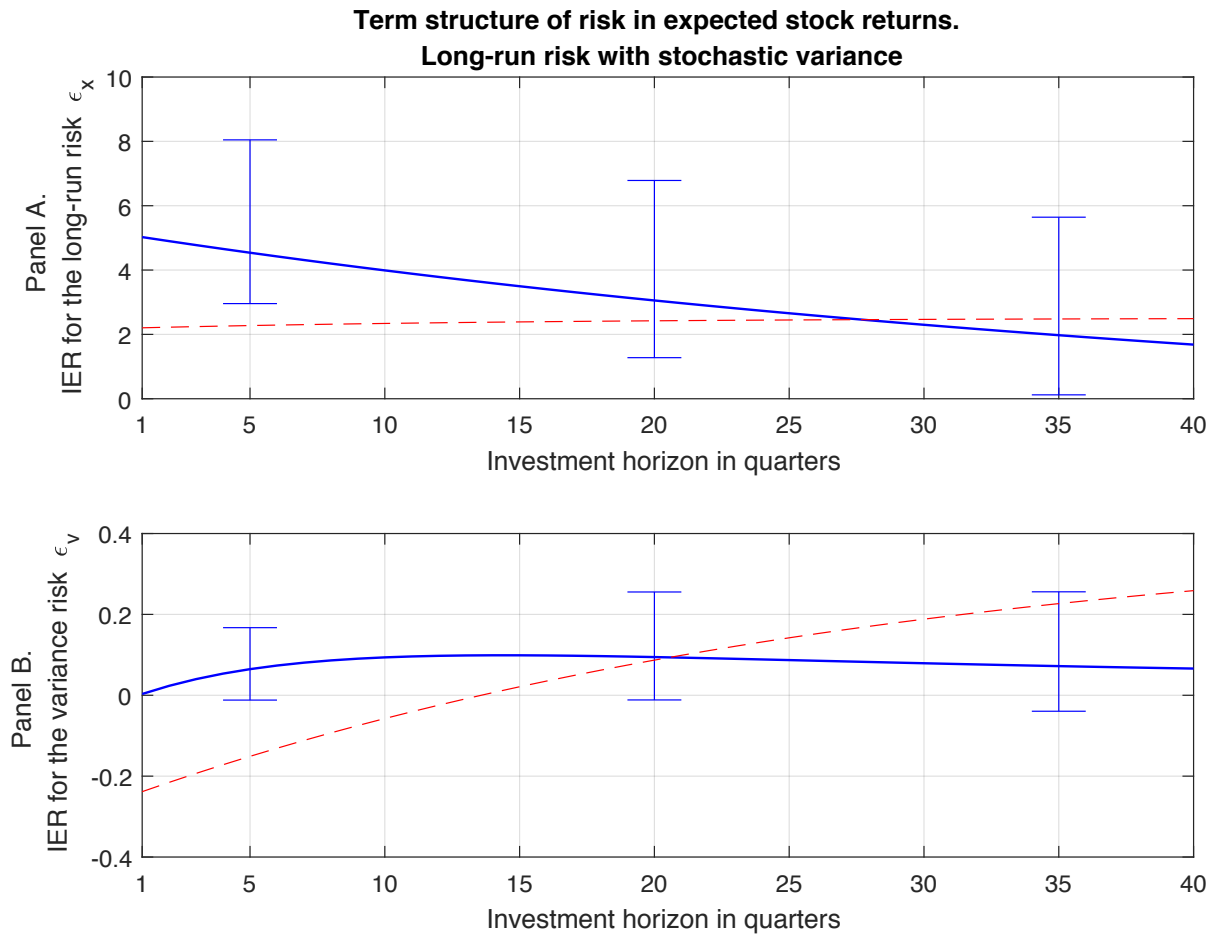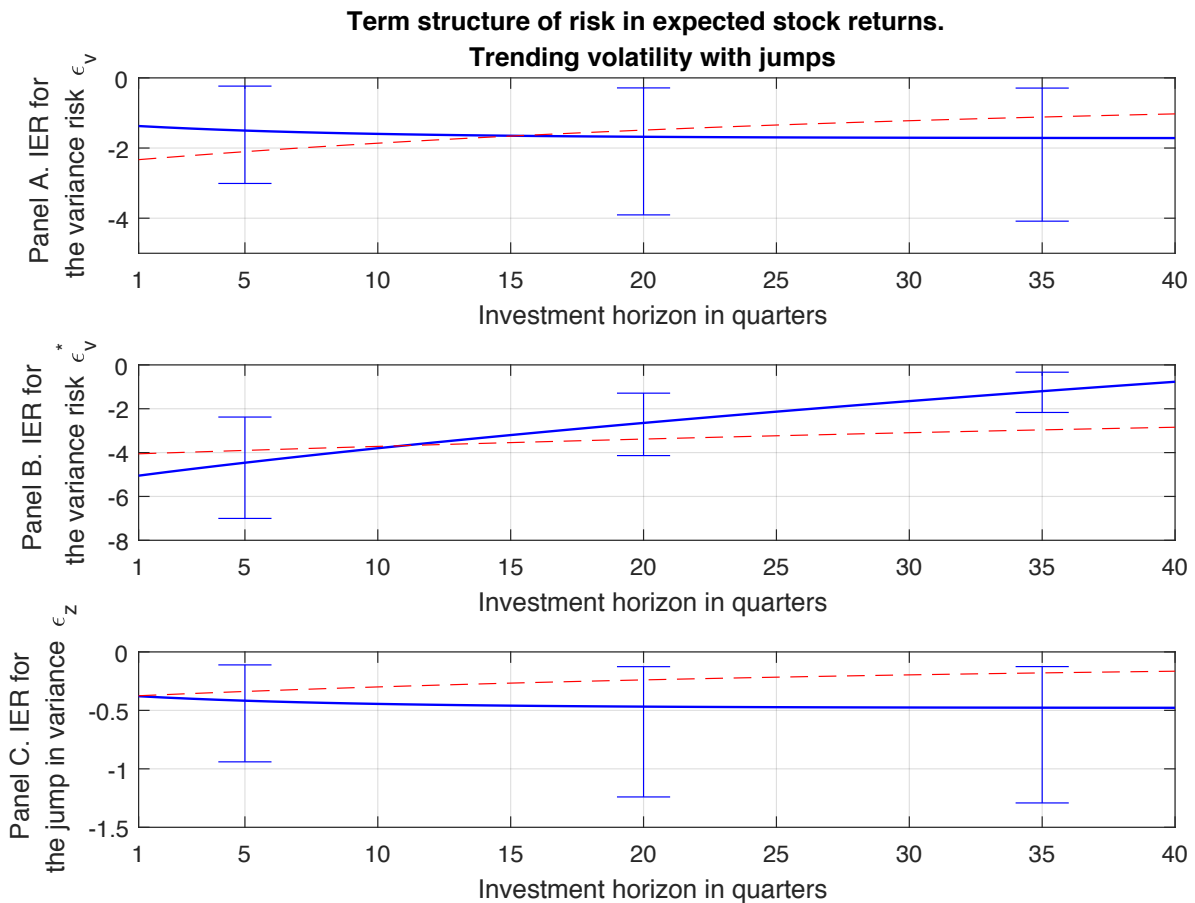
Figure 2
Term structure of risk in expected stock returns. Bansal and Yaron (2004).
The red dashed lines correspond to the theoretical term structures of risk. The blue solid lines correspond to the empirical term structure of risk. Vertical bars indicate 95% credible intervals for the estimated term structures of risk. The incremental expected returns are scaled by the unconditional standard deviation of the one-period stock returns. Quarterly.
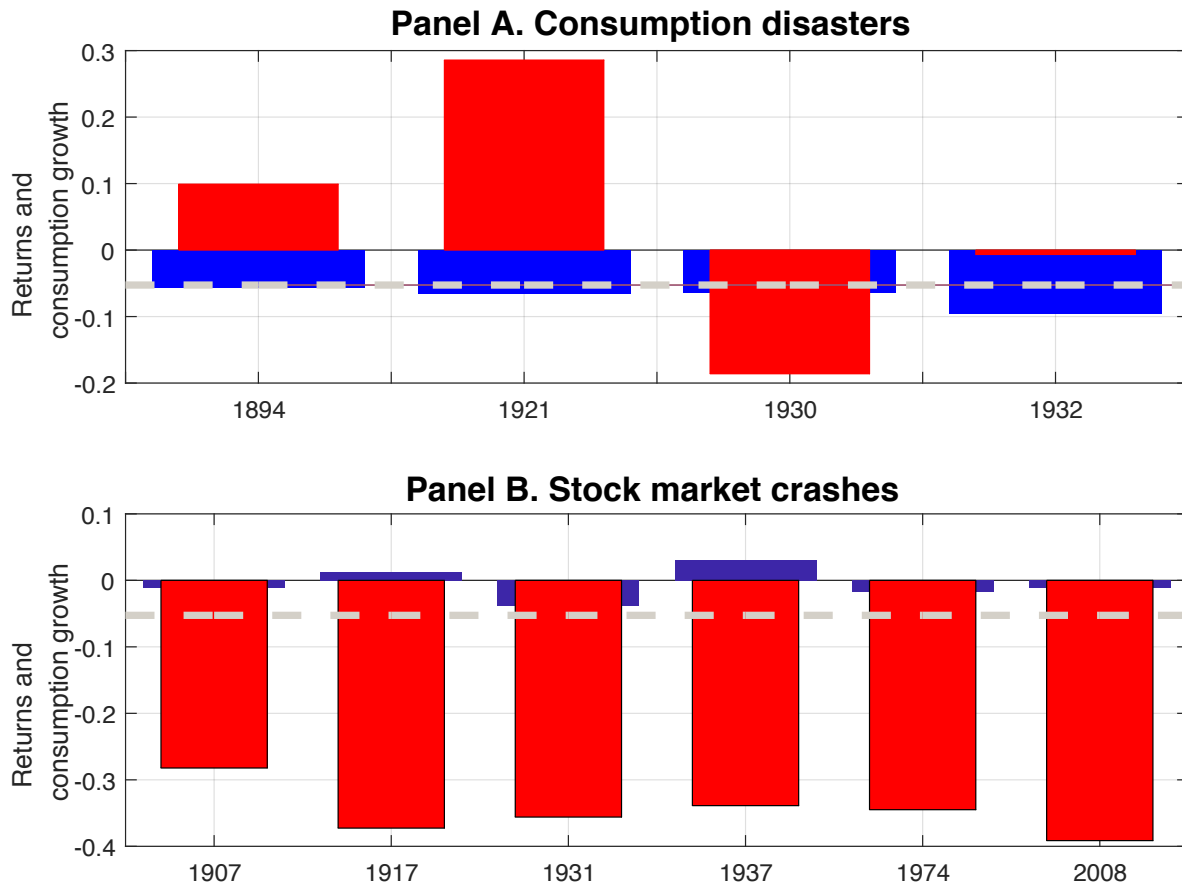
Figure 3
Term structure of risk in expected stock returns. Wachter (2013).
The red dashed lines correspond to the theoretical term structures of risk. The blue solid lines correspond to the empirical term structure of risk. Vertical bars indicate 95% credible intervals for the estimated term structures of risk. The incremental expected returns are scaled by the unconditional standard deviation of the one-period stock returns. Quarterly.
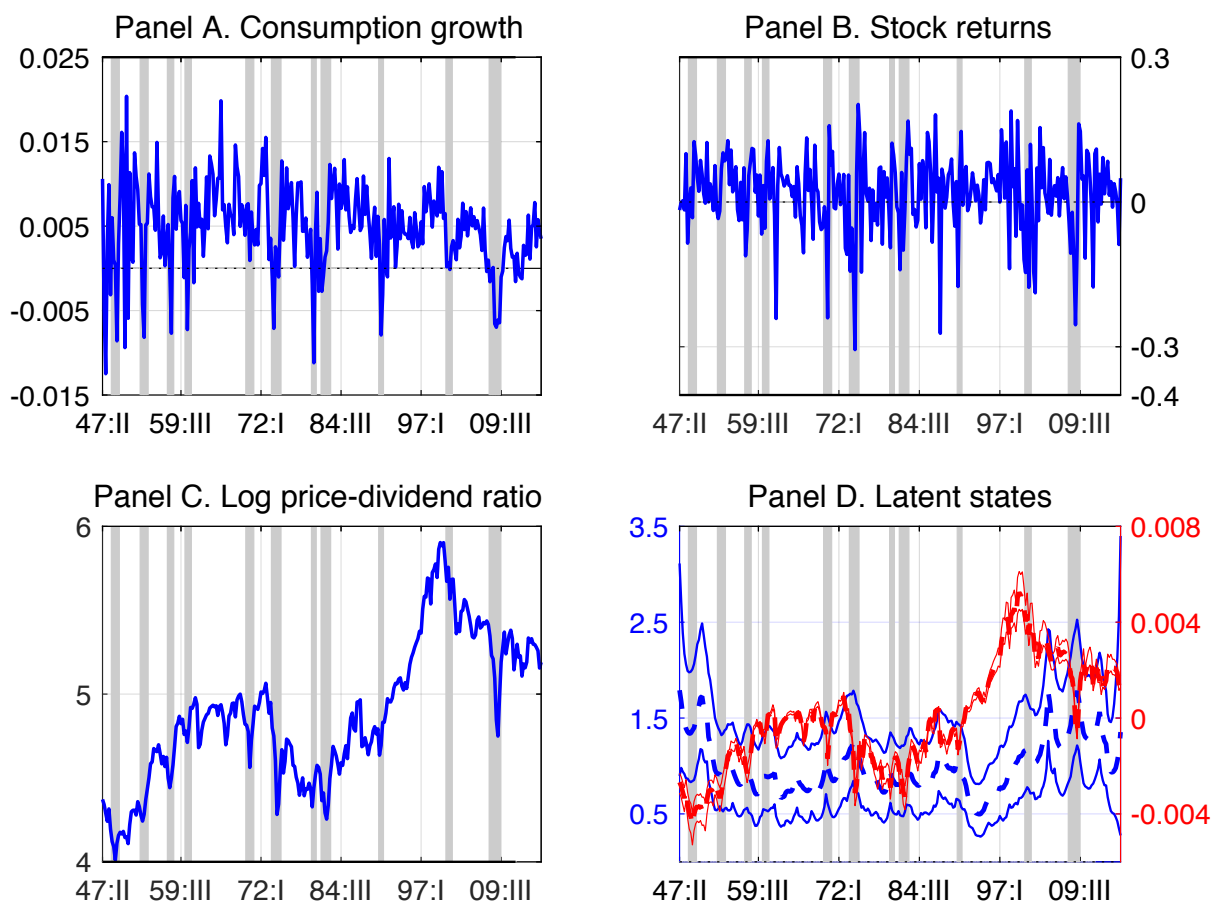
Figure 4
Term structure of risk in expected stock returns. A model in spirit of Drechsler and Yaron (2011).
The red dashed lines correspond to the theoretical term structures of risk. The blue solid lines correspond to the empirical term structure of risk. Vertical bars indicate 95% credible intervals for the estimated term structures of risk. The incremental expected returns are scaled by the unconditional standard deviation of the one-period stock returns. Quarterly.
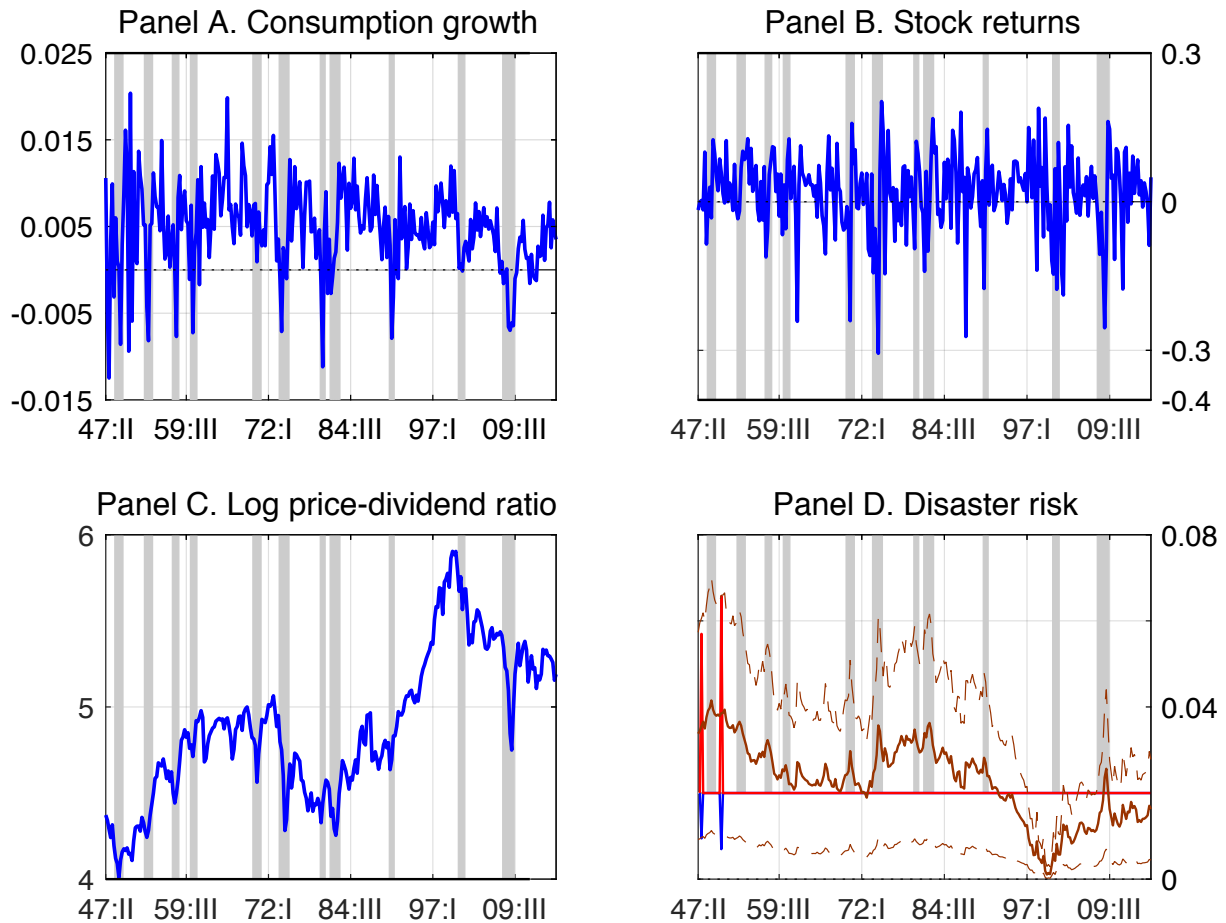
Figure 5
Consumption disasters and stock market crashes.
The red bars correspond to the real one-period stock returns, the blue bars correspons to
the real consumption growth. Panel A shows episodes when the realized real consumption
growth decline more than 2 standard deviations away from its mean. Panel B shows episodes
when the realized real returns on S&P 500 decline more than 2 standard deviations away
from its mean. The dashed grey line indicates a 2 standard deviation move below the mean
of the realized consumption growth. Annual. Data sources are the online data repository
of Robert Shiller `http://www.econ.yale.edu/~shiller/data.htm/` for S&P 500 returns
and the macroeconomic data set of Barro Ursua `https://scholar.harvard.edu/barro/`
`publications/barro-ursua-macroeconomic-data/` augmented by the NIPA tables of the
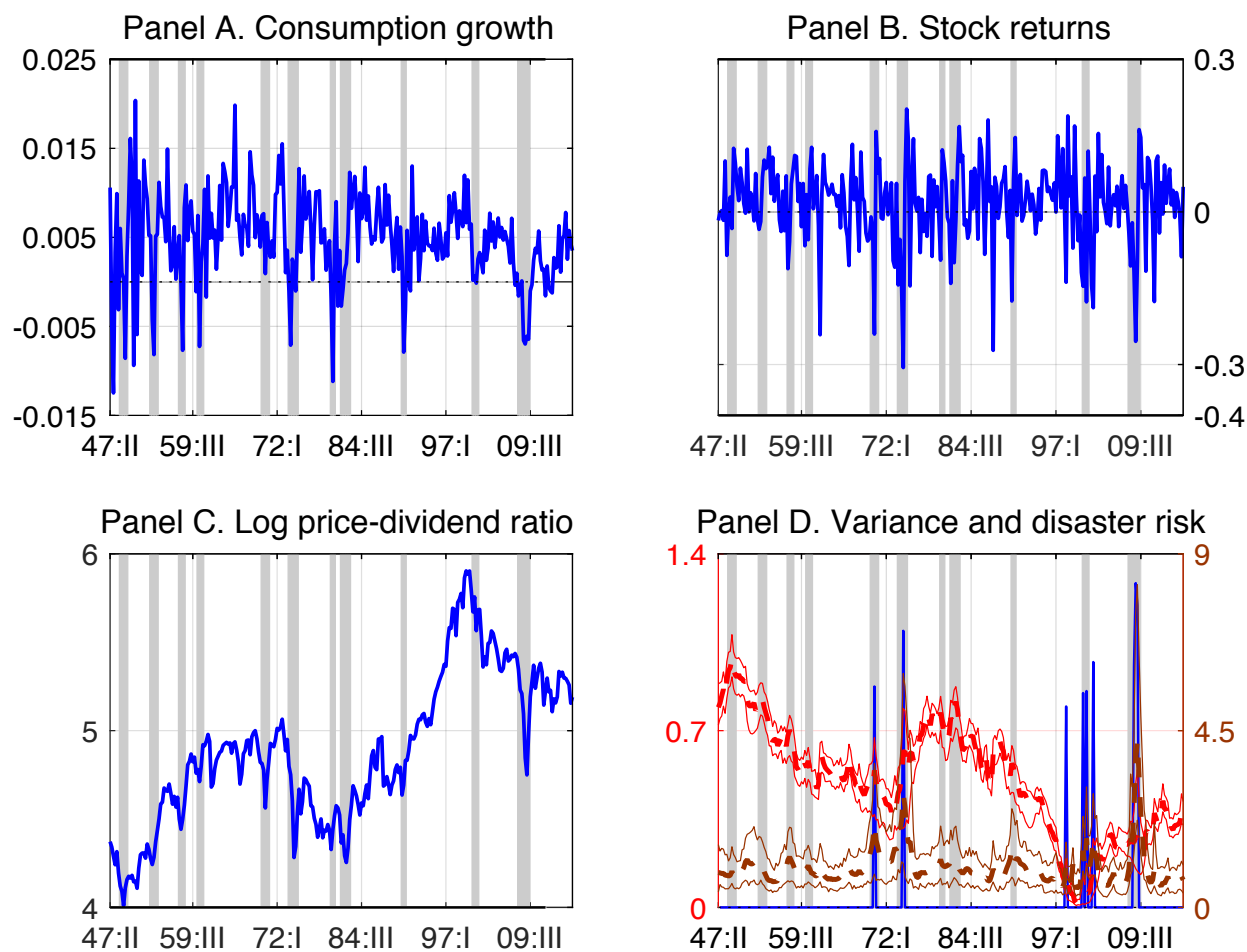Bureau of Economic Analysis for the real consumption data.

Figure 6
Macroeconomy. Bansal and Yaron (2004)
Panels A, B, C display quarterly observations of consumption growth, log stock returns, and log price-dividend ratio, respectively. Panel D displays the mean path of the stochastic variance factor (dashed blue line) with the 95% credible interval (thin solid lines) and the mean path of the expected consumption growth (dashed red line) with the 95% credible interval (thin solid lines). Sample period: second quarter of 1947 to fourth quarter of 2015. Grey bars are the NBER recessions. Quarterly.

Figure 7
Macroeconomy. Wachter (2013)
Panels A, B, C display quarterly observations of consumption growth, log stock returns, and log price-dividend ratio, respectively. Panel D displays consumption disaster risk (blue lines), jump risk in stock returns (red lines). A brown line corresponds to the estimated jump intensity $h_\lambda \lambda_t$, the dashed lines correspond to the 95% credible interval. Sample period: second quarter of 1947 to fourth quarter of 2015. Grey bars are the NBER recessions. Quarterly.

Figure 8
Macroeconomy. Model in spirit of Drechsler and Yaron (2011)
Panels A, B, C display quarterly observations of consumption growth, log stock returns, and log price-dividend ratio, respectively. Panel D displays the mean path of the stochastic variance factor (dashed brown line) with the 95% credible interval (thin brown lines), right axes, and the mean path of the variance factor v_t^* (dashed red line) with the 95% credible interval (thin red lines), left axes. Self exciting jumps (blue bars) are displayed on Panel D, left axes. Sample period: second quarter of 1947 to fourth quarter of 2015. Grey bars are the NBER recessions. Quarterly.

A Appendix

A Estimation output

I use the data displayed in Panels A, B, C of Figure 6 to estimate the unrestricted version of the model with the long-run risk and stochastic variance of consumption growth (22)-(24), the unrestricted version of the model with consumption disasters (26)-(28), and the unrestricted version of the mutli-factor volatility model with jumps in the stochastic variance of consumption growth (29)-(32).

Panel D of Figure 6 displays the estimated stochastic variance factor v_t and the estimated long run risk factor x_t with their 95% credible intervals. The estimated stochastic variance exhibits a significantly higher variance of variance $\sigma_v = 0.178$ and a smaller parameter of the serial correlation $\varphi_v = 0.8008$ than the calibrated model with parameter values $\sigma_v = 0.0653$ and $\varphi_v = 0.9615$, respectively. The component x_t is highly persistent with the serial correlation reaching the level of 0.9767 which is even higher than the parameter value in the original calibration of Bansal and Yaron (2004) ($\varphi_v = 0.9383$ at a quarterly frequency).

Panel D of Figure 7 displays the identified consumption disaster risk in the model of Wachter (2013), along with the estimated jump intensity. The structural model of Wachter (2013) posits that big negative jumps in consumption growth, that is consumption disasters, coincide with even larger negative jumps in dividend growth, and returns. This is the key mechanism of the model which generates high equity risk premium: the representative agent experiences a big loss of wealth at the worst states of the world. At first, the presence of a correlated jump in consumption growth and returns seems like a plausible feature of the data: both time series exhibit negative skewness and excess kurtosis. However, as the data suggest if anything negative jumps in consumption growth coincide with positive jumps in equity returns. This fact puts at risk the economic channel that facilitates matching the level of the risk premium in the structural model.

Panel D of Figure 8 displays the estimated stochastic variance v_t of consumption growth with self-exciting jumps in variance z_{vt} and stochastic long-run mean v_t^* in the model motivated by Drechsler and Yaron (2011). Evidently, this model exhibits much more pronounced spikes in the volatility of consumption growth during the economic recessions compared to those in the model of Bansal and Yaron (2004). This is a result of jumps in variance which happen rarely and tend to arrive during recessions. The long-run mean of the stochastic variance, v_t^*, is highly negatively correlated with the price-dividend ratio. This factor is persistent with the parameter of serial correlation $\varphi_v^* = 0.9881$ and the volatility of volatility $\sigma_v^* = 0.0633$. The stochastic variance itself v_t exhibits the serial correlation of 0.8549 and the volatility of volatility $\sigma_v = 0.2560$. Thus, the stochastic variance is slightly more persistent and volatile in this model compared to the model of Bansal and Yaron (2004).

B Solution of the Bansal and Yaron (2004) model

The model for consumption growth with stochastic variance is

$$
\begin{aligned}
\log g_{t,t+1} &= g + x_t + \gamma_g v_t^{1/2} \varepsilon_{gt+1}, \\
x_{t+1} &= \varphi_x x_t + \gamma_x v_t^{1/2} \varepsilon_{xt+1}, \\
v_{t+1} &= (1 - \varphi_v) + \varphi_v v_t + \sigma_v ((1 - \varphi_v + 2\varphi_v v_t)/2)^{1/2} \varepsilon_{vt+1}.
\end{aligned}
$$

A representative agent has recursive preferences

$$
\begin{aligned}
U_t &= [(1 - \beta)c_t^\rho + \beta \mu_t(U_{t+1})^\rho]^{1/\rho}, \quad &(33) \\
\mu_t(U_{t+1}) &= [E_t(U_{t+1}^\alpha)]^{1/\alpha}.
\end{aligned}
$$

Divide expression (33) by c_t, denote $u_t = U_t/c_t$ and $g_{t,t+1} = c_{t+1}/c_t$ and obtain

$$
u_t = [(1 - \beta) + \beta \mu_t(u_{t+1}g_{t+1})^\rho]^{1/\rho}, \quad (34)
$$

Solve a recursive problem that is a log-linear approximation of the Bellman equation (34)

$$
\log u_t \approx b_0 + b_1 \log \mu_t(g_{t+1}u_{t+1}),
$$

where

$$
\begin{aligned}
b_1 &= \beta e^{\rho \log \mu}/(1 - \beta + \beta e^{\rho \log \mu}), \\
b_0 &= \frac{1}{\rho} \log \left((1 - \beta) + \beta e^{\rho \log \mu}\right) - b_1 \log \mu.
\end{aligned}
$$

Guess the value function

$$
\log u_t = u + p_x x_t + p_v v_t.
$$

Compute

$$
\log u_{t+1} + \log g_{t+1} = u + g + (p_x \varphi_x + 1)x_t + p_v v_{t+1} + \gamma_g v_t^{1/2} \varepsilon_{gt+1} + p_x \gamma_x v_t^{1/2} \varepsilon_{xt+1}
$$

Recall that the cumulant generating function for the variance represented by a gamma autoregressive process of order one is

$$
\kappa(s; v_{t+1}) = \frac{\varphi_v s}{1 - s\sigma_v^2/2} v_t - \frac{(1 - \varphi_v) \log (1 - s\sigma_v^2/2)}{\sigma_v^2/2}.
$$

Compute

$$
\begin{aligned}
\log \mu_t(u_{t+1} g_{t,t+1}) &= u + g - \frac{(1 - \varphi_v) \log (1 - \alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2} + (p_x \varphi_x + 1) x_t \\
&+ \left(\frac{\alpha}{2} \gamma_g^2 + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\varphi_v p_v}{1 - \alpha p_v \sigma_v^2/2} \right) v_t.
\end{aligned}
$$

Hence

$$
\log \mu = u + g + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\alpha}{2} \gamma_g^2 + \frac{p_v \varphi_v}{1 - \alpha p_v \sigma_v^2/2} - \frac{(1 - \varphi_v) \log (1 - \alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2}.
$$

Solve the following system of three equations in three unknowns to verify the guess of the value function

$$
\begin{aligned}
u &= \frac{1}{1 - b_1} \left(b_0 + b_1 g - b_1 \frac{(1 - \varphi_v) \log (1 - \alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2} \right), \\
p_x &= \frac{b_1}{1 - b_1 \varphi_x}, \\
p_v &= b_1 \left(\frac{\alpha}{2} \gamma_g^2 + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\varphi_v p_v}{1 - \alpha p_v \sigma_v^2/2} \right).
\end{aligned}
$$

The quadratic equation for p_v has two roots. I choose root that satisfies the requirement of stochastic stability (Hansen, 2012).

Compute

$$
\begin{aligned}
\log(u_{t+1} g_{t,t+1}) - \log \mu_t(u_{t+1} g_{t,t+1}) &= \frac{(1 - \varphi_v) \log (1 - \alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2} - \left(\frac{\alpha}{2} \gamma_g^2 + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\varphi_v p_v}{1 - \alpha p_v \sigma_v^2/2} \right) v_t \\
&+ p_v v_{t+1} + \gamma_g v_t^{1/2} \varepsilon_{gt+1} + p_x \gamma_x v_t^{1/2} \varepsilon_{xt+1}.
\end{aligned}
$$

The pricing kernel is

$$
\begin{aligned}
\log m_{t,t+1} &= \log \beta + (\rho - 1) \log g_{t,t+1} + (\alpha - \rho)(\log (u_{t+1} g_{t,t+1}) - \log \mu_t(u_{t+1} g_{t,t+1})) \\
&= \underbrace{\log \beta + (\rho - 1) g + (\alpha - \rho) \frac{(1 - \varphi_v) \log (1 - \alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2}}_{\log m} \\
&+ \underbrace{(\rho - 1)}_{m_x} x_t \underbrace{- (\alpha - \rho) \left(\frac{\alpha}{2} \gamma_g^2 + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\varphi_v p_v}{1 - \alpha p_v \sigma_v^2/2} \right)}_{m_v} v_t \\
&+ \underbrace{(\alpha - \rho) p_v}_{m_{\varepsilon v}} v_{t+1} + \underbrace{(\alpha - 1) \gamma_g}_{m_{\varepsilon g}} v_t^{1/2} \varepsilon_{gt+1} + \underbrace{(\alpha - \rho) p_x \gamma_x}_{m_{\varepsilon x}} v_t^{1/2} \varepsilon_{xt+1} \\
&= \log m + m_x x_t + m_v v_t + m_{\varepsilon g} v_t^{1/2} \varepsilon_{gt+1} + m_{\varepsilon x} v_t^{1/2} \varepsilon_{xt+1} + m_{\varepsilon v} v_{t+1}.
\end{aligned}
$$

The model for the dividends is

$$\log d_{t,t+1} \;=\; d + \mu_x x_t + \gamma_d v_t^{1/2}\varepsilon_{dt+1}.$$

Guess that the price-dividend ratio is

$$\log pd_t \;=\; q_0 + q_x x_t + q_v v_t.$$

The log-linearized return is

$$
\begin{aligned}
\log r_{t+1} \;=\;& \kappa_0 + \kappa_1 \log pd_{t+1} + \log d_{t,t+1} - \log pd_t \\
\;=\;& \underbrace{k_0 + (k_1 - 1)q + \log d}_{\log r} + \underbrace{(\mu_x + q_x(k_1\varphi_x - 1))}_{r_x} x_t \underbrace{-q_v}_{r_v} v_t \\
&+ \underbrace{k_1 q_v}_{r_{\varepsilon v}} v_{t+1} + \underbrace{k_1 q_x \gamma_x}_{r_{\varepsilon x}} v_t^{1/2}\varepsilon_{xt+1} + \underbrace{\gamma_d}_{r_{\varepsilon d}} v_t^{1/2}\varepsilon_{dt+1},
\end{aligned}
$$

where

$$
\begin{aligned}
\kappa_0 \;&=\; \log(1 + pd) - \frac{\log(pd) \cdot pd}{1 + pd}, \\
\kappa_1 \;&=\; \frac{pd}{1 + pd}, \\
pd \;&=\; E(pd_t).
\end{aligned}
$$

Or in compact form

$$\log r_{t+1} \;=\; \log r + r_x x_t + r_v v_t + r_{\varepsilon x} v_t^{1/2}\varepsilon_{xt+1} + r_{\varepsilon d} v_t^{1/2}\varepsilon_{dt+1} + r_{\varepsilon v} v_{t+1}.$$

Use the law of one price $E_t[m_{t,t+1}r_{t+1}] = 1$ to obtain

$$\log r + \log m - \frac{(1 - \varphi_v)\log(1 - (r_{\varepsilon v} + m_{\varepsilon v})\sigma_v^2/2)}{\sigma_v^2/2} = 0$$

$$r_x + m_x = 0,$$

$$r_v + m_v + \frac{1}{2}m_{\varepsilon g}^2 + \frac{1}{2}(r_{\varepsilon x} + m_{\varepsilon x})^2 + \frac{1}{2}r_{\varepsilon d}^2 + \frac{\varphi_v(r_{\varepsilon v} + m_{\varepsilon v})}{1 - (r_{\varepsilon v} + m_{\varepsilon v})\sigma_v^2/2} = 0.$$

Solve for q, q_x, and q_v:

$$q_x = \frac{\mu_x + \rho - 1}{1 - k_1\varphi_x},$$

$$-q_v + \frac{\varphi_v(k_1 q_v + m_{\varepsilon v})}{1 - (k_1 q_v + m_{\varepsilon v})\sigma_v^2/2} + D = 0,$$

$$q = \frac{1}{1 - k_1}\left(k_0 + \log d + \log m - \frac{(1 - \varphi_v)\log(1 - (r_{\varepsilon v} + m_{\varepsilon v})\sigma_v^2/2)}{\sigma_v^2/2}\right),$$

41

where

$$D = m_v + m_{\varepsilon g}^2/2 + (r_{\varepsilon x} + m_{\varepsilon x})^2/2 + r_{\varepsilon d}^2/2.$$

The quadratic equation for q_v has two roots. I choose one that satisfies the requirement of stochastic stability (Hansen, 2012).

C   Solution of the Wachter (2013) model

A representative agent is averse to consumption risk and has a recursive utility

$$U_t = [(1 - \beta)c_t^\rho + \beta\mu_t(U_{t+1})^\rho]^{1/\rho}, \tag{35}$$
$$\mu_t(U_{t+1}) = [E_t(U_{t+1}^\alpha)]^{1/\alpha}.$$

The joint data generating process for consumption growth $\log g_{t,t+1}$ and dividend growth $\log d_{t,t+1}$ with consumption disasters z_{t+1} featuring time-varying arrival intensity λ_t is

$$
\begin{aligned}
\log g_{t,t+1} &= g + \gamma_g \varepsilon_{gt+1} - z_{t+1}, \\
j_{gt+1}|\lambda_t &\sim \text{Bernoulli}(h_\lambda \lambda_t), \\
z_{t+1}|j_{gt+1} &\sim \text{Gamma}(j_{gt+1}, \theta_g), \\
\lambda_{t+1} &= (1 - \varphi_\lambda) + \varphi_\lambda \lambda_t + \sigma_\lambda [((1 - \varphi_\lambda) + 2\varphi_\lambda \lambda_t)/2]^{1/2} \varepsilon_{\lambda t+1}, \\
\log d_{t,t+1} &= d + \varphi_d \gamma_g \varepsilon_{gt+1} + \gamma_d \varepsilon_{dt+1} - \varphi_d z_{t+1},
\end{aligned}
$$

where z_{t+1} is a consumption disaster modeled as a Bernoulli mixture of gammas with its central component j_{gt} as a Bernoulli random variable and with a θ_g as an average disaster size. The mean of j_{gt} is a jump arrival rate $h_\lambda E(\lambda_t) = h_\lambda$. The normalized jump intensity λ_t follows the scalar autoregressive process of order one $\lambda_t \sim \mathcal{ARG}(1)$ with the scale parameter $\sigma_\lambda^2/2$, the degrees of freedom $2(1 - \varphi_\lambda)/\sigma_\lambda^2$, and the serial correlation φ_λ. The normal shocks $\varepsilon_{gt+1} \sim \mathcal{N}(0, 1)$ and $\varepsilon_{dt+1} \sim \mathcal{N}(0, 1)$ are orthogonal to each other. The nonnormal shock $\varepsilon_{\lambda t+1}$ has a mean of zero and standard deviation of one.

Divide expression (35) by c_t, denote $u_t = U_t/c_t$ and $g_{t,t+1} = c_{t+1}/c_t$ and obtain

$$u_t = [(1 - \beta) + \beta\mu_t(u_{t+1}g_{t+1})^\rho]^{1/\rho}, \tag{36}$$

Solve a recursive problem that is a log-linear approximation of the Bellman equation (34)

$$\log u_t \approx b_0 + b_1 \log \mu_t(g_{t+1}u_{t+1}),$$

where

$$b_1 = \beta e^{\rho \log \mu}/(1 - \beta + \beta e^{\rho \log \mu}),$$

$$b_0 = \frac{1}{\rho} \log \left((1 - \beta) + \beta e^{\rho \log \mu}\right) - b_1 \log \mu.$$

Guess the value function

$$\log u_t = u + p_\lambda \lambda_t.$$

Compute

$$\log u_{t+1} + \log g_{t,t+1} = (u + g) + \gamma_g \varepsilon_{gt+1} + p_\lambda \lambda_{t+1} - z_{t+1},$$

$$\log \mu_t(u_{t+1} g_{t,t+1}) = (u + g) + \frac{\alpha \gamma_g^2}{2} + \frac{\varphi_\lambda \lambda_t p_\lambda}{1 - \alpha p_\lambda \sigma_\lambda^2/2} - \frac{(1 - \varphi_\lambda) \log (1 - \alpha p_\lambda \sigma_\lambda^2/2)}{\alpha \sigma_\lambda^2/2}$$

$$- \frac{\theta_g h_\lambda \lambda_t}{1 + \alpha \theta_g}.$$

Solve for the parameters of the value function.

There is a quadratic equation for p_λ:

$$C_2 p_\lambda^2 + C_1 p_\lambda + C_0 = 0,$$

where

$$C_2 = \alpha \sigma_\lambda^2/2,$$
$$C_1 = b_1 \varphi_\lambda - 1 - A \alpha \sigma_\lambda^2/2,$$
$$C_0 = A,$$
$$A = -\frac{b_1 h_\lambda \theta_g}{1 + \alpha \theta_g}.$$

Finally

$$\log u = \frac{1}{1 - b_1} \left(b_0 + b_1 g + \frac{\alpha b_1 \gamma_g^2}{2} - \frac{b_1 (1 - \varphi_\lambda) \log (1 - \alpha p_\lambda \sigma_\lambda^2/2)}{\alpha \sigma_\lambda^2/2} \right).$$

Compute

$$\log u_{t+1} g_{t,t+1} - \log \mu_t(u_{t+1} g_{t,t+1}) = \gamma_g \varepsilon_{gt+1} + p_\lambda \lambda_{t+1} - z_{t+1} - \frac{\varphi_\lambda p_\lambda \lambda_t}{1 - \alpha p_\lambda \sigma_\lambda^2/2} - \frac{\alpha \gamma_g^2}{2}$$

$$+ \frac{(1 - \varphi_\lambda) \log (1 - \alpha p_\lambda \sigma_\lambda^2/2)}{\alpha \sigma_\lambda^2/2} + \frac{\theta_g h_\lambda \lambda_t}{1 + \alpha \theta_g}.$$

43

The pricing kernel is

$$
\begin{aligned}
\log m_{t,t+1} &= \log \beta + (\rho - 1)\log g_{t,t+1} + (\alpha - \rho)(\log u_{t+1}g_{t,t+1} - \log \mu_t(u_{t+1}g_{t,t+1})) \\
&= m + m_\lambda \lambda_t + m_{\varepsilon\lambda}\lambda_{t+1} + m_{\varepsilon g}\varepsilon_{gt+1} + m_z z_{t+1},
\end{aligned}
$$

where

$$
\begin{aligned}
m &= \log \beta + (\rho - 1)g + (\alpha - \rho)\frac{(1 - \varphi_\lambda)\log\left(1 - \alpha p_\lambda \sigma_\lambda^2/2\right)}{\alpha \sigma_\lambda^2/2} - \frac{\alpha \gamma_g^2(\alpha - \rho)}{2}, \\
m_\lambda &= -(\alpha - \rho)\left(\frac{\varphi_\lambda p_\lambda}{1 - \alpha p_\lambda \sigma_\lambda^2/2} - \frac{\theta_g h_\lambda}{1 + \alpha\theta_g}\right), \\
m_{\varepsilon g} &= (\alpha - 1)\gamma_g, \\
m_z &= -(\alpha - 1), \\
m_{\varepsilon\lambda} &= (\alpha - \rho)p_\lambda.
\end{aligned}
$$

Guess that the price-dividend ratio is

$$
\log pd_{t+1} = q_0 + q_\lambda \lambda_{t+1}.
$$

The log return is

$$
\begin{aligned}
\log r_{t,t+1} &= k_0 + k_1 \log pd_{t+1} + \log d_{t,t+1} - \log pd_t \\
&= k_0 + k_1 q_0 - q_0 + \log d \\
&\quad - q_\lambda \lambda_t \\
&\quad + k_1 q_\lambda \lambda_{t+1} \\
&\quad + \varphi_d \gamma_g \varepsilon_{gt+1} \\
&\quad + \gamma_d \varepsilon_{dt+1} \\
&\quad - \varphi_d z_{t+1} \\
&= r + r_\lambda \lambda_t + r_{\varepsilon\lambda}\lambda_{t+1} + r_{\varepsilon g}\varepsilon_{gt+1} + r_z z_{t+1} + r_{\varepsilon d}\varepsilon_{dt+1}.
\end{aligned}
$$

where

$$
\begin{aligned}
r &= k_0 + k_1 \log q_0 - \log q_0 + \log d, \\
r_\lambda &= -q_\lambda, \\
r_{\varepsilon\lambda} &= k_1 q_\lambda, \\
r_{\varepsilon g} &= \varphi_d \gamma_g, \\
r_{\varepsilon d} &= \gamma_d, \\
r_z &= -\varphi_d.
\end{aligned}
$$

44

Use the law of one price

$$E_t(r_{t,t+1}m_{t+1}) = 1,$$

to solve for q_0 and q_λ. The two equations are

$$r + m - \frac{(1 - \varphi_\lambda) \log\left(1 - (m_{\varepsilon\lambda} + r_{\varepsilon\lambda})\sigma_\lambda^2/2\right)}{\sigma_\lambda^2/2} + \frac{1}{2}(m_{\varepsilon g} + r_{\varepsilon g})^2 + \frac{1}{2}r_{\varepsilon d}^2 = 0,$$

$$m_\lambda + r_\lambda + \frac{\varphi_\lambda(m_{\varepsilon\lambda} + r_{\varepsilon\lambda})}{1 - (m_{\varepsilon\lambda} + r_{\varepsilon\lambda})\sigma_\lambda^2/2} + \frac{(m_{zg} + r_{zg})h_\lambda\theta_g}{1 - (m_{zg} + r_{zg})\theta_g} = 0.$$

Parameter q_λ is the stochastically stable solution to the quadratic equation

$$C_0 q_\lambda^2 + C_1 q_\lambda + C_2 = 0,$$

where

$$
\begin{aligned}
A &= m_\lambda + \frac{(m_{zg} + r_{zg})h_\lambda\theta_g}{1 - (m_{zg} + r_{zg})\theta_g}, \\
C_0 &= k_1\sigma_\lambda^2/2, \\
C_1 &= k_1\varphi_\lambda + m_{\varepsilon\lambda}\sigma_\lambda^2/2 - 1 - k_1 A\sigma_\lambda^2/2, \\
C_2 &= \varphi_\lambda m_{\varepsilon\lambda} + A - A m_{\varepsilon\lambda}\sigma_\lambda^2/2.
\end{aligned}
$$

Also

$$q_0 = \frac{1}{1 - k_1}\left(k_0 + d + m - \frac{(1 - \varphi_\lambda)\log\left(1 - (m_{\varepsilon\lambda} + r_{\varepsilon\lambda})\sigma_\lambda^2/2\right)}{\sigma_\lambda^2/2} + \frac{1}{2}(m_{\varepsilon g} + r_{\varepsilon g})^2 + \frac{1}{2}r_{\varepsilon d}^2\right).$$

D Solution of the Drechsler and Yaron (2011) model

The model for consumption growth with stochastic variance v_t which has a time-varying long run mean driven by v_t^* and self-exciting jumps z_v is

$$
\begin{aligned}
\log g_{t,t+1} &= g + \gamma_g v_t^{1/2}\varepsilon_{gt+1}, \\
v_{t+1} &= (1 - \tilde{\varphi}_v)v_t^* + (1 - \varphi_v)v + \varphi_v v_t + \sigma_v(((1 - \varphi_v)v + 2\varphi_v v_t)/2)^{1/2}\varepsilon_{vt+1} + z_{vt+1}, \\
v_{t+1}^* &= (1 - \varphi_v^*)v^* + \varphi_v^* v_t^* + \sigma_v^*(((1 - \varphi_v^*)v^* + 2\varphi_v^* v_v^*)/2)^{1/2}\varepsilon_{vt+1}^*, \\
j_{vt+1}|v_t &\sim \text{Poisson}(h_v v_t), \\
z_{vt+1}|j_{vt+1} &\sim \text{Gamma}(j_{vt+1}, \theta_v),
\end{aligned}
$$

where $\tilde{\varphi}_v = \varphi_v + \theta_v h_v$, $v^* = 1/2$, and $v = \frac{1 - \tilde{\varphi}_v}{(1 - \varphi_v)/2}$, so that $Ev_t = 1$ and $Ev_t^* = 1/2$.

A representative agent has recursive preferences

$$
\begin{aligned}
U_t &= [(1-\beta)c_t^\rho + \beta\mu_t(U_{t+1})^\rho]^{1/\rho}, \\
\mu_t(U_{t+1}) &= [E_t(U_{t+1}^\alpha)]^{1/\alpha}.
\end{aligned}
\tag{37}
$$

Divide expression (37) by c_t, denote $u_t = U_t/c_t$ and $g_{t,t+1} = c_{t+1}/c_t$ and obtain

$$
u_t = [(1-\beta) + \beta\mu_t(u_{t+1}g_{t+1})^\rho]^{1/\rho},
\tag{38}
$$

Solve a recursive problem that is a log-linear approximation of the Bellman equation (34)

$$
\log u_t \approx b_0 + b_1 \log \mu_t(g_{t+1}u_{t+1}),
$$

where

$$
\begin{aligned}
b_1 &= \beta e^{\rho \log \mu}/(1-\beta+\beta e^{\rho \log \mu}), \\
b_0 &= \frac{1}{\rho}\log\left((1-\beta)+\beta e^{\rho \log \mu}\right) - b_1 \log \mu.
\end{aligned}
$$

Guess the value function

$$
\log u_t = u + p_v v_t + p_v^* v_t^*.
$$

Compute

$$
\log u_{t+1} + \log g_{t,t+1} = u + g + p_v v_{t+1} + \gamma_g v_t^{1/2}\varepsilon_{gt+1} + p_v^* v_{t+1}^*
$$

Compute

$$
\begin{aligned}
\log \mu_t(u_{t+1}g_{t+1}) &= u + g - \frac{(1-\varphi_v)v\log\left(1-\alpha p_v \sigma_v^2/2\right)}{\alpha\sigma_v^2/2} - \frac{(1-\varphi_v^*)v^*\log\left(1-\alpha p_v^*\sigma_v^{*2}/2\right)}{\alpha\sigma_v^{*2}/2} \\
&+ \frac{\alpha}{2}\gamma_g^2 v_t + \frac{\varphi_v p_v v_t}{1-\alpha p_v \sigma_v^2/2} + \frac{p_v \theta_v h_v v_t}{1-\alpha p_v \theta_v} + \frac{\varphi_v p_v^* v_t^*}{1-\alpha p_v^* \sigma_v^{*2}/2} + p_v(1-\tilde{\varphi}_v)v_t^*.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\log \mu &= u + g + \frac{\alpha}{2}\gamma_g^2 - \frac{(1-\varphi_v)v\log\left(1-\alpha p_v \sigma_v^2/2\right)}{\alpha\sigma_v^2/2} - \frac{(1-\varphi_v^*)v^*\log\left(1-\alpha p_v^*\sigma_v^{*2}/2\right)}{\alpha\sigma_v^{*2}/2} \\
&+ \frac{p_v \varphi_v}{1-\alpha p_v \sigma_v^2/2} + \frac{p_v \theta_v h_v}{1-\alpha p_v \theta_v} + p_v(1-\tilde{\varphi}_v)v^* + \frac{p_v^* \varphi_v^* v^*}{1-\alpha p_v^* \sigma_v^{*2}}.
\end{aligned}
$$

Solve the following system of three equations in three unknowns to verify the guess of the value function

$$u = \frac{1}{1-b_1}\left(b_0 + b_1 g - b_1 \frac{(1-\varphi_v)v\log\left(1-\alpha p_v \sigma_v^2/2\right)}{\alpha \sigma_v^2/2} - b_1 \frac{(1-\varphi_v^*)v^*\log\left(1-\alpha p_v^*\sigma_v^{*2}/2\right)}{\alpha \sigma_v^{*2}/2}\right),$$

$$p_v = b_1\left(\frac{\alpha}{2}\gamma_g^2 + \frac{\varphi_v p_v}{1-\alpha p_v \sigma_v^2/2} + \frac{p_v \theta_v h_v}{1-\alpha p_v \theta_v}\right),$$

$$p_v^* = b_1\left(p_v(1-\tilde{\varphi}_v) + \frac{p_v^* \varphi_v^*}{1-\alpha p_v^*\sigma_v^{*2}/2}\right).$$

The pricing kernel is

$$\log m_{t,t+1} = \underbrace{\log\beta + (\rho-1)g + \frac{(\alpha-\rho)(1-\varphi_v)v\log\left(1-\alpha p_v\sigma_v^2/2\right)}{\alpha\sigma_v^2/2} + \frac{(\alpha-\rho)(1-\varphi_v^*)v^*\log\left(1-\alpha p_v^*\sigma_v^{*2}/2\right)}{\alpha\sigma_v^{*2}/2}}_{m}$$

$$+ \underbrace{\left(-\frac{(\alpha-\rho)\varphi_v p_v}{1-\alpha p_v\sigma_v^2/2} - \frac{(\alpha-\rho)p_v\theta_v h_v}{1-\alpha p_v\theta_v} - \frac{\alpha(\alpha-\rho)}{2}\gamma_g^2\right)v_t}_{m_v}$$

$$+ \underbrace{\left(-\frac{(\alpha-\rho)p_v^*\varphi_v^*}{1-\alpha p_v^*\sigma_2^{*2}/2} - (\alpha-\rho)p_v(1-\tilde{\varphi}_v)\right)v_t^*}_{m_v^*}$$

$$+ \underbrace{(\alpha-1)\gamma_g\, v_t^{1/2}\varepsilon_{gt+1}}_{m_{\varepsilon g}} + \underbrace{(\alpha-\rho)p_v\, v_{t+1}}_{m_{\varepsilon v}} + \underbrace{(\alpha-\rho)p_v^*\, v_{t+1}^*}_{m_{\varepsilon v}^*}.$$

The model for the dividends is

$$\log d_{t,t+1} = d + \mu_v v_t + \gamma_{\varepsilon v}^* v_{t+1}^* + \gamma_{\varepsilon d} v_t^{1/2}\varepsilon_{dt+1},$$

Guess that the price-dividend ratio is

$$\log pd_t = q + q_v v_t + q_v^* v_t^*.$$

The log-linearized return is

$$\log r_{t+1} = \kappa_0 + \kappa_1 \log pd_{t+1} + \log d_{t,t+1} - \log pd_t$$
$$= \underbrace{k_0 + (k_1-1)q + d}_{r} + \underbrace{(\mu_v - q_v)}_{r_v} v_t + \underbrace{(-q_v^*)}_{r_v^*} v_t^*$$
$$+ \underbrace{k_1 q_v}_{r_{\varepsilon v}} v_{t+1} + \underbrace{(k_1 q_v^* + \gamma_{\varepsilon v}^*)}_{r_{\varepsilon v}^*} v_{t+1}^* + \underbrace{\gamma_d}_{r_{\varepsilon d}} v_t^{1/2}\varepsilon_{dt+1},$$

47

where

$$
\begin{aligned}
\kappa_0 &= \log(1+pd) - \frac{\log(pd) \cdot pd}{1+pd}, \\
\kappa_1 &= \frac{pd}{1+pd}, \\
pd &= E(pd_t).
\end{aligned}
$$

Or in compact form

$$
\log r_{t+1} = r + r_v v_t + r_v^* v_t^* + r_{\varepsilon d} v_t^{1/2} \varepsilon_{dt+1} + r_{\varepsilon v} v_{t+1} + r_{\varepsilon v}^* v_{t+1}^*.
$$

Use the law of one price $E_t[m_{t,t+1} r_{t,t+1}] = 1$ to obtain three equations in three unknown q, q_v, q_v^*

$$
m + r - \frac{(1-\varphi_v)v \log\left(1 - (m_{\varepsilon v} + r_{\varepsilon v})\sigma_v^2/2\right)}{\sigma_v^2/2} - \frac{(1-\varphi_v^*)v^* \log\left(1 - (m_{\varepsilon v}^* + r_{\varepsilon v}^*)\sigma_v^{*2}/2\right)}{\sigma_v^{*2}/2} = 0,
$$

$$
m_v + r_v + \frac{m_{\varepsilon g}^2}{2} + \frac{r_{\varepsilon d}^2}{2} + \frac{(m_{\varepsilon v} + r_{\varepsilon v})\varphi_v}{1 - (m_{\varepsilon v} + r_{\varepsilon v})\sigma_v^2/2} + \frac{m_{\varepsilon v}\theta_v h_v}{1 - m_{\varepsilon v}\theta_v} = 0,
$$

$$
m_v^* + r_v^* + (1-\tilde{\varphi}_v)(m_{\varepsilon v} + r_{\varepsilon v}) + \frac{(m_{\varepsilon v}^* + r_{\varepsilon v}^*)\varphi_v^*}{1 - (m_{\varepsilon v}^* + r_{\varepsilon v}^*)\sigma_v^{*2}/2} = 0.
$$

E   Decomposition of the term structure of risk

TBC

F   Calibration of the structural models

Table 1
Calibrations of the leading equity premium models

Quarterly calibrations of discrete-time models with modeling ingredients from Bansal and Yaron (2004), Wachter (2013), and Drechsler and Yaron (2011). A model in spirit of Bansal and Yaron (2004): I choose parameters to match means, unconditional volatility, and persistence of the variables implied by the original monthly calibration of Bansal and Yaron (2004). I posit that the stochastic variance follows the autoregressive process of order one. The calibration inputs are $g = 0.0045$, $\gamma_g = 0.0135$, $\varphi_x = 0.9383$, $\gamma_x = 0.0010$, $\varphi_v = 0.9615$, $\sigma_v = 0.0653$, $d = 0.0045$, $\mu_x = 6$, $\gamma_d = 0.0653$, $\alpha = -9$, $\rho = 1/3$, $\beta = 0.998$, $\kappa_0 = 0.0449$, $\kappa_1 = 0.9923$. A model in spirit of Wachter (2013): I posit that consumption disasters are binomial mixtures of exponentials, whereas the stochastic disaster intensity is an autoregressive gamma process of order one. The calibration inputs are $g = 0.063$, $\gamma_g = 0.01$, $h_\lambda = 0.0075$, $\varphi_\lambda = 0.9802$, $\sigma_\lambda = 0.1743$, $\varphi_d = 2.6$, $d = 0.0163$, $\gamma_d = 0$, $\theta_g = 0.2$, $\alpha_p = -2$, $\rho = 0$, $\beta = 0.997$, $\kappa_0 = 0.0449$, $\kappa_1 = 0.9923$. A model in spirit of Drechsler and Yaron (2011): I cut off the long-run risk channel from the model and posit that the stochastic volatility of consumption growth as well as its long-run mean follow two independent autoregressive gamma processes or order one. The jump in the volatility of consumption growth is modelled as a Poisson mixture of exponentials. The jump arrival rate is a linear function of the volatility of consumption growth. The calibration inputs are $g = 0.0045$, $\gamma_g = 0.0108$, $\varphi_v^* = 0.985$, $\sigma_v^* = 0.1025$, $v^* = 1/2$, $\varphi_v = 0.955$, $\theta_v = 0.8$, $h_v = 0.025$, $\sigma_v = 0.175$, $v = (1 - \varphi_v - \theta_v h_v)/(v \cdot (1 - \varphi_v))$, $d = 0.0175$, $\mu_v = -0.004$, $\gamma_{\varepsilon v}^* = -0.0125$, $\gamma_d = 0.05$, $\alpha = -9$, $\beta = 0.9985$, $\rho = 1/3$, $\kappa_0 = 0.0449$, $\kappa_1 = 0.9923$. In percent.

	BY (2004)	W (2013)	DY (2011)	Data
Log equity premium	1.1785	1.0356	1.0122	1.4971
Std dev of equity return	10.1507	7.6654	10.1828	8.1712
Risk-free rate	0.3125	0.2755	0.3156	0.2768
Std dev of the risk-free rate	0.1993	0.5474	0.1412	0.5455
Mean consumption growth	0.4500	0.4750	0.4500	0.4750
Std dev of consumption growth	1.3821	2.6458	1.0808	0.5067
Mean dividend growth	0.4500	1.2350	0.7250	0.6821
Std dev of dividend growth	6.7823	6.8557	5.0472	2.3355