Skewness Risk and Bond Prices

Francisco Ruge-Murcia*

July 2015

Abstract

This paper uses extreme value theory to the study the implications of skewness risk for nominal loan contracts in a production economy. Productivity and inflation innovations are drawn from generalized extreme value (GEV) distributions. The model is solved using a third-order perturbation and estimated by the simulated method of moments. Results show that the U.S. data reject the hypothesis that innovations are drawn from normal distributions and favor instead the alternative that they are drawn from asymmetric distributions. Estimates indicate that skewness risk accounts for 12 percent of the bond risk premia, reduces bond yields by approximately 55 basis points, and involves a risk-adjustment factor worth 2.05 cents for a 5-year bond that pays 1 dollar at maturity.

JEL Classification: G12, E43, E44

Key Words: Extreme value theory, GEV distribution, skewness risk, skewness smile, nonlinear asset pricing, simulated method of moments.

*I received helpful comments from Julien Bengui, Taeyoung Doh, Larry Epstein, Silvia Gonçalves, Jinill Kim, Federico Ravenna, anonymous referees, and seminar participants at Boston College, Carleton University, and the Getulio Vargas Foundation. This research received the financial support of the Social Sciences and Humanities Research Council of Canada. Correspondence: Department of Economics, McGill University, 855 Sherbrooke Street West, Montréal, Québec, Canada, H3A 2T7. E-mail: francisco.ruge-murcia@mcgill.ca
1 Introduction

This paper uses extreme value theory to study the pricing of nominal loan contracts in an environment where agents face skewness risk. Extreme value theory is a branch of statistics concerned with extreme deviations from the median of probability distributions. Results due to Fisher and Tippett (1928) and Gnedenko (1943) show that extreme realizations can be characterized by asymmetric statistical distributions with different rates of decay in their long tail. Since agents are exposed to large realizations from this tail, they are subject to skewness risk. Then, the issue is what effect this source of risk has on households’ behavior and on asset prices, above and beyond the variance risk examined in previous literature. In particular, I focus on the relation between skewness risk and bond prices in a production economy populated by agents with recursive preferences (Epstein and Zin, 1989).

I show that a third-order perturbation of the policy functions that solve the dynamic model explicitly captures the contribution of skewness to bond prices, yields, and risk premia, and permits the construction of model-based estimates of the effects of skewness risk. The model is estimated by the simulated method of moments (SMM) using quarterly U.S. data. Among the estimated parameters are those of the distributions that generate inflation and productivity innovations. Based on these estimates, I statistically show that the data reject the hypothesis that shock innovations are drawn from normal distributions and favor instead the alternative that they are drawn from asymmetric distributions. In particular, the data prefer a specification where inflation innovations are drawn from a positively skewed distribution and productivity innovations are drawn from a negatively skewed distribution. Thus, in a statistical sense, skewness is significant.

Following earlier work on extreme value theory, I model the shock innovations using the generalized extreme value (GEV) distribution due to Jenkinson (1955), which encompasses the limiting distributions identified by Fisher and Tippett (1928). I also study a version of the model with skew normal innovations. These two distributions have flexible forms, permit both positive and negative skewness, and are shown to fit the data better than the normal distribution. In particular, the models with asymmetric innovations can account for the skewed realizations that characterize consumption growth and bond yields. For the preferred specification with GEV innovations, results indicate skewness risk reduces yields by, e.g., 0.59 percentage points (pp) for the 3-month bond and 0.54 pp for the 5-year bond; that bond prices contain a skewness-risk adjustment factor that ranges from 0.15 cents for

---

1In the model there is a rental market for capital but no stock market to trade claims on capital returns per se. I focus on bond prices in order to keep the scope of this project manageable, but examine stock returns in a related paper (Ruge-Murcia, 2014).

[1]
the 3-month bond to 2.05 cents for the 5-year bond, and respectively constitute 0.15% and 2.66% of their price; and that the contribution of skewness risk to the bond premia varies from 12.4% for the 6-month bond to 11.6% for the 5-year bond. Hence, the effects of skewness risk are economically significant as well.

Previous literature that uses extreme value theory follows a partial equilibrium approach where the risk manager takes returns as given (see the survey by Rocco, 2011). One difficulty with this approach is that returns are not independently and identically distributed (i.i.d.) as required by the theory, and a pre-whitening procedure has to be applied to the data beforehand. Instead, this paper uses a general equilibrium approach that makes the more plausible assumption that innovations to structural shocks are i.i.d. More generally, this paper contributes to the literature concerned with the role of higher-order moments in asset pricing. Kraus and Litzenberger (1976, 1983) extend the capital asset pricing model (CAPM) to incorporate the effect of skewness on valuations. Harvey and Siddique (2000) study the role of the co-skewness with the aggregate market portfolio and find a negative correlation between co-skewness and mean returns. Using the CAPM, Kapadia (2006) and Chang et al. (2013) find that skewness risk has a negative effect on excess returns in a cross-section of stock returns and options. Colacito et al. (2013) study the implications of time-varying skewness in an endowment economy. Compared with this literature, I employ an equilibrium asset-pricing model where consumption is endogenous and structural estimation attempts to reconcile bond prices with macroeconomic aggregates.

As the research on disasters (e.g., Barro, 2006), this paper is concerned with the asymmetry of consumption. I provide statistical evidence that even in the relatively calm, post-WWII U.S., agents face the possibility of substantial decreases in consumption. These decreases are not as dramatic as disasters but they occur more frequently, are primarily associated with the business cycle, and have non-negligible effects on the bond market. This finding is important because work based on option prices (Backus et al., 2011) suggests that more frequent, moderate consumption “disasters” of magnitude comparable to recessions play an important role in explaining U.S. stock returns. Moreover, while that literature focuses only on consumption, this paper also considers asymmetries in inflation that are important for pricing nominal assets. Andreasen (2012) and Gourio (2012) study asset pricing in calibrated disaster economies using the shock formulation in Barro (2006), but applied to structural disturbances rather than to consumption directly. This paper complements their work by formally estimating the parameters of a production economy and providing statistical evidence on the degree of skewness of the innovations and the magnitude of skewness risk.

The paper is organized as follows. Section 2 describes a production economy subject to extreme productivity and inflation shocks. Section 3 describes the data and econometric
method used to estimate the model and reports parameter estimates and measures of fit. Section 4 uses impulse-response analysis to study the dynamics of the model. Section 5 quantifies the contribution of skewness risk to bond prices, yields, and risk premia. Finally, Section 6 concludes.

2 Bond Pricing

This section describes a production economy subject to extreme productivity and inflation shocks, and characterizes bond pricing in this environment. Time is assumed to be discrete.

2.1 The Economy

Consumers are identical, infinitely-lived, and their total number is normalized to be 1. The representative consumer has recursive preferences (Epstein and Zin, 1989),

\[ U_t = \left( (1 - \beta) (c_t)^{1-1/\psi} + \beta \left( E_t \left( U_{t+1}^{1-\gamma} \right) \right)^{(1-1/\psi)/(1-\gamma)} \right)^{1/(1-1/\psi)}, \]

where \( \beta \in (0, 1) \) is the discount factor, \( c_t \) is consumption, \( E_t \) is the expectation conditional on information available at time \( t \), \( \psi \) is the intertemporal elasticity of substitution, and \( \gamma \) is the coefficient of relative risk aversion. As it is well known, recursive preferences decouple elasticity of substitution from risk aversion when \( \gamma = 1/\psi \), and encompass preferences with constant relative risk aversion when \( \gamma = 1/\psi \). Previous literature that employs recursive preferences in consumption-based models of bond pricing include Epstein and Zin (1991), Gregory and Voss (1991), Piazzesi and Schneider (2007), Le and Singleton (2010), van Binsbergen et al. (2012), Andreasen (2012), Rudebusch and Swanson (2012), and Doh (2013).

The consumer is endowed with \( h \) units of time (say, hours) per period, which she can supply in a competitive labor market.

Financial assets are physical capital and zero-coupon nominal bonds with maturities \( \ell = 1, \ldots, L \). Bonds are equally liquid regardless of their maturity and can be costlessly traded in a secondary market. The capital stock follows the law of motion

\[ k_{t+1} = (1 - \delta)k_t + x_t, \]

where \( k_t \) is capital, \( x_t \) is investment, and \( \delta \in (0, 1] \) is the depreciation rate. Investment beyond that required to replace depreciated capital involves a convex cost proportional to the capital stock. This cost is represented using the function

\[ \Phi_t = \frac{\phi}{2} \left( \frac{x_t}{k_t} - \delta \right)^2 k_t, \]
where $\phi \geq 0$ is a constant parameter. The consumer’s budget constraint is

$$c_t + x_t + \sum_{\ell=1}^{L} \frac{Q^\ell_t B^\ell_t}{P_t} = \frac{W_t h_t}{P_t} + \frac{D_t k_t}{P_t} + \sum_{\ell=1}^{L} \frac{Q^{\ell-1}_t B^{\ell-1}_t}{P_t} - \Phi_t,$$

(2)

where $Q^\ell_t$ and $B^\ell_t$ are, respectively, the nominal price and quantity of bonds with maturity $\ell$, $Q^0_t = 1$, $W_t$ is the nominal wage, $h_t$ is hours worked, $D_t$ is the nominal rental rate of capital, and $P_t$ is the price level. Prices are denominated in terms of a unit called money but the economy is cashless otherwise. The normalization $Q^0_t = 1$ implies that bonds pay one unit of money at maturity.

Output is produced by identical firms whose total number is normalized to be 1. The representative firm rents profit-maximizing quantities of capital and labor from consumers, and combines them using the stochastic technology

$$y_t = z_t k^\alpha_t h^{1-\alpha}_t,$$

(3)

where $y_t$ is output, $z_t$ is an aggregate productivity shock, and $\alpha \in (0, 1]$ is a parameter.

Economic agents take as given the joint process of aggregate inflation and productivity. Define

$$A_t = \begin{bmatrix} \ln (\pi_t) - \ln(\pi) \\ \ln (z_t) - \ln(z) \end{bmatrix},$$

(4)

where $\pi_t = P_t/P_{t-1}$ is the gross inflation rate and variables without time subscript denote steady-state values. Assume that the time-series process of inflation and productivity can be well approximated by the vector autoregression (VAR),

$$A_t = \rho A_{t-1} + \epsilon_t,$$

(5)

where $\rho$ is a $2 \times 2$ matrix with roots outside the unit circle, and

$$\epsilon_t = \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix}$$

(6)

is a vector of innovations. The VAR is a convenient device to represent statistically mechanisms by which inflation (or, more generally, monetary policy) can affect real variables and vice versa. Although the VAR is admittedly a reduced-form, it has the advantages of being flexible and general. For example, it includes as a special case the specification in van Binsbergen et al. (2012) where inflation depends on current and past productivity innovations. The innovations in $\epsilon_t$ are assumed to be contemporaneously independent and serially uncorrelated, with mean zero, constant conditional variance, and non-zero skewness. The latter assumption relaxes the usual restriction of zero skewness implicit in most of the previous
literature (for example, through the assumption of normal innovations) and allows me to examine the relation between skewness risk and bond prices. In the empirical part of this paper, I consider two asymmetric distributions for the innovations, namely the generalized extreme value (GEV) and the skew normal distributions.

In equilibrium, bonds are in zero net supply and the labor and goods markets clear:

\[
B^\ell_t = 0, \\
h_t = h, \\
c_t + x_t = y_t - \Phi_t,
\]

where \(h\) is the time endowment and \(\ell = 1, 2, \ldots, L\).

2.2 Bond Yields

Following the literature, the gross yield of the \(\ell\)-period bond is defined as

\[
i^\ell_t = (Q^\ell_t)^{-1/\ell}.
\]  

The risk premium is derived from the Euler equation

\[
Q^\ell_t = \beta E_t \left( \left( \frac{v_{t+1}}{w_t} \right)^{1/\psi - \gamma} \left( \frac{c_{t+1}}{c_t} \right)^{-1/\psi} \frac{Q^{\ell-1}_{t+1}}{\pi_{t+1}} \right) \\
= Q^1_tE_t (Q^{\ell-1}_{t+1}) + R_{t,t},
\]

where \(v_t = \max U_t\) is the value function, \(w_t = E_tv_{t+1}\) is certainty-equivalent future utility, and

\[
R_{t,t} \equiv \beta \text{cov}_t \left( \left( \frac{v_{t+1}}{w_t} \right)^{1/\psi - \gamma} \left( \frac{c_{t+1}}{c_t} \right)^{-1/\psi} \frac{Q^{\ell-1}_{t+1}}{\pi_{t+1}} \right)
\]  

(8)

is the risk premium. The risk premium depends on the maturity and may be positive or negative according to the sign of the covariance between the pricing kernel and \(Q^{\ell-1}_{t+1}/\pi_{t+1}\). This observation highlights the fact that, in general, consumption-based asset-pricing models do not restrict the sign or monotonicity of the bond risk premia. The same is true in the market-segmentation hypothesis by Culbertson (1957) and the preferred-habitat theory by Modigliani and Sutch (1966), but they rely on a strong preference by investors for particular maturities and implicitly rule out arbitrage.

The premium is negative when the covariance in (8) is negative. A negative premium is a discount on the longer-term bond in the sense that buying a \(\ell\)-period bond at time \(t\) at price \(Q^\ell_t\) is cheaper than buying a one-period bond at time \(t\) at price \(Q^1_t\) and a \((\ell - 1)\)-period bond at time \(t + 1\) at expected price \(Q^{\ell-1}_{t+1}\). Up to an approximation, this implies that the yield
of the $\ell$-period bond is larger than the weighted average yield of the 1- and $(\ell - 1)$-period bonds and the yield curve is, therefore, upward slopping.\footnote{This mechanism is different from the one outlined by Hicks (1939, ch. 13) where long-term bonds are “less liquid” than short-term bonds and, hence, a premium is required to induce traders to hold the former. Bansal and Coleman (1996) assume that short-term bonds provide indirect transaction services and the compensation for those services lowers the nominal return of short-term bonds compared with long-term bonds.}

### 2.3 Solution

Since this model does not have an exact analytical solution, I use a perturbation method to obtain an approximate solution. The method involves taking a third-order expansion of the policy functions around the deterministic steady state and characterizing the local dynamics (see Jin and Judd, 2002). An expansion of (at least) third-order is necessary to capture the effect of skewness in the policy functions, to identify the preference parameters, and to generate a time-varying premium.\footnote{As it is well known, first-order solutions feature certainty equivalence and imply that traders are indifferent to the higher-order moments of the shocks. Second-order solutions capture the effect of the variance, but not of the skewness, on the policy functions and imply a constant risk premium. In order to efficiently compute the solution for the prices of long-term bonds, I use the perturbation-on-perturbation method due to Andreasen and Zabczyk (2015) and implement it using an adaptation of their codes.} Caldera et al. (2012) show that for models with recursive preferences, a third-order perturbation is as accurate as projection methods (e.g., value function iteration) in the range of interest while being much faster computationally. The latter is an important advantage for this project because estimation requires solving and simulating the model in each iteration of the routine that optimizes the statistical objective function.

A policy function takes the general form $f(s_t, \sigma)$ where $s_t$ is a vector of state variables and $\sigma$ is a perturbation parameter. For this model, the state variables are aggregate capital, productivity, and inflation. That is, $s_t = [k_t, z_t, \pi_t]'. The goal is to approximate $f(s_t, \sigma)$ using a third-order polynomial expansion around the deterministic steady state where $s_t = s$ and $\sigma = 0$. Using tensor notation, this approximation can be written as

$$
[f(s_t, \sigma)]^j = [f(s, 0)]^j + [f_s(s, 0)]^j_\alpha [(s_t - s)]^\alpha \\
+ (1/2)[f_{ss}(s, 0)]^j_\alpha \beta [(s_t - s)]^\alpha [(s_t - s)]^\beta \\
+ (1/6)[f_{sss}(s, 0)]^j_\alpha \beta \gamma [(s_t - s)]^\alpha [(s_t - s)]^\beta [(s_t - s)]^\gamma \\
+ (1/2)[f_{\sigma s}(s, 0)]^j_\alpha \beta [\sigma][\sigma] \\
+ (1/2)[f_{\sigma \sigma}(s, 0)]^j_\alpha [(s_t - s)]^\alpha [\sigma][\sigma] \\
+ (1/6)[f_{\sigma \sigma \sigma}(s, 0)]^j_\alpha [\sigma][\sigma][\sigma],
$$

\footnote{Caldera et al. (2012) show that for models with recursive preferences, a third-order perturbation is as accurate as projection methods (e.g., value function iteration) in the range of interest while being much faster computationally. The latter is an important advantage for this project because estimation requires solving and simulating the model in each iteration of the routine that optimizes the statistical objective function.}
where \( f(s_t, \sigma)^j \) refers to the \( j \)-th variable in the model (say, the price of the 1-year bond) and elements like \( f_s(s, 0)^j \) are coefficients that depend on structural parameters. As we can see, the policy function includes linear, quadratic, and cubic terms in the state variables, one constant and one time-varying term in the variance, and one constant term in the skewness. The latter term, that is, \( (1/6)[f_{\sigma\sigma}(s, 0)]^j[\sigma][\sigma][\sigma] \), is a risk-adjustment factor due to the skewness of the shocks and it plays a central role in this project. In the special case where the distribution of the innovations is symmetric—and, hence, skewness is zero—this term is zero. In the more general case where the distribution is asymmetric, this term may be positive or negative depending on the sign of the skewness and the values of other structural parameters. In quantifying the contribution of skewness risk to bond prices, yields, and risk premia, I use the observation that the policy functions make explicit the dependence of the variables on the skewness of productivity and inflation.

An alternative solution method involves assuming that the arguments inside the expectations operator in the Euler equation are jointly lognormal and conditionally homoskedastic in order to obtain a linear pricing function with a constant risk-adjustment factor that is proportional to the variance (see, e.g., Jerman, 1998). Instead, the pricing function under (9) is nonlinear, and the risk-adjustment factor is time-varying and depends on both the variance and the skewness of the innovations. In related research, Martin (2013) relaxes the assumption of lognormality in Espetstein and Zin (1991) and expresses the model variables as polynomial functions of consumption growth cumulants. Andreasen (2012) shows that policy functions for risk premia depend only on the terms proportional to the variance and skewness and, thus, asymmetric innovations affect the term premia up to a constant. The latter observation holds more generally for the other variables in this model because the skewness of the innovations is time invariant.

3 Estimation

This section describes the data and method used to estimate the model, reports estimates of the parameters, and evaluates the fit of the model in terms of unconditional moments and the yield curve.

3.1 Data

The model is estimated using 224 quarterly observations of the growth rate of consumption, the growth rate of investment, the inflation rate, the 3-month and 6-month Treasury-Bill rates, and the 1-year, 2-year 3-year, 4-year, and 5-year yields on discount bonds constructed
by Fama and Bliss. The sample period is from 1959Q1 to 2014Q4. Except for the Fama-Bliss series, the raw data were taken from the FRED database available at the Web site of the Federal Reserve Bank of St. Louis (www.stls.frb.org). The Fama-Bliss series are yields on fully taxable, non-callable, non-flower discount bonds and were obtained from the Center for Research in Security Prices (www.crsp.com).

Consumption is measured by personal consumption expenditures on non-durable goods and services. Investment is measured by the sum of private nonresidential fixed investment and personal consumption expenditures on durable goods. Consumption and investment were converted into real per-capita terms by dividing by the consumer price index (CPI) for all urban consumers and by the estimate of civilian noninstitutional population produced by the Bureau of Labor Statistics (BLS). Civilian noninstitutional population is defined as persons older than 15 years of age who are not inmates of institutions or on active duty in the Armed Forces. The inflation rate is measured by the gross quarterly percentage change in the CPI. All yields are quoted as a net annual rate and were transformed into a gross quarterly rate. The raw data on interest rates, the consumer price index, and population are monthly and were converted to a quarterly frequency by averaging the observations for the three months in each quarter. Except for the interest rates and population, the raw data are seasonally adjusted at the source.

3.2 Distributions

Since extreme value theory is concerned with unusual events, the choice of statistical distribution can be important because different distributions have different rates of decay in their long tail. Under the Fisher-Tippett theorem (Fisher and Tippett, 1928) the maxima of a stochastic series converges in distribution to one of three possible extreme value distributions, namely the Gumbel, the Fréchet, and the Weibull distributions. Jenkinson (1955) proposes a generalized extreme value distribution (GEV) that can represent the above-mentioned distributions in an unified way and that for this reason is widely used in extreme value analysis. The GEV distribution is characterized by three parameters: a location, a scale and a shape parameter. The inverse of the shape parameter is known as the tail index because it controls the thickness of the tail of the distribution. Depending on whether the shape parameter is zero, larger than zero, or smaller than zero, the distribution corresponds to either the Gumbel, the Fréchet, or the Weibull distribution. The GEV distribution is asymmetric, allows for both positive and negative skewness, and its support may be bounded above or below.

As a complement to the GEV, I estimate the model using the skew normal distribution. Since the skew normal does not belong to the family of extreme value distributions, it is not
typically used in extreme value analysis. However, this asymmetric distribution is attractive for two reasons. First, as the GEV distribution, it is a three-parameter distribution (in this case, a location, a scale, and a correlation parameter) that can accommodate both positive and negative skewness. Skewness is positive when the correlation parameter is positive and vice versa. Second, it nests the normal distribution as a special case when the correlation parameter is zero. This means that it is straightforward to test the hypothesis that innovations are drawn from a normal (symmetric) distribution against the alternative that they are drawn from a skew normal (asymmetric) distribution. This simply involves performing the two-sided $t$-test of the hypothesis that the correlation parameter is zero against the alternative that it is different from zero. A drawback of the skew normal distribution is that its skewness is bounded between $-0.995$ and $+0.995$, corresponding respectively to the limit cases where the correlation parameter is $-1$ and $+1$. Below, I report that this constraint does not bind for productivity innovations but that it does for inflation innovations.

Finally, I estimate a benchmark version of the model where innovations are drawn from a normal distribution.

### 3.3 Method

The model is estimated by the simulated method of moments (SMM). The SMM estimator minimizes the weighted distance between the unconditional moments predicted by the model and those computed from the data, where the former are computed on the basis of artificial data simulated from the model. Lee and Ingram (1991) and Due and Singleton (1993) show that SMM delivers consistent and asymptotically normal parameter estimates under general regularity conditions. Other estimators (e.g., maximum likelihood) have these desirable properties, but they tend to be less computationally efficient and not as robust to misspecification. $^4$ Ruge-Murcia (2012) explains in detail the application of SMM for the estimation of non-linear dynamic models and provides Monte-Carlo evidence on its small-sample properties.

More formally, define $\theta \in \Theta$ to be a $q \times 1$ vector of structural parameters with $\Theta \subset \mathbb{R}^q$, $\mathbf{m}_t$ to be a $p \times 1$ vector of empirical observations on variables whose moments are of our interest, and $\mathbf{m}_s(\theta)$ to be the synthetic counterpart of $\mathbf{m}_t$ whose elements are obtained from the stochastic simulation of the model. The SMM estimator, $\hat{\theta}$, is the value that solves

$$\min_{\theta} \mathbf{M}(\theta) \mathbf{W} \mathbf{M}(\theta) \mathbf{W},$$

$^4$For example, see the evidence reported in Ruge-Murcia (2007) for linear general equilibrium models.
where

\[ M(\theta) = (1/T) \sum_{t=1}^{T} m_t - (1/T^\lambda) \sum_{t=1}^{\lambda T} m_t(\theta), \]

\( T \) is the sample size, \( \lambda \) is a positive constant, and \( W \) is a \( q \times q \) weighting matrix. Under the regularity conditions in Duffie and Singleton (1993),

\[ \sqrt{T}(\hat{\theta} - \theta) \rightarrow N(0,(1 + 1/\lambda)(J'W^{-1}J)^{-1}J'W^{-1}SW^{-1}J(J'W^{-1}J)^{-1}), \]

(11)

where

\[ S = \lim_{T \to \infty} Var \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} m_t \right), \]

(12)

and \( J = E(\partial m_t(\theta)/\partial \theta) \) is a finite Jacobian matrix of dimension \( p \times q \) and full column rank.

In this application, the weighting matrix is the diagonal of the inverse of the matrix with the long-run variance of the moments. That is, \( W \) has diagonal entries equal to those of \( S^{-1} \) and non-diagonal entries equal to zero. This weighting matrix is attractive because it makes the objective function scale-free and gives a larger weight to the moments that are more precisely estimated.\(^5\)

\( S \) is computed using the Newey-West estimator with a BARLETT kernel and bandwidth given by the integer of \( 4(T/100)^{2/9} \). The number of simulated observations is 20 times larger than the sample size (that is, \( \lambda = 20 \)), but results are robust to using other values of \( \lambda \). The simulation starts at the mean of the ergodic distribution of the variables and, in order to attenuate the effect of starting values on the results, the simulated sample contains 100 additional “training” observations that are discarded for the purpose of computing the moments. The dynamic simulations of the non-linear model are based on the pruned version of the solution. I use the pruning scheme in Andreasen and Zabczyk (2015), but using the unpruned solution delivers similar results as those reported here.

The estimated parameters are the discount factor (\( \beta \)), the intertemporal elasticity of substitution (\( \psi \)), the coefficient of relative risk aversion (\( \gamma \)), the capital adjustment-cost parameter (\( \phi \)), the rate of inflation in the deterministic steady state (\( \pi \)), and the parameters of the VAR and of the innovation distributions of productivity and inflation. The unconditional mean of the productivity process (\( z \)) and time endowment (\( h \)) are normalized to be 1. The production function parameter (\( \alpha \)) is fixed to 0.35, which is the average capital share of total income in the National Income and Product Accounts (NIPA) during the sample period. The depreciation rate is fixed to 0.25, which corresponds to an annual rate of 10%. The moments used to estimate these parameters are the variances, covariances, first-order autocovariances, and third-order moments of all series, plus the unconditional means of inflation and the nominal interest rates.

\(^5\)See Ruge-Murcia (2012, sec. 4.3) for a comparison of the efficiency of the SMM estimator under this and alternative weighting matrices, including the optimal one.
3.4 Identification

Local identification requires

\[ \text{rank}\left\{ E\left( \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right) \right\} = q, \]  

where (with some abuse of the notation) \( \theta \) is the point in the parameter space \( \Theta \) where the rank condition is evaluated. I verified that this condition is satisfied at the optimum \( \hat{\theta} \) for all versions of the model.

3.5 Estimates

Estimates and asymptotic standard errors of the parameters are reported in Table 1. Notice that estimates of the preference parameters are similar across distributions. The discount factor ranges from 0.971 for the skew normal distribution to 0.977 for the normal distribution. The intertemporal elasticity of substitution (IES) is about 0.046 for the two asymmetric distributions and 0.055 for the normal distribution. These estimates are statistically different from both 0 and 1, and in line with values reported in previous literature. For example, Hall (1988) reports estimates between 0.07 and 0.35; Epstein and Zin (1991) report estimates between 0.18 and 0.87 depending on the measure of consumption and instruments used; and Vissing-Jørgensen (2002) reports estimates between 0.30 and 1 depending on the households’ asset holdings.

The coefficient of relative risk aversion is 34 for the GEV and skew normal distributions and 28 for the normal distribution. These estimates are smaller than the estimate of 79 reported by van Binsbergen et al. (2010) and values employed by calibration studies that use Epstein-Zin preferences (e.g., Rudebusch and Swanson, 2012). A hypothesis of interest is \( \gamma = 1/\psi \), meaning that the coefficient of relative risk aversion equals the inverse of the IES and, thus, consumers’ preferences are well approximated by a specification with constant relative risk aversion (CRRA). Since the \( t \)-test delivers \( p \)-values above 0.05 in all cases, the null hypothesis cannot be rejected at the 5% level. However, in interpreting this result, it is important to keep in mind that Monte-Carlo results reported by Ruge-Murcia (2012) suggest that for moderately persistent shock processes, statistical inference based on the SMM estimator tends to be conservative in small samples.

The capital adjustment cost parameter is 19 for the GEV and skew normal distributions, and 13 for the normal distribution. These estimates imply elasticities of investment with respect to the price of installed capital of \( 1/\delta \phi = 2.1 \) and 3.1, respectively. These elasticities are one order of magnitude larger than the values of 0.34 and 0.28 reported by Kim (2000) and Christiano, Eichenbaum, and Evans (2001), respectively, and play a key role in accounting...
for the volatility of investment growth in the data. For instance, the model with GEV innovations predicts a standard deviation of 3.7 percentage points for quarterly investment growth, which is quantitatively close to the 4.1 percentage points in the U.S. data.

The shape parameter (GEV) and the correlation parameter (skew normal) of productivity innovations are negative and imply a skewness equal to $-0.9$ (GEV) and $-0.31$ (skew normal). Note that a negative shape parameter is a necessary condition for the negative skewness of a GEV distribution, while a negative correlation parameter is a sufficient condition for the negative skewness of a skew normal distribution. Since the correlation parameter is statistically different from zero, the null hypothesis that productivity innovations are drawn from a normal distribution can be rejected in favor of the alternative that they are drawn from a skew normal distribution ($p$-value < 0.001). The result that the shape parameter is negative and statistically different from zero means that among extreme value distributions the one that best describes productivity innovations is the Weibull distribution.

Figure 1 plots the estimated probability density function (PDF) of productivity innovations under the three models. The negative skewness of the GEV and skew normal distributions (thick line) is apparent in this figure: the distributions have more mass in the left tail, and less mass in the right tail, than a normal distribution with the same standard deviation (thin line). Thus, loosely speaking, large negative productivity surprises are more likely than positive ones and a bond buyer faces the risk of large unexpected decreases in productivity and, hence, output and consumption, during the bond holding period.

The shape parameter (GEV) and the correlation parameter (skew normal) of inflation innovations are positive and imply a skewness equal to 1.5 (GEV) and 0.995 (skew normal). A positive shape (resp. correlation) parameter is a sufficient condition for the positive skewness of a GEV (resp. skew normal) distribution. Since the correlation parameter is on the upper boundary of the parameter space, the skewness of 0.995 is the largest possible for a skew normal distribution. The results that the shape parameter is positive but not statistically different from zero means that among extreme value distributions the one that best describes inflation innovations is the Gumbell distribution. The finding that inflation innovations follow a Gumbell distribution while productivity innovations follow a Weibull distribution illustrates the advantage of using the flexible GEV distribution in this project.

Figure 2 plots the PDF of inflation innovations under the three models. The PDF of the positively-skewed GEV and skew normal distributions (thick line) have more probability mass in the right tail, and less mass in the left tail, than a normal distribution with the same standard deviation (thin line). Thus, large positive inflation surprises can happen sometimes, but large negative ones are unlikely. This implies that, for a given variance, the buyer of a nominal bond faces the risk of extreme realizations from the right tail of the inflation
distribution, which reduce the real payoff of the bond with maturity equal to 1 period and the real price of bonds with maturity longer than 2 periods. The use of a vector autoregression as a modeling device to represent the joint process of productivity and inflation induces additional dynamic effects. In particular, since the coefficient of lagged inflation in the productivity equation is negative and quantitatively large (albeit not significant), a current positive inflation shock induces future decreases in productivity and, consequently, output, and consumption. As pointed out by Piazzesi and Schneider (2007), a negative correlation between current inflation and future consumption makes long-term bonds risky in the sense that their real payoff is low when consumption is low. As we will see below, the positive skewness of inflation innovations magnifies this effect.

Table 2 compares the unconditional moments predicted by the model with those of the U.S. data using two standard measures of fit. The measures are the root mean-square error (RMSE) and the mean absolute error (MAE). Notice that the model with GEV innovations delivers the smallest RMSE and MAE, followed by the model with skew normal innovations. The main reason why these two models outperform the model with normal innovations is that they predict unconditional skewness for investment and interest rates that are quantitatively closer to those found in the data (see Section 3.7 below).

In summary, the estimates and statistical tests reported here show that the data reject the hypothesis that productivity and inflation innovations are drawn from normal distributions and favor instead the alternative that they are drawn from asymmetric distributions. In particular, the data prefer a specification where productivity innovations are drawn from a negatively skewed distribution and inflation innovations are drawn from a positively skewed distribution.

### 3.6 The Yield Curve

Figure 3 compares the yield curves predicted by model with the one computed from the data. The horizontal axis are maturities and the vertical axis are yields in percent per year. Notice that all predicted yield curves are upward sloping regardless of the shock distribution. Consumption-based models of the term structure predict a downward-sloping yield curve when long-term bonds act as insurance. For example, in Backus et al. (1989) and den Haan (1995) interest rates are low during recessions when consumption is also low. In the data, however, the correlation between consumption and interest rates is negative and statistically significant. For instance, the correlation between the detrended consumption of non-durable goods and services, and the 3-month, 1-year, and 5-year interest rates on discount bonds is $-0.19$, $-0.20$, and $-0.29$, respectively, and the hypothesis that the correlation is zero can be
rejected at the 5% significance level in all cases. Instead, this model captures the mechanism in Piazzesi and Schneider (2007) whereby positive inflation surprises induce future decreases in consumption and, thus, the real payoff on long-term bonds is low when consumption is low. As a result, consumers demand a premium for holding long-term bonds and the term structure is upward sloping. The positive skewness of inflation innovations amplifies this mechanism because consumers face the possibility of extreme realizations from the upper tail of the distribution of inflation innovations that induce large drops in future productivity and consumption, and large increases in the marginal utility. In contrast, large realizations from the lower tail that induce the converse effects are unlikely (see Figure 2).

Figure 3 also shows that the yield curves predicted by the models with asymmetric shocks (GEV and skew normal) are steeper, and fit the U.S yield curve better, than the one predicted by the model with normal shocks. This impression is statistically confirmed by the RMSE and MAE reported in Panel B of Table 2. For example, the RMSE of the models with GEV and normal innovations are 0.165 and 0.237, respectively. Since parameter estimates, other than those of the innovation distributions, are quantitatively similar across the three specifications (see Table 1), this suggests that skewness risk accounts for a substantial part of the difference between the yield curves in Figure 3. Quantitative evidence on the contribution of skewness risk to bond yields is reported in Section 5.2.

3.7 Unconditional Skewness

Skewness is a prominent feature of U.S. economic data. Figure 4 plots the distributions of the variables used to estimate this model and it clearly shows that consumption and investment growth are negatively skewed, while inflation and all bond yields are positively skewed. (The rates in this figure are annual rates.) All large consumption and investment growth drops are associated with recessions. It is also apparent from this figure that the zero-lower bound plays a role in the asymmetry of the yield distribution. Table 3 reports estimates of the skewness of the variables, which are \( \hat{\alpha} = 0.47 \) for consumption growth, \( \hat{\alpha} = 0.62 \) for investment growth, 0.82 for inflation, and larger than +0.4 for all interest rates. Plotting skewness as a function of the maturity reveals the skewness “smile” in Figure 5 whereby skewness is highest for the 3-month bond, declines monotonically reaching at minimum for maturities from one to two years, and increases monotonically thereafter.\(^6\)

\(^6\)For a sample ending in 2008:Q4, the skewness of interest rates is larger than for the full sample and ranges from 0.76 for the 3-month bond to 0.87 for the 5-year bond, while the skewness of consumption and investment growth are roughly the same. Thus, it seems likely that restricting the sample to the pre-Financial Crises period would deliver larger quantitative estimates of the importance of skewness risk than those reported here. The finding of a skewness smile is also robust to using the pre-Crisis sample.
Table 4 reports the $p$-values of the Jarque-Bera test of the hypothesis that the data follow a normal distribution. The Jarque-Bera test is a goodness-of-fit test based on sample estimates of skewness and excess kurtosis, both of which are should be zero if the data are normal. Since the value is below 0.05 in all cases, the null hypothesis that the data are normally distributed can be rejected at the 5% level. This result suggests that the normal distribution is a poor approximation to the unconditional distribution of the data.

In what follows, I use an artificial sample generated using each version of the model to compute the skewness and to test the hypothesis that the simulated data follow a normal distribution. Consider first the sample generated from the model with normal innovations. Table 3 shows that this version of the model predicts limited skewness and, in some cases, of sign opposite to that of the data. Table 4 shows that the hypothesis that consumption growth and inflation are normally distributed cannot be rejected at the 5% level. Thus, the model with normal shocks has limited success in accounting for the departures from Gaussianity observed in the data. This result is not trivial because the model is nonlinear and, thus, it could potentially generate asymmetric outcomes even if shocks are symmetric.

In contrast, the models with GEV and skew normal innovations predict quantitatively large skewness of the same sign as that of the data, and the hypothesis of normality can be rejected at the 5% level for all series, as it is for the U.S. data. Thus, relaxing the assumption of symmetrically distributed shocks and adopting, instead, more general asymmetric shock distributions is helpful in accounting for the skewed realizations observed in the data. Notice, however, that the predicted skewness increases monotonically with the maturity and, thus, the estimated model cannot account for the skewness smile reported in Figure 5.

4 Dynamics

In order to develop intuition that will be useful in understanding the asset-pricing properties of the model, this section uses impulse-response analysis to study the dynamic effects of inflation and productivity shocks on consumption, investment, inflation, and bond yields. Since the model is non-linear, the effects of a shock depend on its sign, size, and timing (Koop et al., 1996). For this reason, I study shocks in the 5th and 95th percentiles of the distribution. These shocks have different sign and, in the case of asymmetric distributions, different size. Regarding timing, shocks are assumed to occur when the system is at the stochastic steady state—i.e., when all variables are equal to the unconditional mean of their ergodic distribution. Responses are reported in Figures 6 and 7 for each of the distributions, with the vertical axis in percentage deviations from the stochastic steady state and the horizontal axis in periods.
Figure 6 plots the responses to inflation shocks by inflation, productivity, consumption, investment, and selected bond yields. The upper, middle, and lower rows correspond to the normal, skew normal, and GEV distributions, respectively. Recall that inflation and productivity innovations are orthogonal by assumption and the off-diagonal elements of $\rho$ in (5) are non-zero (see Table 1). Thus, an inflation shock will propagate both through its effect on productivity (via the VAR) and through the endogenous actions of economic agents. As we can see in Figure 6, a positive inflation shock induces a large and protracted negative response by productivity which prompts agents to reduce consumption and investment. Nominal interest rates increase because expected inflation increases and because the decrease in investment and capital (not shown) pushes the real interest rate up. This is the graphical representation of the mechanism discussed by Piazzesi and Schneider (2007), whereby positive inflation surprises induce future decreases in consumption and increases in yields.

A negative inflation shock produces qualitatively the reverse effects. In the case of the normal distribution, these effects are approximately the mirror image of those of a positive shock. This result primarily reflects the symmetry of the normal distribution. In contrast, for the asymmetric GEV and skew normal distributions, the effects of a negative shock are quantitatively smaller than those of a positive shock. The size (in absolute value) of shocks in the 5th and 95th percentiles is not the same when the distribution is asymmetric, but the point here is that the likelihood of these two realizations is the same and this matters for bond pricing. In particular, the concavity of the utility function implies that the positive inflation shock increases the marginal utility of consumption by more than the equally-likely negative shock decreases marginal utility. Although this observation also holds in the case of a symmetric distribution, the effects are magnified in the case of an asymmetric distribution.

Figure 6 also shows that positive inflation shocks increase, and negative shocks decrease, bond yields. As before, there is an asymmetry in that the magnitude of the responses is larger in the former than in the latter case. Moreover, the effect is larger for short- than for long-term yields and decreases monotonically with the maturity. This implies that inflation shocks alter the slope of the yield curve: negative shocks increase the slope, while positive shocks decrease the slope and can potentially make the yield curve downward sloping. In this sense, inflation acts like the slope factor in factor models of the term structure.

Figure 7 plots the responses to productivity shocks. A positive productivity shock induces an increase in consumption, investment, and bond yields, and a moderate decrease in inflation. A negative shock has qualitatively the converse effects. However, in the case of the GEV and skew normal distributions, the effects of the negative shock are quantitatively larger than those of a positive shock. Notice that, conditional on the shock size, the initial effect of productivity shocks is similar for all yields. Thus, productivity shocks push the
whole yield curve up or down, just like the level factor does in factor models.

The result that productivity acts like a level factor while inflation acts like a slope factor is robust to assuming that the real effects of inflation are due to price rigidity (Wu, 2006), to dropping capital as an argument of the production function, and to modeling output as a stochastic endowment.\(^7\) The explanation for this result is straightforward: the estimated process for productivity is a much more persistent than that of inflation (see Table 1). Since productivity shocks are persistent, long-term rates react as strongly as short-term rates to a productivity shock. In contrast, since inflation shocks are transitory, short-term rates react more strongly than long-term rates to an inflation shock. It is important to stress that the finding that productivity is more persistent than inflation is not a figment of the model. Estimates of the autocorrelation of the U.S. productivity and CPI inflation computed directly from the data are 0.98 and 0.76, respectively.\(^8\)

In general, the factors obtained using principal components are statistical objects without a structural interpretation. In order to examine the relation between statistical factors and structural shocks, Onatski and Ruge-Murcia (2013) apply factor analytic techniques to artificial data generated from a large general equilibrium model with aggregate shocks and idiosyncratic sectoral shocks. They report that despite the pervasiveness of the aggregate shocks, the principal components analysis has a hard time replicating the macroeconomic factor space.

Research that attempts to build a link between latent and macroeconomic factors in term structure models include Ang and Piazzesi (2003), Diebold et al. (2006), and Rudebusch and Wu (2008). Ang and Piazzesi, and Diebold et al. specify factor models where the state vector consists of a mix of latent variables and macroeconomics variables. In the former case, the macroeconomic variables are the first principal components of various inflation and output measures. In the latter case, the macroeconomic variables are inflation, capacity utilization, and the Federal Funds rate, with possible feedback from latent variables to macro variables. Finally, Rudebusch and Wu build a model where two latent factors are respectively driven by the central bank’s inflation target and reaction function.

\(^7\)This claim is based on results reported in previous versions of this paper, which are available upon request.

\(^8\)The data used to make these calculations are the CPI inflation rate and the detrended total factor productivity series constructed by John Fernald and available from the Web site of the Federal Reserve Bank of San Francisco (www.frbsf.org/economic-research/total-factor-productivity-tfp). The frequency of the latter series is quarterly and the sample runs from 1947Q2 to 2014Q2.
5 Extreme Events and Skewness Risk

Skewness risk arises in this model because, for a given variance, consumers face the possibility of extreme realizations from the upper tail of the distribution of inflation innovations and from the lower tail of the distribution of productivity innovations. This section quantifies the contribution of this source of risk to bond premia and its effect on yields. This section also computes the adjustment for skewness risk contained in bond prices. These calculations are based on the policy functions that solve the asset pricing model, where skewness enters as an additively separate component.

5.1 Composition of the Bond Premia

Bond premia depend only on the higher-order moments of the innovations and, hence, a third-order approximation to their policy functions consist only of the last three terms in (9) (Andreasen, 2012). An advantage of focusing on policy functions is that they make explicit the dependence of bond premia on the variance and skewness of the innovations:

\[
[f(s_t, \sigma)]^2 = (1/2)[f_{\sigma\sigma}(s, 0)]^2[\sigma][\sigma]
\]

\[
+ (1/2)[f_{\sigma\sigma\sigma}(s, 0)]^2[(s_t - s)]^2[\sigma][\sigma]
\]

\[
+ (1/6)[f_{\sigma\sigma\sigma}(s, 0)]^2[\sigma][\sigma][\sigma].
\]

In the general case where the distribution is asymmetric, the bond premia depend on both the variance and skewness of the innovations (all three terms in (14)). In the special case where the innovation distribution is symmetric, and skewness is zero, the bond premia depend only on the variance of the innovations (i.e., the first two terms in the right-hand side of (14)). Then, the key difference is the term \((1/6)[f_{\sigma\sigma\sigma}(s, 0)]^2[\sigma][\sigma][\sigma]\), which is non-zero in the former case and zero in the latter case.

Figure 8 plots the bond premia predicted by the model for maturities from 3 months to 5 years under each of the distributions considered. The figure was constructed using the mean of the ergodic distribution of the bond premium for each maturity and distribution. Since the premia are derived from the Euler equation of a nominal bond that pays one unit of money (say, one dollar) at maturity, the units of the premia are in units of money as well.\(^9\) In Figure 8 premia are expressed in cents. The thick line is the total bond premia (including skewness risk) and the thin line is the part of the bond premia due to variance risk alone. The distance between the thin and think lines is the contribution of skewness risk.

\(^9\)This is also true because for bond premia, I carried out the perturbation in levels rather than in logarithms. This was necessary because bond premia are zero in the deterministic steady state.
risk to the bond premia. Put differently, the thick line in Figure 8 is the ergodic mean of (14) for each bond maturity and the thin line is the ergodic mean of its variance component, 
\[ (1/2)[f_{s\sigma}(s, 0)]^2[\sigma][\sigma] + (1/2)[f_{s\alpha\sigma}(s, 0)]^2[(s_t - s)]^2[\sigma][\sigma] \]. The distance between both lines is, therefore, 
\[ (1/6)[f_{s\alpha\sigma}(s, 0)]^2[\sigma][\sigma][\sigma] \]. Since this term is zero in the case of the normal distribution, there is no thick line in this panel: All risk is variance risk.

Notice that the premia is negative, meaning that the \( \ell \)-period bond is sold at a discount \textit{vis a vis} the strategy of buying a 1-period bond today and using the proceeds to buy a \( \ell - 1 \)-period bond tomorrow at an uncertain price. As a result, the yield of the \( \ell \)-period bond is larger than the weighted average of the yields of the 1- and \( \ell - 1 \)-period bonds and the yield curve is upward sloping, as we saw in Figure 3. Overall, the models with asymmetric shocks deliver substantially larger premia than the model with normal shocks, and this difference is amplified by skewness risk. For example, the bond premia for a 5-year bond is \(-0.16 \) cents under the normal distribution, but \(-0.26 \) and \(-0.27 \) under the skew normal and GEV distribution. These results suggest that asymmetric shock distributions are helpful in accounting for the magnitude of the observed bond premia.

The contribution of skewness risk to the bond premia varies from 4.9\% for the 6-month bond to 4.7\% for the 5-year bond under the skew normal distribution. Thus, this contribution appears to be relatively stable across maturities. Under the GEV distribution, the contribution is substantially larger and varies from 12.4\% for the 6-month bond to 11.6\% for the 5-year bond. In the latter case, the 90\% confidence interval ranges from 1.1\% to 68.5\% for the 6-month bond and from 1.6\% to 33.4\% for the 5-year bond.\footnote{These confidence intervals are based on estimates of the bond premia computed using 100 random draws from the unconditional distribution of the estimates. Results for other maturities are available upon request.} Thus, skewness risk constitutes a small but non-negligible proportion of the bond risk premia.

### 5.2 Effect of Skewness Risk on Yields

In order to quantify the effect of skewness risk on bond yields, I compute the mean of their ergodic distribution under two risk scenarios, that is, with only variance risk and with both variance and skewness risk. The former mean is computed imposing the restriction that the skewness of all innovations is zero. The latter mean is computed with skewness equal to the estimated values reported in Table 1. (Variances are the same in both scenarios and equal to the values in Table 1). As before, the contribution of skewness risk is simply the difference between these two means and in terms of the policy function (9) corresponds to the size of 
\[ (1/6)[f_{s\alpha\sigma}(s, 0)]^2[\sigma][\sigma][\sigma] \].

Table 5 reports the results of these calculation and supports the following conclusions.
First, skewness risk reduces the yields of bonds of all maturities. To see this note that all mean yields (in percent at the annual rate) are lower with than without skewness risk. This yield reduction reflects the risk, above and beyond variance risk, that arises from the possibilities that i) an extreme productivity innovation, drawn from the left tail of its negatively-skewed distribution, may unexpectedly increase the marginal utility of consumption, or ii) an extreme inflation innovation, drawn from the right tail of its positively-skewed distribution, may reduce the real payoff of bonds and induce a decline in future consumption. Faced with skewness risk, consumers would attempt accumulate precautionary savings in the form of capital and bond holdings, pushing bond prices up and yields down.

Second, the reduction in yields associated with skewness risk is substantial. In the case where innovations are skew normal, the reduction varies from 0.25 percentage points (pp) for the 3-month bond to 0.22 pp for the 5-year bond. In the case where innovations are GEV, the reduction varies from 0.59 pp for the 3-month bond to 0.54 pp for the 5-year bond. In order to put these figures in perspective, recall that adjustments to the Federal Funds rate target are typically of size 0.25 pp. Kapadia (2006) and Chang et al. (2013) also find that skewness risk has a negative effect on excess returns in a cross-section of stock returns and options. In particular, Chang et al. find that stocks with high exposure to innovations in market skewness have lower average return and the spread in average returns between the fifth and the first quantiles of portfolios sorted by skewness beta is negative and statistically significant. Thus, even though the modeling approach and asset class are different, the results found here are in qualitative agreement with those reported in the finance literature. Finally, the proportion of the yield reduction due to the skewness of inflation and productivity shocks is relatively stable across maturities and respectively equal to 33% and 67% under the skew normal distribution, and to 28% and 72% under the GEV distribution.

5.3 Risk Adjustment in Bond Prices

An alternative way to describe the importance of skewness risk is to quantify the adjustment factor in bond prices due to the skewness of productivity and inflation innovations. Again, this adjustment factor can be read directly from the decision rules and corresponds to the term \((1/6)[f_{\sigma\sigma}(s, 0)]^2[\sigma][\sigma][\sigma]\). Figure 9 plots the adjustment factor (thick line) predicted by the three versions of the model. The factor is expressed in cents. The version with normally distributed innovations predicts zero adjustment factor because, by construction, this model does not price skewness risk. The version with skew normal innovations predicts a relatively small adjustment factor that ranges from 0.06 cents for the 3-month bond to 0.83 cents for the 5-year bond. Since the mean price of these bonds is 98.8 cents and 75.4 cents,
respectively, this means that the adjustment factor for skewness risk constitutes 0.06\% and 1.1\% of their price. Notice that this percentage is increasing in the maturity. The version with GEV innovations predicts relatively large adjustment factors that range from 0.15 cents for the 3-month bond and 2.05 cents for the 5-year bond, and respectively constitute 0.15\% and 2.66\% of their price. Figure 9 also decomposes the adjustment factor into the parts due to productivity and inflation, and shows that the former constitutes approximately 67\% of the factor for all maturities when innovations are skew normal and 71\% when innovation are GEV.

6 Conclusions

This paper uses tools of extreme value theory to study the implications of skewness risk for bond pricing and returns in a production economy. For a given variance, the possibility of extreme realizations from the long tail of the inflation and productivity distributions affects prices, returns and the pricing kernel used by consumers to evaluate payoffs. Quantitative magnitudes of these effects are computed based on parameters estimated from U.S. data. These estimates show that inflation innovations are drawn from a positively skewed distribution, that productivity innovations are drawn from a negatively skewed distribution, and that the hypotheses that they are drawn from normal distributions are rejected by the data. Results indicate that for the preferred specification with GEV innovations skewness risk accounts for approximately 12 percent of the bond risk premia, reduces bond yields by approximately 55 basis points, and involves a price adjustment factor that ranges from 0.15 cents for the 3-month bond to 2.05 cents for the 5-year bond. Thus, skewness is both statistically significant and economically important.

Results reported in this paper are relevant for two streams of the literature. First, for the literature on the role of higher-order moments on valuations, this paper shows that in a flexible-price production economy most of the bond premia is driven by consumption risk and that skewness risk is quantitatively important. Second, for the literature on disasters, this paper shows that even in the relatively calm, post-WWII U.S., agents face the possibility of substantial consumption decreases associated with recessions, as well as of positive inflation surprises. In line with results based on option-prices (Backus et al., 2011), these moderate but frequent consumption disasters have quantitatively important implications for bond pricing.
### Table 1: Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>Distribution</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GEV</td>
<td>Skew Normal</td>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>s.e.</td>
<td>Estimate</td>
<td>s.e.</td>
<td>Estimate</td>
</tr>
<tr>
<td><strong>A. Preferences and Costs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount rate</td>
<td>0.9718*</td>
<td>0.0223</td>
<td>0.9713*</td>
<td>0.0192</td>
<td>0.9769*</td>
</tr>
<tr>
<td>IES</td>
<td>0.0457*</td>
<td>0.0225</td>
<td>0.0461*</td>
<td>0.0173</td>
<td>0.0548*</td>
</tr>
<tr>
<td>Risk aversion</td>
<td>34.7873*</td>
<td>17.9647</td>
<td>34.7867*</td>
<td>14.8608</td>
<td>28.2539</td>
</tr>
<tr>
<td><strong>B. Productivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Productivity lag</td>
<td>0.9526*</td>
<td>0.0375</td>
<td>0.9523*</td>
<td>0.0507</td>
<td>0.9615*</td>
</tr>
<tr>
<td>Inflation lag</td>
<td>-0.3074</td>
<td>0.2306</td>
<td>-0.3077</td>
<td>0.3403</td>
<td>-0.3386*</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td>-0.8369*</td>
<td>0.4119</td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>-0.6030*</td>
<td>0.2725</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale</td>
<td>0.0084*</td>
<td>0.0013</td>
<td>0.0109</td>
<td>0.0052</td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0077</td>
<td>0.0081</td>
<td>0.0081*</td>
<td>0.0015</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.9040</td>
<td>-0.3098</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C. Inflation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steady state</td>
<td>1.0095*</td>
<td>0.0013</td>
<td>1.0096*</td>
<td>0.0016</td>
<td>1.0093*</td>
</tr>
<tr>
<td>Inflation lag</td>
<td>0.8483*</td>
<td>0.2678</td>
<td>0.8477*</td>
<td>0.3840</td>
<td>0.7663*</td>
</tr>
<tr>
<td>Productivity lag</td>
<td>-0.0151</td>
<td>0.0341</td>
<td>-0.0147</td>
<td>0.0460</td>
<td>-0.0168</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>0.0554</td>
<td>0.6871</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale</td>
<td>0.0023*</td>
<td>0.0009</td>
<td>0.0049</td>
<td>0.0054</td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0032</td>
<td>0.0030</td>
<td>0.0041*</td>
<td>0.0019</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>1.5154</td>
<td>0.9945</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* The table reports parameters values estimated under each of the three distributions considered. s.e. denotes asymptotic standard error. The superscript * denotes the rejection of the null hypothesis that the true parameter value is zero. The standard deviation and skewness of the GEV and skew normal innovations are those implied by the parameters of the distribution.
Table 2: Measures of Fit

<table>
<thead>
<tr>
<th></th>
<th>Distribution</th>
<th>Skew</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GEV</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>A. Moments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root mean square error</td>
<td>0.212</td>
<td>0.514</td>
</tr>
<tr>
<td>Mean absolute error</td>
<td>0.078</td>
<td>0.107</td>
</tr>
<tr>
<td><strong>B. Yield curve</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root mean square error</td>
<td>0.165</td>
<td>0.166</td>
</tr>
<tr>
<td>Mean absolute error</td>
<td>0.143</td>
<td>0.138</td>
</tr>
</tbody>
</table>

*Note:* The table reports measures of fit for the unconditional moments and the yield curve under each of the distributions considered.
Table 3: Unconditional Skewness

<table>
<thead>
<tr>
<th>Series</th>
<th>U.S. Data</th>
<th>Skew</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GEV</td>
<td>Normal</td>
</tr>
<tr>
<td>Consumption growth</td>
<td>-0.473</td>
<td>-0.712</td>
<td>-0.323</td>
</tr>
<tr>
<td>Investment growth</td>
<td>-0.615</td>
<td>-0.582</td>
<td>-0.125</td>
</tr>
<tr>
<td>Inflation rate</td>
<td>0.828</td>
<td>0.166</td>
<td>0.147</td>
</tr>
<tr>
<td>3-Month T-Bill rate</td>
<td>0.569</td>
<td>0.088</td>
<td>0.269</td>
</tr>
<tr>
<td>6-Month T-Bill rate</td>
<td>0.477</td>
<td>0.097</td>
<td>0.282</td>
</tr>
<tr>
<td>1-Year rate</td>
<td>0.410</td>
<td>0.112</td>
<td>0.305</td>
</tr>
<tr>
<td>2-Year rate</td>
<td>0.393</td>
<td>0.134</td>
<td>0.336</td>
</tr>
<tr>
<td>3-Year rate</td>
<td>0.400</td>
<td>0.148</td>
<td>0.356</td>
</tr>
<tr>
<td>4-Year rate</td>
<td>0.443</td>
<td>0.159</td>
<td>0.369</td>
</tr>
<tr>
<td>5-Year rate</td>
<td>0.497</td>
<td>0.166</td>
<td>0.379</td>
</tr>
</tbody>
</table>

Note: The table reports the skewness of actual U.S. series and of artificial data simulated from the model under each of the distributions considered.
Table 4: Jarque-Bera Tests

<table>
<thead>
<tr>
<th>Series</th>
<th>U.S. Data</th>
<th>GEV Skew</th>
<th>Normal Skew</th>
<th>Normal Skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption growth</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.1489</td>
</tr>
<tr>
<td>Investment growth</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.0216</td>
</tr>
<tr>
<td>Inflation rate</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.0658</td>
</tr>
<tr>
<td>3-Month T-Bill rate</td>
<td>0.0042</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>6-Month T-Bill rate</td>
<td>0.0130</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>1-Year rate</td>
<td>0.0327</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>2-Year rate</td>
<td>0.0437</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>3-Year rate</td>
<td>0.0408</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>4-Year rate</td>
<td>0.0272</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>5-Year rate</td>
<td>0.0169</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

*Note:* The table reports *p*-values of the Jarque-Bera test of the null hypothesis that the data follows a normal distribution.
Table 5: Bond Yields

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3-Month</th>
<th>6-Month</th>
<th>1-Year</th>
<th>2-Year</th>
<th>3-Year</th>
<th>4-Year</th>
<th>5-Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With variance risk only</td>
<td>5.095</td>
<td>5.120</td>
<td>5.168</td>
<td>5.257</td>
<td>5.342</td>
<td>5.426</td>
<td>5.509</td>
</tr>
<tr>
<td>B. Skew Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With variance risk only</td>
<td>5.235</td>
<td>5.266</td>
<td>5.329</td>
<td>5.459</td>
<td>5.592</td>
<td>5.725</td>
<td>5.859</td>
</tr>
<tr>
<td>With variance and skewness risk</td>
<td>4.989</td>
<td>5.021</td>
<td>5.087</td>
<td>5.222</td>
<td>5.360</td>
<td>5.499</td>
<td>5.637</td>
</tr>
<tr>
<td>Contribution of skewness risk</td>
<td>-0.246</td>
<td>-0.245</td>
<td>-0.242</td>
<td>-0.237</td>
<td>-0.232</td>
<td>-0.227</td>
<td>-0.222</td>
</tr>
<tr>
<td>Of which:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Productivity</td>
<td>-0.166</td>
<td>-0.165</td>
<td>-0.163</td>
<td>-0.159</td>
<td>-0.156</td>
<td>-0.152</td>
<td>-0.149</td>
</tr>
<tr>
<td>Inflation</td>
<td>-0.081</td>
<td>-0.080</td>
<td>-0.079</td>
<td>-0.078</td>
<td>-0.076</td>
<td>-0.074</td>
<td>-0.073</td>
</tr>
<tr>
<td>%</td>
<td>32.520</td>
<td>32.653</td>
<td>32.645</td>
<td>32.911</td>
<td>32.759</td>
<td>33.040</td>
<td>32.883</td>
</tr>
<tr>
<td>C. GEV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With variance risk only</td>
<td>5.570</td>
<td>5.599</td>
<td>5.659</td>
<td>5.782</td>
<td>5.908</td>
<td>6.035</td>
<td>6.162</td>
</tr>
<tr>
<td>With variance and skewness risk</td>
<td>4.976</td>
<td>5.009</td>
<td>5.075</td>
<td>5.211</td>
<td>5.348</td>
<td>5.487</td>
<td>5.625</td>
</tr>
<tr>
<td>Contribution of skewness risk</td>
<td>-0.594</td>
<td>-0.591</td>
<td>-0.584</td>
<td>-0.571</td>
<td>-0.559</td>
<td>-0.548</td>
<td>-0.536</td>
</tr>
<tr>
<td>Of which:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Productivity</td>
<td>-0.428</td>
<td>-0.426</td>
<td>-0.421</td>
<td>-0.411</td>
<td>-0.403</td>
<td>-0.394</td>
<td>-0.386</td>
</tr>
<tr>
<td>%</td>
<td>72.054</td>
<td>72.081</td>
<td>72.089</td>
<td>71.979</td>
<td>72.093</td>
<td>71.898</td>
<td>72.015</td>
</tr>
<tr>
<td>Inflation</td>
<td>-0.166</td>
<td>-0.165</td>
<td>-0.163</td>
<td>-0.160</td>
<td>-0.157</td>
<td>-0.154</td>
<td>-0.150</td>
</tr>
<tr>
<td>%</td>
<td>27.946</td>
<td>27.919</td>
<td>27.911</td>
<td>28.021</td>
<td>27.907</td>
<td>28.102</td>
<td>27.985</td>
</tr>
</tbody>
</table>

Note: The table reports mean bond yields under two risk scenarios. The contribution of skewness risk is the difference in mean yields between the two scenarios and it is decomposed in the parts due to each shock. Yields are expressed in percent at the annual rate.
References


Figure 1: Estimated Probability Density Function of Productivity Innovations
Figure 2: Estimated Probability Density Function of Inflation Innovations
Figure 3: Actual and Predicted Yield Curves
Figure 4: Skewness in the U.S. Data
Figure 5: The Skewness Smile
Figure 6: Responses to Inflation Shocks
Figure 7: Responses to Productivity Shocks

Inflation  Productivity  Consumption  Investment  3-Month Yield  1-Year Yield  3-Year Yield  5-Year Yield

Normal

Skew Normal

GEV

95th  5th
Figure 8: Bond Premia

Normal

Skew Normal

GEV

Without skewness risk

With skewness risk
Figure 9: Skewness Risk Adjustment in Bond Prices