Skewness Risk and Bond Prices

Francisco Ruge-Murcia*

July 2014

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 modeled using the generalized extreme value (GEV) distribution. 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 productivity and inflation innovations are drawn from a normal distribution and favor instead the alternative that they are drawn from an asymmetric distribution. Estimates indicate that skewness risk accounts for 10 percent of the risk premia and has a price of 0.6 percent per year. Despite the fact that bonds are nominal, most of the priced risk is consumption risk.

JEL Classification: G12, E43, E44

Key Words: Extreme value theory, GEV distribution, skewness risk; risk premia; nonlinear dynamic models; simulated method of moments.

*I received helpful comments from Julien Bengui, Taeyoung Doh, Larry Epstein, Silvia Gonçalves, Jinill Kim, Federico Ravenna and seminar participants at Boston College 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, University of Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec) H3C 3J7, Canada.

1 Introduction

This paper uses extreme value theory to the 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 here 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 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 three 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 thick tails and skewed realizations that characterize consumption growth and bond yields.

Results indicate that skewness risk has a price of 0.6% per year, and represent about 10% of the risk premia with a 90% confidence interval ranging from 0.8% to 19.6%. Thus, the effect of skewness is economically significant as well. Despite the fact that bonds are nominal and that the inflation variance and skewness are substantial, I find that consumption risk carries a much larger price than inflation risk and constitutes most of the risk premia.

Previous literature that uses extreme value theory follows a partial equilibrium approach whereby 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 the 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. (2012) find that skewness risk has a negative effect on excess returns in a cross-section of stock returns and option data, respectively. Colacito et al. (2012) study the implications of time-varying skewness in an endowment economy.[1] Compared with that literature, I employ an equilibrium model of asset pricing where consumption is an endogenous choice variable and the structural estimation of the model attempts to reconcile bond prices with macroeconomic aggregates.

As earlier research on disasters (e.g., Rietz, 1988, Barro, 2006, and Barro et al., 2013), 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 bond prices and risk premia. 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 studies asymmetries in inflation, which can be 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, hence, the magnitude of skewness risk.

[1] Golec and Tamirkin (1998) and Brunnermeier et al. (2007) examine the role of preferences in accounting for the importance of skewness. Golec and Tamarkin estimate the payoff function of bettors in horse tracks and find that it is decreasing in the variance but increasing in the skewness, so that at high odds, bettors are willing to accept poor mean returns and variance because the skewness is large. Brunnermeier et al. develop a model of optimal beliefs which predicts that traders will overinvest in right-skewed assets.

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

Output is produced by identical firms whose total number is normalized to 1. The representative firm uses the technology

$$y_t = A_t \left(h_t\right)^{\alpha}, \tag{1}$$

where y_t is output, A_t is an aggregate productivity shock, h_t is labor input, and $\alpha \in (0, 1]$ is a parameter. The productivity shock follows the process

$$\ln(A_t) = (1 - \rho)\ln(A) + \rho \ln(A_{t-1}) + \zeta_t, \tag{2}$$

where $\rho \in (-1, 1)$, $\ln(A)$ is the unconditional mean of $\ln(A_t)$, and ζ_t is an independent and identically distributed (i.i.d.) innovation with mean zero 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.

Profit maximization implies that the marginal productivity of labor equals the real wage. That is,

$$\alpha A_t \left(h_t\right)^{\alpha-1} = W_t / P_t, \tag{3}$$

where W_t is the nominal wage and P_t is the price level. Prices are denominated in terms of a unit called money but the economy is cashless otherwise.

2.2 Consumption

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

$$U_t = \left((1-\beta)\,(c_t)^{1-1/\psi} + \beta \left(E_t \left(U_{t+1}^{1-\gamma} \right) \right)^{(1-1/\psi)/(1-\gamma)} \right)^{1/(1-1/\psi))}, \tag{4}$$

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, ψ is the intertemporal elasticity of substitution, and γ is the coefficient of relative risk aversion. As it is well known, recursive preferences decouple elasticity of substitution from risk aversion when $\gamma \neq 1/\psi$, 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 Schnider (2007), Le and Singleton (2010), van Binsbergen et al. (2012), Andreasen (2012), Rudebusch and Swanson (2012), and Doh (2013). The consumer supplies a fixed time endowment, h, in a competitive labor market.

Financial assets are zero-coupon nominal bonds with maturities $\ell = 1, \ldots, L$. All bonds are equally liquid regardless of their maturity and can be costlessly traded in a secondary market. The consumer's budget constraint is

$$c_t + \sum_{\ell=1}^{L} \frac{Q_t^{\ell} B_t^{\ell}}{P_t} = \frac{hW_t}{P_t} + \sum_{\ell=1}^{L} \frac{Q_t^{\ell-1} B_{t-1}^{\ell}}{P_t}, \tag{5}$$

where Q_t^{ℓ} and B_t^{ℓ} are, respectively, the nominal price and quantity of bonds with maturity ℓ, and $Q_t^0 = 1$. The latter normalization simply means that bonds pay one unit of currency at maturity. Consumers take as given the aggregate inflation rate, $\Pi_t = P_t/P_{t-1}$. The inflation rate follows the process

$$\ln(\Pi_t) = (1-\mu)\ln(\Pi) + \mu \ln(\Pi_{t-1}) + \eta_t, \tag{6}$$

where $\mu \in (-1,1)$, $\ln(\Pi)$ is the unconditional mean of $\ln(\Pi_t)$, and η_t is an i.i.d. innovation with mean zero and non-zero skewness. Because prices are flexible, assuming a process for inflation is equivalent to, but more parsimonious than, specifying a process for the rate of money growth and a mechanism (e.g., money in the utility function) that incites consumers to hold money even though it is dominated in rate of return by bonds.

The Euler equations that characterize the consumer's utility maximization are

$$\frac{Q_t^{\ell}}{P_t} = \beta E_t \left(\left(\frac{v_{t+1}}{w_t} \right)^{1/\psi-\gamma} \left(\frac{c_{t+1}}{c_t} \right)^{-1/\psi} \left(\frac{Q_{t+1}^{\ell-1}}{P_{t+1}} \right) \right), \text{ for } \ell = 1, 2, \ldots L, \tag{7}$$

where

$$v_t \equiv \max_{\left\{c_t, B_t^1, ..., B_t^L\right\}} U_t$$

and

$$w_t \equiv E_t v_{t+1}$$

are the value function and the certainty-equivalent future utility, respectively. As usual, the Euler equations (7) compare the marginal cost of acquiring an additional unit of the financial asset with the discounted expected marginal benefit of keeping the asset till next period.

2.3 Equilibrium

The equilibrium is an allocation for the consumer $\mathcal{C} = \left\{c_t, \left(B_t^\ell\right)_{\ell=1,..L}\right\}_{t=0}^{\infty}$, an allocation for the firm $\mathcal{Y} = \{y_t, h_t\}_{t=0}^{\infty}$ and a price system $\mathcal{P} = \left\{(Q_t^\ell)_{\ell=1,..L}, W_t\right\}_{t=0}^{\infty}$ such that given the price system: (i) the allocation $\mathcal{C}$ solves the trader's problem; (ii) the allocation $\mathcal{Y}$ solves the firms's problem; (iii) the goods market clears: $C_t = Y_t$; (iv) the labor market clears: $H_t = H$; and (v) bonds are in zero net supply: $B_t^\ell = 0$ for all $\ell = 1, 2, ..., L$. In this definition, C_t, Y_t, and H_t are the aggregate counterparts of c_t, y_t, and h_t, respectively.

2.4 Bond Yields

Following the literature, the gross yield of the ℓ-period bond is

$$i_t^\ell = \left(Q_t^\ell\right)^{-1/\ell}, \tag{8}$$

for $\ell = 1, 2, ..., L$. The bond risk premium (or bond premium for short) is the component of the long-term bond price that accounts for the risk involved in holding this bond compared with a sequence of rolled-over shorter-term bonds. The risk arises because future bond prices are not known in advance and, so, the strategy of rolling over shorter-term bonds entails a gamble. I derive the bond premium recursively from the Euler equations

$$Q_t^\ell = Q_t^1 E_t\left(Q_{t+1}^{\ell-1}\right) + \beta cov_t\left(\left(\frac{V_{t+1}}{W_t}\right)^{1/\psi-\gamma}\left(\frac{C_{t+1}}{C_t}\right)^{-1/\psi}, \frac{Q_{t+1}^{\ell-1}}{\Pi_{t+1}}\right), \tag{9}$$

for $\ell = 2, ...L$, where V_t and W_t are, respectively, the aggregate counterparts of v_t and w_t, and I have used the fact that for the one-period bond

$$Q_t^1 = \beta E_t\left(\left(\frac{V_{t+1}}{W_t}\right)^{1/\psi-\gamma}\left(\frac{C_{t+1}}{C_t}\right)^{-1/\psi}\frac{1}{\Pi_{t+1}}\right). \tag{10}$$

As in Ljungqvist and Sargent (2004, ch. 13.8), the premium on the ℓ-period bond is defined as

$$\Gamma_{\ell,t} \equiv \beta cov_t \left(\left(\frac{V_{t+1}}{W_t} \right)^{1/\psi-\gamma} \left(\frac{C_{t+1}}{C_t} \right)^{-1/\psi}, \frac{Q_{t+1}^{\ell-1}}{\Pi_{t+1}} \right). \tag{11}$$

Notice that the premium depends on the maturity and may be positive or negative according to the sign of the covariance between the pricing kernel and $Q_{t+1}^{\ell-1}/\Pi_{t+1}$. This observation highlights the fact that, in general, consumption-based asset-pricing models do not restrict the sign or monotonicity of the bond 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 (11) is negative. A negative premium is a discount on the longer-term bond in the sense that buying a ℓ-period bond at time t at price Q_t^ℓ is cheaper than buying a one-period bond at time t at price Q_t^1 and a $(\ell-1)$-period bond at time $t+1$ at expected price $Q_{t+1}^{\ell-1}$. Up to an approximation, this implies that the yield of the ℓ-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.[2] Conversely, the premium is positive when the covariance in (11) is positive and the yield curve is downward sloping.

2.5 Solution

Since this model does not have an exact analytical solution, I use a perturbation method to obtain an approximate solution (see Jin and Judd, 2002). This method involves taking a third-order expansion of the policy functions around the deterministic steady state and characterizing the local dynamics. 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.[3] Caldera et al. (2012) show that for models with recursive preferences, a third-order perturbation is as accurate as projection methods (e.g., Chebychev polynomials and value function iteration) in the range of interest while being

[2] 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. Instead, in this model all bonds are equally liquid and can be traded without penalty before their maturity. 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.

[3] 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 deliver a constant risk premium. For the implementation of the third-order perturbation, I use the codes described in Ruge-Murcia (2012).

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 $f(x_t, \sigma)$ where x_t is a vector of state variables and σ is a perturbation parameter. For this model, the state variables are aggregate productivity and inflation. That is, $x_t = [A_t \ \Pi_t]'$. The goal is to approximate $f(x_t, \sigma)$ using a third-order polynomial expansion around the deterministic steady state where $x_t = x$ and $\sigma = 0$. Using tensor notation, this approximation can be written as

$$\begin{aligned}
[f(x_t,\sigma)]^j &= [f(x,0)]^j + [f_x(x,0)]^j_a[(x_t - x)]^a \\
&\quad +(1/2)[f_{xx}(x,0)]^j_{ab}[(x_t - x)]^a[(x_t - x)]^b \\
&\quad +(1/6)[f_{xxx}(x,0)]^j_{abc}[(x_t - x)]^a[(x_t - x)]^b[(x_t - x)]^c \\
&\quad +(1/2)[f_{\sigma\sigma}(x,0)]^j[\sigma][\sigma] \\
&\quad +(1/2)[f_{x\sigma\sigma}(x,0)]^j_a[(x_t - x)]^a[\sigma][\sigma] \\
&\quad +(1/6)[f_{\sigma\sigma\sigma}(x,0)]^j[\sigma][\sigma][\sigma],
\end{aligned} \tag{12}$$

where elements like $[f_x(x,0)]^j_a$ are coefficients that depend nonlinearly on structural parameters.[4] 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. In the special case where the distribution of the innovations is symmetric—and, hence, skewness is zero—the latter 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.[5] As shown by Andreasen (2012, p. 300), 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. This observation holds more generally for the other variables in the model because the conditional skewness of the innovations is time invariant.

[4] In writing (12), I have used the intermediate results $[f_{x\sigma}(x,0)]^j_a = [f_{\sigma x}(x,0)]^j_a = 0$ (Schmitt-Grohé and Uribe, 2004, p. 763), $[f_{xx\sigma}(x,0)]^j_{ab} = [f_{\sigma xx}(x,0)]^j_{ab} = [f_{x\sigma x}(x,0)]^j_{ab} = 0$ (Ruge-Murcia, 2012, p. 936), and $[f_{\sigma\sigma x}(x,0)]^j_a = [f_{\sigma x\sigma}(x,0)]^j_a = [f_{x\sigma\sigma}(x,0)]^j_a$ by Clairaut's theorem.

[5] An alternative solution strategy used in earlier literature is to assume 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. For example, see Hansen and Singleton (1983), Campbell (1996) and Jerman (1998). However, since the adjustment factor is proportional to the variance only, this approach assumes away the contribution of higher-order moments, like the skewness, to the risk premia. Martin (2012) relaxes the assumption of lognormality and expresses variables as polynomial functions of consumption growth cumulants. Similarly, the perturbation method that I use here allows higher-order moments of productivity and inflation to affect the model variables.

3 Estimation

This section describes the econometric method and data used to estimate the model, and reports estimates of the structural parameters under three different statistical distributions of the innovations.

3.1 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 Duffie and Singleton (1993) show that SMM delivers consistent and asymptotically normal parameter estimates under general regularity conditions. 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 $\Theta \subset \Re^q$, $\mathbf{m}_t$ to be a $p \times 1$ vector of empirical observations on variables whose moments are of our interest, and $\mathbf{m}_\iota(\theta)$ to be the synthetic counterpart of $\mathbf{m}_t$ whose elements are obtained from the stochastic simulation of the model. The SMM estimator, $\widehat{\theta}$, is the value that solves

$$\min_{\{\theta\}} \mathbf{M}(\theta)'\mathbf{W}\mathbf{M}(\theta), \tag{13}$$

where

$$\mathbf{M}(\theta) = (1/T)\sum_{t=1}^{T}\mathbf{m}_t - (1/\lambda T)\sum_{\iota=1}^{\lambda T}\mathbf{m}_\iota(\theta),$$

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

$$\sqrt{T}(\widehat{\theta} - \theta) \rightarrow N(\mathbf{0},(1 + 1/\lambda)(\mathbf{J}'\mathbf{W}^{-1}\mathbf{J})^{-1}\mathbf{J}'\mathbf{W}^{-1}\mathbf{S}\mathbf{W}^{-1}\mathbf{J}(\mathbf{J}'\mathbf{W}^{-1}\mathbf{J})^{-1}), \tag{14}$$

where

$$\mathbf{S} = \lim_{T\rightarrow\infty} Var\left((1/\sqrt{T})\sum_{t=1}^{T}\mathbf{m}_t\right), \tag{15}$$

and $\mathbf{J} = E(\partial\mathbf{m}_\iota(\theta)/\partial\theta)$ 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, which was computed using the Newey-West estimator

with a Barlett kernel and bandwidth given by the integer of $4(T/100)^{2/9}$.[6] The number of simulated observations is 5 times larger than the sample size (that is, $\lambda = 5$). Using $\lambda = 10$ and $\lambda = 20$ deliver similar results but at a higher computational cost, which makes unfeasible the bootstrap procedure used to compute the standard errors of the parameter estimates. The dynamic simulations of the non-linear model are based on the pruned version of the solution, as suggested by Kim et al. (2008).[7]

The estimated parameters are the discount factor (β), the intertemporal elasticity of substitution (ψ), the coefficient of relative risk aversion (γ), the rate of inflation in the deterministic steady state (Π), and the autoregressive coefficients and parameters of the innovation distributions of productivity and inflation. The unconditional mean of the productivity process and time endowment were normalized to 1. The moments used to estimate these parameters are the variances, covariances, skewness, and first-order autocovariances of all five data series, plus the unconditional means of inflation and the nominal interest rates.

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. In practice, however, researchers use a highly flexible distribution, namely the generalized extreme value distribution (GEV) due to Jenkinson (1955). 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. These distributions are the Gumbel, Fréchet, and Weibull distributions, also known as extreme value distribution of type I, II, and III, respectively. The three distributions can be represented in a unified way using the GEV distribution. 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. There are values of the shape parameter for which the mean and variance of the distribution do not exist—in particular, the mean is not defined when

[6] In order to examine the robustness of results to using a different weighting matrix, I also estimated the model using $\mathbf{W} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix and found that parameter estimates were very similar to those reported here.

[7] Kim et al. discuss only the pruning of models approximated up to a second order, but the extension of their ideas to the third-order approximation is straightforward.

this parameter larger than or equal to 1, and the variance is not defined when it is larger than or equal to 0.5—, but this turns out to be not empirically relevant here.

As a complement to the GEV, I also 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 symmetric (normal) distribution against the alternative that they are drawn from an asymmetric (skew normal) 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. However, a drawback of the skew normal distribution is that its skewness is bounded between -1 and $+1$. Although in some empirical applications this constraint may not bind, we will see that in this case it does for productivity innovations.

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

3.3 Data

The model is estimated using quarterly observations of the growth rate of consumption, the inflation rate, and the 3-month, 6-month, and 12-month Treasury-Bill rates for the period 1960Q1 to 2001Q2.[8] Treasury Bills are ideal for this analysis because, like the bonds in the model, they are zero-coupon bonds with negligible default risk. 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*). Consumption is measured by personal consumption expenditures on non-durable goods and services, which were converted into real per-capita terms by dividing by the quarterly average of the Consumer Price Index (CPI) for all urban consumers and by the quarterly average of the mid-month U.S. population estimate produced by the Bureau of Economic Analysis (BLS). The inflation rate is measured by the (gross) quarterly percentage change in the CPI. Since a period in the model is one quarter, the 3-month Treasury Bill rate serves as the empirical counterpart of the one-period nominal interest rate. The 6-month

[8] The latter date was determined by the availability of the 12-month Treasury Bill rate, which was not reported by the Federal Reserve from 27 August 2001 to 2 June 2008.

and 12-month rates are the two-period and four-period interest rates, respectively. Rather than averaging the Treasury Bill rates over the quarter, I used the observations for the first trading day of the second month of each quarter (February, May, August and November). The original interest rate series, which are quoted as a net annual rate, were transformed into a gross quarterly rate. Except for the nominal interest rates, all raw data are seasonally adjusted at the source.

3.4 Identification

It is usually difficult to verify that parameters are globally identified, but local identification simply requires that

$$rank\left\{E\left(\frac{\partial \mathbf{m}_{\iota}(\theta)}{\partial \theta}\right)\right\} = q, \tag{16}$$

where (with some abuse of the notation) θ is the point in the parameter space Θ where the rank condition is evaluated. I verified that this condition is indeed satisfied at the optimum $\widehat{\theta}$ for the three versions of the model.

3.5 Parameter Estimates

Estimates of the parameters are reported in Table 1. Standard errors are reported in parenthesis and were computed using a block bootstrap. The size of the block was set to the integer of $4(T/100)^{2/9}$, where T is the sample size, and the number of replications was 99. First, notice that estimates of the preference parameters are remarkably similar across innovation distributions. The discount factor is 0.99, which implies an annualized gross real interest rate of 1.04 in the deterministic steady state. The intertemporal elasticity of substitution (IES) is statistically different from both 0 and 1. Estimates—between 0.13 and 0.16 depending on the distribution—are in line with values reported in earlier 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 about 50, which is of the same order of magnitude as the estimate of 79 reported by van Binsbergen et al. (2010) and the values employed by calibration studies that use Epstein-Zin preferences (e.g., Tallarini, 2000, Andreasen, 2012, and Rudebusch and Swanson, 2012).[9]

[9] Rudebusch and Swanson (2012, p. 123) discuss possible explanations for why representative-agent models, like this one, require high levels of risk aversion to match asset prices (e.g., larger consumption volatility for asset holders and model uncertainty).

The inflation rate is mildly persistent and, since the correlation parameter (skew normal) and shape parameter (GEV) are positive, inflation innovations are positively skewed. That is, their distribution has a longer tail on the right than on the left side, and the median is below the mean. Furthermore, because the correlation parameter is statistically different from zero, the null hypothesis that inflation innovations are drawn from a normal distribution can be rejected in favor of the alternative that they are drawn from an asymmetric skew normal distribution with positive skewness (p-value < 0.001). The Lagrange Multiplier (LM) test of this hypothesis (p-value < 0.001) supports the same conclusion. Note that since the shape parameter is positive and statistically different from zero, the GEV distribution corresponds, more precisely, to a Fréchet distribution.

Figure 1 plots the estimated cumulative distribution function (CDF) of the inflation innovations under the three models. Note that the CDF of the skew normal and the GEV 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, loosely speaking, large positive inflation surprises can happen sometimes, but large negative ones are unlikely. This means 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 price of bonds with maturity longer than 2 periods and the real payoff of the bond with maturity equal to 1 period.

The productivity process is very persistent, and the correlation parameter (skew normal) and shape parameter (GEV) are negative. Productivity innovations are negatively skewed: their distribution has a longer tail on the left than on the right side, and the median is above the mean. Since the correlation parameter is on the boundary of the parameter space, I use the LM test rather than the (Wald-type) t-test to evaluate the hypothesis that innovations are drawn from a normal distribution against the alternative that they are drawn from a skew normal distribution. Since the p-value is below 0.001, the hypothesis can be rejected at standard significance levels. Note that the shape parameter is negative and statistically different from zero, meaning that the GEV distribution corresponds to a Weibull distribution. The finding that productivity innovations follow a Weibull distribution but inflation innovations follow a Fréchet distribution illustrates the advantage of using the flexible GEV distribution in this empirical project.

Figure 2 plots the estimated CDF of the productivity innovations under the three models. In the cases of the skew normal and GEV distributions, the CDF has more mass in the left tail, and less mass in the right tail, a normal distribution with the same standard deviation (thin line). Thus, large negative productivity surprises are more probable 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. By a large decline in consumption, I mean, for example, a reduction in consumption 5 percentage points below the mean, which can take place with probability 3.8% in the model with GEV innovations. In the quarterly U.S. data, the proportion of observations 5 percentage points below trend is 4.2% and these observations are primarily associated with recessions.

The literature on consumption disasters also features a negatively skewed distribution from the combination of a normal distribution for non-disaster periods and a Bernoulli distribution for disasters (e.g., see Barro, 2006) . The results in this section imply that even outside disaster episodes, consumers still face negative consumption skewness due to business cycle fluctuations—in addition, to positive inflation skewness. This finding is important because recent work based on option prices (Backus et al., 2011) suggests that more frequent, moderate consumption disasters of magnitude comparable to recessions play a key role in explaining U.S. stock returns. Backus et al. report (p. 1995) that for their option-based model, the skewness of consumption is -0.28 and the probability of a three standard deviation drop in consumption is about 1%, while in this model the skewness of consumption is -0.31 and the probability is 0.33%.

In summary, the estimates and statistical tests reported above 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 inflation innovations are drawn from a positively skewed distribution, and productivity innovations are drawn from a negatively skewed distribution. Having established the statistical significance of skewness, sections 4 and 5 below will examine its economic implications for bond pricing.

3.6 Fit of the Model

Figure 3 reports the fit of the three versions of the model by comparing actual and predicted moments. The horizontal axes are moments computed from U.S. data while the vertical axes and dots are moments predicted by the model. The straight line is the 45 degree line and, thus, if a model were to perfectly fit the data moments, all dots would be on this line.

We can see in this figure that the models with asymmetric innovations fit the data better than the model with normal innovations. This impression is statistically confirmed by the root mean-square errors (RMSE), which are 0.15, 0.10 and 0.075 for the models with normal, skew normal and GEV innovations, respectively. Most of the difference comes from the fact that the model with normal innovations cannot match the unconditional skewness of the data. (This issue is further explored in section 4.1.) For instance, this model predicts that

the skewness of consumption growth and inflation is basically zero and the corresponding dots are, therefore, far from the 45 degree line. (The dots corresponding to these two moments are indicated in figure 3.) Note in the left panel of this figure that the vertical distance between these two dots and the 45 degree line is large. In contrast, the models with asymmetric innovations predict negative skewness in consumption growth and positive skewness in inflation and interest rates, as in the data, and the vertical distance between their dots and the 45 degree line is correspondingly small.

Among the models with asymmetric innovations, the model with GEV innovations has lower RMSE than the model with skew normal innovations. Although both models predict similar (negative) skewness in consumption growth, the former predicts higher (positive) skewness for inflation and interest rates. This difference is partly due to the constraint in the range of possible values of skewness under the skew normal distribution (from -1 to $+1$). Free from this constraint, the model with GEV innovations delivers predictions about skewness that are much closer to those observed in the data.

Figure 4 through 6 report the fit of the U.S. data series under each of the innovation distributions. (Notice that the rates of inflation and nominal interest in these figures are expressed in quarterly rates.) All three version of the model fit consumption equally well and, broadly speaking, track well the dynamics of inflation and the nominal interest rate. However, the model with GEV innovations captures better the spikes in the data and delivers a lower RMSE than the alternative models. For example for the inflation rate, the RMSE of this model is 0.41, compared with 0.46 and 0.43 for the models with skew normal and normal innovations.

Overall, results in this section suggest that the version of the model with GEV innovations delivers the best model fit, both in terms of moments and of actual data series. For this reason, in what follows I treat this version as the preferred specification.

4 Extreme Events and Skewness Risk

This section explores further the finding that there are departures from Gaussianity in the data that are better explained by the model with extreme value innovations, links the asymmetry of the GEV distribution to skewness risk, quantifies the contribution of skewness risk to the bond premia and its effect on bond yields, and reports the price of skewness risk implied by the model.

4.1 Unconditional Skewness

Skewness is a prominent feature of U.S. economic data. Figure 7 shows that the unconditional distribution of inflation and of the 3- and 12-month Treasury-Bill rates are positively skewed, while that of consumption growth is negatively skewed. All large consumption growth declines are associated with recessions. Table 2 reports estimates of the skewness of these variables which, as we can see, are larger than $+1$ in the case of inflation and interest rates and about -0.6 in the case of consumption growth. Finally, Table 3 reports the p-values of the Jarque-Bera test of the hypothesis that the data follow a normal distribution. This goodness-of-fit test is based on sample estimates of the skewness and excess kurtosis, both of which are should be zero if the data were normal. Given the plots in Figure 7 and skewness estimates in Table 2, it is not surprising that p-values are well below 0.05 for all series and, hence, the hypotheses can be rejected. As a whole, this evidence 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 of the ergodic distribution of consumption growth, inflation and the nominal interest rates, and to test the hypothesis that the artificial data follows a normal distribution. Consider first the sample generated from the model with normal innovations. Table 2 shows that the unconditional skewness predicted by this model is basically zero, while Table 3 shows that the hypothesis that the artificial data follow a normal distribution cannot be rejected at the 5% level for any variable.

In contrast, for the samples generated from the models with skew normal and GEV innovations, the unconditional skewness predicted by these models (see Table 2) are quantitatively large and of the same sign as the actual data—though in general the skewness of consumption growth is larger than the data, while that of inflation and of the interest rates are smaller than the data. Moreover, the hypothesis of normality can be rejected at the 5% level in all cases as it is for the U.S. data (see Table 3).

4.2 The Contribution of Skewness Risk to the Bond Premia

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. The former reduce the real payoffs and prices of nominal bonds. The latter increase the marginal utility consumption and, thus, the kernel used by consumers to value financial assets. Since the productivity and inflation processes are serially correlated, these effects are persistent.

As it was pointed out above, in the special case where the innovation distribution is

symmetric—and skewness is, therefore, zero—the bond premia depends only on the variance of the innovations. In the more general case where the distribution is asymmetric, the bond premia depends on both the variance and skewness of the innovations. In terms of the policy functions (see equation (12)), the key difference is the term $(1/6)[f_{\sigma\sigma\sigma}(x,0)]^j[\sigma][\sigma][\sigma]$, which is zero in the former case and non-zero in the latter case.

Figure 8 plots the bond premia predicted by the three versions of the model and maturities ranging from 1 to 8 periods. 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 for a nominal bond that pays one unit of currency (say, one dollar) at maturity, the units of the premia are in units of currency as well.[10] In Figure 8, premia are expressed in cents. The thin line is the part of the bond premia due to the variance risk only. The thick line is the total bond premia, including skewness risk. Since the latter part is zero in the case of the normal distribution, there is no thick line in this panel: all risk is variance risk. The distance between the thin and think lines is the contribution of skewness risk to the bond premia. In terms of the policy functions in (12), the distance between zero and the thin line is $(1/2)[f_{\sigma\sigma}(x,0)]^j[\sigma][\sigma] + (1/2)[f_{x\sigma\sigma}(x,0)]^j_a[(x_t - x)]^a[\sigma][\sigma]$, while the distance between the thin and think lines is $(1/6)[f_{\sigma\sigma\sigma}(x,0)]^j[\sigma][\sigma][\sigma]$. All other terms are zero because the premia depend only on the higher-order moments of the innovations.

Notice that the premia is negative, meaning that the ℓ-period bond is sold at a discount *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 ℓ-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 slopping. Overall, the models with asymmetric shocks deliver large premia than the model with normal shocks, and this difference is amplified by skewness risk. Furthermore, skewness risk accounts for about 6% of the bond premia when innovations follow a skew normal distribution and about 10% when innovations follow a GEV distribution. The corresponding figure when innovation follow a normal distribution is zero. For the preferred specification with GEV innovations, the 90% confidence interval for the skewness contribution ranges from 0.8% to 19.6%.[11] Hence, these results suggest that skewness risk constitutes a non-negligible proportion of the bond risk premia.

[10] This is also true because for the bond premia (only), I took the expansion in levels, rather than in logs. This was necessary because the expansion was taken around the deterministic steady state, where the bond premia is zero and its logarithm is not defined.

[11] This interval corresponds to the two-period bond but results are similar for longer maturities. The interval was constructed using a block bootstrap with 99 replications and block size given by the integer of $4(T/100)^{2/9}$. Given that the time-consuming solution of the model prevents me from using a larger number of replications and longer maturities in the bootstrap, these results are best interpreted as indicative only.

4.3 Effects of Skewness Risk on Yields

In order to quantify the effect of skewness risk on bond yields, I compute the mean of the ergodic distribution of bond yields for each version of the model. Table 4 reports certainty-equivalent yields, yields when there is only variance risk, and yields when there is both variance and skewness risk. All yields are in percent at the quarterly rate.

Comparing the certainty-equivalent and other yields, we can see that risk-averse consumers in a stochastic economy induce higher bond prices and lower yields. For example in the case with normal innovations, the yield of the 1-period bond is 0.57 percentage points lower than the certainty-equivalent yield, while the yield of the 8-period is 0.54 points lower.

Comparing yields with both risks and with variance risk only, we can see that skewness risk induces even lower yields. This additional yield decrease primarily reflects the extra risk that an extreme productivity innovation, drawn from the left tail of its negatively-skewed distribution, may unexpectedly increase the marginal utility of consumption next period. It also reflects the risk that an extreme inflation innovation drawn from the right tail of its positively-skewed distribution may reduce the real payoff/price of this nominal asset. Quantitatively, skewness risk decreases the bond yield of a 1-period bond by 0.026 percentage points (that is, 10.1 annual basis points) in the model with skew normal innovations, and by 0.051 percentage points (20.4 annual basis points) in the model with GEV innovations. Results for the other maturities are similar.

4.4 The Price of Skewness Risk

Table 5 reports the price of variance and skewness risk for selected bond maturities, distinguishing between the risk due to productivity and the risk due to inflation. Three observations follow from Table 5. First, concerning productivity, variance risk carries a negative price or discount, while skewness risk carries a positive price. The former is negative because consumers react to variance risk by building up precautionary savings, which in equilibrium reduces bond prices. The latter is positive because in a market with a skewed productivity process, nominal bonds may or may not be helpful for the purpose of consumption smoothing depending on the sign of the skewness. If skewness is positive, the bond payoff may arrive at a time consumption is exceptionally high and, thus, the consumer would demand compensation in the form of a higher yield. However, if skewness is negative, the bond payoff may arrive at a time consumption is exceptionally low and, thus, the consumer would accept a lower yield. The latter is the empirically relevant case here and means that both variance and skewness risk induce lower bond yields. Quantitatively the price of productivity skewness risk decreases slowly with the maturity. In the case with skew normal innovations, the

price is about 0.1% return at the quarterly rate (0.4 at the annual rate), while in the case with GEV innovations, it is 0.15% at the quarterly rate (0.6 at the annual rate).

Second, concerning inflation, variance risk carries a negative price, while skewness risk carries a positive price. In this case, positive inflation skewness may induce exceptionally large inflation realizations which reduce the real payoff and prices of bonds. In order to make bonds attractive in this market, the yield should be higher than in a market with no or negative inflation skewness. Quantitatively, price of inflation skewness risk is small.

Finally, comparing the prices of productivity and inflation risk shows that the former carries a much larger price than the latter. Since estimates of the second- and third-order moments of productivity and inflation innovations are of the same order of magnitude (see Table 1), we can conclude that, as a rough approximation, all risk priced in this model is productivity (or consumption) risk.

5 Dynamics

In this section, I use impulse-response analysis to study the dynamic effects of inflation and productivity shocks on bond yields. Since the model is non-linear, the effects of a shock depend on its sign, size, and timing (see Gallant et al.,1993, and Koop et al., 1996). For this reason, I consider shocks of different sign and size, and assume that they 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. In particular, I study innovations in the 5th, 25th, 75th and 95th percentiles plus the median innovation. Of course, the size (in absolute value) of innovation in the 5th and 95th (and in the 25th and 75th) percentiles are not same when the distribution is asymmetric, but the point is that the likelihood of these two realizations is the same. Responses are reported in Figures 9 and 10, with the vertical axis denoting percentage deviation from the stochastic steady state and the horizontal axis denoting periods (quarters). Yields are in percent at the quarterly rate.

Figure 9 plots the response of yields to inflation shocks drawn from the each of the distributions. Positive (resp. negative) inflation shocks increase (resp. decrease) bond yields, but the magnitude of the response decreases monotonically with the maturity. In the case of the normal distribution, the effect of a positive shock is the mirror image of the negative shock of the same magnitude. This result is, of course, due to the symmetry of the normal distribution. In the case of the skew normal and GEV distributions, there is an asymmetry in that the positive shock in the 95th percentile delivers a larger response than the equally likely negative shock in the 5th percentile. This asymmetry is reversed for the smaller shocks in the 25th and 75th percentiles. Finally, since the median is less than the

mean (zero), the median inflation shock reduces bond yields.

The dynamic effects of inflation shocks are driven by the persistence of inflation: a current, positive shock means higher inflation in future periods, and, as a result, consumers rationally expect a lower real bond prices and payoffs in the future. When inflation is serially uncorrelated, a current shock has no effect on yields because the forecasts of future payoffs and prices are unchanged. In this sense, what matters for bond yields in this model is anticipated, rather than unanticipated, inflation.

Figure 10 plots the response of yields to productivity shocks drawn from the each of the distributions. Positive productivity shocks reduce yields because consumers faced with increased output would attempt to intertemporally smooth consumption by saving. Since this is not possible in the aggregate, bond prices must increase and yields decrease to induce consumers to optimally consume the current output. As before, responses to normal shocks are symmetric, but those to skew normal and GEV shocks are asymmetric. In particular, negative shocks at the 5th and 25th percentiles induce larger effects than the equally-likely, but positive, shocks at the 95th and 75th percentiles. Since the median is larger than the mean (zero), the median productivity shock reduces bond yields.

As in the case of inflation, the dynamic effects depend crucially on the persistence of the productivity process. When productivity is positively autocorrelated, a current positive shock signals above-average output and consumption in future periods as well, and, so, the effect described above take place in both the current and future periods. In contrast, when productivity is serially uncorrelated, the effects take place only in the current period because, by construction, the economy will return to steady state next period.

From these figures, we can see that, conditional on the shock size, the initial effect of productivity shocks is similar for all yields. Hence, productivity shocks push the whole yield curve up or down, just like the level factor does in factor models of the term structure. In contrast, inflation shocks have a larger effect on short- than on long-term yields. As a result, inflation shocks change the slope of the yield curve. In this way, inflation acts like the slope factor in factor models.[12] However, while statistical factors have no structural interpretation, in this model the level shifts and slope changes of the yield curve are empirically associated with macroeconomic variables, respectively productivity and inflation.[13]

[12] A similar result is reported by Wu (2006) based on a calibrated New Keynesian model

[13] Previous research that attempts to build a link between latent and macroeconomic factors in the U.S. term structure 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. Rudebusch and Wu build a model where the slope and level factors depend on the central bank's inflation target and reaction function.

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 skewness risk has a price of 0.6% per year and accounts for about 10% of the risk premia with a 90% confidence interval between 0.8% and 19.6%. 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.

In current and future work, I extend the model here to incorporate nominal frictions. This is important because one of the reasons inflation skewness plays a limited in this paper is that prices and wages are completely flexible and inflation has no real effects. I also currently use extreme value theory to study of optimal monetary policy: just as engineers seek to design structures that withstand extreme events like earthquakes or hurricanes, economists are concerned with designing monetary policy that optimally reacts to extreme events like the oil shocks in the 1970s or the financial shocks associated with the Great Recession.

Table 1
Parameter Estimates

	Distribution		
		Skew	
Parameter	Normal	Normal	GEV
Preferences			
Discount factor	0.990*	0.990*	0.988*
	(0.003)	(0.005)	(0.008)
IES	0.165*	0.158*	0.134*
	(0.040)	(0.037)	(0.043)
Risk aversion	49.775*	49.774*	50.198*
	(0.041)	(5.9593)	(0.191)
Inflation			
Steady state	1.011*	1.011*	1.011*
	(0.001)	(0.001)	(0.001)
AR coefficient	0.635*	0.565*	0.747*
	(0.145)	(0.096)	(0.084)
Correlation parameter	0	0.994*	–
		(0.310)	
Shape parameter	–	–	0.251*
			(0.105)
Scale parameter	–	0.011*	0.002*
		(0.003)	(0.001)
Standard deviation	0.006*		
	(0.002)		
Productivity			
AR coefficient	0.970*	0.970*	0.971*
	(0.001)	(0.024)	(0.001)
Correlation parameter	0	−1	–
Shape parameter	–	–	−0.770*
			(0.237)
Scale parameter	–	0.010*	0.007*
		(0.001)	(0.001)
Standard deviation	0.006*		
	(0.001)		

Note: The table reports SMM estimates of the parameters under each distribution of the model innovations. The figures in parenthesis are standard errors computed using a block bootstrap with 99 replications. During the estimation, the location parameter of the skew normal and GEV distributions was adjusted so that the mean of the innovations is zero. The superscript $*$ denotes significance at the 5% level.

Table 2
Unconditional Skewness

Series	U.S. Data	Distribution: Normal	Distribution: Skew Normal	Distribution: GEV
Consumption growth	−0.606	−0.007	−0.948	−1.273
Inflation rate	1.226	−0.027	0.629	1.650
3-Month T-Bill rate	1.222	0.024	0.330	0.611
6-Month T-Bill rate	1.146	0.022	0.293	0.512
12-Month T-Bill rate	1.011	0.016	0.270	0.427

Note: The table reports the skewness of actual U.S. series and of artificial data simulated from the model under each of the three distributions considered.

Table 3
Jarque-Bera Tests

Series	U.S. Data	Distribution: Normal	Distribution: Skew Normal	Distribution: GEV
Consumption growth	< 0.001	0.346	< 0.001	< 0.001
Inflation rate	< 0.001	0.559	< 0.001	< 0.001
3-Month T-Bill rate	< 0.001	0.532	0.003	< 0.001
6-Month T-Bill rate	< 0.001	0.436	< 0.001	< 0.001
12-Month T-Bill rate	< 0.001	0.327	< 0.001	< 0.001

Note: The table reports p-values of the Jarque-Bera test of the hypothesis that the data follows a normal distribution.

Table 4
Bond Yields

	Maturity				
	1	2	4	6	8
	Normal Distribution				
Certainty-equivalent	2.051	2.051	2.05	2.051	2.051
With variance risk	1.480	1.485	1.494	1.503	1.512
	Skew Normal Distribution				
Certainty-equivalent	2.083	2.083	2.083	2.083	2.083
With variance risk	1.478	1.483	1.493	1.503	1.513
With variance and skewness risk	1.453	1.458	1.468	1.479	1.490
Contribution of skewness risk	−0.026	−0.026	−0.025	−0.024	−0.024
	GEV Distribution				
Certainty-equivalent	2.273	2.273	2.273	2.273	2.273
With variance risk	1.500	1.508	1.522	1.536	1.549
With variance and skewness risk	1.449	1.458	1.474	1.489	1.503
Contribution of skewness risk	−0.051	−0.050	−0.048	−0.047	−0.046

Note: The table reports the ergodic mean of bond yields of selected maturities under different risk scenarios. Yields are expressed in percent return at the quartely rate.

Table 5
The Price of Risk

	Maturity				
	1	2	4	6	8
	Normal Distribution				
Productivity					
Variance risk	−1.436	−1.421	−1.392	−1.365	−1.339
Inflation					
Variance risk	−0.005	−0.009	−0.016	−0.021	−0.025
	Skew Normal Distribution				
Productivity					
Variance risk	−1.510	−1.494	−1.463	−1.434	−1.406
Skewness risk	0.103	0.102	0.099	0.098	0.094
Inflation					
Variance risk	−0.005	−0.009	−0.014	−0.017	−0.020
Skewness risk ($\times 10^{-3}$)	0.017	0.040	0.085	0.115	0.135
	GEV Distribution				
Productivity					
Variance risk	−1.944	−1.922	−1.881	−1.842	−1.805
Skewness risk	0.150	0.148	0.144	0.140	0.136
Inflation					
Variance risk	−0.005	−0.010	−0.021	−0.031	−0.038
Skewness risk ($\times 10^{-3}$)	0.017	0.053	0.162	0.283	0.392

Notes: The table reports the price of different risk sources under each of the three statistical distributions considered. Prices are expressed in percent return at the quarterly rate.

References

[1] Andreasen, M. M., 2012. On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-linear DSGE Models. *Review of Economic Dynamics* 15, pp. 295-316.

[2] Ang, A., Piazzesi, M., 2003. A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics* 50, pp. 723-744.

[3] Backus, D., Chernov, M., Martin, I., 2011. Disasters Implied by Equity Index Options. *Journal of Finance* 66, pp. 1969-2012.

[4] Bansal, R., Coleman, J. W., 1996. A Monetary Explanation of the Equity Premium, Term Premium, and Risk-Free Rate Puzzles. *Journal of Political Economy* 104, pp. 1135-1171

[5] Barro, R. J., 2006. Rare Disasters and Asset Markets in the Twentieth Century. *Quarterly Journal of Economics* 121, pp. 823-866.

[6] Barro, R. J., Nakamura, E., Steinsson, J., Ursua, J., 2013. Crises and Recoveries in an Empirical Model of Consumption Disasters. *American Economic Journal: Macroeconomics* 5, pp. 35-74.

[7] Brunnermeier, M. K., Gollier, C., Parker, J. A., 2007. Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns. *American Economic Review* 97, pp. 159-165.

[8] Caldera, D., Fernández-Villaverde, J., Rubio-Ramírez, J. F., 2012. Computing DSGE Models with Recursive Preferences. *Review of Economic Dynamics* 15, pp. 188–206

[9] Campbell, J. Y., 1996. Understanding Risk and Return. *Journal of Political Economy* 104, pp. 298-345.

[10] Chang, B. Y., Christoffersen, P., Jacobs, K., 2012. Market Skewness Risk and the Cross-Section of Stock Returns. *Journal of Financial Economics*, forthcoming.

[11] Culbertson, J. M., 1957. The Term Structure of Interest Rates. *Quarterly Journal of Economics* 71, pp. 485-517.

[12] Diebold, F. X., Rudebusch, G., Aruoba, B., 2006. The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. *Journal of Econometrics* 131, pp. 309-338.

[13] Doh, T., 2013. Long-Run Risks in the Term Structure of Interest Rates: Estimation. *Journal of Applied Econometrics* 28, pp. 478-497.

[14] Duffie, D., Singleton, K. J., 1993. Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica* 61, pp. 929-952.

[15] Fisher, R. A., Tippett, L. H. C., 1928. On the Estimation of the Frequency Distributions of the Largest or the Smallest Member of a Sample. *Proceedings of the Cambridge Philosophical Society* 24, pp. 180-190.

[16] Epstein, L., Zin, S., 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57, pp. 937-969.

[17] Epstein, L., Zin, S., 1991. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns. *Journal of Political Economiy* 99, pp. 263-286.

[18] Gallant, A. R., Rossi, P. E., Tauchen, G., 1993. Nonlinear Dynamic Structures. *Econometrica* 61, pp. 871-908.

[19] Gnedenko, B. V., 1943. Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics* 44, pp. 423-453.

[20] Golec, J, Tamarkin, M., 1998. Bettors Love Skewness, Nor Risk, at the Horse Track. *Journal of Political Economy* 106, pp. 205-225.

[21] Gregory, A., Voss, G. M., 1991. The Term Structure of Interest Rates: Departures from Time-Separable Expected Utility. *Canadian Journal of Economics* 24, pp. 923-939.

[22] Gurio, F., 2012, Disaster Risk and Business Cycles. *American Economic Review* 102, pp. 2734–2766.

[23] Hall, R. E., 1988. Intertemporal Substitution in Consumption. *Journal of Political Economy* 96, pp. 339-357.

[24] Hansen, L. P., Singleton, K., 1982. Generalized Instrumental Variable Estimation of Nonlinear Rational Expectations Models. *Econometrica* 50, pp. 1269-1286.

[25] Hansen, L. P., Singleton, K., 1983. Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns. *Journal of Political Economy* 91, pp. 249-265.

[26] Harvey, C. R., Siddique, A., 2000. Conditional Skewness and Asset Pricing Tests. *Journal of Finance* 55, pp. 1263-1295.

[27] Hicks, J., 1939. *Value and Capital*, Oxford: Clarendon Press.

[28] Jenkinson, A.F., (1955). The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements. *Quarterly Journal of the Royal Meteorological Society* 81, pp. 158-171.

[29] Jermann, U. J., 1998. Asset Pricing in Production Economies. *Journal of Monetary Economics* 41, pp. 257-275.

[30] Jin, H., Judd, K. L., 2002. Perturbation methods for general dynamic stochastic models. Hoover Institution, Mimeo.

[31] Kapadia, N., 2006. The Next Microsoft? Skewness, Idyosincratic Volatility, and Expected Returns. Rice University, Mimeo.

[32] Kim. J., Kim, S., Schaumburg E., Sims, C., 2008. Calculating and Using Second-Order Accurate Solutions of Discrete-Time Dynamic Equilibrium Models. *Journal of Economic Dynamics and Control* 32, pp. 3397-3414.

[33] Koop, G., Pesaran, M. H., Potter, S., 1996. Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics* 74, pp. 119-147.

[34] Kraus, A., Litzenberger, R. H., 1976. Skewness Preference and the Valuation of Risky Assets. *Journal of Finance* 31, pp. 1085-1100.

[35] Kraus, A., Litzenberger, R. H., 1983. On the Distributional Conditions for a Consumption Oriented Three Moment CAPM. *Journal of Finance* 38, pp. 1381-1391.

[36] Le, A., Singleton, K. J., 2010. An Equilibrium Term Structure Model with Recursive Preferences. *American Economic Review* 100, pp. 1-5.

[37] Lee, B.-S., Ingram, B. F., 1991. Simulation Estimation of Time-Series Models. *Journal of Econometrics* 47, pp. 195-205.

[38] Ljungqvist, L., Sargent, T., 2004. *Recursive Macroeconomic Theory*, Cambridge: MIT Press.

[39] Martin, I., 2012. Consumption-Based Asset Pricing with Higher Cumulants. *Review of Economic Studies*, forthcoming.

[40] Modigliani, F., Sutch, R., 1966. Innovations in Interest Rate Policy. *American Economic Review* 56, pp. 178-197.

[41] Piazzesi, M., Schneider, M., 2007. Equilibrium Yield Curves. In *NBER Macroeconomics Annual 2006*, Acemoglu, D., Rogoff, K., and Woodford M. (eds.), pp. 389-472, Cambridge: MIT Press.

[42] Rietz, T., 1988. The Equity Risk Premium: A Solution. *Journal of Monetary Economics* 22, pp. 117-131.

[43] Rocco, M., 2011. Extreme Value Theory for Finance: A Survey. Bank of Italy, Occasional Paper No. 99.

[44] Rudebusch, G., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy. *Economic Journal* 118, pp. 906-926

[45] Rudebusch, G., Swanson, E., 2012. The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks. *American Economic Journal: Macroeconomics* 4, pp. 105-143.

[46] Ruge-Murcia, F., 2012. Estimating Nonlinear DSGE Models by the Simulated Method of Moments: With an Application to Business Cycles, *Journal of Economic Dynamics and Control* 36, pp. 914-938.

[47] Schmitt-Grohé, S., Uribe, M., 2004. Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function. *Journal of Economic Dynamics and Control* 28, pp. 755-775.

[48] Tallarini, T. D., 2000. Risk-Sensitive Real Business Cycle Models. *Journal of Monetary Economics* 45, pp. 507-532.

[49] van Binsbergen, J., Fernández-Villaverde, J., Koijen, R, Rubio-Ramirez, J., 2010. The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences. Stanford University, Mimeo.

[50] Vissing-Jørgensen, A., 2002. Limited Asset Market Participation and the Elasticity of Intertemporal Substitution. *Journal of Political Economy* 110, pp. 825-853.

[51] Wu, T., 2006. Macro Factors and the Affine Term Structure of Interest Rates. *Journal of Money, Credit and Banking* 38, pp. 1847-1876.

Figure 1: Estimated CDF of Inflation Innovations

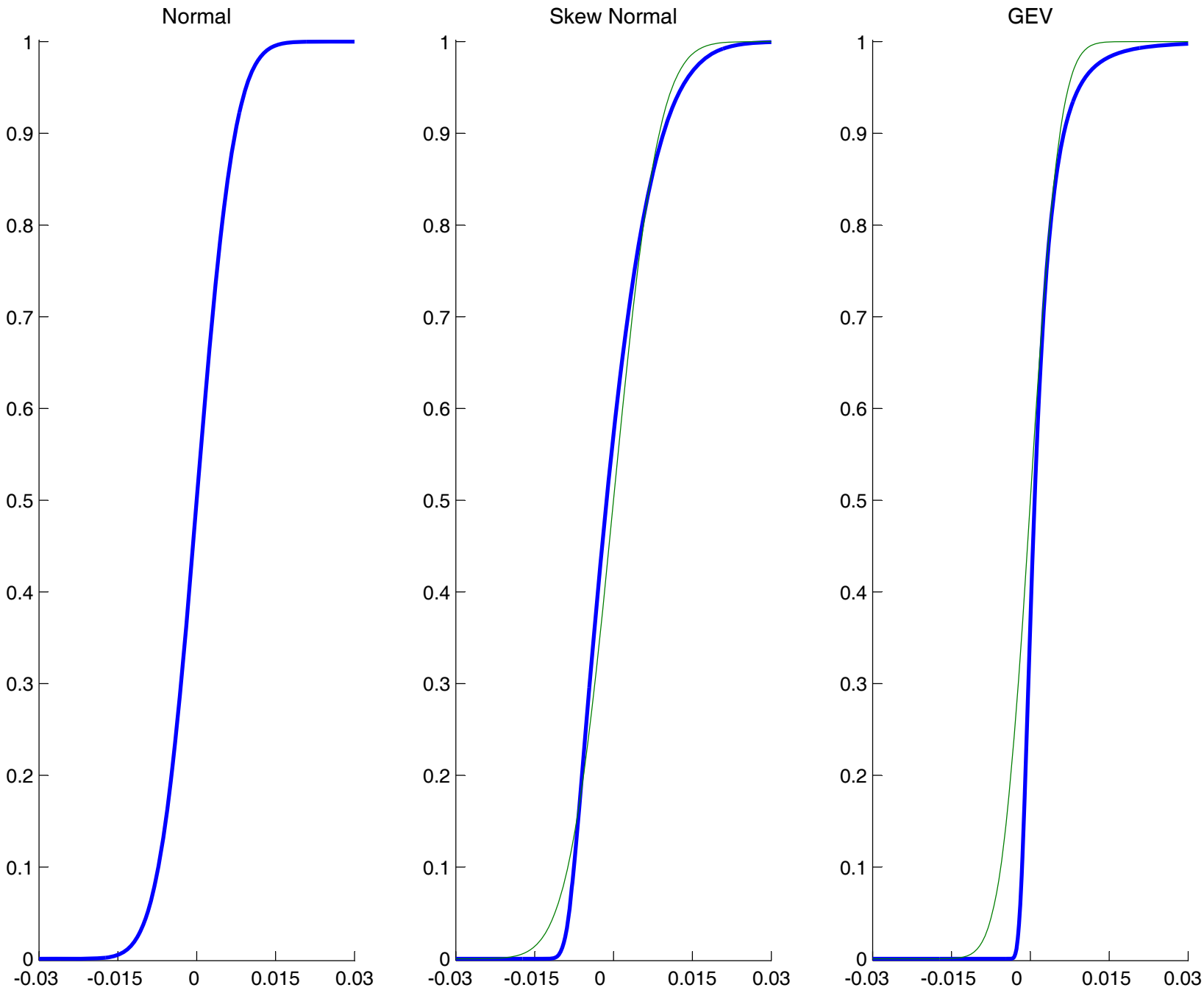

Figure 2: Estimated CDF of Productivity Innovations

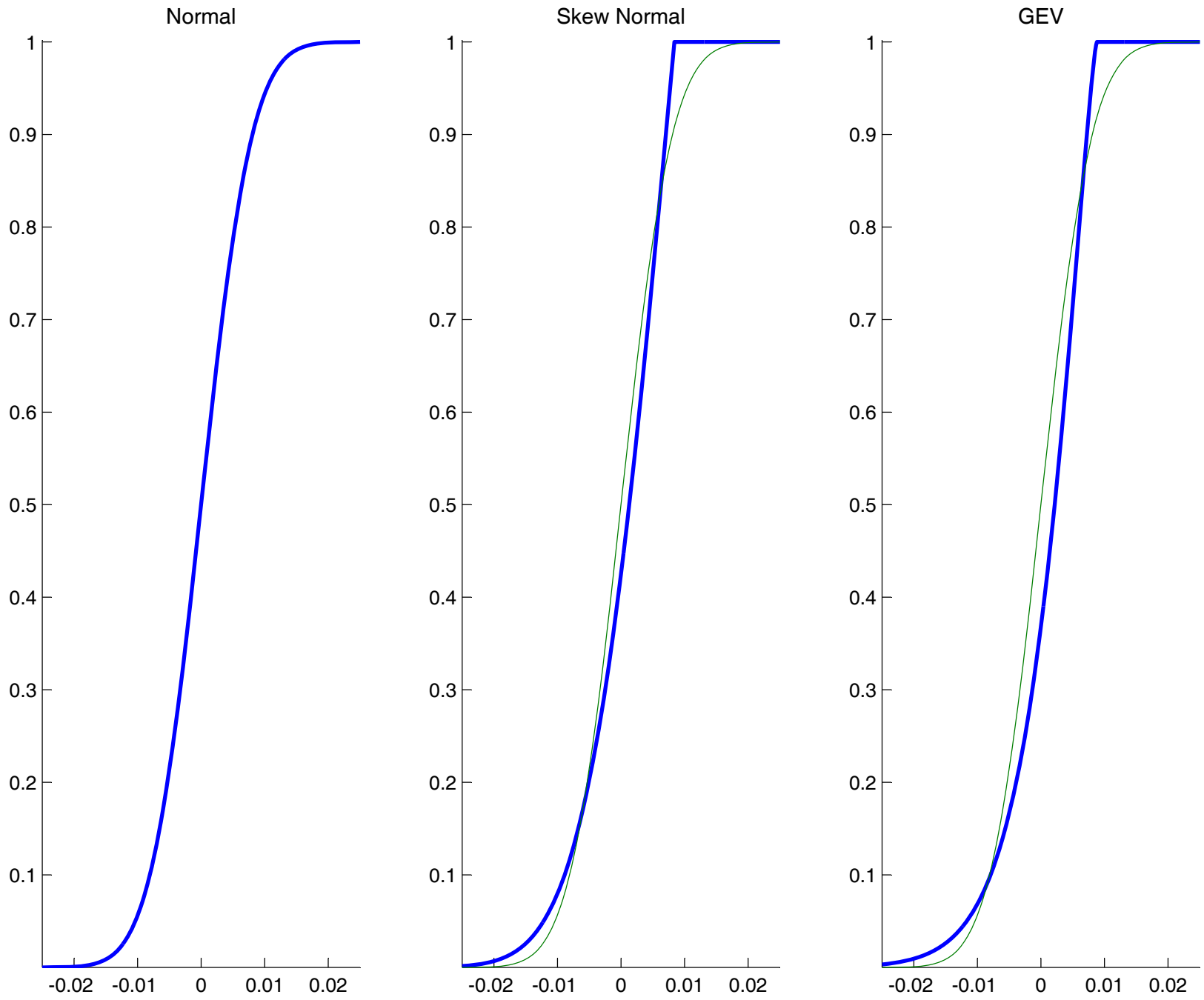

Figure 3: Fit of the Moments

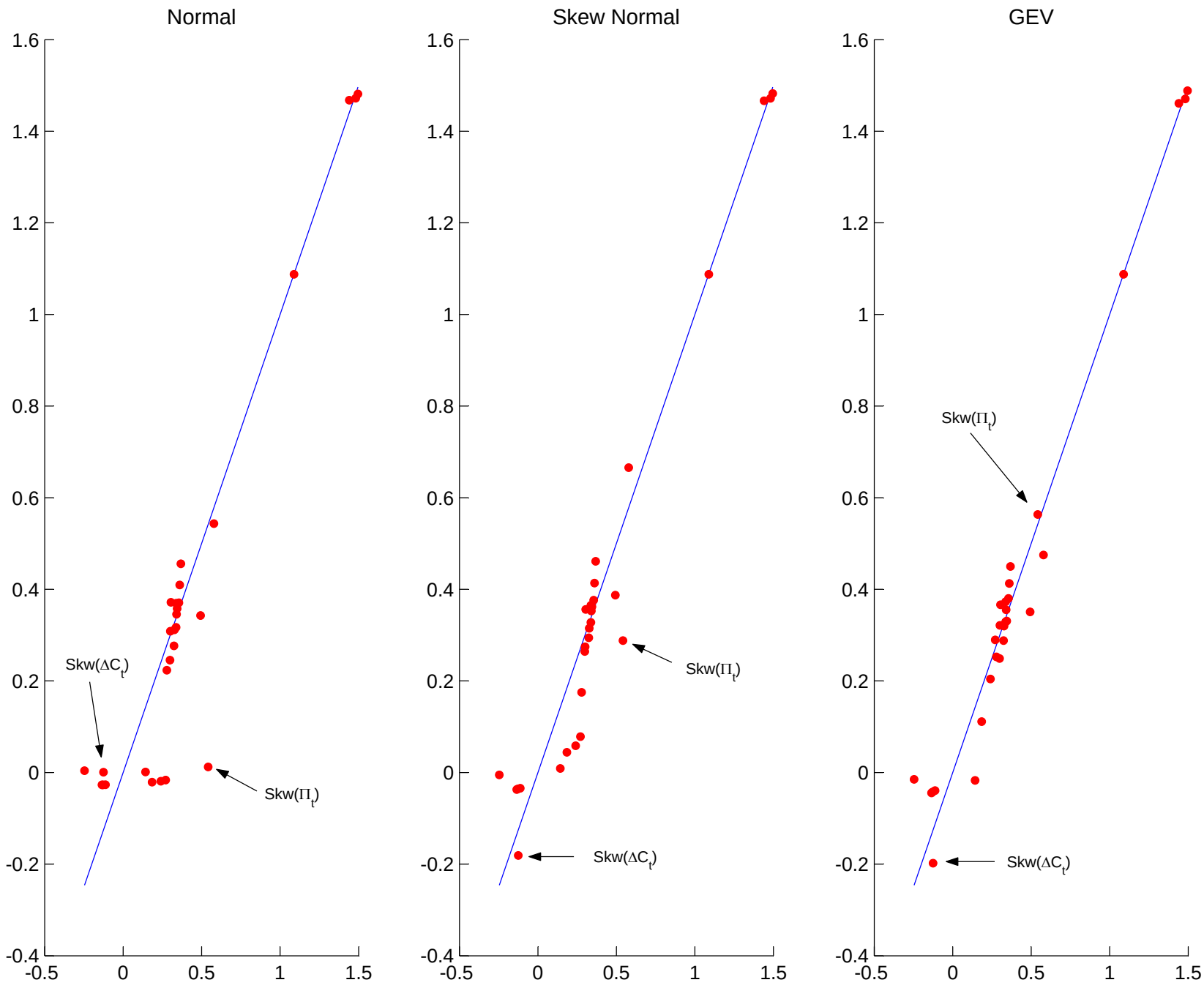

Figure 4: Fit of the Data using Skew Normal Distribution

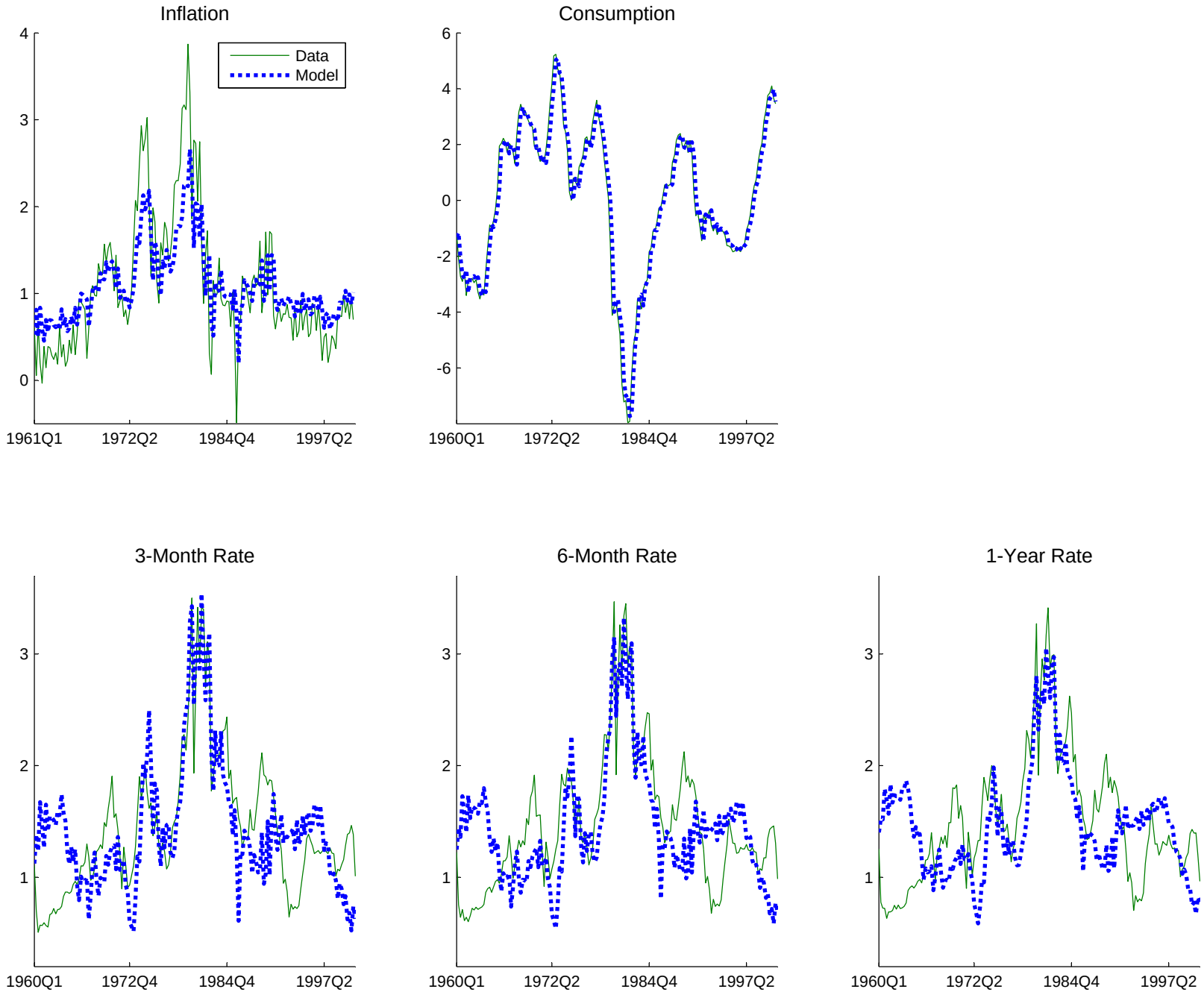

Figure 5: Fit of the Data using GEV Distribution

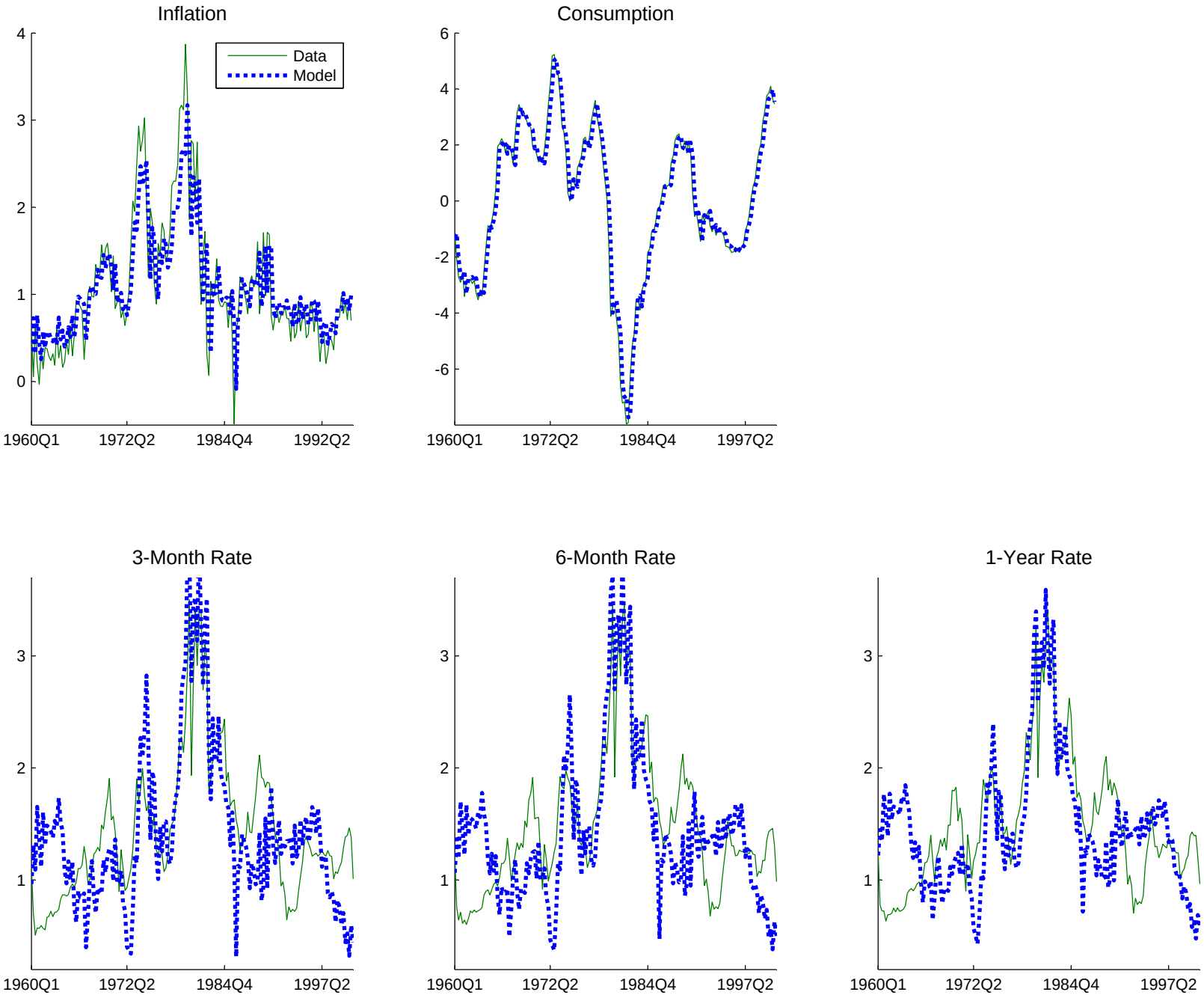

Figure 6: Fit of the Data using Normal Distribution

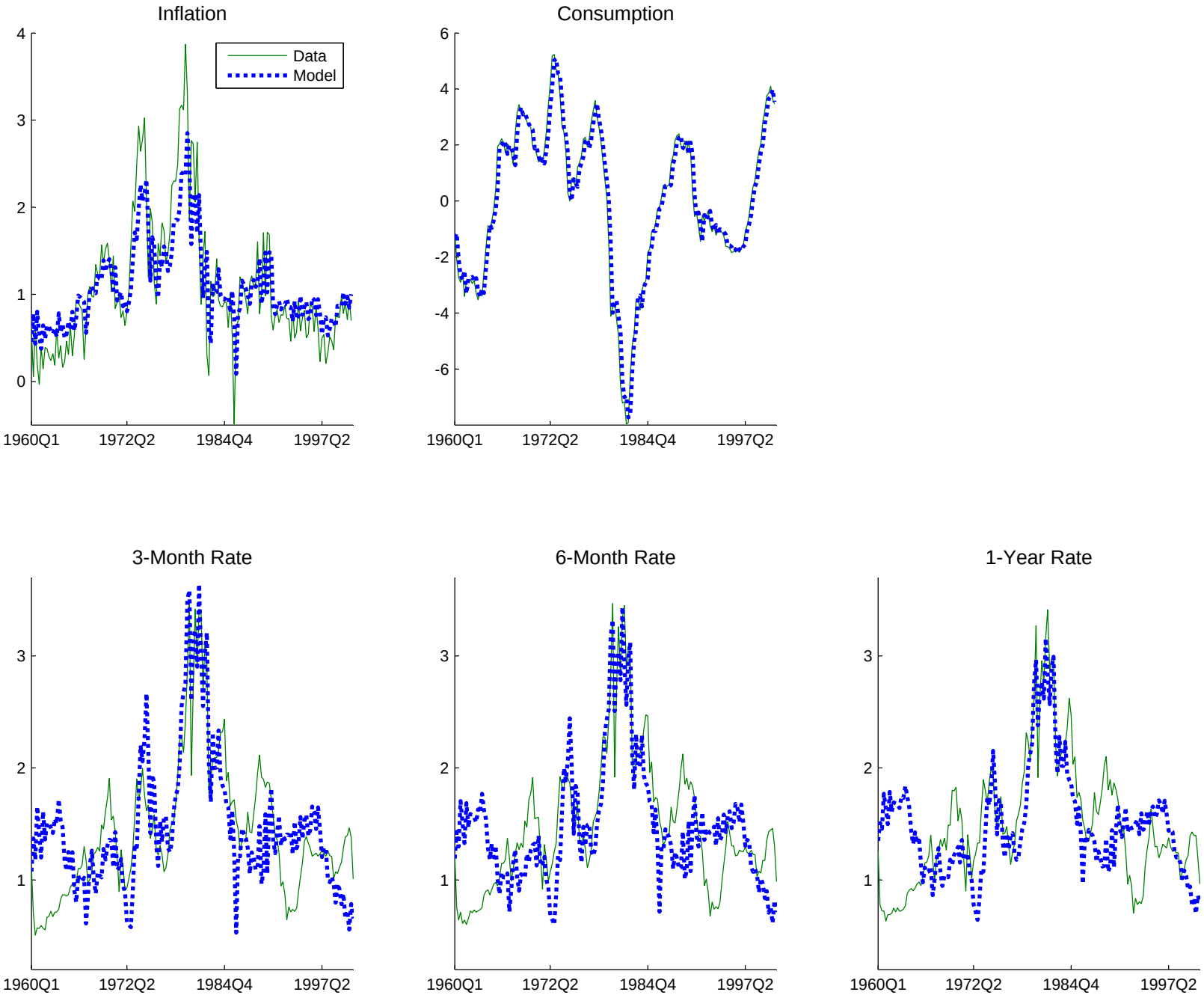

Figure 7: Skewness in U.S. Data

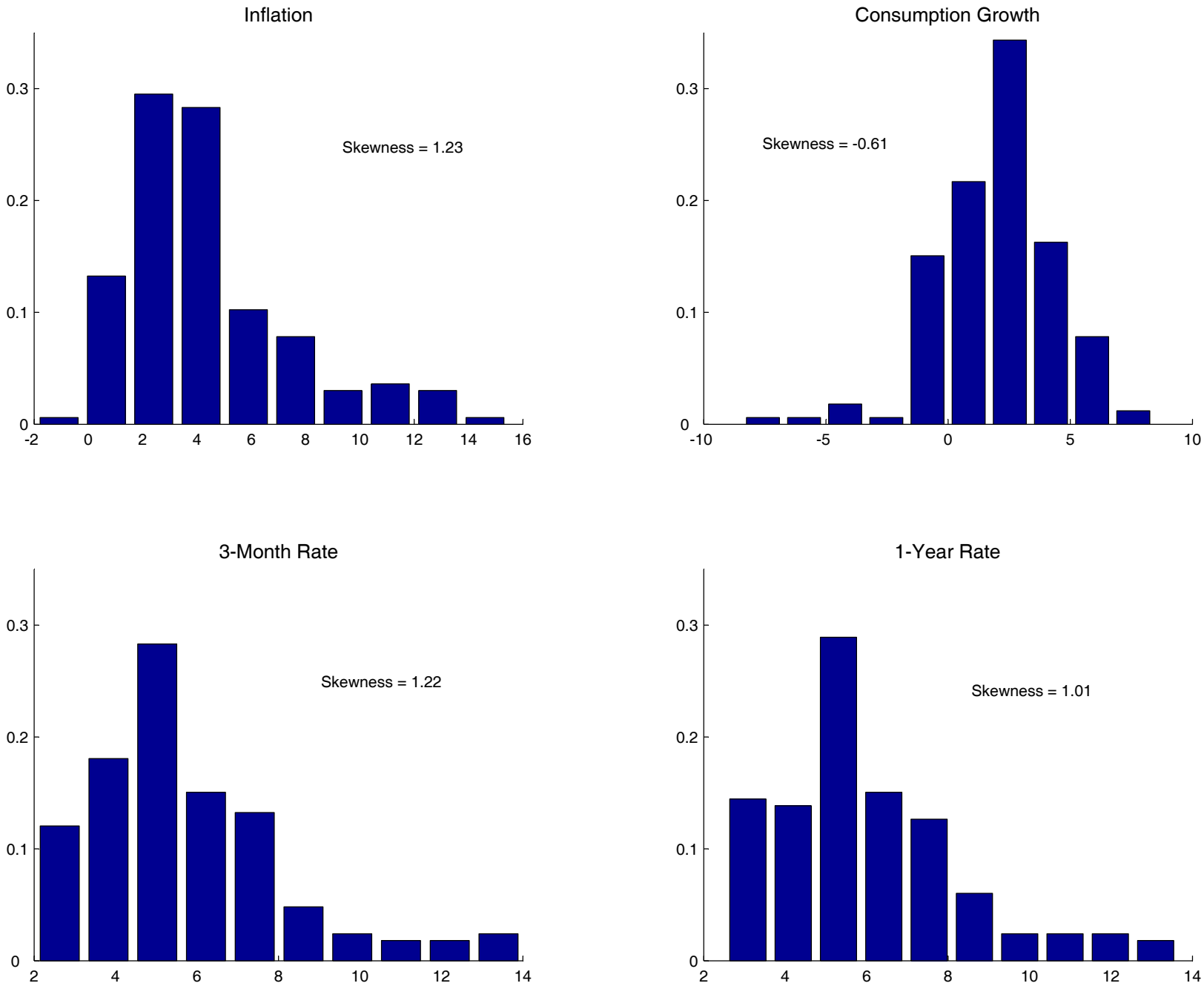

Figure 8: Bond Premia

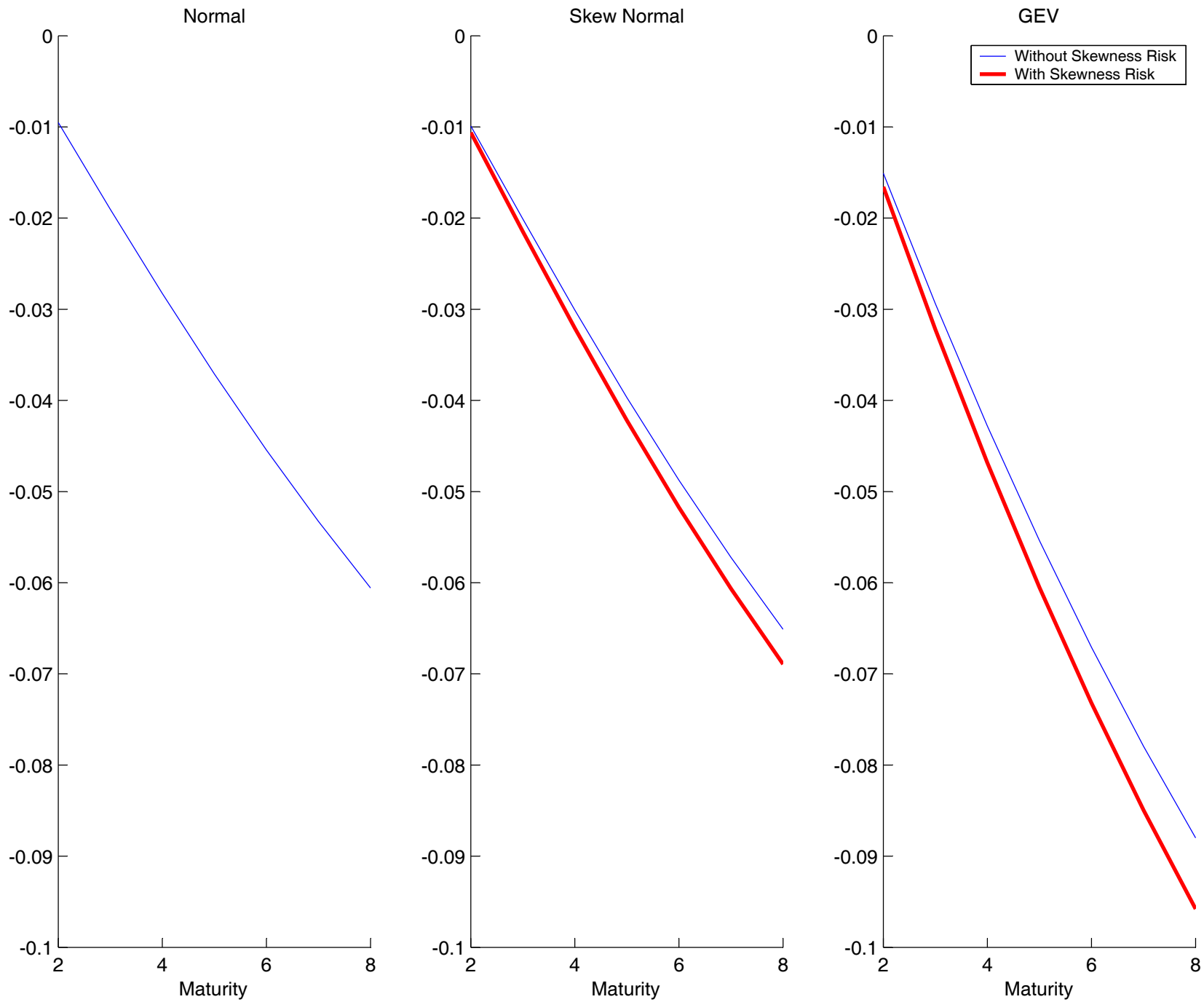

Figure 9: Responses to Inflation Shocks

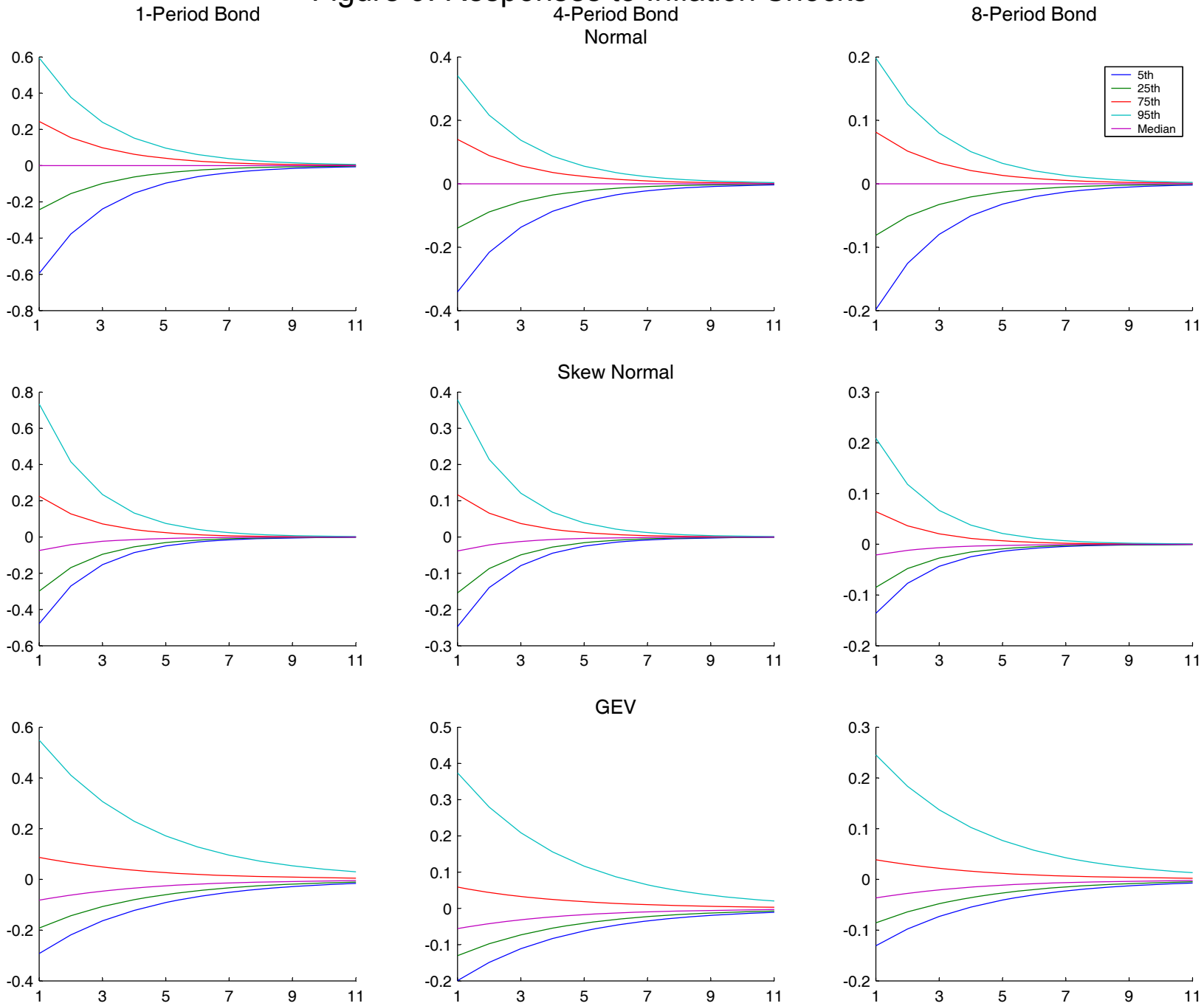

Figure 10: Response to Productivity Shocks