Term structure of risk in expected returns*

Irina Zviadadze†

This Draft: January 7, 2019

Abstract

This article develops an empirical methodology to determine which economic shocks span risk in asset returns and fluctuations in discount rate and cash flow news. A theoretically motivated shock identification scheme in a present-value model identifies economic shocks. The choice of identifying restrictions is based on the properties of the term structure of risk in expected returns in the data and in equilibrium models. Empirically, I relate equity discount rate news and cash flow news to multiple sources of risk in the variance of consumption growth. Both types of news are almost equally important for the aggregate market risk.

JEL Classification Codes: C32, C52, G12

Keywords: incremental expected return, incremental expected dividend, permanent and transient shocks

*I am grateful to Anna Kleymenova, Ravi Bansal, Gurdip Bakshi, Geert Bekaert, Jaroslav Borovička, Nina Boyarchenko, Sofya Budanova, Anna Cieslak, Mikhail Chernov, Riccardo Colacito, Richard Crump, Magnus Dahlquist, Greg Duffee, Michael Gallmeyer, Francisco Gomes, Lars Hansen, Christian Julliard, Ralph Koijen, Howard Kung, Philippe Mueller, Alessia Paccagnini, Anna Pavlova, Julien Penasse, Lukas Schmid, George Skiadopoulos, Andrea Tamoni, Fabio Trojani, Stijn van Nieuwerburgh, Raman Uppal, and Amir Yaron for helpful discussions and thank participants at the AFA, Adam Smith Workshop, Bank of England, BI-SHoF, Early Career Women in Finance Conference, Fed Board, FRB of New York, First Asset Pricing Conference at LT@UniTo, Frontiers in Macrofinance at JHU Carey, Fuqua, Gothenburg University, 7th International Moscow Finance Conference, Indiana, Kenan-Flagler, Luxembourg School of Finance, MFS, Maynooth University, McGill, Piraeus University, SED, SFS Cavalcade, Stockholm School of Economics, Sveriges Riksbank, Toulouse Financial Econometrics Conference, Toulouse School of Economics, and University College Dublin for comments. I especially thank Anh Le for making his code that computes the density of the autoregressive gamma process, available on his website. I gratefully acknowledge the financial support from the Jan Wallander and Tom Hedelius Foundation and the Swedish House of Finance. This paper was previously titled "Term structure of risk in macrofinance models." The latest version of the paper is available at `sites.google.com/site/irinazviadadzessite/home/research/zviadadze_ts.pdf`

†Address: Stockholm School of Economics, Department of Finance, Drottninggatan 98, 111 60 Stockholm, Sweden, and CEPR, Email: Irina.Zviadadze@hhs.se, Web: sites.google.com/site/irinazviadadzessite.

1 Introduction

The question of what drives asset returns has been the focus of finance researchers for several decades. The seminal work of Campbell (1991) introduced simple but powerful methods to measure the relative importance of discount rate news and cash flow news for explaining the risk in asset returns. As asset prices and macroeconomic indicators follow common business cycles, fluctuations in asset returns, discount rate news, and cash flow news must reflect macroeconomic sources of risk. In this paper, I ask what these sources of risk are.

The theoretical asset pricing literature provides various answers to this question. Models that rely on different types of systematic shocks (long-run consumption risk, volatility risk, preference shocks, risk aversion shocks, economic disasters, demographic risk, learning-induced shocks, etc.) have similar success in matching high magnitudes of risk premia, presence of return predictability, and low correlations between returns and cash flows. As a result, they provide similar decomposition of risk in asset returns in terms of cash flow news and discount rate news. This makes it difficult to distinguish different economic explanations of unexpected variation in asset returns.

The objective of this paper is to propose an empirical methodology that can deliver new evidence suitable for supporting or disqualifying competing economic shocks spanning risk in asset returns. The methodology has two stages: (i) identification of orthogonal aggregate shocks and (ii) their subsequent characterization. As key characteristics of aggregate shocks, I use empirical properties of multiperiod risk sensitivities of expected buy-and-hold returns. I construct a term structure of multiperiod risk sensitivities and label it a *term structure of risk in expected returns*. The signs of the level and of the slope of the term structure of risk in the data and in equilibrium models show promise in distinguishing alternative theories of the risk-return tradeoff in asset markets.

As an application, I explore which sources of consumption risk can explain variation in stock returns and fluctuations in cash flow news and discount rate news. I find that a theory based on multiple sources of risk in the variance of consumption growth can account for the salient properties of the term structure of risk in expected returns. Three shocks, which are originated in the variance of consumption growth, are responsible for 94% of risk in the one-period stock returns: a jump in the variance of consumption growth (jump in variance), a regular shock in the variance of consumption growth (regular variance shock), and a shock that drives the long-run mean of the variance in consumption growth (variance trend shock). The variance trend shock spans time-variation in discount rate news, whereas the jump in variance, the regular variance shock, and the direct dividend shock span time-variation in discount rate news. Cash flow news and discount rate news contribute about equally to the unexpected variation in the one-quarter stock returns.

My methodology identifies shocks by modeling log asset returns, cash flow growth, and a predictive variable, implied by a present-value identity (Campbell and Shiller, 1989), jointly with economic states in a state-space model.[1] The choice of economic states and types of shocks that feed these states represent a shock identification scheme. Many alternative theories rely on importance of latent states such as a macroeconomic volatility or a persistent component in expected consumption growth, whose estimation is challenging. The observation equation in the state-space model, which relates the observable predictive variable to the state vector, provides valuable statistical power for identification of latent states and subsequently aggregate shocks. Different statistical properties of alternative economic shocks, such as frequency of their arrival (rare shocks versus regular shocks), persistence (shocks originated in a persistent state variable or in a random walk), and size (small versus large shocks, one-sided versus two-sided shocks), facilitate shock identification.

My methodology delivers informative description of the aggregate shocks by quantifying their relative importance for risk in asset returns and cash flow growth at alternative investment horizons. I use an entropy-based measure of risk, and therefore, I can characterize both normal and nonnormal shocks in a unified fashion. I formalize these quantitative descriptions in new metrics labeled the *incremental expected return* ($\mathcal{IER}$) and the *incremental expected dividend* ($\mathcal{IED}$). A collection of $\mathcal{IER}$s for alternative horizons constitutes a term structure of risk in expected returns. A collection of $\mathcal{IED}$s for alternative horizons constitutes a term structure of risk in expected cash flow growth.

A theory of dynamic value decomposition (for a comprehensive review, see Borovička and Hansen (2016) and references therein) and especially the toolkit of shock elasticities of Borovička

[1]Prominent examples of a present-value model in equities, foreign exchange, and fixed income respectively are Campbell and Shiller (1989), Froot and Ramadorai (2005), and Campbell and Ammer (1993).

and Hansen (2014) inspire the design of the $\mathcal{IER}$s and $\mathcal{IED}$s. The $\mathcal{IER}$s and $\mathcal{IED}$s advance methods of dynamic value decomposition to nonnormal shocks, thereby representing a unique toolbox for quantifying nonnormal and heteroskedastic shocks across alternative horizons. Put simply, an introduction of these metrics opens up a door for analyzing quantitative characteristics of alternative sources of risk premia in different asset markets across multiple horizons.

The empirical analysis, which illustrates how my methodology works for examining the risk-return tradeoff in the aggregate U.S. stock market, proceeds in two steps. In the first step, I set up a reduced-form state-space model of the joint dynamics of stock returns, dividend growth, consumption growth, and the price-dividend ratio. I take into account the parameter restrictions implied by Campbell and Shiller's (1989) loglinearization of a gross stock return. I impose three distinct shock identification schemes on the innovations of the state-space model. These three identification schemes differ with respect to the choice of state variables, which span time-variation in the price-dividend ratio, and with respect to the choice of shocks, which span unexpected variation in the state variables.

I consider three identification schemes that are based on aggregate shocks, which play a prominent role in leading asset pricing theories. In the first identification scheme, labeled "Long-Run Risk," I hypothesize that shocks in the expected consumption growth and in the variance of consumption growth are the main sources of time-variation in the price-dividend ratio. In the second scheme, labeled "Consumption Disasters," I hypothesize that the time-varying probability of rare consumption disasters is the only source of time-variation in the price-dividend ratio. In the third scheme, labeled "Consumption Uncertainty," I assume that multiple sources of risk in the variance of consumption growth, such as a regular variance shock, a variance trend shock, and occasional jumps, span time variation in the price-dividend ratio.

In the second step, I evaluate which of the three identification schemes are consistent with implications of structural models that feature the same state-space representation, and therefore the same shocks. To this end, I follow the idea of Christiano, Eichenbaum, and Evans (1999), who select realistic identifying assumptions for macroeconomic shocks by comparing the theoretical and empirical *impulse response functions* ($\mathcal{IRF}$s). The $\mathcal{IRF}$s characterize multiperiod implications of normal homoskedastic shocks, whereas sources of risk premia, the focus here, are nonnormal. As a result, instead of $\mathcal{IRF}$s for log returns, I use $\mathcal{IER}$s, which are developed specifically for characterizing nonnormal shocks across alternative horizons in the presence of heteroskedasticity. I select realistic shock-identifying assumptions by comparing the theoretical and empirical $\mathcal{IER}$s. The empirical state-space model, which I estimate, implies the empirical term structures of risk, whereas the structural models in Bansal and Yaron (2004), Wachter (2013), and Drechsler and Yaron (2011) imply the theoretical term structures of risk.

Of the three evaluated identifications, only "Consumption Uncertainty" produces shocks whose properties are comparable both in the data and in the model. Armed with this identification, I decompose unexpected variation in the one-period stock returns and analyze the sources of cash flow news and discount rate news. I find that 94% and 6% of risk in the one-period stock returns, respectively, relate to the multiple shocks in the variance of consumption growth and a direct dividend shock. The term structures of the $\mathcal{IER}$s and $\mathcal{IED}$s indicate that at horizons longer than 10 years, revisions of multiperiod discount rates reflect exclusively risk exposures of returns to the variance trend risk, whereas revisions of multiperiod expected cash flow growth reflect exclusively risk exposures of dividends to the regular variance shock, jump in variance, and the direct dividend shock. Taken together, these findings have important implications for macrofinance modeling: both cash flow news and discount rate news, as represented by multiple sources of risk in the variance of consumption growth, play a crucial role in aggregate stock market risk. Cash flow news account for 53%, whereas discount rate news account for 47% of risk in the one-period aggregate stock returns.

The analysis reveals discrepancies between the empirical and theoretical $\mathcal{IER}$s associated with the long-run risk and time-varying consumption disasters. In Bansal and Yaron (2004), the slope of the term structure of the long-run risk is positive (longer holding period returns have higher sensitivities to the shock), whereas in the data it is negative (shorter holding period returns have higher sensitivities to the shock). In Wachter (2013), the level of the term structure of the disaster shock is negative (returns drop upon arrival of a consumption disaster), whereas in the data it is positive (returns rally upon arrival of a consumption disaster). The cross-equation restrictions inherent to the structural models cause a tension between the empirical and theoretical term structures of risk.

Given that "Consumption Uncertainty" is the realistic shock-identifying scheme, I decompose the joint variation in discount rate news and cash flow news into transitory and permanent components and characterize their economic drivers. The transient component in stock prices and dividends is the variance trend shock. The permanent components are the jump in the variance of consumption growth, the regular variance shock, and the direct dividend shock. To the best of my knowledge, this is the first study that provides an economic interpretation of the transitory and permanent components in stock prices and dividends, thereby complementing the influential evidence of Cochrane (1994).

Two sources of permanent risk in stock prices and dividends, the regular variance shock and the jump in variance, resemble the asset pricing innovation of Bryzgalova and Julliard (2018) to which consumption growth responds slowly. As a result, my empirical findings have at least two interesting implications for the interaction of macroeconomy and asset markets. First, I uncover macroeconomic underpinnings of the common source of risk in the cross section of bonds and stock returns that Bryzgalova and Julliard (2018) identify. Second, my results suggest a refinement of the definition of the long-run risk in consumption growth. A common innovation in the first and higher-order moments of log consumption growth holds promise for understanding the risk-return tradeoff in asset markets.

Last but not least, I identify the economic sources of predictability. The identification "Consumption Uncertainty" suggests that the variance trend shock drives predictability in gross returns, whereas the regular variance shock and the jump in variance drive predictability in dividend growth. Consistent with Binsbergen and Koijen (2010), I find that the persistence of expected dividend growth is lower than the persistence of expected log returns. As my empirical facts pinpoint the economic sources of return and dividend predictability, they complement the important empirical evidence of Binsbergen and Koijen (2010) in a nontrivial way.

Related literature

The methodology in this paper is inspired by the influential literature in applied macroeconomics that uses techniques of impulse response function matching for testing and estimating stochastic general equilibrium models. See, for example, Christiano, Eichenbaum, and Evans (2005), Fernandez-Villaverde, Rubio-Ramirez, and Schorfheide (2016), Giacomini (2013), and Guerron-Quintana, Inoue, and Kilian (2017) and references therein. I combine the ideas in this literature with recents advances in the theory of dynamic value decomposition (Borovička and Hansen, 2014; Borovička, Hansen, Hendricks, and Scheinkman, 2011; Borovička, Hansen, and Scheinkman, 2014; Hansen, 2012), whose methods I extend to nonnormal shocks. As a result, I develop an empirical approach that can characterize how alternative sources of risk premia propagate in expected returns and cash flows across alternative horizons.

Trajectories of how alternative shocks affect discount rates and expected cash flows are informative about the multiperiod interaction of cash flow risk exposures with prices of risk. As a result, term structures of risk in expected returns and cash flow growth are useful moments for understanding the multiperiod risk-return tradeoff in asset markets. In this regard, my approach complements the recent literature on the term structure of risk premia in different asset classes (for example, Backus, Boyarchenko, and Chernov, 2018 and Binsbergen and Koijen, 2017). The current focus is on the term structure of expected returns on the same asset but at different holding periods, rather than on the term structure of expected returns on short-term and long-term assets, as in the aforementioned papers.

My paper is also related to the literature on predictability of returns and cash flows from the perspective of the present-value identity (for a comprehensive review see Koijen and Nieuwerburgh (2011) and references therein). Binsbergen and Koijen (2010), Piatti and Trojani (2017), and Rytchkov (2012) use a latent variable approach within present-value models to quantify the magnitude of return predictability and dividend growth predictability. They specify exogenous models for expected returns and expected cash flow growth with multiple shocks but do not relate them to macroeconomic risk. My focus is different. While I also exploit present-value relationships, I relate time variation in predictive variables to multiple sources of consumption risk. As a result, latent states in my approach are necessarily economic state variables featured in leading asset pricing models.

The paper is organized as follows. Section II introduces the methodology for identifying aggregate shocks and describing their quantitative properties in the term structure of expected returns and expected cash flow growth. Section III sets up a measurement exercise for examining aggregate sources of risk in the term structure of expected stock returns and expected dividends and discusses empirical findings. Section IV provides critical assessment of the methodology. Section V concludes. Five appendices contain supplementary material: explicit formulation of the empirical model under different identification schemes, description of the estimation output, solutions to the equilibrium models that provide theoretical term structures of risk, analysis of theoretical term structures of risk, and discussion of how the empirical $\mathcal{IER}$s and $\mathcal{IED}$s relate to the cross section of prices on dividend strips of different maturities. The Online Appendix introduces three tractable cases of shock elasticities for nonnormal shocks, provides a full description of the $\mathcal{IER}$s and $\mathcal{IED}$s in the equilibrium and empirical models, and sketches the estimation algorithm.

2 Methodology

One of the central questions in asset pricing is which shocks drive returns. Empirical and theoretical asset pricing literatures pursue independent agendas to provide an answer to the question.

The focus of the empirical literature is on the dynamics of normal homoskedastic innovations in discount rates and expected cash flow growth. Researchers use implications of vector autoregressions (VARs) (i) to analyze how cash flow news and discount rate news contribute to the variance of log returns and (ii) to decompose innovations in prices and cash flows into transient and permanent components. The prominent example of the former is Campbell (1991); the prominent example of the latter is Cochrane (1994). The cash flow news and discount rate news are endogenous functions of aggregate structural shocks; the latter is the object of interest of the theoretical literature.[2]

The theoretical literature has a wide range of structural models describing how different economic shocks interact with preferences. Researchers evaluate plausibility of alternative risk mechanisms by calibrating the corresponding models to a standard set of moments in the data. Given that the existing models match the moments similarly well, it is hard to select a preferred model.

[2]I use the terms structural shocks and economic shocks interchangeably.

Nevertheless, one message from this analysis emerges forcefully: nonnormal and/or heteroskedastic shocks play a special role in asset markets and macroeconomy. Nonnormal shocks justify high risk premia and the presence of heavy tails in the distribution of asset returns and macroeconomic fundamentals; heteroskedastic shocks serve as sources of time-varying risk premia.[3]

To make further progress, I use lessons from both literatures and design an empirical methodology for examining properties of alternative structural shocks in the term structures of expected returns and cash flow growth. The methodology has two stages: (i) identification of economic shocks and (ii) characterization of economic shocks in the term structure of discount rates and expected cash flow growth. It applies universally to normal and nonnormal shocks and allows for heteroskedasticity. Thus, the methodology suits the analysis of alternative sources of time-varying risk premia well. As an output, the methodology (i) decomposes risk in asset returns, as well as fluctuations in cash flow news and discount rate news into orthogonal structural shocks; (ii) identifies economic sources of transitory and permanent risks in asset prices and cash flows; and (iii) relates predictability to economic sources of risk. In short, the methodology provides an economic interpretation of the key aspects of risk in asset returns.

A starting point of my approach is the present-value identity that relates an observable valuation ratio to multiperiod discount rates and expected multiperiod log cash flow growth. Campbell and Shiller (1989), Lettau and Ludvigson (2001), Froot and Ramadorai (2005), Gourinchas and Rey (2007), and Campbell and Ammer (1993) among others use present-value identities in different asset classes. As any predictive variable is an endogenous function of economic state variables, its innovation is an endogenous function of orthogonal structural shocks. As a result, shock identification is a necessary step towards quantifying the role of individual structural shocks on multiperiod discount rates and expected cash flows.

I address the identification problem by modeling the joint dynamics of log returns $\log r_{t,t+1}$, log cash flow growth $\log g^d_{t,t+1}$, and the predictive variable $\log \delta_t$ in a state-space model and imposing identifying restrictions. The vector $Y_t = (\log r_{t-1,t}\ \log g^d_{t-1,t},\ y_t')'$ combines the log return and the

[3] In discrete-time asset pricing models, shocks, which drive stochastic variance or time-varying intensity of jumps, can also be nonnormal.

log cash flow growth with a vector of state variables y_t (to be defined shortly) and follows an autoregressive process with a matrix of shocks W_{t+1}. The observation equation is a linear mapping between the predictive variable $\log \delta_t$ and the state vector y_t,

$$
\begin{aligned}
Y_{t+1} &= A + BY_t + W_{t+1}, &(1)\\
\log \delta_t &= a + b' y_t + w_t, &(2)
\end{aligned}
$$

where $w_t \sim \mathcal{N}(0, \sigma_\delta^2)$ is an observation error orthogonal to the shocks in W_t. The shocks in W_{t+1} can be either normal or nonnonrmal and may exhibit heteroskedasticity. A choice of the state vector y_t and a vector of structural shocks ε_t such that $W_t = W(\varepsilon_t)$, serves as shock-identifying assumptions.[4]

Macro-finance theory guides the choice. First, through the first-order optimality conditions, equilibrium models explicitly specify economic state variables (for example, expected consumption growth, demographic uncertainty, default intensity, etc.) which span time variation in the predictive variable or serve as alternative predictive variables.[5] Second, different structural models advocate shocks of different size (crashes versus regular movements), persistence (shocks originated in a persistent variable, random walk, or white noise), and frequency of arrival (rare versus regular events) as main drivers of expected returns and cash flows. The shocks may affect variables of interest contemporaneously or with a lag, positively or negatively, and may or may not have a non-zero permanent impact. As a result, every structural model as a collection of state variables and aggregate shocks with specific characteristics serves as an identification scheme.

A classic VAR of an asset return, cash flow growth, and a predictive variable in logs is a special case of the state-space model given in expressions (1)-(2). The two are equivalent if (i) $a = 0$, $b = 1$, $\sigma_\delta = 0$ and (ii) shocks in W_{t+1} are normal and homoskedastic. The present-value identity implies that there are two distinct (orthogonal) shocks in the VAR. Standard identifying restrictions

[4]If the system (1)-(2) is only subject to normal and homoskedastic shocks, then W is a linear multivariate mapping between W_t and ε_t.

[5]The system (1)-(2) does not imply that variation in $\log \delta_t$ spans return and cash flow predictability. The vector y_t can include alternative observable predictive variables that have zero correlation with the predictive variable implied by the present-value identity. In such a case, the corresponding elements of the vectors a and b are null. Similarly, in the absence of observable states, all but one latent states can have explanatory power for expected returns and cash flow growth over and above that of the variable $\log \delta_t$.

from macroeconomics, such as zero contemporaneous restrictions (Sims, 1980), long-run restrictions (Blanchard and Quah, 1989), or sign restrictions (Uhlig, 2005), can identify such shocks.[6] The shock identification implies an explicit representation of the innovations in log returns, $\log r_{t,t+1} - E_t \log r_{t,t+1}$, as a linear combination of (orthogonal) structural shocks, thereby characterizing which shocks contribute most to the risk in the one-period log returns.

Alternatively, the classic VAR also serves as a base model for quantifying the impact of cash flow news and discount rate news on the contemporaneous innovations in log returns via a famous decomposition of Campbell (1991)[7]

$$\underbrace{\log r_{t,t+1} - E_t \log r_{t,t+1}}_{\text{innovations in log returns}} \approx \underbrace{(E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \log g^d_{t+j,t+j+1}}_{\text{cash flow news}} - \underbrace{(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \log r_{t+j,t+j+1}}_{\text{discount rate news}} . \quad (3)$$

This decomposition does not require shock identification. Cash flow news are compounded, scaled (by powers of ρ, where ρ is a constant of loglinear approximation) innovations in the cash flow growth; discount rate news are compounded scaled innovations in the log expected returns. In practice, cash flow news and discount rate news can be correlated, in contrast to structural shocks. As a result, the conditional variance $Var_t(\log r_{t,t+1})$, as a measure of risk in log returns, is a sum of the conditional variances of cash flow news and of discount rate news and the non-zero covariance term between them. The presence of the covariance term complicates an economic interpretation of the main driver of risk in stock returns.

In the context of the classic VAR, a famous apparatus of $\mathcal{IRF}$s is a natural bridge between characterization of risk in one-period log returns in terms of structural shocks and in terms of discount rate and cash flow news. I fix the notation and recall the definition of the $\mathcal{IRF}$ before spelling out the relationship explicitly.

The $\mathcal{IRF}$ of the future one-period log return $\log r_{t+j,t+j+1}$ associated with some type of

[6]For example, innovations in the VAR can be spanned by: (i) a shock that drives the observable state variable $\log \delta_t$, and a cash flow shock that has no permanent impact on asset prices or (ii) a shock that drives the observable state variable $\log \delta_t$, and a shock that has no contemporaneous impact on asset returns. These structural shocks do not have macroeconomic interpretation.

[7]For exchange rates and bonds, the decomposition is exact.

risk ε_i measures the sensitivity of the first conditional moment of the log return distribution $E_t(\log r_{t+j,t+j+1})$ to the shock ε_{it+1},

$$\mathcal{IRF}(\log r_{t+j,t+j+1}, \varepsilon_{it+1}) = E_{t+1}(\log r_{t+j,t+j+1}|\varepsilon_{it+1}) - E_t(\log r_{t+j,t+j+1}).$$

The *cumulative* impulse response function of the τ-period log return $\log r_{t,t+\tau}$ is the sum of the $\mathcal{IRF}$s of the one-period log returns $\log r_{t+j,t+j+1}$ over $j = \overline{0, \tau - 1}$, or alternatively the sensitivity of the τ-period discount rate $E_t(\log r_{t,t+\tau})$ to the shock ε_{it+1},

$$\mathcal{IRF}(\log r_{t,t+\tau}, \varepsilon_{it+1}) = \sum_{j=0}^{\tau} \mathcal{IRF}(\log r_{t+j,t+j+1}, \varepsilon_{it+1}) = E_{t+1}(\log r_{t,t+\tau}|\varepsilon_{it+1}) - E_t(\log r_{t,t+\tau}). \tag{4}$$

Due to linearity of the VAR, the revision of a future one-period discount rate is the sum of the impulse response functions of the corresponding one-period discount rate with respect to all shocks in the economic environment:

$$\underbrace{(E_{t+1} - E_t) \log r_{t+j,t+j+1}}_{\text{revision of } E_t(\log r_{t+j,t+j+1})} = \sum_{\varepsilon_{it+1}} \underbrace{[E_{t+1}(\log r_{t+j,t+j+1}|\varepsilon_{it+1}) - E_t(\log r_{t+j,t+j+1})]}_{\mathcal{IRF} \text{ of the discount rate associated with } \varepsilon_{it+1}}.$$

A similar relationship holds for the revision of a future cash flow growth

$$\underbrace{(E_{t+1} - E_t) \log g^d_{t+j,t+j+1}}_{\text{revision of } E_t(\log g^d_{t+j,t+j+1})} = \sum_{\varepsilon_{it+1}} \underbrace{\left[E_{t+1}(\log g^d_{t+j,t+j+1}|\varepsilon_{it+1}) - E_t(\log g^d_{t+j,t+j+1})\right]}_{\mathcal{IRF} \text{ of the cash flow growth associated with } \varepsilon_{it+1}}.$$

The decomposition of Campbell (1991), formulated in expression (3), shows that the revisions of discount rates and cash flow growth are the elements of the cash flow news and discount rate news. As a result, risk in log returns can be represented as a sum of cash flow and discount rate news or alternatively as a scaled sum of $\mathcal{IRF}$s of log returns and log cash flow growth across all types of shocks arriving at each period of time. Thus, identification of structural shocks is a natural starting point for the joint analysis of risk sensitivities of log returns, multiperiod discount rates, and expected cash flow growth.

In an economic environment with a time-varying risk premium, that is, with nonnormal and heteroskedastic shocks, risk affects returns not only directly but also indirectly. Structural shocks propagate through the time-varying variance, skewness, kurtosis, and all the other higher-order moments of the log return distribution. Classic homoskedastic VARs cannot describe these dynamics, and consequently $\mathcal{IRF}$s cannot capture how higher-order shocks contribute to risk in asset returns. These limitations prompt a new approach.

My new approach builds on two important extensions of the classic paradigm of impulse response functions implied by VARs. First, I use multivariate state-space models (1)-(2) with nonnormal and heteroskedastic shocks to estimate a realistic distribution of log returns. Second, I examine how every structural shock contributes to risk in multiperiod log returns, by measuring the sensitivity of $\log E_t r_{t,t+\tau}$ to the shock for alternative τ. As the log expected return is a scaled sum of all conditional cumulants κ_{jt} of the log return distribution,

$$\log E_t r_{t,t+\tau} = \sum_{j=1}^{\infty} \kappa_{jt}(\log r_{t,t+\tau})/j! = \underbrace{E_t(\log r_{t,t+\tau})}_{\kappa_{1t}(\log r_{t,t+\tau})} + \underbrace{\underbrace{\frac{Var_t(\log r_{t,t+\tau})}{2!}}_{\kappa_{2t}(\log r_{t,t+\tau})/2!} + \frac{\kappa_{3t}(\log r_{t,t+\tau})}{3!} + \cdots}_{\text{entropy of } \log r_{t,t+\tau}},$$

the risk sensitivity of $\log E_t r_{t,t+\tau}$ captures how the shock affects *all* the cumulants of the log return distribution in one number. This is in contrast to the classic cumulative $\mathcal{IRF}$, which measures the impact of individual shocks only on the first cumulant of the log return distribution, $E_t \log r_{t,t+\tau}$, as in (4). The difference between the $\mathcal{IER}$ and $\mathcal{IRF}$ is in the impact of shocks on the entropy of the log return.

I label the risk sensitivity of log expected return $\log E_t r_{t,t+\tau}$ to an individual shock the *Incremental Expected Return* ($\mathcal{IER}$). I use the term incremental because the $\mathcal{IER}$ quantifies the incremental effect of the individual shock on expected returns, taking into account the presence of and interaction among multiple sources of risk in the economic environment. I measure the $\mathcal{IER}$ associated with a shock ε_{it+1} as a difference in log expected returns, when the economic environment does and does not experience an injection of an additional amount of risk Δ_i at time $t+1$,

$$\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{it+1}) = \log E_t(\tilde{r}_{t,t+\tau} | \tilde{\varepsilon}_{it+1} = \varepsilon_{it+1} + \Delta_i) - \log E_t(r_{t,t+\tau}). \tag{5}$$

The log return $\log \tilde{r}_{t,t+\tau}$ is subject to the sequence of the same shocks as the log return $\log r_{t,t+\tau}$

with the exception of the shock $\tilde{\varepsilon}_{it+1} = \varepsilon_{it+1} + \Delta_i$, which exceeds ε_{it+1} by a fixed amount of risk Δ_i.

In the environment with normal and homoskedastic shocks

$$\log E_t r_{t,t+\tau} \quad = \quad E_t \log r_{t,t+\tau} + c(\tau),$$

where $c(\tau)$ is a horizon-specific constant variance of log returns. As the variance is constant and all the higher-order moments are zero, the risk sensitivity of the log expected return $\log E_t r_{t,t+\tau}$ is equal to the risk sensitivity of the expected log return $E_t \log r_{t,t+\tau}$. Thus, in this case the $\mathcal{IER}$ coincides with the $\mathcal{IRF}$ for log returns. In the environment with normal *heteroskedastic* shocks, the $\mathcal{IER}$ coincides with the shock return elasticity of Borovička and Hansen (2014), but not the $\mathcal{IRF}$. The shock return elasticity measures the *marginal* sensitivity of $\log E_t r_{t,t+\tau}$ to an individual shock, and therefore by construction it accounts for nonlinear interaction across shocks.

Similar to the shock elasticity, the $\mathcal{IER}$ is related to the nonlinear impulse responses that Gallant, Rossi, and Tauchen (1993), Koop, Pesaran, and Potter (1996), and Gourieroux and Jasiak (2005) develop. These nonlinear impulse responses differ in the definition of a primitive shock, implementation, and applications.[8] My contribution is in adopting a notion of a nonlinear impulse response for asset pricing applications and in providing an explicitly tractable characterization of the sensitivities of expected multiperiod returns to both normal and *nonnormal* shocks.

To the best of my knowledge, this paper is the first to discuss how to characterize the impact of nonnormal shocks on discount rates, thereby extending the methods of dynamic value decomposition (Borovička, Hansen, and Scheinkman, 2014; Hansen, 2012) to the case of nonnormal sources of risk. This extension is useful for the asset pricing literature as it allows analysis of horizon-dependent sensitivities of expected returns to alternative sources of time-varying risk premia. The only predecessor of this extension is the paper of Borovička, Hansen, Hendricks, and Scheinkman

[8] Gallant, Rossi, and Tauchen (1993) examine the effect of the perturbation of an economic variable rather than the effect of the perturbation of a shock on the dynamic system. Koop, Pesaran, and Potter (1996) construct stochastic impulse responses. Conditional on a shock, the average impulse response of Koop, Pesaran, and Potter (1996) coincides with the impulse response of Gallant, Rossi, and Tauchen (1993). Gourieroux and Jasiak (2005) show how to construct independent and identically distributed innovations for general nonlinear processes and examine the effects of perturbations of these type of innovations on the dynamic system.

(2011), which explores the sensitivity of expected cash flows and expected returns to the risk of regime shifts. Being inspired by the approach of marginal sensitivities, I extend the notion of shock elasticity to the case of nonnormal shocks. In the example that follows shortly, I discuss why and how the $\mathcal{IER}$ and shock elasticity differ for nonnormal shocks.

In a similar way, I measure the risk sensitivity of the multiperiod expected cash flow growth. I label the corresponding metric an incremental expected dividend ($\mathcal{IED}$)

$$\mathcal{IED}(g^d_{t,t+\tau}, \varepsilon_{it+1}) \quad = \quad \log E_t(\tilde{g}^d_{t,t+\tau} | \tilde{\varepsilon}_{it+1} = \varepsilon_{it+1} + \Delta_i) - \log E_t(g^d_{t,t+\tau}). \tag{6}$$

I collect $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{it+1})$ and $\mathcal{IED}(g^d_{t,t+\tau}, \varepsilon_{it+1})$ for alternative τ and refer to them as a term structure of risk ε_{it+1} in expected returns and in expected dividends, respectively.

The level and the shape of term structures of risk are informative moments about the multiperiod risk-return tradeoff in the marketplace. They reflect (i) whether a positive shock shifts realized and expected returns and cash flows up or down, (ii) whether the corresponding impact is horizon-dependent, and (iii) which shock exerts the largest impact at different horizons. In empirical applications, term structures of risk play at least two broad important roles. First, they economically describe risk in asset returns by identifying economic shocks, which span fluctuations in discount rate news and cash flow news, drive predictability in returns and cash flow growth, and have permanent impact on asset prices and cash flows. Second, they represent new moments that can be used for evaluating economic mechanisms in macro-based asset pricing models, that is, for testing, calibrating, and estimating structural models.

Before I turn to an empirical application, I illustrate how to compute the $\mathcal{IER}$ for normal and nonnormal shocks in a simple example below.

Example. Basics on the incremental expected return

I posit that the log return follows a jump-diffusion model with stochastic variance v_t and stochastic jump intensity λ_t

$$\begin{aligned}\log r_{t,t+1} &= r + \mu_\lambda \lambda_t + \xi v_t^{1/2} \varepsilon_{rt+1} + \gamma z_{t+1}, \\ v_{t+1} &= (1-\varphi_v) + \varphi_v v_t + \sigma_v \varepsilon_{vt+1}, \\ \lambda_{t+1} &= (1-\varphi_\lambda) + \varphi_\lambda \lambda_t + \sigma_\lambda \varepsilon_{\lambda t+1} + z_{t+1},\end{aligned}$$

where jump z_{t+1} is a Poisson mixture of Gamma distributions

$$z_{t+1}|p_{t+1} \sim \Gamma(p_{t+1}, \theta), \quad \text{and} \quad p_{t+1} \sim \text{Poisson}(h_\lambda \lambda_t),$$

shocks ε_{rt+1}, ε_{vt+1} and $\varepsilon_{\lambda t+1}$ are standard normal random variables, independent of each other, over time, and of jumps z_{t+1}. The central stochastic component of jump risk p_{t+1} controls how many jumps of average size θ arrive per period of time.[9]

I compute the $\mathcal{IER}$s in two steps. In the first step, I represent the logarithm of the expected multiperiod return as the log expected value of an exponential function of the state variables at time t and shocks at time $t+1$

$$\begin{aligned}\log E_t r_{t,t+\tau} &= \log E_t(r_{t,t+1} \cdot E_{t+1}(r_{t+1,t+\tau})) \\ &= \log E_t(\exp(c_0(\tau) + c_v(\tau) v_t + c_\lambda(\tau)\lambda_t + d_r(\tau) v_t^{1/2} \varepsilon_{rt+1} + d_v(\tau)\varepsilon_{vt+1} + d_\lambda(\tau)\varepsilon_{\lambda t+1} \\ &+ d_z(\tau) z_{t+1})),\end{aligned} \tag{7}$$

[9] Strictly speaking, an autoregressive process is not a suitable choice for modeling nonnegative random variables such as the variance factor or jump intensity, as it does not eliminate the possibility of a negative realization of the variable of interest. In the context of this illustrative example, such a modeling choice is for simplicity and without loss of generality. In empirical work, I model persistent nonnegative random variables as autoregressive gamma processes (Gourieroux and Jasiak, 2006; Le, Singleton, and Dai, 2010).

where

$$
\begin{aligned}
c_0(\tau) &= r + \mathcal{A}_0(\tau-1) + \mathcal{A}_v(\tau-1)(1-\varphi_v) + \mathcal{A}_\lambda(\tau-1)(1-\varphi_\lambda), \\
c_v(\tau) &= \mathcal{A}_v(\tau-1)\varphi_v, \\
c_\lambda(\tau) &= \mu_\lambda + \mathcal{A}_\lambda(\tau-1)\varphi_\lambda, \\
d_r(\tau) &= \xi, \\
d_v(\tau) &= \mathcal{A}_v(\tau-1)\sigma_v, \\
d_\lambda(\tau) &= \mathcal{A}_\lambda(\tau-1)\sigma_\lambda, \\
d_z(\tau) &= \gamma + \mathcal{A}_\lambda(\tau-1).
\end{aligned}
$$

Horizon-specific $\mathcal{A}_0(\tau)$, $\mathcal{A}_v(\tau)$, and $\mathcal{A}_\lambda(\tau)$ define the term structure of expected returns

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

and satisfy the recursive system

$$
\begin{aligned}
\mathcal{A}_0(\tau) &= r + \mathcal{A}_0(\tau-1) + \mathcal{A}_v(\tau-1)(1-\varphi_v) + \mathcal{A}_\lambda(\tau-1)(1-\varphi_\lambda) + \mathcal{A}_v^2(\tau-1)\sigma_v^2/2 \\
&+ \mathcal{A}_\lambda^2(\tau-1)\sigma_\lambda^2/2, \\
\mathcal{A}_\lambda(\tau) &= \mu_\lambda + \theta h_\lambda(\gamma + \mathcal{A}_\lambda(\tau-1))/(1-(\gamma+\mathcal{A}_\lambda(\tau-1))\theta) + \mathcal{A}_\lambda(\tau-1)\varphi_\lambda, \qquad (8) \\
\mathcal{A}_v(\tau) &= \xi^2/2 + \mathcal{A}_v(\tau-1)\varphi_v \qquad (9)
\end{aligned}
$$

with initial conditions $\mathcal{A}_0(1) = r$, $\mathcal{A}_v(1) = \xi^2/2$, and $\mathcal{A}_\lambda(1) = \mu_\lambda + \gamma\theta h_\lambda/(1-\gamma\theta)$. This step is instrumental for separating multiperiod effects of shocks arriving at time $t+1$ from the effects of shocks arriving at any other time and is inspired by the temporal entropy decomposition of Borovička and Hansen (2014) (see Section 3).

In the second step, I compare the expected return in the underlying economic environment with the expected return in the environment that experiences an injection of an additional amount of risk next period. Specifically, I use expression (7) to measure how an additional amount of risk Δ_i, associated with a specific structural shock ε_{it+1}, shifts the trajectory of multiperiod expected returns

$$
\begin{aligned}
\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{rt+1}) &= \log E_t(\tilde{r}_{t,t+\tau}|\tilde{\varepsilon}_{rt+1} = \varepsilon_{rt+1} + \Delta_r) - \log E_t(r_{t,t+\tau}) \\
&= \log E_t(\exp(c_0(\tau) + c_v(\tau)v_t + c_\lambda(\tau)\lambda_t + d_r(\tau)v_t^{1/2}(\varepsilon_{rt+1} + \Delta_r) + d_v(\tau)\varepsilon_{vt+1} \\
&+ d_\lambda(\tau)\varepsilon_{\lambda t+1} + d_z(\tau)z_{t+1})) - \log E_t(\exp(c_0(\tau) + c_v(\tau)v_t + c_\lambda(\tau)\lambda_t \\
&+ d_r(\tau)v_t^{1/2}\varepsilon_{rt+1} + d_v(\tau)\varepsilon_{vt+1} + d_\lambda(\tau)\varepsilon_{\lambda t+1} + d_z(\tau)z_{t+1})) = d_r(\tau)v_t^{1/2}\Delta_r. \quad (10)
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{vt+1}) &= \log E_t(\tilde{r}_{t,t+\tau}|\tilde{\varepsilon}_{vt+1} = \varepsilon_{vt+1} + \Delta_v) - \log E_t(r_{t,t+\tau}) = d_v(\tau)\Delta_v, \quad (11) \\
\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{\lambda t+1}) &= \log E_t(\tilde{r}_{t,t+\tau}|\tilde{\varepsilon}_{\lambda t+1} = \varepsilon_{\lambda t+1} + \Delta_\lambda) - \log E_t(r_{t,t+\tau}) = d_\lambda(\tau)\Delta_\lambda. \quad (12)
\end{aligned}
$$

For a normal shock, that does not affect the future variance or other higher-order cumulants of the log return distribution, the $\mathcal{IER}$ coincides with the $\mathcal{IRF}$. For example,

$$
\mathcal{IRF}(r_{t,t+\tau}, \varepsilon_{r,t+1}) = E_t(\log r_{t,t+\tau}|\varepsilon_{rt+1} = \Delta_r) - E_t(\log r_{t,t+\tau}) = d_r(\tau)v_t^{1/2}\Delta_r,
$$

that is the same as $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{rt+1})$ given in expression (10).

For a normal shock that drives the variance and/or other higher-order cumulants of the log return distribution, the $\mathcal{IER}$ is distinct from the $\mathcal{IRF}$. For example,

$$
\mathcal{IRF}(r_{t,t+\tau}, \varepsilon_{v,t+1}) = E_t(\log r_{t,t+\tau}|\varepsilon_{vt+1} = \Delta_v) - E_t(\log r_{t,t+\tau}) = 0,
$$

whereas $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{vt+1}) \neq 0$ (see expression (11)). The $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{vt+1})$ captures the sensitivity of the conditional variance of returns to the variance shock over and above the risk sensitivity of the expected log return reflected in the cumulative $\mathcal{IRF}(r_{t,t+\tau}, \varepsilon_{vt+1})$. In its turn, the $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{vt+1})$ is equal to the shock elasticity of Borovička and Hansen (2014). As the shock elasticity is a marginal sensitivity to normal risk, it accounts for any type of nonlinearity in models with normal shocks. In the context of the current example, the presence of the stochastic variance v_t introduces nonlinearity in the interaction between the shock ε_{vt+1} and shocks ε_{rt+j} for $j \geq 1$.

Computation of the sensitivity of expected returns to jumps involves comparison of expected returns in the baseline environment and in the environment with some additional amount of jump risk. The prior literature has faced the challenge to define an additional amount of jump risk in the context of nonlinear impulse responses. Jump risk is defined as an interaction of two random variables: (i) one random variable controls how many jumps occur per period of time and (ii) the other variable controls the size of the jump. The core of the problem is in the question of whether extra jump risk has to be related to an increase in the probability of a jump arrival or to an increase in the average size of a jump. See Backus (2014) for a discussion on this issue.

I overcome this challenge by using an insight from Gourieroux and Jasiak (2006) and represent the jump shock z_{t+1} as a stochastic process

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

I bundle together the risks associated with the jump arrival and jump size into one random variable ε_{zt+1}. I infer the sensitivity of expected returns $\log E_t r_{t,t+\tau}$ to the crash risk z_{t+1} by measuring the sensitivity of the expected returns to the shock ε_{zt+1}. This is a legitimate and sensible exercise because the shock ε_{zt+1} fully characterizes jump uncertainty. The shocks $\varepsilon_{zt+1} \sim \mathcal{D}(0,1)$ have nonzero moments of order higher than 1.[10] As a result,

$$\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{zt+1}) = \log E_t(\tilde{r}_{t,t+\tau} | \tilde{\varepsilon}_{zt+1} = \varepsilon_{zt+1} + \Delta_z) - \log E_t(r_{t,t+\tau}) = (2h\theta^2\lambda_t)^{1/2} d_z(\tau)\Delta_z.$$

The choice of Δ_r, Δ_v, Δ_λ, and Δ_z is a matter of normalization and depends on an application of interest. They can be constant or state-dependent. Here $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{rt+1})$ and $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{zt+1})$ are state-dependent because the corresponding Δ_r and Δ_z are constant, whereas the shocks ε_{rt+1} and ε_{zt+1} enter the data-generating process for the log return with time-varying variances. For an unconditional analysis, the state-dependent $\mathcal{IER}$s can be evaluated at the long-run means of the jump intensity and the variance, that is, setting $\lambda_t = 1$ and $v_t = 1$.

In the definition of the $\mathcal{IER}$ for the jump risk, there is one step in which I ignore the asymmetry of the distribution of ε_{zt+1}: I do not take into account that symmetric shifts of a random variable ε_{zt+1} by Δ_z to the right and to the left are associated with different probabilities. To account

[10]Notation $\mathcal{D}$ stands for a probability distribution function different from the standard normal distribution.

for the asymmetry, I extend the notion of the shock elasticity of Borovička and Hansen (2014) to nonnormal shocks (see Online Appendix). As the shock elasticities are marginal sensitivities, they are associated with the effect of an infinitesimal change in the amount of risk on the variables of interest. For the purpose of my empirical exercise, the toolkit of the $\mathcal{IER}$s and $\mathcal{IED}$s can be used without loss of generality. The concept of generalized shock elasticities has advantage in a problem of dissecting multiperiod risk premia into contributions of individual shocks that arrive at different times. Such an exercise is not the focus of this paper.

In applied research, a natural object of interest is risk sensitivity of discount rates and cash flow growth of an infinite investment horizon. Understanding whether or not the cumulative effect of risk on asset prices and/or cash flows is zero leads to an identification of a transitory or permanent component, respectively, in asset prices and cash flows.[11] This is interesting on its own but also important for empirical analyses of policy shocks on asset markets.

The infinite-horizon risk sensitivity of a discount rate and cash flow growth is the limiting case of the multiperiod $\mathcal{IER}$s and $\mathcal{IED}$s, respectively. In this example, the infinite-horizon risk sensitivities of expected returns are

$$\begin{aligned}
\mathcal{IER}(r_{t,t+\infty}, \varepsilon_{rt+1}) &= d_r^* v_t^{1/2} \Delta_r, \\
\mathcal{IER}(r_{t,t+\infty}, \varepsilon_{vt+1}) &= d_v^* \Delta_v, \\
\mathcal{IER}(r_{t,t+\infty}, \varepsilon_{\lambda t+1}) &= d_\lambda^* \Delta_\lambda, \\
\mathcal{IER}(r_{t,t+\infty}, \varepsilon_{zt+1}) &= d_z^* \Delta_z,
\end{aligned}$$

where $d_r^* = \xi$, $d_v^* = \mathcal{A}_v^* \sigma_v$, $d_\lambda^* = \mathcal{A}_\lambda^* \sigma_\lambda$, $d_z^* = \gamma + \mathcal{A}_\lambda^*$, and A_v^* and A_λ^* solve the following system of two equations

$$\mathcal{A}_\lambda = \theta h_\lambda(\gamma + \mathcal{A}_\lambda)/(1 - (\gamma + \mathcal{A}_\lambda)\theta) + \mathcal{A}_\lambda \varphi_\lambda + \mu_\lambda, \tag{13}$$

$$\mathcal{A}_v = \xi^2/2 + \mathcal{A}_v \varphi_v. \tag{14}$$

The system represented in expressions (13)-(14) is a steady state counterpart of the system given by expressions (8)-(9).

[11] Here I follow the lead of Cochrane (1994) and label the shock ε_{it+1} a transitory component in asset prices or cash flows, if its impact on the discount rate or cash flow growth of an infinite horizon is zero. Otherwise, the shock is a permanent component in asset prices or cash flows.

For correct identification of transient and permanent components in asset prices and cash flows, it is imperative to account for the impact of higher-order risk on expected returns and cash flow growth. In the presence of shock heteroskedasticity, the infinite-horizon $\mathcal{IRF}$s of log returns and cash flow growth cannot identify permanent and transitory shocks in asset prices and cash flows.

For example, consider the limiting $\mathcal{IER}$ and the cumulative infinite-horizon $\mathcal{IRF}$ associated with the intensity shock $\varepsilon_{\lambda t+1}$

$$\begin{aligned}\mathcal{IER}(r_{t,t+\infty},\varepsilon_{\lambda t+1}) &= A^*_\lambda \sigma_\lambda \Delta_\lambda, \\ \mathcal{IRF}(r_{t,t+\infty},\varepsilon_{\lambda t+1}) &= (\mu_\lambda + \gamma\theta h_\lambda)\sigma_\lambda \Delta_\lambda/(1-\varphi_\lambda),\end{aligned}$$

where $\mathcal{A}^*_\lambda$ is a stable root (Hansen and Scheinkman, 2012) of the quadratic equation

$$\mathcal{A}^2_\lambda(1-\varphi_\lambda)\theta + \mathcal{A}_\lambda(\theta h_\lambda - (1-\varphi_\lambda)(1-\gamma\theta) - \mu_\lambda\theta) + \mu_\lambda(1-\gamma\theta) + \gamma\theta h_\lambda = 0.$$

The necessary condition for the infinite-horizon $\mathcal{IER}$ to be equal to zero is

$$\mu_\lambda = -\frac{\theta h_\lambda \gamma}{1-\theta\gamma},$$

whereas the necessary condition for the cumulative infinite-horizon $\mathcal{IRF}$ to be equal to zero is

$$\mu_\lambda = -\gamma\theta h_\lambda.$$

The conditions are different unless there is no contemporaneous exposure of the log return to the jump risk $\gamma = 0$. As a result, if (i) $\mu_\lambda = -\gamma\theta h_\lambda$ and (ii) a researcher uses the $\mathcal{IRF}$, then she concludes that the intensity shock is transient, whereas in fact it is permanent.

In macroeconomics, it is customary to identify aggregate shocks by alluding to the sensitivities of the infinite-horizon macroeconomic variables to shocks, also known as long-run restrictions (Blanchard and Quah, 1989; King and Watson, 1991). As standard macroeconomic models feature normal shocks, the long-run restrictions are formulated in terms of the infinite-horizon $\mathcal{IRF}$s. With an introduction of the apparatus of the $\mathcal{IER}$s and $\mathcal{IED}$s, the long-run restrictions can be used to identify economic sources of risk premia in asset prices and cash flows. This opens up an avenue for a new line of empirical work in finance.

3 Empirical term structure of risk in expected stock returns

In this section, I set up an econometric framework to examine how multiple sources of consumption risk affect the term structure of expected aggregate stock returns and cash flow growth. First, I identify candidate sources of consumption risk and study how they interact with expected returns and cash flows at alternative horizons. Second, I use implications of the economic theory to select a realistic set of shock-identifying assumptions. Armed with the realistic shock identification, I describe which economic shocks span risk in asset returns and fluctuations in discount rate news and cash flow news; determine permanent and transient shocks; and identify economic sources of return and dividend predictability.

A Setup

Motivated by the present-value identity described in Campbell and Shiller (1989), I model jointly the dynamics of the log gross stock return $\log r_{t,t+1}$, the log dividend growth $\log g^d_{t,t+1}$ and the log price-dividend ratio $\log \delta_t$ in a state-space model (1)-(2). I identify alternative macroeconomic shocks that are known as the leading sources of the level and/or time-variation in the equity risk premium in the asset pricing theory. I explore three compelling hypotheses of the sources of stock price fluctuations. The first hypothesis, "Long-Run Risk," relies on the importance of shocks that drive the long-run growth prospects of the macroeconomy. The second hypothesis, "Consumption Disasters," highlights the importance of rare large negative shocks in consumption growth, also known as disasters, along with the shocks that drive the time-varying probability of disaster arrival. The third hypothesis, "Consumption Uncertainty," emphasizes the role of alternative shocks in the variance of consumption growth: regular variance shock, jump in variance, and the shock that drives the long-run mean of the variance (variance trend shock). These different variance shocks may reflect different aspects of macroeconomic variance risk, for example, those related to business-cycle variation and tightness of financial conditions, as in Adrian and Rosenberg (2008).

Different hypotheses imply different composition of the state vector y_t in the model (1)-(2).

As all the hypotheses rely on the multiple sources of consumption risk, the log consumption growth $\log g_t^c$ is the common component in y_t. A persistent component in the expected consumption growth x_t and the conditional variance of consumption growth v_t are the other two components in y_t in the "Long-Run Risk," $y_t = (\log g_t^c,\ x_t,\ v_t)'$. The time-varying probability of consumption disasters λ_t is the only additional component in y_t in the "Consumption Disasters," $y_t = (\log g_t^c, \lambda_t)'$. Finally, the variance of consumption growth v_t and its long-run trend v_t^* complement the state vector y_t in the "Consumption Uncertainty," $y_t = (\log g_t^c,\ v_t,\ v_t^*)'$.

I translate the hypotheses into three distinct shock identification schemes and apply them to the reduced-form innovations of the model (1)-(2). I label the shock identification schemes in the same way as the hypotheses that underlie them. In the "Long-Run Risk," the vector of innovations W_{t+1} is mapped into a vector of shocks $\varepsilon_{t+1} = (\varepsilon_{dt+1},\ \varepsilon_{ct+1},\ \varepsilon_{xt+1},\ \varepsilon_{vt+1})'$, where $\varepsilon_{ct+1} \sim \mathcal{N}(0,1)$ is a direct shock in consumption growth, $\varepsilon_{xt+1} \sim \mathcal{N}(0,1)$ is a shock into the persistent component of expected consumption growth x_{t+1}, $\varepsilon_{vt+1} \sim \mathcal{D}(0,1)$ is a variance shock, and $\varepsilon_{dt+1} \sim \mathcal{N}(0,1)$ is a direct shock in the log dividend growth. The stochastic variance v_t follows the scalar autoregressive gamma process of order 1, $v_t \sim \mathcal{ARG}(1)$,

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

and therefore the shock ε_{vt+1} is nonnormal. The components of the vector ε_{t+1} are orthogonal to each other.

In the "Consumption Disasters," the vector of innovations W_{t+1} is mapped into a vector of shocks $\varepsilon_{t+1} = (\varepsilon_{dt+1},\ \varepsilon_{ct+1},\ \varepsilon_{\lambda t+1},\ \varepsilon_{zt+1}^c)'$. The shocks ε_{dt+1} and ε_{ct+1} are defined as above. The shock $\varepsilon_{zt+1}^c \sim \mathcal{D}(0,1)$ is a consumption disaster risk, and the shock $\varepsilon_{\lambda t+1} \sim \mathcal{D}(0,1)$ is a disaster intensity risk. The shock ε_{zt+1}^c reflects unpredictable variation in rare large negative jumps in consumption growth, also known as consumption disasters. Mathematically, the jump z_{t+1}^c (the negative of the consumption disaster) is a Poisson mixture of gammas: $z_{t+1}^c \sim \Gamma(j_{t+1}^c, \theta_c)$. Its central ingredient j_{t+1}^c is a Poisson random variable $j_{t+1}^c \sim \text{Poisson}(h_\lambda \lambda_t)$, which controls how many jumps of average size θ_c arrive per period of time

$$z_{t+1}^c \quad = \quad \theta_c h_\lambda \lambda_t + (2\theta_c^2 h_\lambda \lambda_t)^{1/2}\varepsilon_{zt+1}^c.$$

The (scaled) disaster probability λ_t follows the scalar autoregressive gamma process of order 1, $\lambda_t \sim \mathcal{ARG}(1)$,

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

Naturally, the shocks ε^c_{zt+1} and $\varepsilon_{\lambda t+1}$ are nonnormal. The components of the vector ε_{t+1} are independent of each other.

In the "Consumption Uncertainty," the vector of innovations W_{t+1} is mapped into a vector of shocks $\varepsilon_{t+1} = (\varepsilon_{dt+1},\ \varepsilon_{ct+1},\ \varepsilon_{vt+1},\ \varepsilon^*_{vt+1},\ \varepsilon^v_{zt+1})'$. The shocks ε_{dt+1}, ε_{ct+1}, and ε_{vt+1} have the same economic interpretation as before, while the shock $\varepsilon^*_{vt+1} \sim \mathcal{D}(0,1)$ drives the time-variation in the trend of the stochastic variance of consumption growth. The shock ε^v_{zt+1} reflects unpredictable variation in the jump z^v_{t+1}. The jump in variance z^v_{t+1} is modeled as a Poisson mixture of gammas: $z^v_{t+1}|j^v_{t+1} \sim \Gamma(j^v_{t+1}, \theta_v)$. Its central ingredient j^v_{t+1} is a Poisson random variable, $j^v_{t+1} \sim \text{Poisson}(h_v v_t)$, which controls how many jumps of average size θ_v arrive in the consumption variance per period of time

$$z^v_{t+1} \;\; = \;\; \theta_v h_v v_t + (2\theta_v^2 h_v v_t)^{1/2} \varepsilon^v_{zt+1}.$$

The (scaled) stochastic variance of consumption growth follows the scalar autoregressive gamma process of order 1 with jumps z^v_{t+1} and the time-varying long-run trend driven by v^*_t, which itself follows an autoregressive gamma process of order 1

$$\begin{aligned} v_{t+1} &= (1 - \tilde{\varphi}_v)v^*_t + (1 - \varphi_v)v + \varphi_v v_t + \sigma_v(((1 - \varphi_v)v + 2\varphi_v v_t)/2)^{1/2} \varepsilon_{vt+1} + z^v_{t+1}, \\ v^*_{t+1} &= (1 - \varphi^*_v)v^* + \varphi^*_v v^*_t + \sigma^*_v(((1 - \varphi^*_v)v^* + 2\varphi^*_v v^*_t)/2)^{1/2} \varepsilon^*_{vt+1}. \end{aligned}$$

Naturally, the shocks ε_{vt+1}, ε^*_{vt+1}, and ε^v_{zt+1} are nonnormal. The elements of the vector ε_{t+1} are independent of each other.

Every identification scheme implies that all the cumulants of the conditional distributions of log consumption growth and log stock returns are time-varying. Naturally, the sources of time-variation are identification specific. For example, according to the "Long-Run Risk," the stochastic variance of consumption growth v_t is the source of heteroskedasticity in the innovations W_{t+1}. Alternatively, the identification "Consumption Disasters" implies that the disaster intensity factor λ_t accounts for the heteroskedasticity. Finally, the "Consumption Uncertainty" implies that the two variance factors v_t and v_t^* drive time-variation in the conditional moments of the innovations W_{t+1}. The identification-specific state vector y_t spans time-variation in the first conditional moments of log consumption growth and log stock returns. As a result, the basic properties of consumption growth and stock returns are similar across the identification schemes, yet their economic underpinnings are different.

The state-space system (1)-(2) features the cross-equation restrictions implied by Campbell-Shiller's (1989) loglinearization of a gross stock return

$$\log r_{t,t+1} \quad \approx \quad \kappa_0 + \kappa_1 \log \delta_{t+1} - \log \delta_t + \log g^d_{t,t+1},$$

where κ_0 and κ_1 are constants of the loglinear approximation. As a result, the log return equation does not have any independent source of variation beyond the shocks that drive the state vector y_t and cash flow growth $\log g^d_{t-1,t}$. There are no other restrictions in the system. Appendix A explicitly formulates the state-space model (1)-(2) in the context of the aforementioned identification schemes.

Identification of structural shocks in the state-space model (1)-(2) is similar to but not exactly the same as shock identification in structural VARs in macroeconomics. The stark difference is in statistical properties of structural shocks that work as identifying restrictions. I formulate identifying restrictions in terms of the size, frequency of arrival, and persistence of structural shocks. For example, ε_{xt+1} is a shock that arrives every period, is relatively small, and originates in the persistent component of the expected consumption growth. Alternatively, z^c_{t+1} is a shock that arrives infrequently, is positive and large, and originates in the observable consumption growth. In macroeconomics, structural shocks are normal disturbances identified by contemporaneous restrictions (Sims, 1980), long-run restrictions (Blanchard and Quah, 1989), or sign restrictions (Uhlig, 2005). I borrow contemporaneous zero restrictions from macroeconomics to identify the direct consumption shock ε_{ct+1} and the direct dividend shock ε_{dt+1}; both are normal as defined above.

B Data and preliminaries

I conduct my empirical analysis on a sample of quarterly U.S. real consumption growth, real stock returns, and price-dividend ratios from the second quarter of 1947 through the fourth quarter of 2015. The National Income and Product Accounts tables of the Bureau of Economic Analysis provide consumption and price data. Real consumption is measured as per capita expenditure on nondurable goods and services deflated by the corresponding price index (PCE). The Center for Research in Security Prices provides monthly value-weighted return data with and without dividends on the value-weighted portfolio of all NYSE, Amex, and NASDAQ stocks. Quarterly dividends are constructed by aggregating monthly cash flows implied by the difference in gross returns with and without dividends. To remove seasonality in dividend payments, the measure of quarterly dividends in quarter t is the average of the dividends in quarters $t-3$, $t-2$, $t-1$, and t. The nominal returns and dividends are converted to real variables by the PCE deflator. The log growth rates of consumption and cash flows represent the consumption growth and the dividend growth, respectively. This procedure of variable construction is standard and follows Bansal, Dittmar, and Lundblad (2005) and Hansen, Heaton, and Li (2008), among others.

I estimate the joint dynamics of consumption growth, stock returns, dividend growth, and the price-dividend ratio specified in expressions (1)-(2) under three different identification schemes. I use the Bayesian MCMC methods, as they allow me to identify latent states (stochastic variance of consumption growth, intensity of consumption disasters, time-varying trend in the stochastic variance of consumption growth) and jumps (consumption disasters and jumps in the variance of consumption growth) in the data. Appendix B provides the description of the estimation output; the Online Appendix discusses the details of the estimation procedure.

Before discussing the empirical properties of identified shocks, I highlight several intrinsic features of the $\mathcal{IER}$s which are important for the future analysis. First, as the $\mathcal{IER}$s measure the effect of a shock arriving at time $t+1$ on the *multiperiod* expected returns $\log(E_t r_{t,t+\tau})$, they quantify a *cumulative* effect of risk on asset prices from $t+1$ to $t+\tau$. Second, the one-period $\mathcal{IER}$ ($\tau = 1$) measures the effect of a shock on the one-period *realized* log return.[12] Therefore,

[12] When computing the $\mathcal{IER}$s, I condition on amount of risk Δ_i, which I add to the economic environment.

the one-period $\mathcal{IER}$ signals whether a specific shock is good (positive return sensitivity) or bad (negative return sensitivity). I define the slope of the term structure of risk as the *absolute value* of the difference between the 10-year horizon $\mathcal{IER}$ and the one-quarter horizon $\mathcal{IER}$. As a result, a negative (positive) slope manifests higher risk sensitivity at shorter (longer) horizons.

The $\mathcal{IER}$s and $\mathcal{IED}$s scale up and down depending on the choice of normalization for a risk perturbation Δ_i (see formulas (5) and (6)). Normalization determines two characteristics of risk sensitivities: (i) state-dependence and (ii) magnitude. As the empirical focus of this paper is on the unconditional properties of structural shocks, I compute state-independent risk sensitivities. I add an additional amount of risk to the economic environment by perturbing a state variable but I assign the effect of this perturbation to a specific shock that feeds the state variable. I set Δ_i to be equal to the ratio of one standard deviation of the perturbed state variable to one standard deviation of the log return (log dividend growth) in the case of the $\mathcal{IER}$ ($\mathcal{IED}$). For example, for the $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{vt+1})$ in the "Long-Run Risk," $\tilde{v}_{t+1} = v_{t+1} + \Delta_v$, where $\Delta_v = [Var(v_t)/Var(\log r_{t,t+1})]^{1/2}$; whereas for the $\mathcal{IER}(r_{t,t+\tau}, \varepsilon^c_{zt+1})$ in the "Consumption Disasters," $z^c_{t+1} = \tilde{z}^c_{t+1} + \Delta^c_z$, where $\Delta^c_z = [Var(z^c_t)/Var(\log r_{t,t+1})]^{1/2}$.

C Empirical findings

I use an estimated state-space model (1)-(2) under mulitple shock identification schemes to analyze properties of economic shocks in the term structure of expected returns and dividend growth. The implied term structures of risk are empirical moments that equilibrium models featuring the same economic shocks are expected to match.

Independently of shock-identifying restrictions, the empirical term structure of expected buy-and-hold stock returns has a significant negative slope (Table I). The slope is the difference between the long-term and short-term per-period expected returns. The significant slope reveals the presence of the multivariate mean-reversion in returns (Cochrane, 2001). The negative sign indicates that shorter holding period returns exhibit higher sensitivity to at least one economic shock in each identification scheme. As I quantitatively describe the $\mathcal{IER}$s for alternative structural shocks, I can identify this shock in every identification scheme, thereby revealing the economic source of return predictability.

Different identification schemes feature alternative economic shocks, and therefore imply different sources of return predictability. A natural question is how to choose the right one, or alternatively how to distinguish across shock-identifying assumptions. The economic theory guides the choice of a realistic shock identification scheme. I follow the idea of Christiano, Eichenbaum, and Evans (1999), who select realistic identifying restrictions for macroeconomic shocks by comparing theoretical and empirical impulse response functions. I choose shock-identifying assumptions that imply similar term structures of the $\mathcal{IER}$s in the data and in equilibrium models. I rely on the $\mathcal{IER}$s rather than the $\mathcal{IRF}$s for log returns because I describe the properties of nonnormal shocks in the presence of heteroskedasticity, which $\mathcal{IRF}$s cannot quantify. The empirical properties of the term structure of risk in expected returns are informative moments for distinguishing different theories of the risk-return tradeoff in the aggregate stock market. As the data generating process for log returns is an endogenous outcome of equilibrium models, the comparison of empirical and theoretical term structures of risk corresponds to an implicit test of cross-equation restrictions.[13]

Every identification scheme is based on a set of identifying restrictions formulated in terms of the collection of state variables and shocks characterized by size, frequency of arrival, and persistence. There are families of equilibrium models that share similar state-space representations with similar underlying shocks. I characterize the *theoretical* term structures of risk associated with the economic shocks of interest through the lens of the most parsimonious equilibrium models that imply or nest the identification schemes "Long-Run Risk," "Consumption Disasters," and "Consumption Uncertainty."

The class of equilibrium models that motivates the "Long-Run Risk" identification scheme includes the models of Bansal and Yaron (2004), Bansal and Shaliastovich (2013), and Colacito and Croce (2011). I use the model of Bansal and Yaron (2004) to describe the theoretical term

[13]The term structure of risk in expected cash flow growth is less informative about the properties of competing theories of the risk-return tradeoff in asset markets. In endowment economies, a cash flow growth process is specified exogenously, and therefore, model-implied term structures of risk in expected cash flow growth does not encode interaction between preferences and macroeconomic risk.

structures of risk for the long-run risk ε_{xt+1} and the variance shock ε_{vt+1}. The equilibrium models of Du (2013), Farhi and Gabaix (2016), Gabaix (2012), Gourio (2013), Tsai (2014), and Wachter (2013) represent theoretical settings in support of the hypothesis "Consumption Disasters." I theoretically characterize the shocks featured in the "Consumption Disasters" through the lens of the model in Wachter (2013). The equilibrium models of Benzoni, Collin-Dufresne, and Goldstein (2011), Branger, Rodrigues, and Schlag (2018), Drechsler (2013), and Drechsler and Yaron (2011) have multiple sources of variance risk as in the identification "Consumption Uncertainty." I use a parsimonious version of the model in Drechsler and Yaron (2011) without the long-run risk x_t to characterize theoretical term structures of risk for the multiple sources of the variance risk.

I rely on the original calibrations of the equilibrium models (Table II) to describe quantitatively theoretical term structures of risk in expected returns and expected dividends. As I use the restricted version of the model in Drechsler and Yaron (2011), I slightly modify the original calibration to guarantee that the restricted model successfully reproduces the key macroeconomic and asset pricing moments. I compare the signs of the levels and the signs of the slopes of the theoretical and empirical term structures of risk in expected returns. These two simple metrics parsimoniously and informatively summarize the properties of gross return predictability in the models and in the data. I complete the characterization of shocks by also comparing the empirical and theoretical term structures of risk in expected dividends with respect to the same metrics of the level and the slope.

Identification "Consumption Uncertainty."

Figure 1 illustrates the term structures of risk for the multiple sources of variance risk featured in the identification "Consumption Uncertainty" in the data and in the equilibrium model. For every shock, the implications of the equilibrium model are in line with the empirical properties of the multiperiod risk sensitivities of expected returns. As a result, the identification "Consumption Uncertainty" passes the hurdle of a realistic shock identification scheme. I use this identification scheme for an economic analysis of the risk-return tradeoff in the aggregate stock market.

Figure 1 and Table III show that the variance trend shock ε^*_{vt+1} is the sole source of return predictability. Upon arrival of a positive shock ε^*_{vt+1} that increases the long-run mean of the stochastic variance of consumption growth, the realized log return decreases, whereas the expected future one-period stock returns slightly increase. As a result, the cumulative effects of the shock on the expected multiperiod returns are negative but less so than the contemporaneous effect of the shock on the one-period realized return. Hence, there is a negative and significant slope of the term structure of ε^*_{vt+1} in expected returns. The term structure of the $\mathcal{IER}$s for the regular variance shock ε_{vt+1} and the term structure of the $\mathcal{IER}$s for the jump in variance ε^v_{zt+1} have negative levels and insignificant slopes.

As the dividend process is part of the state-space model (1)-(2), I also describe the term structure of risk in expected dividends. Figure 2 illustrates that the $\mathcal{IED}$s for the regular variance shock ε_{vt+1} and for the jump in variance ε^v_{zt+1} are negative, whereas the $\mathcal{IED}$s for the variance trend shock ε^*_{vt+1} are positive for short horizons and insignificant for medium-term and long-term horizons. Table III shows that the term structures of shocks ε_{vt+1} and ε^v_{zt+1} exhibit significant positive slopes. As a result, the regular variance shock and the jump in variance are sources of dividend predictability and their individual impacts on expected dividends are monotonically increasing with an investment horizon. As it is substantially more difficult to detect dividend predictability than return predictability, the credible intervals for the slopes of the term structure of risk in expected dividends are much wider than those for the term structures of risk in expected returns.

Based on the nature of economic shocks which drive return and dividend predictability, the trend variance factor v^*_t (shock ε^*_{vt} drives return predictability) spans time-variation in expected returns, whereas the stochastic variance factor v_t (shocks ε_{vt} and ε^v_{zt} drive dividend predictability) spans time-variation in expected dividends. Therefore, as in Binsbergen and Koijen (2010), expected returns exhibit higher persistence than expected dividends: the trend variance factor v^*_t has an annual persistence equal to 0.94, whereas the variance factor v_t has an annual persistence equal to 0.53.[14] The analysis of the permanent effects of the multiple sources of risk in the variance of consumption growth on stock prices and dividends shows that the trend variance shock

[14]The 95% credible interval for the estimate of the annual persistence in expected returns is [0.9201, 0.9536], whereas the 95% credible interval for the estimate of the annual persistence in expected dividends is [0.4273, 0.6322]. Binsbergen and Koijen (2010) report the annual persistence of expected returns of 0.932, and the annual persistence of expected dividend growth of 0.354.

is a transitory component, whereas the other two variance shocks are the permanent components (Table IV). This evidence results in identifying restrictions for the source of return predictability in the aggregate stock market – zero permanent effects on stock prices and dividends – that can be utilized in empirical work.

The identification "Consumption Uncertainty," similar to other identification schemes, features the direct shock in consumption growth ε_{ct+1} and the direct shock in dividend growth ε_{dt+1}. The extant asset pricing literature suggests the limited role of these shocks in the risk-return tradeoff in the aggregate stock market. Therefore I discuss these shocks briefly and mainly for completeness. The shocks originate in the observable variables and do not drive time-variation in conditional cumulants of the log return and log dividend distributions. Thus, the term structures of these shocks in expected dividends and expected returns are flat. The levels of these term structures are set equal to the exposures of the log consumption growth and log dividend growth to the shocks.

Next, I address the main question, which shocks span risk in stock returns? I start with a one-period decomposition of risk in log returns and log dividend growth into contributions of structural shocks, as measured by average entropy (Table V). The multiple sources of risk in the variance of consumption growth, ε^v_{zt+1}, ε_{vt+1}, and ε^*_{vt+1}, contribute 94%, whereas the direct dividend shock contributes only 6% in the aggregate stock return risk. Despite the sizeable exposure of the log return to the direct dividend shock, the contribution of this shock is small. This is because the shock ε_{dt+1} does not affect the stochastic variance or other higher-order moments of the log return distribution unlike the variance shocks. The direct dividend shock ε_{dt+1} contributes 95%, whereas the trend variance shock contributes the remaining 5% in the risk of quarterly dividend growth. The direct dividend shock ε_{dt+1} and the variance trend shock ε^*_{vt+1} are the sources of the conditional covariance between the log returns and log dividend growth.

It is customary to relate risk in one-period log returns to fluctuations in cash flow news and discount rate news. Figure 3 shows the risk sensitivities of expected returns and dividend growth to different economic shocks at alternative investment horizons.[15] These sensitivities can be viewed as

[15]To facilitate comparison, I rescale the $\mathcal{IER}$s and $\mathcal{IED}$s so that they represent the sensitivities of expected returns and expected dividends to risk perturbations of the same size. I perturb each shock by Δ_i and fix state variables at their unconditional means.

the revisions of multiperiod discount rates and expected cash flow growth adopted to an economic environment with heteroskedastic nonnormal shocks. Revisions in long-term discount rates and long-term expected cash flow growth proxy for discount rate news and cash flow news. As discount rate news do not take into account the contemporaneous effect of the shock on the one-period log return, I accordingly adjust the multiperiod $\mathcal{IER}$s by subtracting the one-period $\mathcal{IER}$ from them.

Panel A of Figure 3 shows that the long-term expected returns are sensitive to the variance trend shock ε_{vt+1}. Panel B shows that the long-term expected cash flows are sensitive to the regular variance shock ε_{vt+1}, the jump in variance ε^{*}_{vt+1}, and the direct dividend growth shock ε_{dt+1}. As a result, the trend variance shock spans exclusively cash flow news, whereas the regular variance shock, jump in variance, and the direct dividend shock span exclusively discount rate news. Taken this evidence and the decomposition of risk in stock returns in Table V, cash flow news and discount rate news contribute in about equal proportions to the risk in stock returns: 53% and 47%, respectively.

To the best of my knowledge, this is the first paper that explicitly interprets economic sources of risk driving cash flow news and discount rate news, thereby complementing the influential evidence in Campbell (1991). The realistic shock-identifying restrictions suggest that the cash flow news and discount rate news should not be viewed as competing aspects of risk in asset returns as they both (i) originate in the variance of consumption growth and (ii) are at the forefront of the risk-return tradeoff in the aggregate stock market.

Identification "Long-Run Risk"

There is a clear discrepancy between theoretical and empirical term structures of risk for shocks featured under the identification "Long-Run Risk." Figure 4 and Table III illustrate that the empirical term structure of the long-run risk ε_{xt+1} has a significant negative slope, whereas the theoretical term structure has a positive slope. In the data, upon arrival of a positive shock ε_{xt+1}, which increases expected future consumption growth, the realized gross return goes up, whereas the one-period expected future returns go down but not enough to offset the contemporaneous effect of the shock on the realized return. In the model, a positive shock ε_{xt+1} positively affects the one-period realized log return and future expected returns. As a result, the sign of return predictability in the model is different from that in the data.

In the data, the variance shock ε_{vt+1}, similar to the long-run risk shock ε_{xt+1}, exhibits the term structure of risk with a significant negative slope. Upon arrival of a positive shock ε_{vt+1} that increases future variance, the realized gross return goes down, whereas the one-period expected future returns go up but not enough to offset the contemporaneous effect of the shock on the realized return. As a result, the level of the term structure of risk is negative across horizons from one quarter to ten years, and the risk sensitivities gradually approach zero. In the model, the risk sensitivities of expected returns to the variance shock changes the sign from negative to positive at horizons longer than fifteen quarters. As the term structure of ε_{vt+1} has the same shape but different level in the model and in the data, the quantitative role of the variance risk is distorted in the equilibrium model. The data suggest either a higher exposure of cash flows to the variance shock and/or a higher price of the variance risk than those in the original calibration of Bansal and Yaron (2004).

The term structure of $\mathcal{IED}$s also exhibits disagreement between the model and the data (Figure 5). First, the expected multiperiod dividends have negative sensitivities to the long-run risk shock ε_{xt+1} and the variance shock ε_{vt+1} in the data but positive risk sensitivities in the model. The slope of the term structure of $\mathcal{IED}$s for ε_{xt+1} is negative in the data, but positive in the model. The absolute values of the slopes are similar in the data and in the model, and therefore the model exhibits a realistic magnitude of dividend predictability but implies the wrong sign of predictability.

As a next step, I investigate whether the tension between the empirical and theoretical term structures of risk in expected returns are pervasive or calibration specific. The main source of tension is in the slope of the term structure of the long-run risk shock ε_{xt+1}: it is positive in the theory but negative in the data.[16] As Appendix D shows, the necessary condition for generating the negative slope in the equilibrium model is to use a calibration with a negative parameter of the intertemporal elasticity of substitution. This condition is economically implausible, and therefore

[16]The difference in the level of the term structure of ε_{vt+1} in the model and in the data can be resolved by introducing a negative exposure of the one-period dividend growth $\log g^d_{t,t+1}$ to the variance shock ε_{vt+1}. Panel B of Figure 5 supports such a mechanism.

the properties of the theoretical $\mathcal{IER}$s for ε_{xt+1} are calibration invariant. This result suggests that going forward it may be fruitful to focus on a different definition of long-run consumption risks, that is, to consider different identification restrictions for the long-run risk shocks. The cointegration-based long-run consumption risks of Bansal, Dittmar, and Kiku (2009) or common persistent innovations in the first and second moments of consumption growth of Backus, Boyarchenko, and Chernov (2018) seem like natural alternatives to explore in future research.

Identification "Consumption Disasters"

Figures 6 and 7 illustrate the theoretical and empirical properties of the $\mathcal{IER}$s and $\mathcal{IED}$s for the shocks featured under the identification "Consumption Disasters." Upon arrival of a positive disaster intensity shock $\varepsilon_{\lambda t+1}$, which increases the probability of a disaster, the realized log return goes down, whereas the expected future one-period returns go slightly up. Taken together, the risk sensitivities of the multiperiod expected returns are negative and decreasing in absolute value with an investment horizon. The theoretical properties of the term structures of $\varepsilon_{\lambda t+1}$ in expected returns are broadly in line with the data, with some quantitative differences at short horizons. Table III shows that the slope of the $\mathcal{IER}$s for the disaster intensity shock is negative and significant, and therefore the disaster intensity shock is the source of return predictability.

In the data, upon arrival of a consumption disaster, the realized return and realized dividend growth slightly increase, whereas future one-period expected returns and dividends do not change. As a result, expected returns and dividend growth of any investment horizon have similar risk sensitivities to ε^c_{zt+1}. Instead, the equilibrium model posits that the consumption disaster coincides with a bigger negative shock in dividend growth, thereby implying that the theoretical $\mathcal{IER}$s and $\mathcal{IED}$s for ε^c_{zt+1} are negative and sizable.

The main source of disagreement between the theoretical and empirical risk sensitivities of expected returns to the shocks featured in the "Consumption disasters" is in the level of the term structure of disaster risk ε^c_{zt+1}.[17] As a next step, I investigate whether it is possible to overcome the tension by recalibrating the equilibrium model. Appendix D provides the necessary condition

[17]The quantitative differences in the term structure of $\varepsilon_{\lambda t+1}$ in the model and in the data can be resolved by recalibrating the equilibrium model to exhibit a higher price of the disaster intensity risk.

to generate positive theoretical $\mathcal{IER}$s. The condition implies a negative leverage parameter that relates the dividend claim to the consumption claim. This necessary condition is economically implausible, and therefore the tension between the term structures of disaster risk ε^{c}_{zt+1} in the data and in the model is pervasive.

Given that the negative skewness of consumption growth and stock returns is a robust feature of the data, equilibrium models with jump risk in consumption growth and stock returns hold promise. The facts documented in this paper suggest that future models of consumption disaster risk should feature less than perfectly correlated arrivals of consumption disasters and return crashes. Muir (2017) discusses several interesting theoretical alternatives that can meet this requirement.

D Macroeconomic implications

Empirical properties of return predictability examined within the present-value model imply that the identification scheme "Consumption Uncertainty" features realistic shock-identifying assumptions. The equilibrium model is consistent with the stylized facts which describe the empirical term structures of risk in expected dividends and returns. Naturally, there could be alternative shock-identifying restrictions, beyond those considered in this paper, which can be classified as realistic. I do not pursue an exhaustive analysis of different identifying assumptions, but instead examine macroeconomic implications of the shocks featured in the identification "Consumption Uncertainty."

Under the identification "Consumption Uncertainty," there is a negative contemporaneous correlation between the expected consumption growth and the conditional variance of consumption growth, implied by a negative parameter B_{34} in Table BIII. Independently Backus, Chernov, and Zin (2014) and Bryzgalova and Julliard (2018) advocate for the same feature in the data. Backus, Chernov, and Zin (2014) analyze multiperiod properties of alternative models of the real pricing kernel and conclude that one solution to generate a realistic slope of the yield curve is to specify the interaction between the conditional mean and variance of consumption growth. Bryzgalova and Julliard (2018) identify a slow reaction of consumption to the common innovations in bonds and stock returns and conclude that the corresponding equilibrium model with stochastic volatility should include a contemporaneous leverage effect, that is a nonzero correlation between the conditional mean and conditional variance of log consumption growth.

Under identification "Consumption Uncertainty," the regular variance shock ε_{vt+1} and jump in variance ε^{v}_{zt+1} are the sources of correlation between the first and second moments of the log consumption growth. These shocks are the prototype of the asset pricing innovation of Bryzgalova and Julliard (2018), to which consumption growth responds slowly. Bryzgalova and Julliard (2018) document the following properties of their asset pricing innovation: (i) consumption reacts to the innovation over the period of two to four years; and (ii) the innovation accounts for about 27% of the time series variation of the consumption process and for about 79% of time series variation in stock returns. Table V shows that the shocks ε_{vt+1} and ε^{v}_{zt+1} explain about 10.65% of the average entropy of the one-period log consumption growth and 46.59% of the average entropy of the one-period log returns. Figure 8 illustrates that consumption growth responds to these shocks slowly with a peak reached at the horizon of 5 years. As a result, the basic properties of the dynamic interaction between asset returns and consumption growth in my empirical analysis and that of Bryzgalova and Julliard (2018) are consistent with each other. This is not a trivial result, because Bryzgalova and Julliard (2018) identify the asset pricing innovation from the cross section of bonds and stock returns, whereas I identify the multiple sources of variance risk from the macro-based present-value model for the aggregate stock market. Differences in empirical settings drive some quantitative differences in empirical results, but the common spirit of the findings is striking.

The economic environment under the identification scheme "Consumption Uncertainty" also features the contemporaneous negative correlation between consumption growth and the market variance. Parameters h_{33}, h_{34}, and h_{35} in Table BIII indicate that the contemporaneous correlation between consumption growth and its variance is negative, whereas the market variance itself is a function of the variance factor of consumption growth v_t and the long-run variance factor v^*_t. The negative correlation generates a more pronounced negative skewness of the consumption growth distribution when the market variance is high, which by itself is a stylized fact according to, for example, Figure 2 in Bekaert and Engstrom (2017).

The contemporaneous correlation between consumption growth and asset returns is small because the sources of variance risk that are priced in the aggregate stock market work similar to the long-run risk of Bansal and Yaron (2004). These shocks affect consumption growth contemporaneously but the main pricing affect comes through the continuation utility. In this respect, my empirical findings suggest a refinement of the definition of the long-run risk. Perhaps, future research should model the persistent long-lived shocks in consumption growth as common innovations in the first and higher-order moments of the log consumption growth distribution.

4 Discussion

There is a wide scope for applications of the empirical approach presented in this paper. First, the current setting can be extended to include alternative aggregate shocks, for example, shocks that drive the surplus-consumption ratio as in Campbell and Cochrane (1999) or shocks that pertain to the "bad environment-good environment" process for the consumption growth of Bekaert and Engstrom (2017), etc. Here I choose the hypotheses "Long-Run Risk," "Consumption Disasters," and "Consumption Uncertainty" for two reasons. First, they provide excellent examples of shocks that generate similar properties of observable macroeconomic and financial variables, but represent fundamentally different risk channels. Second, as the term structures of risk are tractable, these cases are illustrative and pedagogical.

Furthermore, the proposed measurement procedure for the fundamental sources of risk can be successfully used in other asset markets, such as the fixed income market or the foreign exchange market. The empirical analyses of Froot and Ramadorai (2005) and Campbell and Ammer (1993), which use the present-value identities for currencies and bonds, respectively, can serve as a starting point for quantifying the term structures of risk in expected foreign exchange returns and expected bond returns. The dynamics of cross sectional risk premia are also of interest and can be analyzed by measuring the term structures of risk in the cross sections of returns and cash flows. Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2008) and Zviadadze (2017) are examples that follow this route.

The empirical term structures of risk summarize properties of the data through the lens of a flexible empirical model. A natural question is whether the empirical results are robust with respect to the choice of the state vector and observable variables spanned by the elements of the state vector. As the ultimate goal is to make progress towards understanding *economically interpretable* channels driving the fluctuations in asset prices and cash flows, the economic theory guides the choice. A realistic empirical identification scheme would have implications consistent with the implications of the underlying equilibrium model. Only the number of alternative risk mechanisms advocated in the theory limit the number of possible identification schemes.

A related question is, what is the best unit of observation in the analysis of risk sensitivities across different horizons, a gross return or an excess return? My first choice is to examine the risk sensitivities of expected *gross* returns. The data exhibit similar properties of gross and excess return predictability (see, for example, Cochrane, 2008). In equilibrium models, a different number of state variables may drive expected gross returns and expected excess returns. For example, in the model of Bansal and Yaron (2004) the long-run risk x_t and consumption variance v_t drive the expected gross returns, whereas v_t alone drives the expected excess returns. As a result, the equilibrium models that match excess return predictability (a standard moment to target in calibration) may imply a wrong sign of the gross return predictability. Such a shortcoming goes undetected, if the empirical analysis focuses on the term structure of risk in excess returns.

I characterize theoretical term structures of risk through the lens of calibrated equilibrium models. One could question robustness of my empirical results with respect to choices I have to make: (i) the decision interval of a representative agent, (ii) the use of loglinear approximations to solve the equilibrium models of interest, or (iii) the choice of a predictable variable (for example, price-dividend ratio versus price-earnings ratio). But for a simple and intuitive reason, it is unlikely any of these concerns are critical: I rely on very basic metrics that describe the shocks in the data and in equilibrium models: *signs* of the slope and of the level of the term structures of risk. The level indicates whether the shock impacts returns (cash flows) positively or negatively upon its arrival; the slope indicates whether returns (cash flows) with shorter holding periods are more or less sensitive to risk. These implications are robust to the aforementioned choices.

I use empirical properties of return predictability to obtain stylized facts distinguishing alternative identification schemes. Multiperiod properties of return predictability naturally encode a multiperiod risk-return tradeoff in financial markets, that is a multiperiod interaction of risk exposures and associated with them compensations. A recent literature on the term structure of risk premia in asset markets (see Bansal, Miller, and Yaron, 2017; Binsbergen, Brandt, and Koijen, 2012; Binsbergen, Hueskes, Koijen, and Vrugt, 2013; Dahlquist and Hasseltoft, 2013; Dew-Becker, Giglio, Le, and Rodriguez, 2017; Giglio, Maggiori, and Stroebel, 2015; and Gormsen, 2018, among others) use cross section of zero-coupon assets with different maturities to shed light on how risk in different assets is priced across alternative investment horizons. As I do not use a cross section of zero-coupon assets in my estimation, I view my empirical methodology as a complementary approach to this literature. Specifically, I provide an explicit empirical characterization of economically-interpretable shocks in the term structure of expected returns and cash flow growth. This evidence is important for characterizing the risk-return tradeoff in asset markets and for guiding future progress in equilibrium modeling.

As Appendix E shows, the empirical objects, which I measure, do not have direct implications for the term structure of zero-coupon assets, unless I impose assumptions about risk preferences. Introduction of risk preferences implies cross-equation restrictions on the parameters of the state-space model (1)-(2), and therefore the corresponding estimation results lose interpretation of stylized facts. As the latter is an undesirable feature of my empirical analysis, I interpret documented in this paper term structures of risk as complementary shock-based evidence on the multiperiod risk-return tradeoff in the aggregate stock market.

5 Conclusion

This paper develops a measurement procedure for the fundamental sources of risk in the term structure of expected returns and/or expected cash flow growth. I identify alternative structural shocks in the macro-based state-space model that describes the joint evolution of asset returns, cash flows, and economic states. The observation equation that maps the predictive variable implied by the present-value identity into economic states ties together the ingredients of the state-space model and incorporates shock-identifying restrictions. I introduce two metrics, labeled the $\mathcal{IER}$ and $\mathcal{IED}$, in order to quantify an incremental effect of a structural shock on future τ-period buy-and-hold expected returns and cash flow growth. These metrics are applicable to normal and nonnormal shocks that may exhibit heteroskedasticity, and therefore are suitable for describing empirical properties of alternative sources of time-varying risk premia.

As an application, I examine the term structure of aggregate risk in expected stock returns and dividends. The economic theory guides my choice of a realistic identification scheme. The analysis of the term structures of the $\mathcal{IER}$s and $\mathcal{IED}$s suggests that (i) the multiple sources of risk in the variance of consumption growth span 94% of risk in the one-period stock returns, (ii) the variance trend shock spans discount rate news, (iii) the jump in variance, the regular variance shock, and the direct dividend shock span cash flow news. The quantitative results have implications for a longstanding question about the relative importance of cash flow news and discount rate news in the risk of asset returns. Both are at the forefront of the risk-return tradeoff in the aggregate stock market: cash flow news account for 53%, whereas discount rate news account for 47% of the aggregate stock market risk. I leave it for future research to explore microfoundations of the multiple sources of risk in the variance of consumption growth.

References

Adrian, Tobias, and Joshua Rosenberg, 2008, Stock returns and volatility: Pricing the short-run and long-run components of market risk, *Journal of Finance* 63, 2997–3030.

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

———, Nina Boyarchenko, and Mikhail Chernov, 2018, Term structures of asset prices and returns, *Journal of Financial Economics* 129, 1–23.

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

Bansal, Ravi, Robert F. Dittmar, and Dana Kiku, 2009, Cointegration and consumption risks in asset returns, *Review of Financial Studies* 22, 1343–1375.

Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad, 2005, Consumption, dividends, and the cross-section of equity returns, *Journal of Finance* 60, 1639–1672.

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.

Bekaert, Geert, and Eric Engstrom, 2017, Asset return dynamics under habits and bad environment-good environment fundamentals, *Journal of Political Economy* 125, 713–760.

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

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

Binsbergen, Jules H. van, and Ralph S. J. Koijen, 2010, Predictive regressions: A present-value approach, *Journal of Finance* 65, 1439–1471.

——— , 2017, The term structure of returns: facts and theory, *Journal of Financial Economics* 124, 1–21.

Blanchard, Olivier Jean, and Danny Quah, 1989, The dynamic effects of aggregate demand and supply disturbances, *American Economic Review* 79, 655–673.

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

——— , 2016, Term structure of uncertainty in the macroeconomy, in *Handbook of Macroeconomics, vol. 2B* (Elsevier: Amsterdam).

——— , 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.

Branger, Nicole, Paulo Rodrigues, and Christian Schlag, 2018, Level and slope of volatility smiles in long-run risk models, *Journal of Economic Dynamics and Control* 86, 95–122.

Bryzgalova, Svetlana, and Christian Julliard, 2018, Consumption and asset returns, manuscript, June.

Campbell, John Y., 1991, A variance decomposition for stock returns, *The Economic Journal* 101, 157–179.

——— , and John Ammer, 1993, What moves the stock and bond markets? A variance decomposition for long-term asset returns, *Journal of Finance* 48, 3–37.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Campbell, John Y., and Robert J. Shiller, 1989, The dividend-price ratio and expectations of future dividends and discount rates, *Review of Financial Studies* 1, 195–228.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, 1999, Monetary policy shocks: What have we learned and to what end?, in J.B. Taylor, and M. Woodford, ed.: *Handobook of Macroeconomics* (Elsevier Science: Amsterdam).

——— , 2005, Nominal rigidities and the dynamic effects of a shock to a monetary policy, *Journal of Political Economy* 113, 1–45.

Cochrane, John H., 1994, Permanent and transitory components of GNP and stock prices, *Quarterly Journal of Economics* 109, 241–265.

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

——— , 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, 153–181.

Dahlquist, Magnus, and Henrik Hasseltoft, 2013, International bond risk premia, *Journal of International Economics* 90, 17–32.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2017, The price of variance risk, *Journal of Financial Economics* 123, 225–250.

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.

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

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).

Froot, Kenneth A., and Tarun Ramadorai, 2005, Currency returns, intrinsic value, and institutional-investor flows, *Journal of Finance* 60, 1535–1566.

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, Peter E. Rossi, and George Tauchen, 1993, Nonlinear dynamic srtructures, *Econometrica* 61, 871–907.

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* (Emerald Group Publishing Limited).

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel, 2015, Very long run discount rates, *Quarterly Journal of Economics* 130, 1–53.

Gormsen, Niels, 2018, Time variation of the equity term structure, manuscript, December.

Gourieroux, Christian, and Joann Jasiak, 2005, Nonlinear innovations and impulse responses with applications to VaR sensitivity, *Annales d'Economie et de Statistique* 78, 1–31.

———, 2006, Autoregressive gamma processes, *Journal of Forecasting* 25, 129–152.

Gourinchas, Pierre-Olivier, and Helene Rey, 2007, International financial adjustments, *Journal of Political Economy* 115, 665–703.

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.

———, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long-run risk, *Journal of Political Economy* 116, 260–302.

Hansen, Lars P., and José A. Scheinkman, 2012, Recursive utility in a Markov environment with stochastic growth, *Proceedings of the National Academy of Sciences* 109, 11967–11972.

King, Robert G., Charles I. Plosser James H. Stock, and Mark W. Watson, 1991, Stochastic trends and economic fluctuations, *American Economic Review* 81, 819–840.

Koijen, Ralph S. J., and Stijn Van Nieuwerburgh, 2011, Predictability of returns and cash flows, *Annual Review of Financial Economics* 3, 467–491.

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter, 1996, Impulse response analysis in nonlinear multivariate models, *Journal of Econometrics* 74, 119–147.

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.

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

Muir, Tyler, 2017, Financial crises and risk premia, *Quarterly Journal of Economics* 125, 765–809.

Piatti, Ilaria, and Fabio Trojani, 2017, Predictable risks and predictive regression in present-value models, manuscript, July.

Rytchkov, Oleg, 2012, Filtering out expected dividends and expected returns, *Quarterly Journal of Finance* 2, 1250012.

Sims, Christopher A., 1980, Macroeconomics and reality, *Econometrica* 48, 1–48.

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

Uhlig, Harald, 2005, What are the effects of monetary policy on output? Results from an agnostic identification procedure, *Journal of Monetary Economics* 52, 381–419.

Wachter, Jessica, 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* 68, 987–1035.

Zviadadze, Irina, 2017, Term structure of consumption risk premia in the cross-section of currency returns, *Journal of Finance* 72, 1529–1566.

Table I
Slope of the term structure of expected returns

The slopes of the empirical term structures of expected returns in the identification schemes "Long-Run Risk," "Consumption Disasters," and "Consumption Uncertainty." The slope is the difference between the per-period log expected return of the holding period of 10 years and the log expected return of the holding period of 1 quarter, $\log(E_t r_{t,t+40})/40 - \log E_t r_{t,t+1}$. Quarterly. In percent.

	Median	Credible Interval, 95%
"Long-Run Risk"	-0.21	[-0.23, -0.19]
"Consumption Disasters"	-0.09	[-0.13, -0.07]
"Consumption Uncertainty"	-0.19	[-0.29, -0.05]

Table II
Calibrations of theoretical models and data

A model in spirit of Bansal and Yaron (2004):

$$
\begin{aligned}
\log g^c_{t+1} &= g^c + x_t + \gamma_c v_t^{1/2} \varepsilon_{ct+1}, \\
x_{t+1} &= \varphi_x x_t + \gamma_x v_t^{1/2} \varepsilon_{xt+1}, \\
v_{t+1} &= (1-\varphi_v) + \varphi_v v_t + \sigma_v ((1-\varphi_v + 2\varphi_v v_t)/2)^{1/2} \varepsilon_{vt+1}, \\
\log g^d_{t+1} &= g^d + \mu_x x_t + \gamma_d v_t^{1/2} \varepsilon_{dt+1}.
\end{aligned}
$$

The calibration inputs are $g^c = 0.0045$, $\gamma_c = 0.0135$, $\varphi_x = 0.9383$, $\gamma_x = 0.0010$, $\varphi_v = 0.9615$, $\sigma_v = 0.0653$, $g^d = 0.0045$, $\mu_x = 6$, $\gamma_d = 0.0653$, $\alpha = -9$, $\rho = 1/3$, $\beta = 0.998$, $\kappa_0 = 0.0449$, $\kappa_1 = 0.9923$.
A model in spirit of Wachter (2013):

$$
\begin{aligned}
\log g^c_{t+1} &= g^c + \gamma_c \varepsilon_{ct+1} - z^c_{t+1}, \\
j^c_{t+1}|_{\lambda_t} &\sim \text{Poisson}(h_\lambda \lambda_t), \quad z^c_{t+1}|_{j^c_{t+1}} \sim \text{Gamma}(j^c_{t+1}, \theta_c) \\
\lambda_{t+1} &= (1-\varphi_\lambda) + \varphi_\lambda \lambda_t + \sigma_\lambda [((1-\varphi_\lambda) + 2\varphi_\lambda \lambda_t)/2]^{1/2} \varepsilon_{\lambda t+1}, \\
\log g^d_{t+1} &= g^d + \varphi_d \gamma_c \varepsilon_{ct+1} + \gamma_d \varepsilon_{dt+1} - \varphi_d z^c_{t+1}.
\end{aligned}
$$

The calibration inputs are $g^c = 0.063$, $\gamma_c = 0.01$, $h_\lambda = 0.0075$, $\varphi_\lambda = 0.9802$, $\sigma_\lambda = 0.1743$, $\varphi_d = 2.6$, $g^d = 0.0163$, $\gamma_d = 0$, $\theta_c = 0.2$, $\alpha = -2$, $\rho = 0$, $\beta = 0.997$, $\kappa_0 = 0.0449$, $\kappa_1 = 0.9923$.
A model in spirit of Drechsler and Yaron (2011):

$$
\begin{aligned}
\log g^c_{t+1} &= g^c + \gamma_c v_t^{1/2} \varepsilon_{ct+1}, \\
v_{t+1} &= (1-\tilde{\varphi}_v) v^*_t + (1-\varphi_v) v + \varphi_v v_t + \sigma_v (((1-\varphi_v) v + 2\varphi_v v_t)/2)^{1/2} \varepsilon_{vt+1} + z^v_{t+1}, \\
v^*_{t+1} &= (1-\varphi^*_v) v^* + \varphi^*_v v^*_t + \sigma^*_v (((1-\varphi^*_v) v^* + 2\varphi^*_v v^*_v)/2)^{1/2} \varepsilon^*_{vt+1}, \\
j^v_{t+1}|_{v_t} &\sim \text{Poisson}(h_v v_t), \quad z^v_{t+1}|_{j^v_{t+1}} \sim \text{Gamma}(j^v_{t+1}, \theta_v) \\
\log g^d_{t+1} &= g^d + \mu_v v_t + \mu^*_v v^*_t + \gamma^*_{\varepsilon v} v^*_{t+1} + \gamma_{\varepsilon d} v_t^{1/2} \varepsilon_{dt+1} + \gamma_{zd} z^v_{t+1}.
\end{aligned}
$$

The calibration inputs are $g^c = 0.0045$, $\gamma_c = 0.0108$, $\varphi^*_v = 0.995$, $\sigma^*_v = 0.1025$, $v^* = 1/2$, $\varphi_v = 0.955$, $\theta_v = 0.8$, $h_v = 0.025$, $\tilde{\varphi}_v = \varphi_v + \theta_v h_v$, $\sigma_v = 0.18$, $v = (1-\varphi_v - \theta_v h_v)/(v \cdot (1-\varphi_v))$, $g^d = 0.0185$, $\mu_v = -0.0169$, $\mu^*_v = -0.0013$, $\gamma^*_{\varepsilon v} = 0.01$, $\gamma_{\varepsilon d} = 0.025$, $\gamma_{zd} = -0.2$, $\alpha = -9$, $\beta = 0.9985$, $\rho = 1/3$, $\kappa_0 = 0.0449$, $\kappa_1 = 0.9923$. In all models, I model stochastic variance factors and jump intensity factors as autoregressive gamma processes instead of autoregressive processes, as in the original papers. Such a modification does not change the implications of the models but guarantees that stochastic variances and jump intensities are well-defined and never reach negative values. Quarterly. In percent.

	BY	W	DY	Data
Log equity premium	1.1785	1.0356	0.9778	1.4971
Std dev of equity return	10.1507	7.6654	16.5649	8.1712
Mean consumption growth	0.4500	0.4750	0.4500	0.4750
Std dev of consumption growth	1.3821	2.6490	1.0808	0.5067
Mean dividend growth	0.4500	1.2350	0.6000	0.6821
Std dev of dividend growth	6.7601	9.3807	4.6530	2.3355

Table III
Slope of the term structure of risk in expected returns and expected dividends

Panel A presents the slopes of the term structures of risk in the identification "Long-Run Risk;" Panel B presents the slopes of the term structures of risk in the identification "Consumption Disasters;" Panel C presents the slopes of the term structures of risk in the identification "Consumption Uncertainty." The slopes are measured as the difference in the absolute values of the risk sensitivities at the horizon of 40 quarters and 1 quarter: the slope in $\mathcal{IER}$s is $|\mathcal{IER}(r_{t,t+40}, \text{shock}_{t+1})| - |\mathcal{IER}(r_{t,t+1}, \text{shock}_{t+1})|$, the slope in $\mathcal{IED}$s is $|\mathcal{IED}(g^d_{t,t+40}, \text{shock}_{t+1})| - |\mathcal{IED}(g^d_{t,t+1}, \text{shock}_{t+1})|$. The median and 95% credible interval.

	Slope in $\mathcal{IER}s$		*Slope in $\mathcal{IED}s$*	
	Median	Credible Interval, 95%	Median	Credible Interval, 95%
	Panel A. The "Long-Run Risk"			
ε_{xt+1}	-3.23	[-3.53, -2.90]	-0.89	[-1.21, -0.60]
ε_{vt+1}	-0.94	[-1.55, -0.34]	0.84	[-0.24, 2.08]
	Panel B. The "Consumption Disasters"			
$\varepsilon_{\lambda t+1}$	-2.78	[-3.12, -2.49,]	-0.65	[-2.69, 0.75]
ε^c_{zt+1}	0.00	[0.00, 0.00]	0.00	[0.00, 0.00]
	Panel C. The "Consumption Uncertainty"			
ε_{vt+1}	-0.23	[-0.89, 1.40]	4.04	[0.84, 11.70]
ε^*_{vt+1}	-3.35	[-4.58, -2.17]	-3.29	[-6.22, 1.36]
ε^v_{zt+1}	-0.08	[-0.33, 0.40]	1.37	[0.37, 3.27]

Table IV
Permanent impact of shocks of stock prices and dividends

Panel A presents the permanent impacts of structural shocks on stock prices and dividends in the identification "Long-Run Risk;" Panel B presents the permanent impacts of structural shocks on stock prices and dividends in the identification "Consumption Disasters;" Panel C presents the permanent impacts of structural shocks on stock prices and dividends in the identification "Consumption Uncertainty." The median and 95% credible interval.

	Impact on stock prices		*Impact on dividends*	
	Median	Credible Interval, 95%	Median	Credible Interval, 95%
		Panel A. The "Long-Run Risk"		
ε_{xt+1}	-0.91	[-1.21, -0.50]	0.22	[-0.07, 1.11]
ε_{vt+1}	0.08	[-0.58, 0.69]	-2.40	[-4.19, -0.48]
		Panel B. The "Consumption Disasters"		
$\varepsilon_{\lambda t+1}$	2.86	[1.25, 3.94]	-2.26	[-4.46, -0.45]
ε^{c}_{zt+1}	0.04	[0.01, 0.07]	0.13	[0.05, 0.24]
		Panel C. The "Consumption Uncertainty"		
ε_{vt+1}	-1.57	[-3.99, -0.51]	-5.90	[-14.70, -2.04]
ε^{*}_{vt+1}	2.38	[-0.37, 7.18]	-4.26	[-28.08, 6.77]
ε^{v}_{zt+1}	-0.54	[-1.12, -0.21]	-2.03	[-4.13, -0.78]

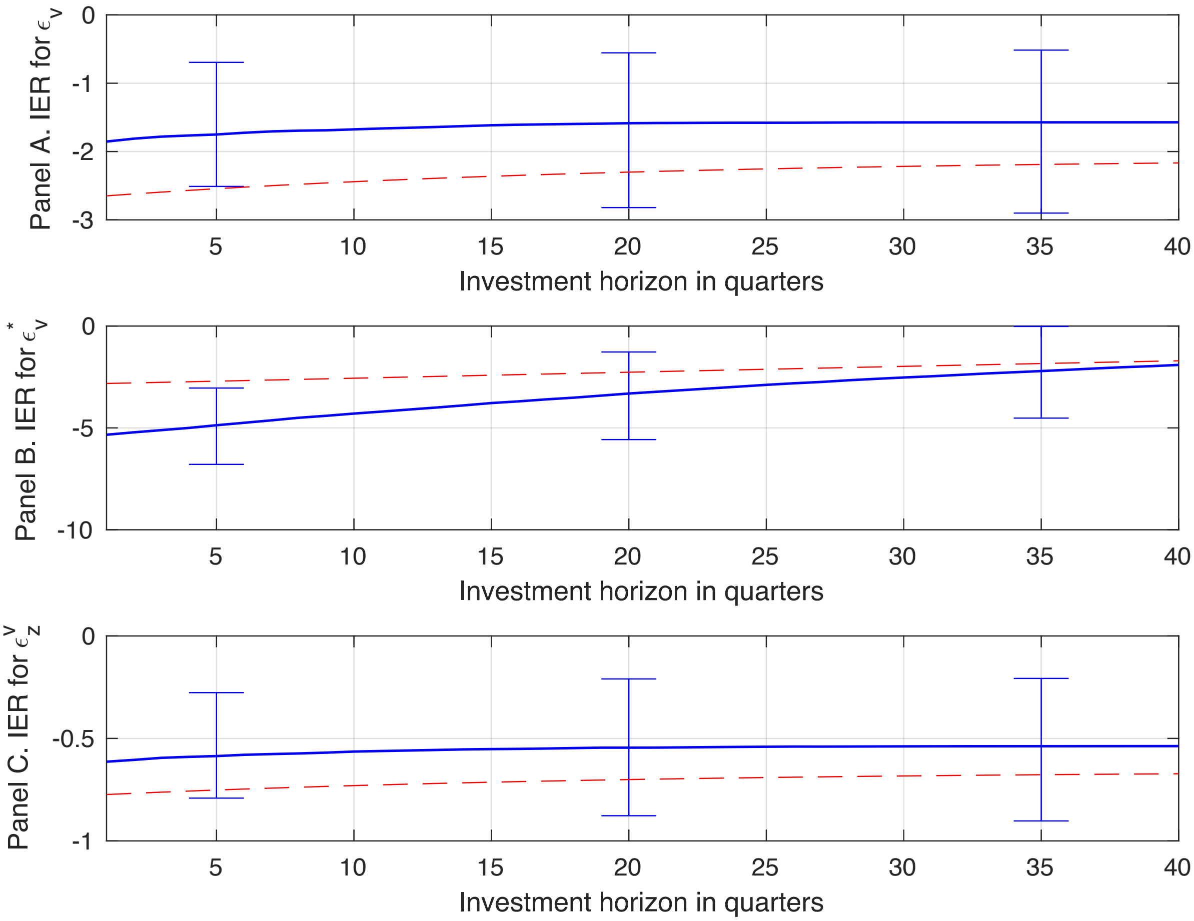
Table V
Entropy decomposition

Decomposition of average entropy of the one-quarter log returns, dividend growth, and consumption growth into contributions of multiple sources of risk in v_{t+1} (ε_{vt+1} and ε^{v}_{zt+1}), variance trend shock (ε^{*}_{vt+1}), direct consumption risk (ε_{ct+1}), and direct dividend risk (ε_{dt+1}). In percent.

	Entropy of $\log r_{t,t+1}$	Entropy of $\log g^{d}_{t,t+1}$	Entropy of $\log g^{c}_{t+1}$
ε_{vt+1} and ε^{v}_{zt+1}	46.59	0.32	10.65
	[18.70, 60.09]	[0.02, 1.87]	[3.62, 22.39]
ε^{*}_{vt+1}	47.37	4.74	2.91
	[34.28, 73.94]	[1.95, 8.60]	[0.78, 5.94]
ε_{ct+1}	0.01	0.12	86.40
	[0.00, 0.06]	[0.00, 0.85]	[76.48, 91.79]
ε_{dt+1}	6.19	94.60	0.00
	[4.60, 8.12]	[90.41, 97.40]	[0.00, 0.00]

Figure 1

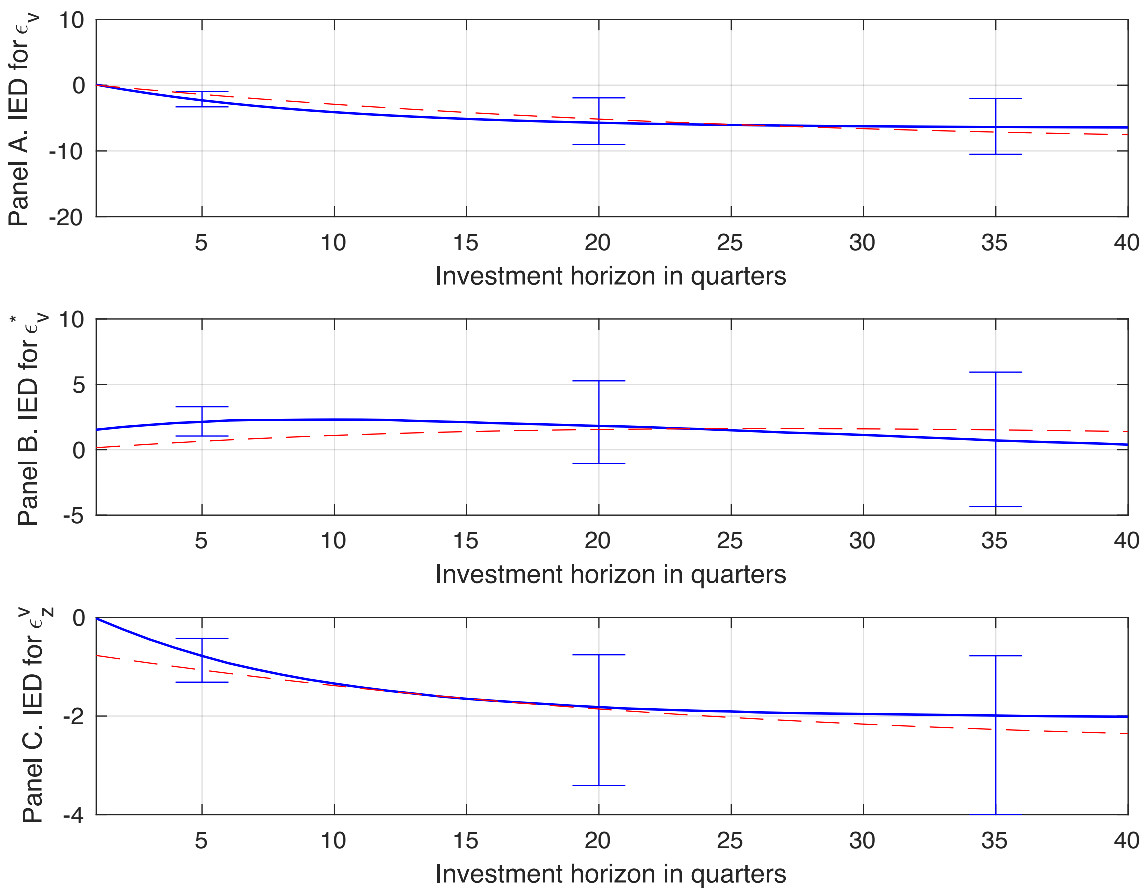
Term structure of risk in expected stock returns. Identification "Consumption Uncertainty."



The blue solid lines correspond to the empirical term structures of risk in expected returns; vertical bars indicate 95% credible intervals. See Table III for statistical significance of the slopes of the term structures of risk. The red dashed lines correspond to the theoretical term structures of risk implied by the parsimonious version of the model of Drechsler and Yaron (2011). Panel A illustrates the term structures of risk in expected returns for the regular variance shock ε_{vt+1}. Panel B illustrates the term structures of risk in expected returns for the variance trend shock ε^*_{vt+1}. Panel C illustrates the term structures of risk in expected returns for the jump in variance ε^v_{zt+1}. Quarterly.

Figure 2

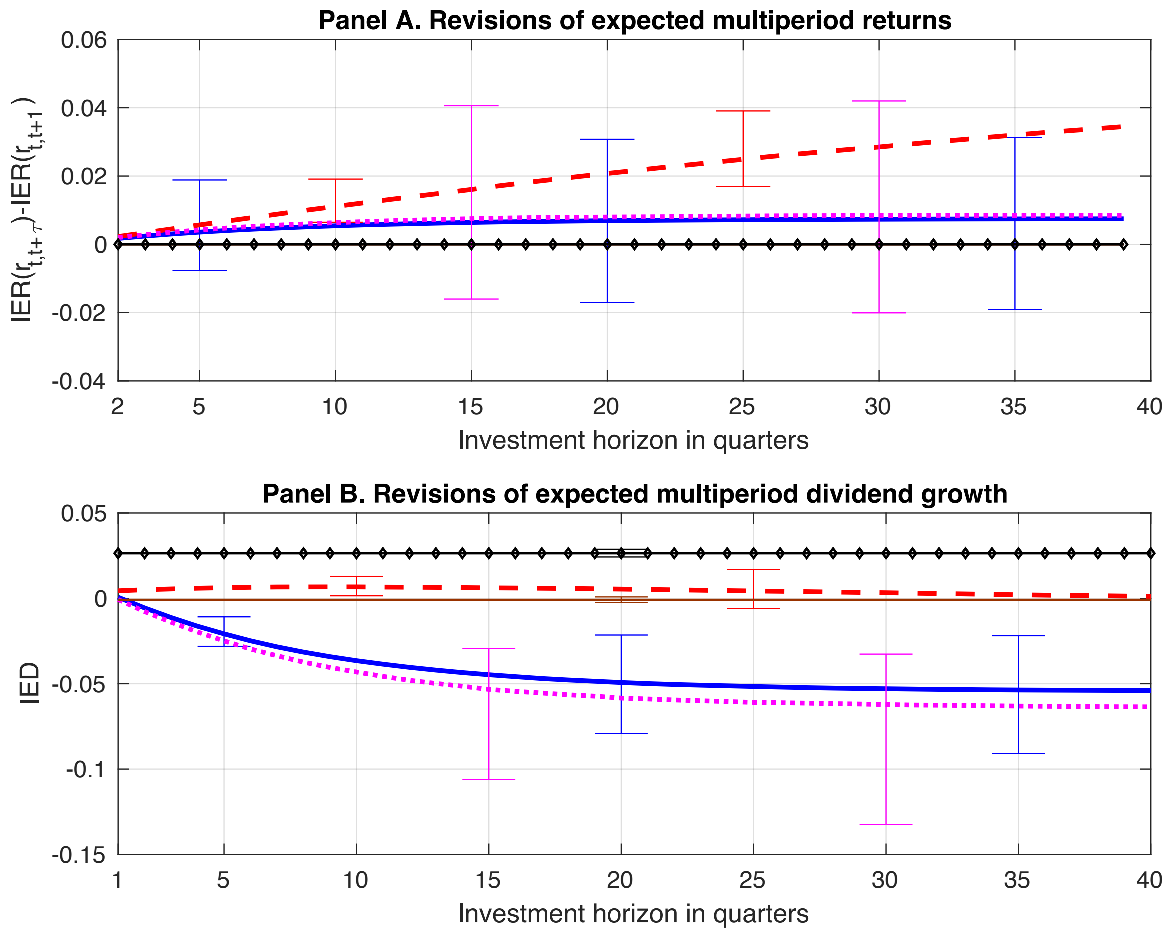
Term structure of risk in expected dividend growth. Identification "Consumption Uncertainty."



The blue solid lines correspond to the empirical term structures of risk in expected dividend growth; vertical bars indicate 95% credible intervals. See Table III for statistical significance of the slopes of the term structures of risk. The red dashed lines correspond to the theoretical term structures of risk implied by the parsimonious version of the model of Drechsler and Yaron (2011). Panel A illustrates the term structures of risk in expected dividend growth for the regular variance shock ε_{vt+1}. Panel B illustrates the term structures of risk in expected dividend growth for the variance trend shock ε^*_{vt+1}. Panel C illustrates the term structures of risk in expected dividend growth for the jump in variance ε^v_{zt+1}. Quarterly.

Figure 3

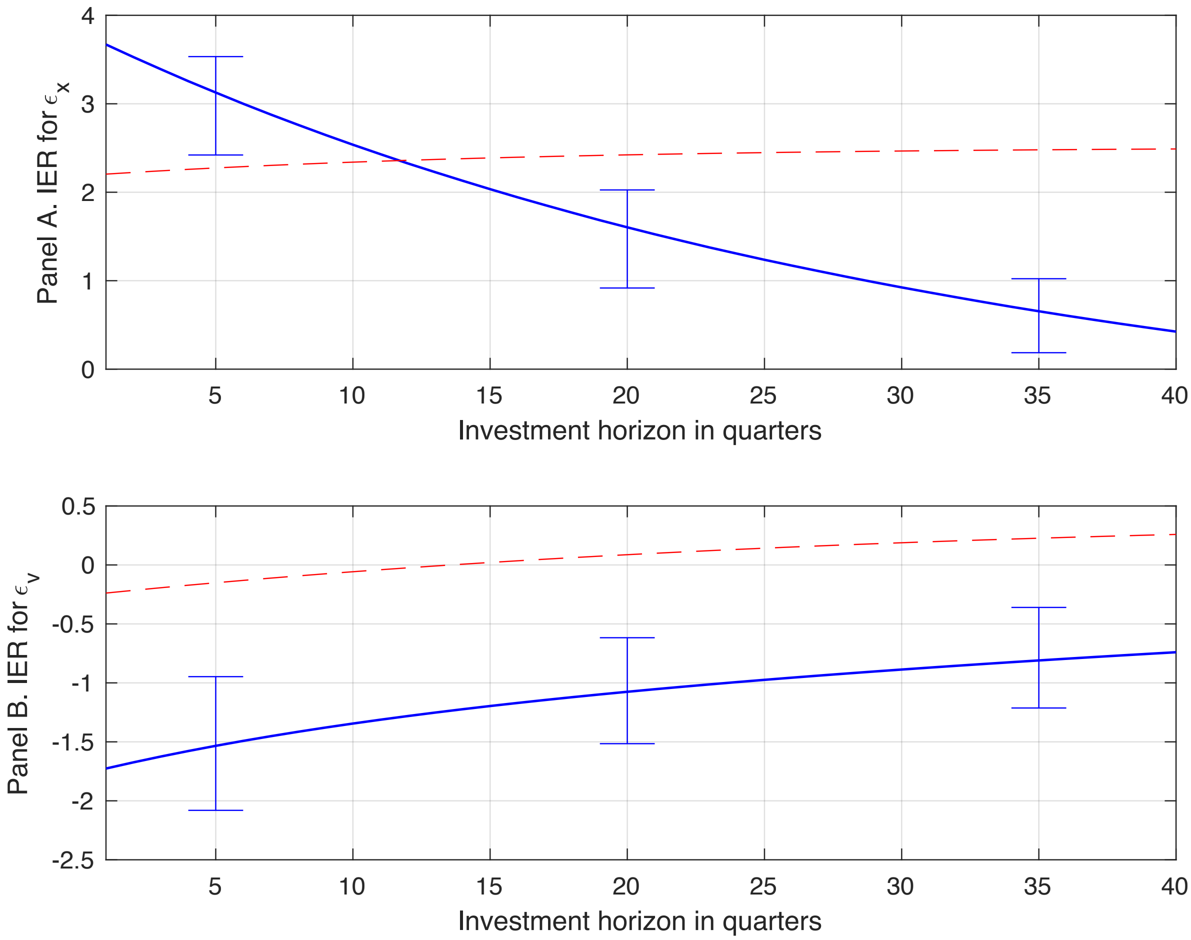
Term structures of revisions of expected returns and dividend growth. Identification "Consumption Uncertainty."



This figure plots the term structures of revisions of expected stock returns (Panel A) and dividend growth (Panel B). The blue thick lines correspond to the revisions associated with the regular variance shock ε_{vt+1}. The red dashed lines correspond to the revisions associated with the variance trend shock ε^{*}_{vt+1}. The magenta dotted lines correspond to the revisions associated with the jump in variance ε^{v}_{zt+1}. The thin brown lines correspond to the revisions associated with the direct consumption shock. The black marked lines correspond to the revisions associated with the direct dividend shock. The $\mathcal{IER}$s and $\mathcal{IED}$s are normalized so that their magnitudes are comparable across the shocks. I set $\Delta_i = 1$ and evaluate $\mathcal{IER}$s and $\mathcal{IED}$s at the unconditional means of the state variables. Quarterly.

Figure 4

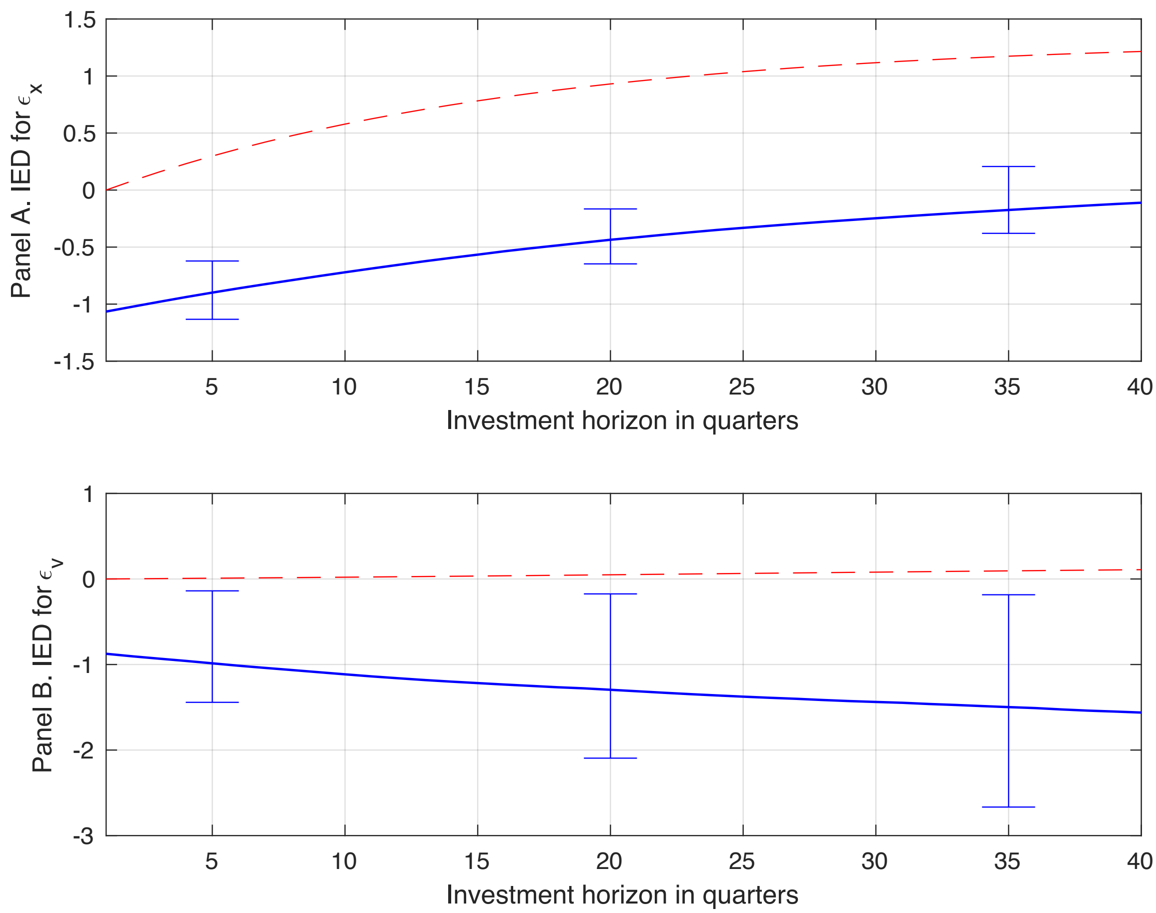
Term structure of risk in expected stock returns. Identification "Long-Run Risk."



The blue solid lines correspond to the empirical term structures of risk in expected returns; vertical bars indicate 95% credible intervals. See Table III for statistical significance of the slopes of the term structures of risk. The red dashed lines correspond to the theoretical term structures of risk implied by the model of Bansal and Yaron (2004). Panel A illustrates the term structures of risk in expected returns for the long-run risk ε_{xt+1}. Panel B illustrates the term structures of risk in expected returns for the variance risk ε_{vt+1}. Quarterly.

Figure 5

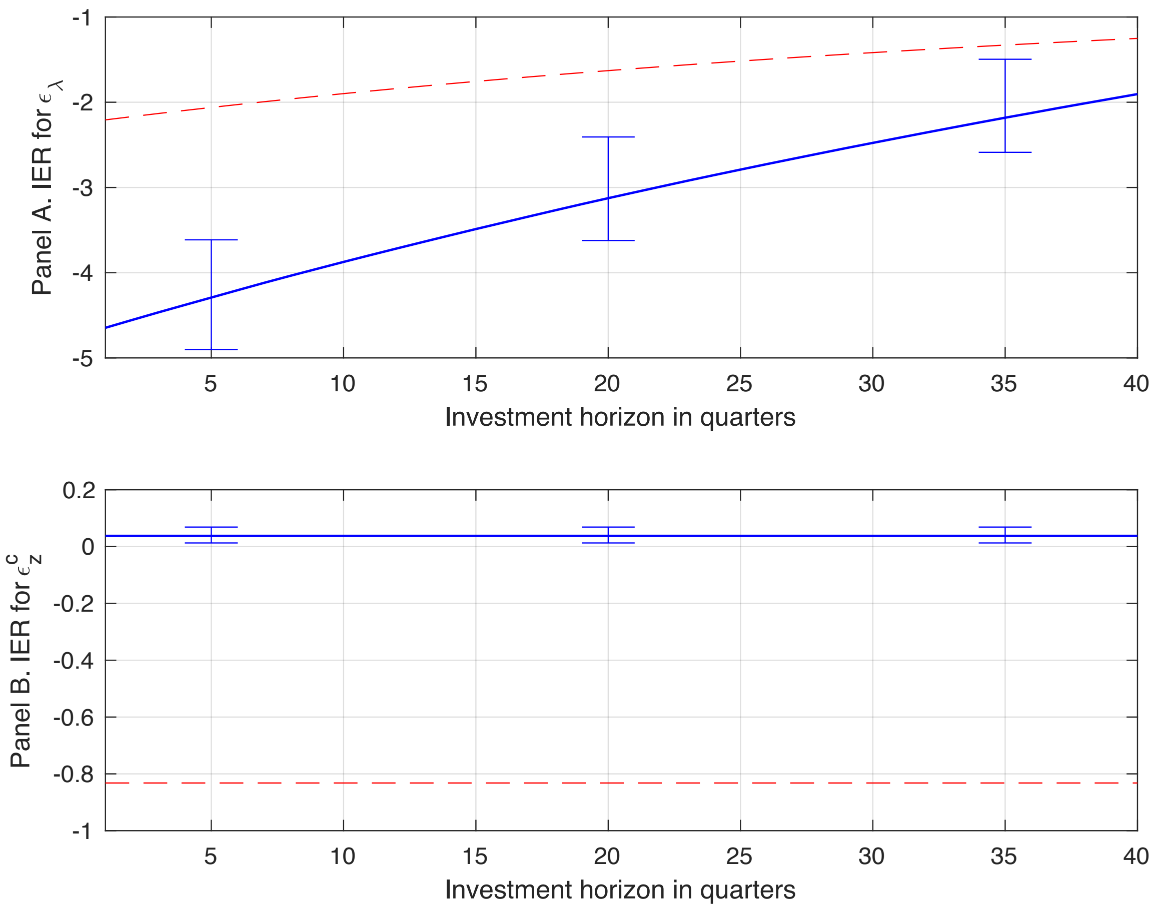
Term structure of risk in expected dividend growth. Identification "Long-Run Risk."



The blue solid lines correspond to the empirical term structures of risk in expected dividend growth; vertical bars indicate 95% credible intervals. See Table III for statistical significance of the slopes of the term structures of risk. The red dashed lines correspond to the theoretical term structures of risk implied by the model of Bansal and Yaron (2004). Panel A illustrates the term structures of risk in expected dividend growth for the long-run risk ε_{xt+1}. Panel B illustrates the term structures of risk in expected dividend growth for the variance risk ε_{vt+1}. Quarterly.

Figure 6

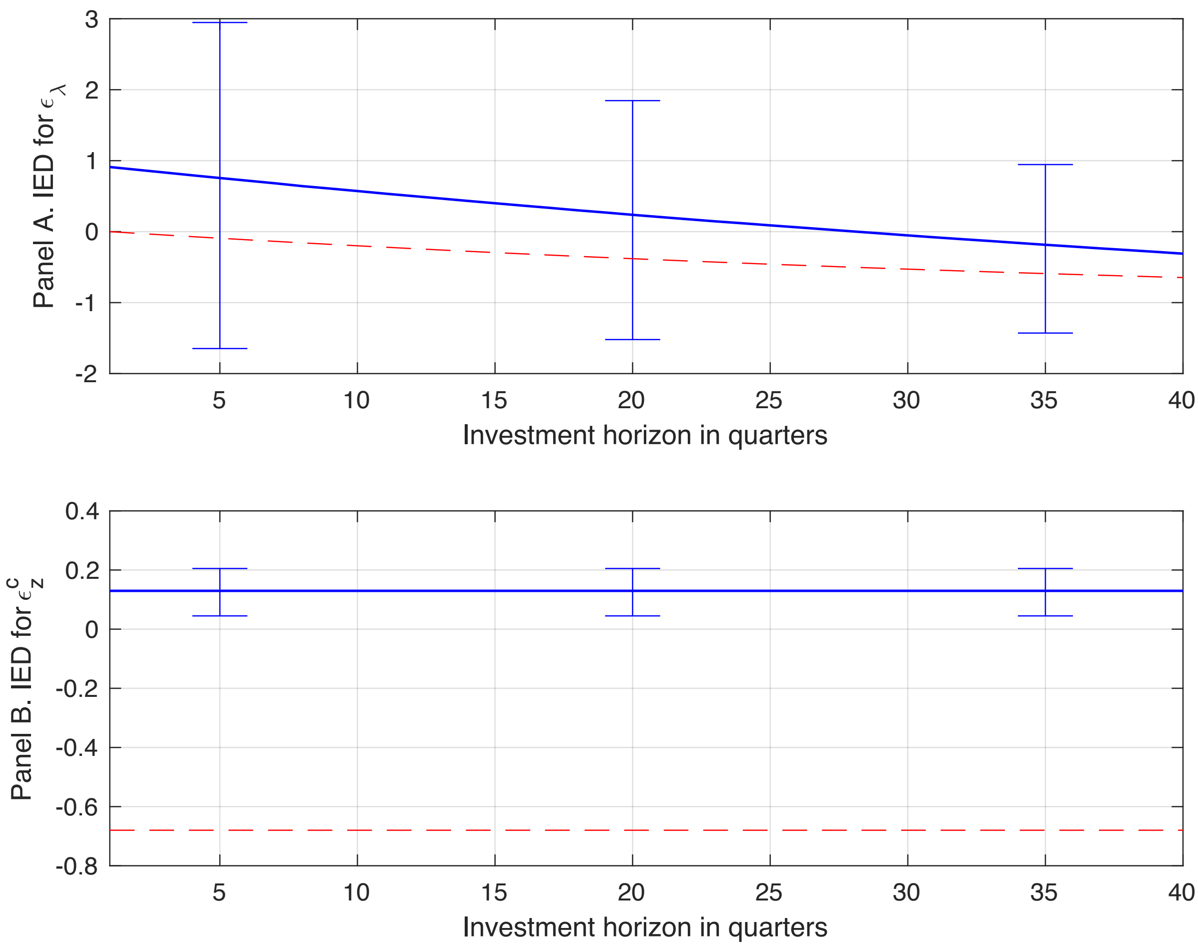
Term structure of risk in expected stock returns. Identification "Consumption Disasters."



The blue solid lines correspond to the empirical term structures of risk in expected returns; vertical bars indicate 95% credible intervals. See Table III for statistical significance of the slopes of the term structures of risk. The red dashed lines correspond to the theoretical term structures of risk implied by the model of Wachter (2013). Panel A illustrates the term structures of risk in expected returns for the disaster intensity shock $\varepsilon_{\lambda t+1}$. Panel B illustrates the term structures of risk in expected returns for the disaster shock ε^c_{zt+1}. Quarterly.

Figure 7

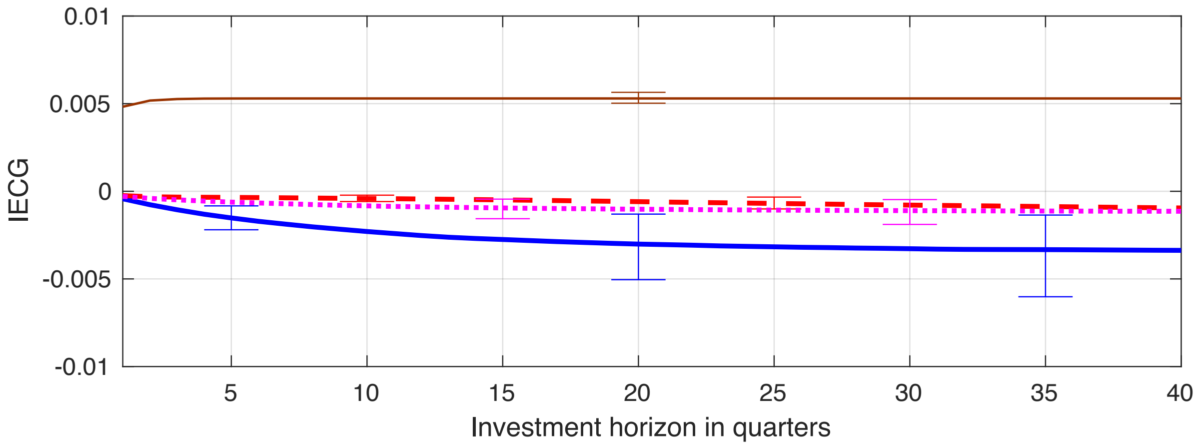
Term structure of risk in expected dividend growth. Identification "Consumption Disasters."



The blue solid lines correspond to the empirical term structures of risk in expected dividend growth; vertical bars indicate 95% credible intervals. See Table III for statistical significance of the slopes of the term structures of risk. The red dashed lines correspond to the theoretical term structures of risk implied by the model of Wachter (2013). Panel A illustrates the term structures of risk in expected dividend growth for the disaster intensity shock $\varepsilon_{\lambda t+1}$. Panel B illustrates the term structures of risk in expected dividend growth for the disaster shock ε^c_{zt+1}. Quarterly.

Figure 8

Term structures of risk in expected consumption growth.



This figure plots the term structures of risk in expected consumption growth $\log E_t g^c_{t,t+\tau}$. IECG stands for the incremental expected consumption growth. The blue thick lines correspond to the term structures of risk associated with the regular variance shock ε_{vt+1}. The red dashed lines correspond to the term structures of risk associated with the variance trend shock ε^*_{vt+1}. The magenta dotted lines correspond to the term structures of risk associated with the jump in variance ε^v_{zt+1}. The thin brown lines correspond to the term structures of risk associated with the direct consumption shock. The sensitivities of expected consumption growth to the shocks are normalized so that their magnitudes are comparable across the shocks. I set $\Delta_i = 1$ and evaluate risk sensitivities of consumption growth at the unconditional means of the state variables. Quarterly.

A Appendix

A The state-space model under different identification schemes

I estimate the state-space model (1)-(2) for the vector $Y_t = (\log r_{t-1,t},\ \log g^d_{t-1,t},\ y_t')'$ and the price-dividend ratio $\log \delta_t$

$$\begin{aligned} Y_{t+1} &= A + BY_t + W_{t+1}, \\ \log \delta_t &= a + b' y_t + w_t \end{aligned}$$

and impose different shock identification schemes. The shock identification schemes are determined by the composition of the state vector y_t and the mapping between the reduced-form innovations W_t and a vector of structural shocks ε_t. The elements of the vector ε_t have the mean of zero and the standard deviation of 1 but do not necessarily follow the standard normal distribution.

A.1 The "Long-Run Risk" identification scheme

In the "Long-Run Risk" identification scheme, $y_t = (\log g^c_t,\ x_t,\ v_t)'$ and $\varepsilon_{t+1} = (\varepsilon_{dt+1},\ \varepsilon_{ct+1},\ \varepsilon_{xt+1},\ \varepsilon_{vt+1})'$. The shocks W_{t+1} are heteroskedastic; the stochastic variance factor v_t is the only source of heteroskedasticity. The mapping between the reduced-form innovations W_{t+1} and structural shocks ε_{t+1} is defined as $W_{t+1} = H(v_t)\varepsilon_{t+1}$, where $H(v_t)$ is a matrix with 5 rows and 4 columns, whose entries depend on the stochastic variance v_t.

The state-space system features the cross-equation restrictions implied by Campbell-Shiller's (1989) linearization of a gross stock return

$$\log r_{t,t+1} \approx \kappa_0 + \kappa_1 \log \delta_{t+1} - \log \delta_t + \log g^d_{t,t+1}.$$

In addition, I impose the following parameter restrictions: $A_4 = B_{41} = B_{42} = B_{43} = B_{45} = 0$ (persistent component x_t has a mean of zero and depends only on its own lag), $B_{34} = 1$ (identification of x_t); $A_5 = 1 - B_{55}$, $B_{51} = B_{52} = B_{53} = B_{54} = 0$ (the stochastic variance factor has a mean of one and follows an autoregressive gamma process of order one); $B_{11} = B_{12} = B_{21} = B_{22} = B_{31} = B_{32} = 0$ (the lagged log return and the lagged dividend growth are not state variables). Furthermore, I find that the parameters b_1, B_{13}, B_{23}, and h_{12} are not statistically significant (the log-price dividend ratio does not contemporaneously load on the consumption growth; the log return and log dividend growth do not depend on the lagged consumption growth; the direct consumption shock does not contemporaneously affect the log return); and that fixing them at zero does not change the implications of the model. For parsimony, I fix them at zero.

The matrix $H(v_t)$ has the following representation

$$H(v_t) = \begin{pmatrix} h_{11}v_t^{1/2} & h_{12}v_t^{1/2} & h_{13}v_t^{1/2} & h_{14}\sigma_v((1-\varphi_v+2\varphi_v v_t)/2)^{1/2} \\ h_{11}v_t^{1/2} & h_{12}v_t^{1/2} & (h_{13}-\kappa_1 b_2 h_{43})v_t^{1/2} & (h_{14}-\kappa_1 b_3-\kappa_1 b_2 h_{44})\sigma_v((1-\varphi_v+2\varphi_v v_t)/2)^{1/2} \\ 0 & h_{32}v_t^{1/2} & h_{33}v_t^{1/2} & h_{34}\sigma_v((1-\varphi_v+2\varphi_v v_t)/2)^{1/2} \\ 0 & 0 & h_{43}v_t^{1/2} & h_{44}\sigma_v((1-\varphi_v+2\varphi_v v_t)/2)^{1/2} \\ 0 & 0 & 0 & \sigma_v((1-\varphi_v+2\varphi_v v_t)/2)^{1/2} \end{pmatrix}.$$

The entry in the third row and first column of the matrix $H(v_t)$ is set to zero, as an identifying assumption for the direct dividend shock ε_{dt+1}: contemporaneously the shock affects the dividend growth, but not the consumption growth. The zero entries in the fourth row are identifying restrictions for the direct dividend and direct consumption shocks (neither affect x_t contemporaneously). The zero entries in the fifth row of the matrix $H(v_t)$ are due to the variance factor following an autoregressive gamma process of order one. The second row of the matrix $H(v_t)$ is not free: it is related to the other rows of the matrix $H(v_t)$, as well as the parameters of the vector b through the cross-equation restrictions implied by Campbell-Shiller's (1989) loglinearization of a gross stock return.

A.2 The "Consumption Disasters" identification scheme

In the "Consumption Disaster" identification scheme, $y_t = (\log g_t^c,\ \lambda_t)'$ and $\varepsilon_{t+1} = (\varepsilon_{dt+1},\ \varepsilon_{ct+1},\ \varepsilon_{\lambda t+1},\ \varepsilon^c_{zt+1})'$. The shocks W_{t+1} are heteroskedastic; the disaster intensity factor λ_t is the only source of heteroskedasticity. The mapping between the reduced-form innovations W_{t+1} and structural shocks ε_{t+1} is defined as $W_{t+1} = H(\lambda_t)\varepsilon_{t+1}$, where $H(\lambda_t)$ is a matrix with 4 rows and 4 columns, whose entries depend on the stochastic jump intensity λ_t.

The state-space system features the cross-equation restrictions implied by Campbell-Shiller's (1989) loginearization of a gross stock return

$$\log r_{t,t+1} \approx \kappa_0 + \kappa_1 \log\delta_{t+1} - \log\delta_t + \log g^d_{t,t+1}.$$

In addition, I impose the following parameter restrictions: $A_4 = 1 - B_{44}$, $B_{41} = B_{42} = B_{43} = 0$, $h_{44} = 0$ (the jump intensity factor has a mean of one and follows an autoregressive gamma process of order one); $B_{11} = B_{12} = B_{21} = B_{22} = B_{31} = B_{32} = 0$ (the lagged log return and the lagged log dividend growth are not state variables); $h_{34} = -1$ (identification of the consumption disaster z^c_{t+1}). Furthermore, I find that

b_1, B_{13} and B_{23} are not statistically significant and that fixing them directly at zero does not change the implications of the model. For parsimony, I fix them at zero.

The matrix $H(\lambda_t)$ has the following representation

$$H(\lambda_t) = \begin{pmatrix} h_{11} & h_{12} & h_{13}\sigma_\lambda((1-\varphi_\lambda+2\varphi_\lambda\lambda_t)/2)^{1/2} & h_{14}(2\theta_c^2 h_\lambda \lambda_t)^{1/2} \\ h_{11} & h_{12} & (h_{13}-\kappa_1 b_2)\sigma_\lambda((1-\varphi_\lambda+2\varphi_\lambda\lambda_t)/2)^{1/2} & h_{14}(2\theta_c^2 h_\lambda \lambda_t)^{1/2} \\ 0 & h_{32} & h_{33}\sigma_\lambda((1-\varphi_\lambda+2\varphi_\lambda\lambda_t)/2)^{1/2} & h_{34}(2\theta_c^2 h_\lambda \lambda_t)^{1/2} \\ 0 & 0 & \sigma_\lambda((1-\varphi_\lambda+2\varphi_\lambda\lambda_t)/2)^{1/2} & 0 \end{pmatrix}.$$

The entry in the third row and first column of the matrix $H(\lambda_t)$ is set to zero, as an identifying assumption for the direct dividend shock ε_{dt+1}: the shock contemporaneously affects the dividend growth but not the consumption growth. The fourth row of the matrix $H(\lambda_t)$ has zero entries, because the jump intensity factor λ_t follows an autoregressive gamma process of order one. The second row of the matrix $H(\lambda_t)$ is not free: it is related to the other rows of the matrix $H(\lambda_t)$, as well as the parameters of the vector b through the cross-equation restrictions implied by Campbell-Shiller's (1989) loglinearization of a gross stock return. Even though there are four shocks and three independent equations, the system is just-identified. The shock ε^c_{zt+1} is identified thanks to the assumption that consumption disasters z^c_{t+1} are rare, big, negative shocks. The identification of the other three shocks is standard: a system of three equations with three independent shocks is just-identified, if it features three zero restrictions. The three zero-restrictions here are $H_{31}(\lambda_t) = H_{41}(\lambda_t) = H_{42}(\lambda_t) = 0$.

A.3 The "Consumption uncertainty" identification scheme

In the "Consumption Uncertainty" identification scheme, $y_t = (\log g^c_t,\ v_t,\ v^*_t)'$, $\varepsilon_{t+1} = (\varepsilon_{dt+1},\ \varepsilon_{ct+1},\ \varepsilon_{vt+1},\ \varepsilon^*_{vt+1},\ \varepsilon^v_{zt+1})'$. The shocks W_{t+1} are heteroskedastic; the stochastic variance factors v_t and v^*_t are the sources of heteroskedasticity. The mapping between the reduced-form innovations W_{t+1} and structural shocks ε_{t+1} is defined as $W_{t+1} = H(v_t, v^*_t)\varepsilon_{t+1}$, where $H(v_t, v^*_t)$ is a matrix with 5 rows and 5 columns, whose entries depend on the stochastic variance v_t and its long-run mean factor v^*_t.

The state-space system features the cross-equation restrictions implied by Campbell-Shiller's (1989) loglinearization of a gross stock return

$$\log r_{t,t+1} \quad \approx \quad \kappa_0 + \kappa_1 \log \delta_{t+1} - \log \delta_t + \log g^d_{t,t+1}.$$

In addition, I impose the following parameter restrictions: $A_4 = (1-B_{44})v$ $(v = (1-B_{44}-\theta_v h_v)/[2(1-B_{44})])$, $B_{41} = B_{42} = B_{43} = 0$, $B_{45} = 1 - B_{44} - \theta_v h_{vv}$ (the variance factor v_t follows an autoregressive gamma process of order one with jumps and time-varying long-run mean); $A_5 = (1 - B_{55})v^*$ ($v^* = 1/2$), $B_{51} = B_{52} = B_{53} = B_{54} = 0$ (the variance factor v_t^* follows an autoregressive gamma process of order one); $B_{11} = B_{12} = B_{21} = B_{22} = B_{31} = B_{32} = 0$ (the lagged log return and the lagged log dividend growth are not state variables). Furthermore, I find that b_1, B_{13}, B_{23}, B_{35} are statistically insignificant and that fixing them at zero does not change the implications of the model. For parsimony, I fix them at zero.

The matrix $H(v_t, v_t^*)$ has the following representation

$$H(v_t, v_t^*) = \begin{pmatrix} h_{11}v_t^{1/2} & h_{12}v_t^{1/2} & H_{13}(v_t, v_t^*) & H_{14}(v_t, v_t^*) & H_{15}(v_t, v_t^*) \\ h_{11}v_t^{1/2} & h_{12}v_t^{1/2} & H_{23}(v_t, v_t^*) & H_{24}(v_t, v_t^*) & H_{25}(v_t, v_t^*) \\ 0 & h_{32}v_t^{1/2} & H_{33}(v_t, v_t^*) & H_{34}(v_t, v_t^*) & H_{35}(v_t, v_t^*) \\ 0 & 0 & H_{43}(v_t, v_t^*) & 0 & H_{45}(v_t, v_t^*) \\ 0 & 0 & 0 & H_{54}(v_t, v_t^*) & 0 \end{pmatrix},$$

where

$$\begin{aligned}
H_{13}(v_t, v_t^*) &= h_{13}\sigma_v(((1-\varphi_v)v + 2\varphi_v v_t)/2)^{1/2}, \\
H_{14}(v_t, v_t^*) &= h_{14}\sigma_v^*(((1-\varphi_v^*)v^* + 2\varphi_v^* v_t^*)/2)^{1/2}, \\
H_{15}(v_t, v_t^*) &= h_{15}(2\theta_v h_v v_t)^{1/2}, \\
H_{23}(v_t, v_t^*) &= (h_{13} - \kappa_1 b_2)\sigma_v(((1-\varphi_v)v + 2\varphi_v v_t)/2)^{1/2}, \\
H_{24}(v_t, v_t^*) &= (h_{14} - \kappa_1 b_3)\sigma_v^*(((1-\varphi_v^*)v^* + 2\varphi_v^* v_t^*)/2)^{1/2}, \\
H_{25}(v_t, v_t^*) &= (h_{15} - \kappa_1 b_2)(2\theta_v h_v v_t)^{1/2}, \\
H_{33}(v_t, v_t^*) &= h_{33}\sigma_v(((1-\varphi_v)v + 2\varphi_v v_t)/2)^{1/2}, \\
H_{34}(v_t, v_t^*) &= h_{34}\sigma_v^*(((1-\varphi_v^*)v^* + 2\varphi_v^* v_t^*)/2)^{1/2}, \\
H_{35}(v_t, v_t^*) &= h_{35}(2\theta_v h_v v_t)^{1/2}, \\
H_{43}(v_t, v_t^*) &= \sigma_v(((1-\varphi_v)v + 2\varphi_v v_t)/2)^{1/2}, \\
H_{45}(v_t, v_t^*) &= (2\theta_v h_v v_t)^{1/2}, \\
H_{54}(v_t, v_t^*) &= \sigma_v^*(((1-\varphi_v^*)v^* + 2\varphi_v^* v_t^*)/2)^{1/2}.
\end{aligned}$$

The entry in the third row and first column of the matrix $H(v_t, v_t^*)$ is set to zero, as an identifying assumption for the direct dividend shock ε_{dt+1}: the shock contemporaneously affects the dividend growth, not the consumption growth. The fourth and fifth rows of the matrix $H(v_t, v_t^*)$ have zero entries because the variance factors follow the autoregressive gamma processes of order one. As a result, the shocks ε_{vt+1}

and ε^{*}_{vt+1} are identified; the shock ε^{v}_{zt+1} is identified thanks to the assumption that jumps z^{v}_{t+1} are relatively rare, big, positive movements in the stochastic variance v_{t+1}. The second row of the matrix $H(v_t, v^*_t)$ is not free: it is related to the other rows of the matrix $H(v_t, v^*_t)$, as well as the parameters of the vector b through the cross-equation restrictions implied by Campbell-Shiller's loglinearization of a gross stock return.

B Estimation output

I use the data displayed in Panels A, B, and C of Figure B1 to estimate the state-space model (1)-(2). I consider three shock identification schemes. Figure B1 and Table BI contain the estimation output from the identification scheme "Long-Run Risk;" Figure B2 and Table BII contain the estimation output from the identification scheme "Consumption Disasters;" Figure B3 and Table BIII contain the estimation output from the identification scheme "Consumption Uncertainty."

Table BI
The state-space model. Identification "Long-Run Risk." Parameter estimates

I estimate

$$
\begin{aligned}
Y_{t+1} &= A + BY_t + W_{t+1}, \\
\log \delta_t &= a + b' y_t + w_t,
\end{aligned}
$$

with parameter restrictions $B_{11} = B_{12} = B_{13} = B_{21} = B_{22} = B_{23} = B_{31} = B_{32} = A_4 = B_{41} = B_{42} = B_{43} = B_{45} = B_{51} = B_{52} = B_{53} = B_{54} = h_{12} = 0$, $A_2 = A_1 - \kappa_0 - \kappa_1 a + a$, $A_5 = 1 - B_{55}$, $B_{55} = \varphi_v$, $B_{34} = 1$, $B_{24} = B_{14} + b_2 - \kappa_1 b_2 B_{44}$, $B_{25} = B_{15} + b_3 - \kappa_1 b_3 \varphi_v$, $b_1 = 0$, $h_{21} = h_{11}$, $h_{22} = h_{12}$, $h_{23} = h_{13} - \kappa_1 b_2 h_{43}$, $h_{24} = h_{14} - \kappa_1 b_3 - \kappa_1 b_2 h_{44}$, which guarantee identification and satisfy regularity conditions and Campbell and Shiller's (1989) loglinear approximation. $\kappa_0 = 0.0499$ and $\kappa_1 = 0.9923$. The median and 95% credible interval. Appendix A.1 contains the full description of the model.

Parameter	Estimate	Credible interval, 95%
A_1	0.0110	(0.0090, 0.0128)
A_3	0.0073	(0.0067, 0.0082)
B_{14}	-6.0708	(-8.6412, -4.6009)
B_{15}	0.0061	(0.0042, 0.0081)
B_{33}	0.1673	(0.1281, 0.1975)
B_{35}	-0.0034	(-0.0041, -0.0029)
B_{44}	0.9687	(0.9531, 0.9756)
B_{55}	0.9786	(0.9764 0.9815)
h_{11}	0.0219	(0.0209, 0.0227)
h_{13}	0.0718	(0.0674, 0.0758)
h_{14}	-0.1611	(-0.2107, -0.1011)
h_{32}	0.0050	(0.0048, 0.0052)
h_{33}	5.49e-4	(1.81e-4, 7.98e-4)
h_{34}	-0.0057	(-0.0089, -0.0030)
h_{43}	4.61e-4	(4.14e-4, 5.32e-4)
h_{44}	7.30e-4	(0.0002, 0.0011)
σ_v	0.1803	(0.1614, 0.1968)
a	5.1495	(5.2030, 5.0892)
b_2	169.77	(150.33, 190.88)
b_3	-0.2603	(-0.3139, -0.2001)
σ_δ	0.0040	(0.0031, 0.0053)

Table BII
The state-space model. Identification "Consumption Disasters." Parameter estimates

I estimate

$$\begin{aligned} Y_{t+1} &= A + BY_t + W_{t+1}, \\ \log\delta_t &= a + b' y_t + w_t, \end{aligned}$$

with parameter restrictions $A_2 = A_1 - \kappa_0 - \kappa_1 a + a$, $B_{11} = B_{12} = B_{13} = B_{21} = B_{22} = B_{23} = B_{31} = B_{32} = B_{41} = B_{42} = B_{43} = h_{44} = 0$, $A_4 = 1 - B_{44}$, $B_{44} = \varphi_\lambda$, $B_{24} = B_{14} + b_2 - \kappa_1 b_2 B_{44}$, $h_{14} = -1$, $h_{24} = h_{14}$, $b_1 = 0$, $h_{21} = h_{11}$, $h_{22} = h_{12}$, $h_{23} = h_{13} - \kappa_1 b_2$, which guarantee identification and satisfy regularity conditions and Campbell and Shiller's (1989) loglinear approximation. $\kappa_0 = 0.0499$ and $\kappa_1 = 0.9923$. The median and 95% credible interval. Appendix A.2 contains the full description of the model.

Parameter	Estimate	Credible interval, 95%
A_1	-0.0029	(-0.0113, -0.0013)
A_3	3.33e-3	(3.31e-3, 3.34e-3)
B_{14}	0.0177	(0.0160, 0.0261)
B_{33}	0.3328	(0.3312, 0.3345)
B_{34}	2.31e-4	(2.99e-5, 4.96e-4)
B_{44}	0.9885	(0.9801 0.9900)
h_{11}	0.0221	(0.0217 , 0.0224)
h_{12}	0.0033	(0.0027 , 0.0039)
h_{13}	-0.9998	(-1.0240, -0.9808)
h_{14}	-2.2810	(-2.6412, -1.9018)
h_{32}	4.37e-3	(4.36e-3, 4.38e-3)
h_{33}	-0.0119	(-0.0124, -0.0114)
σ_λ	0.0587	(0.0527, 0.0660)
θ_c	0.0040	(0.0026, 0.0064)
h_λ	0.0591	(0.0076, 0.0979)
a	5.9514	(5.7987, 6.1179)
b_2	-1.0622	(-1.2288, -0.9075)
σ_δ	0.0045	(0.0033, 0.0062)

Table BIII
The state-space model. Identification "Consumption Uncertainty." Parameter estimates

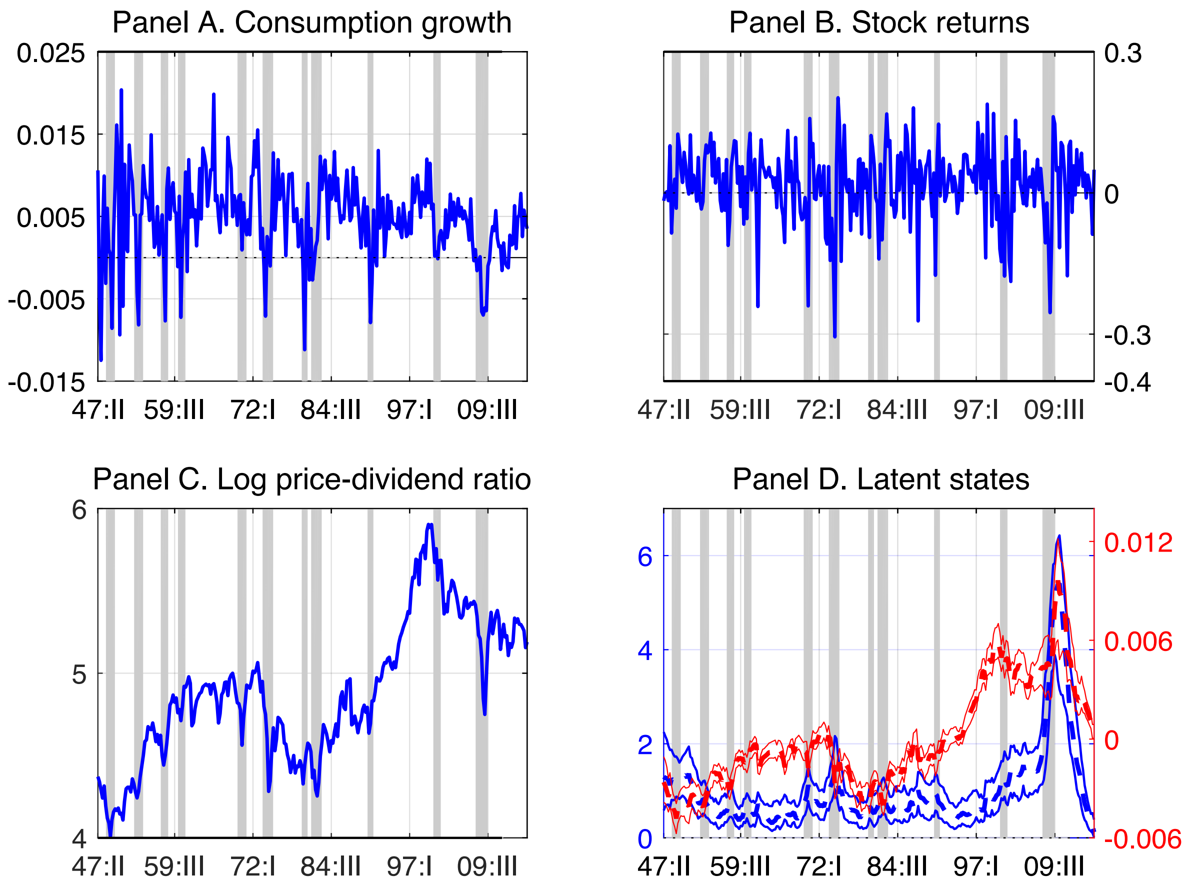
I estimate

$$\begin{aligned} Y_{t+1} &= A + BY_t + W_{t+1}, \\ \log\delta_t &= a + b'y_t + w_t, \end{aligned}$$

with parameter restrictions $B_{11} = B_{12} = B_{13} = B_{21} = B_{22} = B_{23} = B_{31} = B_{32} = B_{41} = B_{42} = B_{43} = B_{51} = B_{52} = B_{53} = B_{54} = b_1 = 0$, $A_2 = A_1 - \kappa_0 - \kappa_1 a + a$, $A_4 = (1 - B_{44})v$, where $v = (1 - B_{45})/(1 - B_{44})/2$, $A_5 = (1 - B_{55})v^*$, where $v^* = 1/2$, $B_{24} = B_{14} + b_2 - \kappa_1 b_2 B_{44}$, $B_{25} = B_{15} + b_3 - \kappa_1 b_3 B_{55}$, $B_{45} = 1 - B_{44} - \theta_v h_v$, $B_{44} = \varphi_v$, $B_{55} = \varphi_v^*$, $h_{21} = h_{11}$, $h_{22} = h_{12}$, $h_{23} = h_{13} - \kappa_1 b_2$, $h_{24} = h_{14} - \kappa_1 b_3$, $h_{25} = h_{15} - \kappa_1 b_2$, $h_{31} = h_{41} = h_{42} = h_{44} = h_{51} = h_{52} = h_{53} = h_{55} = 0$, which guarantee identification and satisfy regularity conditions and Campbell and Shiller's (1989) loglinear approximation. $\kappa_0 = 0.0499$ and $\kappa_1 = 0.9923$. The median and 95% credible interval. Appendix A.3 contains the full description of the model.

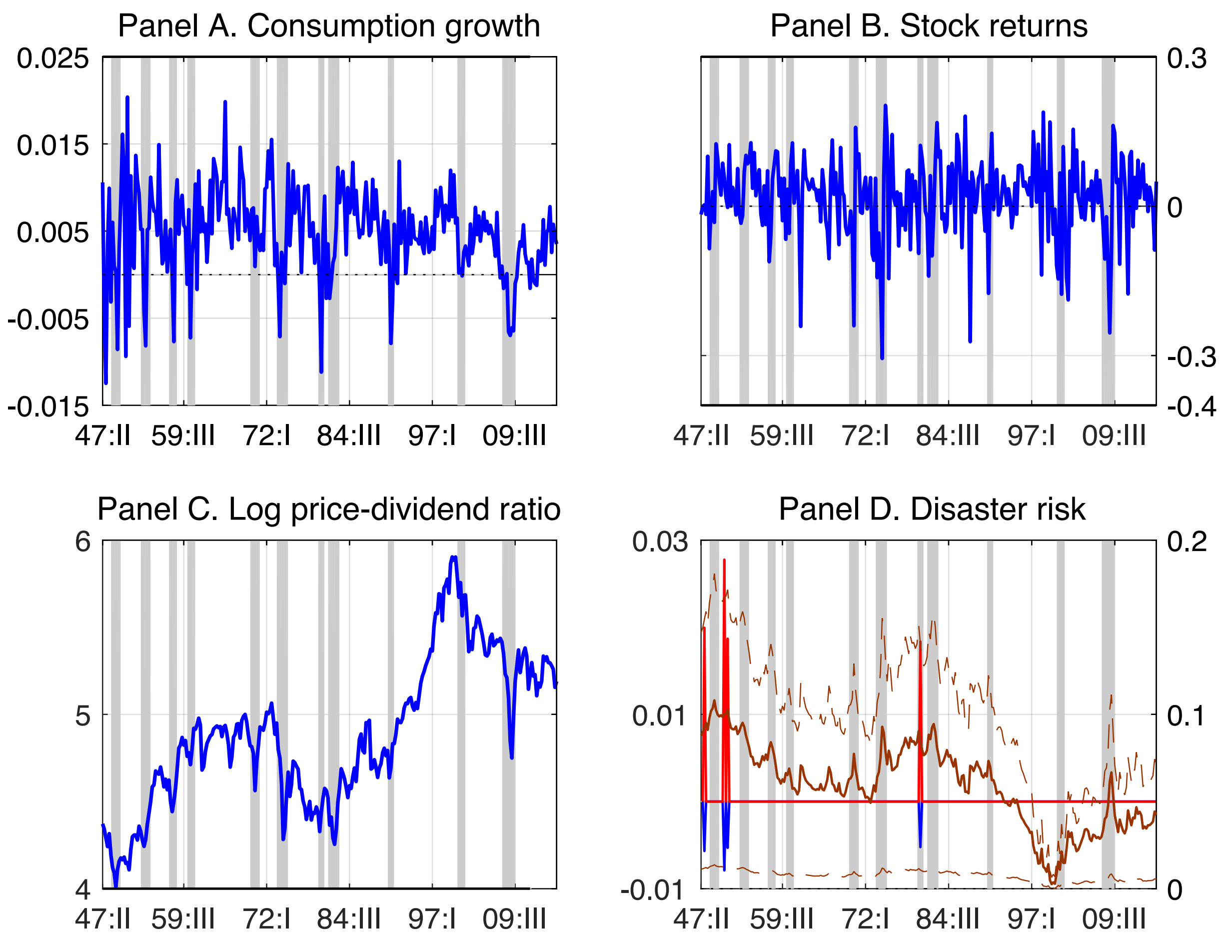
Parameter	Estimate	Credible interval, 95%
A_1	-0.0015	(-0.0180, 0.0106)
A_3	0.0050	(0.0039, 0.0062)
B_{14}	0.0063	(-0.0122, 0.0278)
B_{15}	0.0295	(-0.0056, 0.0666)
B_{33}	0.2710	(0.2199, 0.3004)
B_{34}	-0.0015	(-0.0005, -0.0026)
B_{44}	0.8735	(0.7839 , 0.9283)
B_{55}	0.9857	(0.9692, 0.9941)
h_{11}	0.0264	(0.0242, 0.0288)
h_{12}	-8.85e-4	(-0.0024 0.0008)
h_{13}	-0.2589	(-0.3560, -0.1113)
h_{14}	-1.6197	(-1.8630, -1.4545)
h_{15}	-0.2643	(-0.3622, -0.1301)
h_{32}	0.0048	(0.0046, 0.0051)
h_{33}	-0.0032	(-0.0056, -0.0014)
h_{34}	-0.0196	(-0.0276, -0.0113)
h_{35}	-0.0048	(-0.0070, -0.0026)
σ_v	0.2320	(0.1870, 0.2916)
σ_v^*	0.0648	(0.0586, 0.0720)
θ_v	0.9568	(0.8014, 1.1727)
h_v	0.0199	(0.0132, 0.0353)
a	6.0346	(5.7979, 6.2475)
b_2	-0.2629	(-0.3622, -0.1176)
b_3	-1.7652	(-1.9923, -1.5824)
σ_δ	0.0100	(0.0098, 0.0101)

Figure B1. Data and estimated states in the "Long-Run Risk."



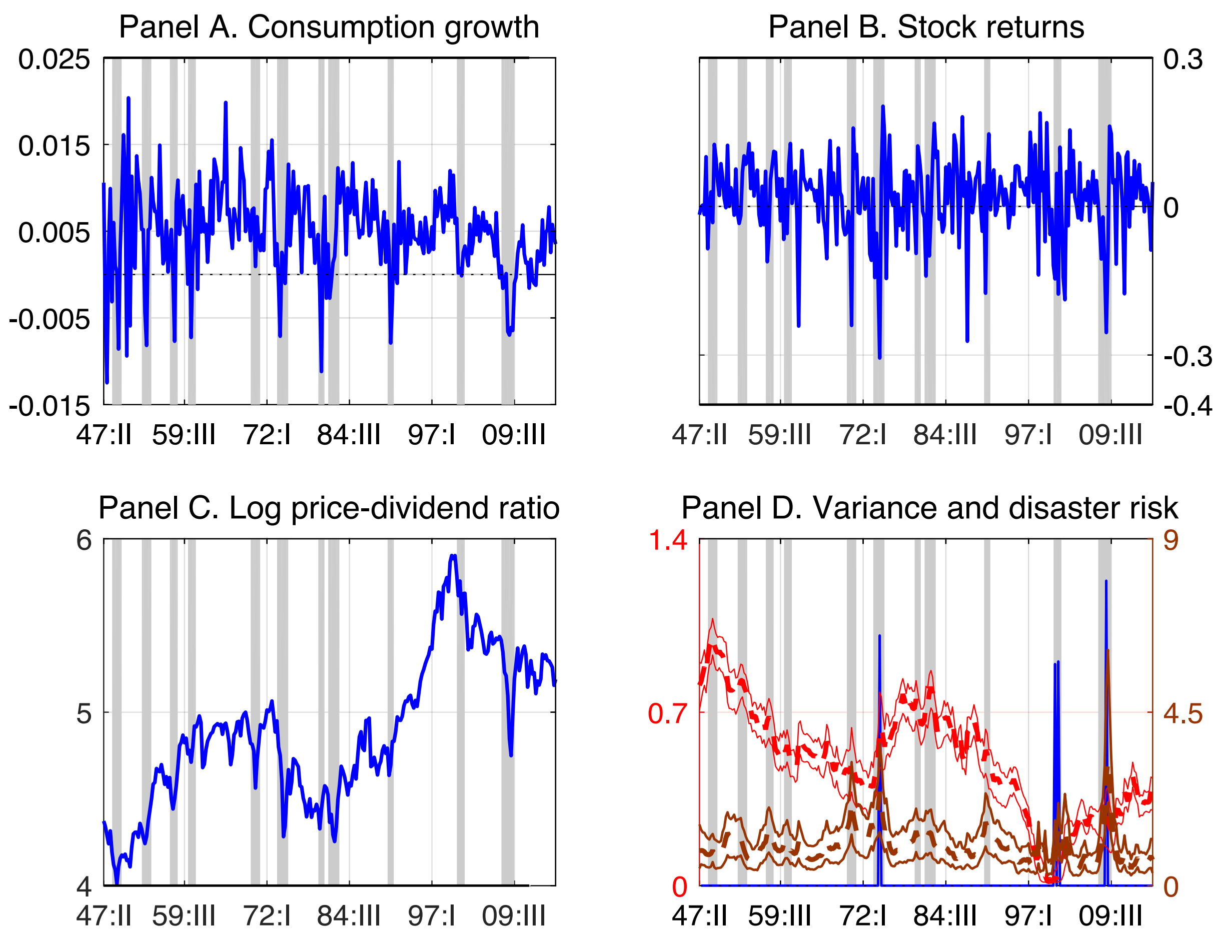
Panels A, B, C display quarterly observations of log consumption growth, log stock returns, and log price-dividend ratio, respectively. Panel D displays the mean path of the stochastic variance factor (dashed blue line) with the 95% credible interval (thin solid lines) and the mean path of the expected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

Figure B2. Data and estimated states in the "Consumption Disasters."



Panels A, B, C display quarterly observations of log consumption growth, log stock returns, and log price-dividend ratio, respectively. Panel D displays consumption disaster risk (blue lines) and 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. The sample period: second quarter of 1947 to fourth quarter of 2015. Grey bars are the NBER recessions. Quarterly.

Figure B3. Data and estimated states in the "Consumption Uncertainty."



Panels A, B, C display quarterly observations of log 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 in variance (blue bars) are displayed on Panel D, left axes. The sample period: second quarter of 1947 to fourth quarter of 2015. Grey bars are the NBER recessions. Quarterly.

C Solutions of the equilibrium models

C.1 Solution of the model of Bansal and Yaron (2004)

The model for consumption growth with stochastic variance is

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

A representative agent has recursive preferences

$$\begin{aligned}
U_t &= [(1-\beta)c_t^\rho + \beta \mu_t (U_{t+1})^\rho]^{1/\rho}, \qquad (15) \\
\mu_t(U_{t+1}) &= [E_t(U^\alpha_{t+1})]^{1/\alpha}.
\end{aligned}$$

I divide expression (15) by c_t, denote $u_t = U_t/c_t$ and $g^c_{t+1} = c_{t+1}/c_t$ and obtain

$$u_t = [(1-\beta) + \beta \mu_t (u_{t+1} g^c_{t+1})^\rho]^{1/\rho}. \qquad (16)$$

Next, I solve a recursive problem that is a log-linear approximation of the Bellman equation (16)

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

where

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

I guess the value function

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

and compute

$$\log u_{t+1} + \log g^c_{t+1} = u + g^c + (p_x \varphi_x + 1) x_t + p_v v_{t+1} + \gamma_c v_t^{1/2} \varepsilon_{ct+1} + p_x \gamma_x v_t^{1/2} \varepsilon_{xt+1}.$$

The cumulant generating function for the variance $v_t \sim \mathcal{ARG}(1)$

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

and therefore,

$$
\begin{aligned}
\log \mu_t(u_{t+1} g^c_{t+1}) &= u + g^c - \frac{(1-\varphi_v)\log(1-\alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2} + (p_x \varphi_x + 1) x_t \\
&+ \left(\frac{\alpha}{2}\gamma_c^2 + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\varphi_v p_v}{1-\alpha p_v \sigma_v^2/2} \right) v_t, \\
\log \mu &= u + g^c + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\alpha}{2}\gamma_c^2 + \frac{p_v \varphi_v}{1-\alpha p_v \sigma_v^2/2} - \frac{(1-\varphi_v)\log(1-\alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2}, \\
\log(u_{t+1} g^c_{t+1}) - \log \mu_t(u_{t+1} g^c_{t+1}) &= \frac{(1-\varphi_v)\log(1-\alpha p_v \sigma_v^2/2)}{\alpha \sigma_v^2/2} - \left(\frac{\alpha}{2}\gamma_c^2 + \frac{\alpha}{2} p_x^2 \gamma_x^2 + \frac{\varphi_v p_v}{1-\alpha p_v \sigma_v^2/2} \right) v_t \\
&+ p_v v_{t+1} + \gamma_c v_t^{1/2} \varepsilon_{ct+1} + p_x \gamma_x v_t^{1/2} \varepsilon_{xt+1}.
\end{aligned}
$$

I solve the following system of three equations in three unknowns u, p_x, and p_v in order to verify the guess of the value function

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

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

Next, I obtain the pricing kernel

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

The model for the dividend growth is

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

I guess that the price-dividend ratio $\log \delta_t$ is

$$\log \delta_t \quad = \quad q_0 + q_x x_t + q_v v_t,$$

and derive the process for the log stock return as

$$\begin{aligned}
\log r_{t,t+1} &= \kappa_0 + \kappa_1 \log \delta_{t+1} + \log g^d_{t,t+1} - \log \delta_t \\
&= \underbrace{k_0 + (k_1 - 1) q_0 + g^d}_{r} + \underbrace{(\mu_x + q_x(k_1 \varphi_x - 1))}_{r_x} x_t \underbrace{-q_v}_{r_v} v_t \\
&+ \underbrace{k_1 q_v}_{r_{\varepsilon v}} v_{t+1} + \underbrace{k_1 q_x \gamma_x}_{r_{\varepsilon x}} v_t^{1/2} \varepsilon_{xt+1} + \underbrace{\gamma_d}_{r_{\varepsilon d}} v_t^{1/2} \varepsilon_{dt+1},
\end{aligned}$$

where

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

or in compact form

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

I use the law of one price $E_t[m_{t,t+1} r_{t,t+1}] = 1$ to obtain three equations in three unknowns q_0, q_x, and q_v

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

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

$$q_0 = \frac{1}{1 - k_1} \left(k_0 + g^d + m - \frac{(1 - \varphi_v) \log(1 - (r_{\varepsilon v} + m_{\varepsilon v}) \sigma_v^2 / 2)}{\sigma_v^2 / 2} \right),$$

where

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

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

C.2 Solution of the model of Wachter (2013)

The model for consumption growth with time-varying consumption disasters is

$$
\begin{aligned}
\log g^c_{t+1} &= g^c + \gamma_c \varepsilon_{ct+1} - z^c_{t+1}, \\
j^c_{t+1}|_{\lambda_t} &\sim \text{Poisson}(h_\lambda \lambda_t), \\
z^c_{t+1}|_{j^c_{t+1}} &\sim \text{Gamma}(j^c_{t+1}, \theta_c), \\
\lambda_{t+1} &= 1 - \varphi_\lambda + \varphi_\lambda \lambda_t + \sigma_\lambda [(1 - \varphi_\lambda + 2\varphi_\lambda \lambda_t)/2]^{1/2} \varepsilon_{\lambda t+1},
\end{aligned}
$$

where $-z^c_{t+1}$ is a consumption disaster. The random variable z^c_{t+1} is modeled as a Poisson mixture of gammas, $z^c_{t+1} \sim \Gamma(j^c_{t+1}, \theta_c)$; its central component $j^c_{t+1} \sim \text{Poisson}(h_\lambda \lambda_t)$ controls how many jumps of average size θ_c arrive per period of time. The mean of j^c_{t+1} 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 $\sigma^2_\lambda/2$, the degrees of freedom $2(1-\varphi_\lambda)/\sigma^2_\lambda$, and the serial correlation φ_λ. The normal shocks $\varepsilon_{ct+1} \sim \mathcal{N}(0,1)$ 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 a standard deviation of one and is orthogonal to ε_{ct+1} and ε_{dt+1}.

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

$$
\begin{aligned}
U_t &= [(1-\beta)c_t^\rho + \beta \mu_t(U_{t+1})^\rho]^{1/\rho}, \qquad (17) \\
\mu_t(U_{t+1}) &= [E_t(U^\alpha_{t+1})]^{1/\alpha}.
\end{aligned}
$$

I divide expression (17) by c_t, denote $u_t = U_t/c_t$ and $g^c_{t+1} = c_{t+1}/c_t$ and obtain

$$
u_t = [(1-\beta) + \beta \mu_t(u_{t+1} g^c_{t+1})^\rho]^{1/\rho}. \qquad (18)
$$

I solve a recursive problem that is a log-linear approximation of the Bellman equation (16)

$$
\log u_t \approx b_0 + b_1 \log \mu_t(g^c_{t+1} u_{t+1}), \qquad (19)
$$

where

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

I guess the value function

$$
\log u_t = u + p_\lambda \lambda_t \qquad (20)
$$

and compute

$$\log(u_{t+1}g^c_{t+1}) = (u + g^c) + \gamma_c\varepsilon_{ct+1} + p_\lambda\lambda_{t+1} - z^c_{t+1},$$

$$\log\mu_t(u_{t+1}g^c_{t+1}) = (u + g^c) + \frac{\alpha\gamma_c^2}{2} + \frac{\varphi_\lambda\lambda_t p_\lambda}{1 - \alpha p_\lambda\sigma^2_\lambda/2} - \frac{(1-\varphi_\lambda)\log(1 - \alpha p_\lambda\sigma^2_\lambda/2)}{\alpha\sigma^2_\lambda/2} - \frac{\theta_c h_\lambda\lambda_t}{1+\alpha\theta_c}, \quad (21)$$

$$\log(u_{t+1}g^c_{t+1}) - \log\mu_t(u_{t+1}g^c_{t+1}) = \gamma_c\varepsilon_{ct+1} + p_\lambda\lambda_{t+1} - z^c_{t+1} - \frac{\varphi_\lambda p_\lambda\lambda_t}{1 - \alpha p_\lambda\sigma^2_\lambda/2} - \frac{\alpha\gamma_c^2}{2}$$

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

I substitute (20) and (21) in equation (19) and solve for the parameters u and p_λ of the value function.

For p_λ, I have the following quadratic equation

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

where

$$\begin{aligned} C_2 &= \alpha\sigma^2_\lambda/2, \\ C_1 &= b_1\varphi_\lambda - 1 - A\alpha\sigma^2_\lambda/2, \\ C_0 &= A, \\ A &= -\frac{b_1 h_\lambda\theta_c}{1+\alpha\theta_c}. \end{aligned}$$

I solve it and choose the root that satisfies the requirement of stochastic stability (Hansen, 2012). Next, I find u

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

The pricing kernel is

$$\begin{aligned} \log m_{t,t+1} &= \log\beta + (\rho - 1)\log g^c_{t+1} + (\alpha - \rho)(\log(u_{t+1}g^c_{t+1}) - \log\mu_t(u_{t+1}g^c_{t+1})) \\ &= m + m_\lambda\lambda_t + m_{\varepsilon\lambda}\lambda_{t+1} + m_{\varepsilon c}\varepsilon_{ct+1} + m_z z^c_{t+1}, \end{aligned}$$

where

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

The model for the dividend growth is

$$
\log g^d_{t,t+1} = g^d + \varphi_d\gamma_c\varepsilon_{ct+1} + \gamma_d\varepsilon_{dt+1} - \varphi_d z^c_{t+1}.
$$

I guess that the price-dividend ratio is

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

and compute the implied process for the log return

$$
\begin{aligned}
\log r_{t,t+1} &= k_0 + k_1\log\delta_{t+1} + \log g^d_{t,t+1} - \log\delta_t\\
&= [k_0 + k_1 q_0 - q_0 + g^d] - q_\lambda\lambda_t + k_1 q_\lambda\lambda_{t+1} + \varphi_d\gamma_c\varepsilon_{ct+1} + \gamma_d\varepsilon_{dt+1} - \varphi_d z^c_{t+1}\\
&= r + r_\lambda\lambda_t + r_{\varepsilon\lambda}\lambda_{t+1} + r_{\varepsilon c}\varepsilon_{ct+1} + r_z z^c_{t+1} + r_{\varepsilon d}\varepsilon_{dt+1},
\end{aligned}
$$

where

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

I use the law of one price $E_t(r_{t,t+1}m_{t,t+1}) = 1$ to obtain two equations in two unknowns q_0 and q_λ

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

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

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

C.3 Solution of the restricted version of the model of Drechsler and Yaron (2011)

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_t^v, is

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

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 $E(v_t) = 1$ and $E(v_t^*) = 1/2$.

A representative agent has recursive preferences

$$
\begin{aligned}
U_t &= [(1-\beta) c_t^\rho + \beta \mu_t (U_{t+1})^\rho]^{1/\rho}, \qquad (22) \\
\mu_t(U_{t+1}) &= [E_t(U_{t+1}^\alpha)]^{1/\alpha}.
\end{aligned}
$$

I divide expression (22) by c_t, denote $u_t = U_t/c_t$ and $g_{t+1}^c = c_{t+1}/c_t$ and obtain

$$
u_t = [(1-\beta) + \beta \mu_t (u_{t+1} g_{t+1})^\rho]^{1/\rho}. \qquad (23)
$$

Next, I solve a recursive problem that is a log-linear approximation of the Bellman equation (16)

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

where

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

I guess the value function

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

and compute

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

$$\log \mu_t(u_{t+1} g^c_{t+1}) = u + g^c - \frac{(1-\varphi_v) v \log(1-\alpha p_v \sigma_v^2/2)}{\alpha\sigma_v^2/2} - \frac{(1-\varphi_v^*) v^* \log(1-\alpha p_v^* \sigma_v^{*2}/2)}{\alpha\sigma_v^{*2}/2}$$
$$+ \frac{\alpha}{2}\gamma_c^2 v_t + \frac{\varphi_v p_v v_t}{1-\alpha p_v \sigma_v^2/2} + \frac{p_v \theta_v h_v v_t}{1-\alpha p_v \theta_v} + \frac{\varphi_v^* p_v^* v_t^*}{1-\alpha p_v^* \sigma_v^{*2}/2} + p_v(1-\tilde{\varphi}_v) v_t^*,$$

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

I solve the following system of three equations in three unknowns u, p_v, and p_v^* in order to verify the guess of the value function

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

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

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

and obtain the model for the pricing kernel

$$\log m_{t,t+1} = \underbrace{\log\beta + (\rho-1) g^c + \frac{(\alpha-\rho)(1-\varphi_v) v \log(1-\alpha p_v \sigma_v^2/2)}{\alpha\sigma_v^2/2} + \frac{(\alpha-\rho)(1-\varphi_v^*) v^* \log(1-\alpha p_v^* \sigma_v^{*2}/2)}{\alpha\sigma_v^{*2}/2}}_{m}$$
$$+ \underbrace{\left(-\frac{(\alpha-\rho)\varphi_v p_v}{1-\alpha p_v \sigma_v^2/2} - \frac{(\alpha-\rho) p_v \theta_v h_v}{1-\alpha p_v \theta_v} - \frac{\alpha(\alpha-\rho)}{2}\gamma_c^2\right)}_{m_v} v_t$$
$$+ \underbrace{\left(-\frac{(\alpha-\rho) p_v^* \varphi_v^*}{1-\alpha p_v^* \sigma_2^{*2}/2} - (\alpha-\rho) p_v (1-\tilde{\varphi}_v)\right)}_{m_v^*} v_t^*$$
$$+ \underbrace{(\alpha-1)\gamma_c}_{m_{\varepsilon c}} v_t^{1/2}\varepsilon_{ct+1} + \underbrace{(\alpha-\rho) p_v}_{m_{\varepsilon v}} v_{t+1} + \underbrace{(\alpha-\rho) p_v^*}_{m_{\varepsilon v}^*} v_{t+1}^*.$$

The model for the log dividend growth is

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

I guess that the log price-dividend ratio $\log \delta_t$ is a linear function of states

$$\log \delta_t \quad = \quad q_0 + q_v v_t + q_v^* v_t^*$$

and derive the implied stochastic process for the log stock return

$$\begin{aligned}
\log r_{t,t+1} &= \kappa_0 + \kappa_1 \log \delta_{t+1} + \log g_{t,t+1}^d - \log \delta_t \\
&= \underbrace{k_0 + (k_1 - 1)q_0 + g^d}_{r} + \underbrace{(\mu_v - q_v)}_{r_v} v_t + \underbrace{(\mu_v^* - q_v^*)}_{r_v^*} v_t^* \\
&+ \underbrace{k_1 q_v}_{r_{\varepsilon v}} v_{t+1} + \underbrace{(k_1 q_v^* + \gamma_{\varepsilon v}^*)}_{r_{\varepsilon v}^*} v_{t+1}^* + \underbrace{\gamma_{\varepsilon d}}_{r_{\varepsilon d}} v_t^{1/2} \varepsilon_{dt+1} + \underbrace{\gamma_{zd}}_{r_z} z_{vt+1},
\end{aligned}$$

where

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

or in compact form

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

I use the law of one price $E_t[m_{t,t+1} r_{t,t+1}] = 1$ to obtain three equations in three unknowns q_0, q_v, q_v^*

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

$$m_v + r_v + \frac{m_{\varepsilon c}^2}{2} + \frac{r_{\varepsilon d}^2}{2} + \frac{(m_{\varepsilon v} + r_{\varepsilon v})\varphi_v}{1 - (m_{\varepsilon v} + r_{\varepsilon v})\sigma_v^2/2} + \frac{(m_{\varepsilon v} + r_{\varepsilon v} + r_z)\theta_v h_v}{1 - (m_{\varepsilon v} + r_{\varepsilon v} + r_z)\theta_v} = 0, \tag{24}$$

$$m_v^* + r_v^* + (1 - \tilde{\varphi}_v)(m_{\varepsilon v} + r_{\varepsilon v}) + \frac{(m_{\varepsilon v}^* + r_{\varepsilon v}^*)\varphi_v^*}{1 - (m_{\varepsilon v}^* + r_{\varepsilon v}^*)\sigma_v^{*2}/2} = 0. \tag{25}$$

I choose the stochastically stable roots (Hansen, 2012) of the cubic equation (24) in q_v and quadratic equation (25) in q_v^*.

D Term structures of risk in expected returns in the models of Bansal and Yaron (2004) and Wachter (2013)

In this appendix, I analyze whether alternative calibrations of the equilibrium models of Bansal and Yaron (2004) and Wachter (2013) can match the empirical properties of the term structures of risk in expected returns.

D.1 The term structure of risk in expected returns in the model of Bansal and Yaron (2004)

As it concerns term structures of risk, the main disagreement of the model in Bansal and Yaron (2004) with the data is in the sign of the slope of the term structure of ε_{xt+1} in expected returns. The sign in the data is negative and significant, whereas it is positive in the original calibration of the equilibrium model. I analyze whether an alternative calibration of the equilibrium model can deliver a negative sign.

Given the representation of stock returns in terms of economic states and shocks

$$\begin{aligned}
\log r_{t,t+1} &= r + r_x x_t + r_v v_t + r_{\varepsilon x} v_t^{1/2} \varepsilon_{xt+1} + r_{\varepsilon d} v_t^{1/2} \varepsilon_{dt+1} + r_{\varepsilon v} v_{t+1}, \\
x_{t+1} &= \varphi_x x_t + \gamma_x v_t^{1/2} \varepsilon_{xt+1}, \\
v_{t+1} &= (1-\varphi_v) + \varphi_v v_t + \sigma_v((1-\varphi_v + 2\varphi_v v_t)/2)^{1/2} \varepsilon_{vt+1},
\end{aligned}$$

the term structure of ε_{xt+1} in expected returns is defined as

$$\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{xt+1}) = (r_{\varepsilon_x}/\gamma_x + \mathcal{A}_x(\tau-1))\Delta_x,$$

where

$$\begin{aligned}
\mathcal{A}_x(1) &= r_x, \\
\mathcal{A}_x(\tau) &= r_x + \mathcal{A}_x(\tau-1)\varphi_x = r_x(1+\varphi_x+\cdots+\varphi_x^{\tau-1}).
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathcal{IER}(r_{t,t+1}, \varepsilon_{xt+1}) &= r_{\varepsilon x}\Delta_x/\gamma_x, \\
\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{xt+1}) &= (r_{\varepsilon x}/\gamma_x + r_x(1+\varphi_x+\cdots+\varphi_x^{\tau-2}))\Delta_x \quad \text{for } \tau > 1.
\end{aligned}$$

The term structure of ε_{xt+1} has a negative slope, $|\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{xt+1})| - |\mathcal{IER}(r_{t,t+1}, \varepsilon_{xt+1})| < 0$, and a positive level $\mathcal{IER}(r_{t,t+\tau}, \varepsilon_{xt+1}) \geq 0$ ($\tau = \overline{1,40}$), if

1. $r_{\varepsilon x}/\gamma_x > 0$,

2. $r_x < 0$,

3. $r_{\varepsilon x}/\gamma_x + r_x(1+\varphi_x+\cdots+\varphi_x^{\tau-2}) \geq 0$ for $\tau = \overline{1,40}$.

Appendix C.1 shows, that the parameters r_x and $r_{\varepsilon x}$ are the following functions of the deep parameters of the model:

$$\begin{aligned} r_x &= \mu_x + q_x(\kappa_1\varphi_x - 1) = \mu_x + \underbrace{\frac{\mu_x+\rho-1}{1-\kappa_1\varphi_x}}_{=q_x}(\kappa_1\varphi_x - 1) = 1-\rho, \\ r_{\varepsilon x} &= \kappa_1 q_x = \kappa_1\frac{\mu_x+\rho-1}{1-\kappa_1\varphi_x}. \end{aligned}$$

The second necessary condition $r_x < 0$ is equivalent to $\rho > 1$. The parameter ρ is related to the parameter of the elasticity of the intertemporal substitution IES $= (1-\rho)^{-1}$, and therefore $\rho > 1$ means IES < 0. A negative value of IES is economically implausible, and therefore, the necessary condition cannot be satisfied.

D.2 The term structure of risk in expected returns in the model of Wachter (2013)

As it concerns term structures of risk, the main disagreement of the model in Wachter (2013) with the data is in the sign of the level of the term structure of ε^c_{zt+1} in expected returns. The sign in the data is positive and significant, whereas it is negative in the original calibration of the equilibrium model. I analyze whether an alternative calibration of the equilibrium model can deliver a positive sign.

Given the representation of log stock returns in terms of economic states and shocks

$$\begin{aligned} \log r_{t,t+1} &= r + r_\lambda\lambda_t + r_{\varepsilon\lambda}\lambda_{t+1} + r_{\varepsilon c}\varepsilon_{ct+1} + r_z z^c_{t+1} + r_{\varepsilon d}\varepsilon_{dt+1}, \\ \lambda_{t+1} &= (1-\varphi_\lambda) + \varphi_\lambda\lambda_t + \sigma_\lambda((1-\varphi_\lambda+2\varphi_\lambda\lambda_t)/2)^{1/2}\varepsilon_{\lambda t+1}, \\ j^c_{t+1}|_{\lambda_t} &\sim \text{Poisson}(h_\lambda\lambda_t), \\ z_{t+1}|_{j^c_{t+1}} &\sim \text{Gamma}(j^c_{t+1},\theta_c), \end{aligned}$$

the term structure of ε^c_{zt+1} in expected returns is

$$\mathcal{IER}(r_{t,t+\tau},\varepsilon^c_{zt+1}) = r_z\Delta_z.$$

Appendix C.2 shows $r_z = -\varphi_d$, and therefore the term structure of disaster risk in expected returns has a positive level $\mathcal{IER}(r_{t,t+\tau}, \varepsilon^c_{zt+1}) \geq 0$, if $\varphi_d < 0$. The negative value of the parameter φ_d implies that the dividend claim is a levered consumption claim with a negative leverage parameter. This condition is economically implausible, and therefore the necessary conditions cannot be satisfied.

E From the term structures of risk in expected returns and dividend growth to the prices of dividend strips

Under different shock-identifying assumptions, I empirically characterize $\log E_t(g^d_{t,t+\tau})$ and $\log E_t(r_{t,t+\tau})$. I denote d_t a dividend payment at time t and notice that the price $p^d_t(\tau)$ of a dividend claim with maturity τ quarters is

$$p^d_t(\tau) = E_t(m_{t,t+\tau} \cdot d_{t+\tau}) = E_t\left(\frac{d_{t,t+\tau}}{r_{t,t+\tau}} \cdot m_{t,t+\tau} r_{t,t+\tau}\right) = E_t\left(\frac{d_{t,t+\tau}}{r_{t,t+\tau}}\right) \cdot E_t(m_{t,t+\tau} r_{t,t+\tau}) + cov_t\left(\frac{d_{t,t+\tau}}{r_{t,t+\tau}}, m_{t,t+\tau} r_{t,t+\tau}\right).$$

By the law of one price, $E_t(m_{t,t+\tau} r_{t,t+\tau}) = 1$, and therefore,

$$p^d_t(\tau) = E_t(m_{t,t+\tau} \cdot d_{t+\tau}) = E_t\left(\frac{d_{t,t+\tau}}{r_{t,t+\tau}}\right) + cov_t\left(\frac{d_{t,t+\tau}}{r_{t,t+\tau}}, m_{t,t+\tau} r_{t,t+\tau}\right).$$

While I can characterize empirically $E_t\left(\frac{d_{t,t+\tau}}{r_{t,t+\tau}}\right)$ and the associated term structure of risk, the conditional covariance $cov_t\left(\frac{d_{t,t+\tau}}{r_{t,t+\tau}}, m_{t,t+\tau} r_{t,t+\tau}\right)$ cannot be characterized without modeling the pricing kernel $m_{t,t+\tau}$. For example, if an agent, who prices assets, has a log utility, than $m_{t,t+1} = 1/r_{t,t+1}$, and therefore the covariance term is equal to 0. However, in the presence of assumptions about risk preferences, internal consistency of the state-space model (1)-(2) implies that the joint dynamics of consumption, returns, and dividends must feature cross-equation restrictions. The implied parameter restrictions drastically change interpretation of the estimated term structures of risk: they are not any longer stylized facts but implications of the specific equilibrium pricing kernel. As a result, I cannot characterize prices on dividend strips $p^d_t(\tau)$ without further assumptions.