Extreme Events and the Fed*

Jinill Kim† Francisco Ruge-Murcia‡

August 2014
Preliminary and Incomplete

Abstract

This paper applies extreme value theory to monetary policy design. The analysis is based on a small-scale New Keynesian model with sticky prices and wages where shocks are drawn from an asymmetric generalized extreme value (GEV) distribution. A non-linear perturbation of the model is estimated by the simulated method of moments. Under the Ramsey policy, a benevolent central bank responds non-linearly and asymmetrically to shocks and induces different ergodic means for inflation and the interest rate compared with the symmetric case. The relative contribution of shock asymmetry and model non-linearity is evaluated using restricted versions of the benchmark model. The analysis sheds some light on the Fed's actions during the recent Financial Crisis.

JEL Classification: E4, E5.
Key Words: Extreme value theory, nonlinear models, skewness risk, monetary policy, third-order perturbation, simulated method of moments.

*The second author thanks the Social Sciences and Humanities Research Council of Canada for financial support.

†Department of Economics, Korea University, Korea. E-mail: jk9n@hotmail.com
‡Department of Economics, McGill University, Canada. E-mail: francisco.ruge-murcia@mcgill.ca

1 Introduction

This paper applies extreme value theory to the study of monetary policy. Extreme value theory is a branch of statistics concerned with extreme events and was primarily developed in engineering, where designers seek to protect structures against infrequent but potentially damaging events like earthquakes or hurricanes.[1] Economies are also subject to extreme shocks—think, for example, of the oil shocks in the 1970s or the financial shocks associated with the Great Recession—and, so it is important to design monetary policy with the possibility of extreme events in mind.

This paper constructs and estimates the model of a dynamic economy where extreme shocks can occasionally happen and, as a result, agents and policy makers face skewness risk. In particular, the analysis is carried out using a small-scale New Keynesian model with sticky prices and wages where shocks are drawn from a generalized extreme value (GEV) distribution. This distribution is widely used in extreme value theory to model the maxima (or minima) of long sequence of random variables. The distribution has three independent parameters, which have implications for location, scale and shape—the first, second and third moments of the distribution. To be consistent with considering three moments of the distribution, we approximate the model dynamics using a third-order perturbation and the model is, therefore, nonlinear. The nonlinear model is estimated by the simulated method of moments (SMM). In order to disentangle the relative contribution of asymmetric shocks and nonlinearity to our results, we also estimate a nonlinear version of the model with normal shocks, and a linear version with GEV shocks. Statistical results show that the data reject the hypothesis that shocks are normally distributed and favor instead an asymmetric distribution. In particular, the data prefer a specification where monetary policy 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. The positive implications of skewness for monetary policy are analyzed using impulse-response analysis and our analysis sheds some light on the Fed's actions during the recent Financial Crisis.

Using the estimated parameters, we pursue a normative analysis in the context of a Ramsey planner. We show that under the Ramsey policy, a benevolent central bank responds asymmetrically to shocks and induces different ergodic means for inflation and interest rates, compared with the symmetric case. Specially, since the shocks are not symmetrically dis-

[1] Key contributions in extreme value theory are those of Fisher and Tippett (1928), Gnedenko(1943), and Jenkinson (1955). For a review of applications of this theory in engineering, meteorology, and insurance, see Embrechts, Klüppelberg, and Mikosch (1997) and Coles (2001).

tributed, the shocks at the 1st and 99th percentiles are asymmetrically located about the mean. We also find that the Ramsey planner responds more strongly to large shocks and to variance and skewness risk than a policy maker that follows a Taylor-type rule. In addition, to deriving the optimal monetary policy response to large shocks, this paper also derives specific policy prescriptions concerning optimal inflation targets. This issue is important because in light of the recent Financial Crisis, Blanchard et al. (2010) propose inflation targets of 4% per year (as opposed to the 2% per year used, for example, by the Bank of Canada) in order to provide a larger buffer zone from the zero-lower bound on interest rates.

Previous research on the positive analysis of monetary policy typically works under the dual assumptions that the propagation mechanism is linear and shocks are symmetric, usually normal. In some normative analysis, it is necessary to go beyond a linear approximation of the model dynamics to avoid spurious welfare implications, but a second-order approximation is consistent with any two-parameter distribution. Since the Normal distribution satisfies this two-degrees-of-freedom specification, the normal is also widely used in normative analysis. This strategy leads to tractable models but, as we argue below, it is unsatisfactory to understand policy responses to extreme events.

The paper is organized as follows. Section 2 presents a small-scale New Keynesian model occasionally subject to extreme shock realizations. Section 3 discusses the estimation method, data and identification, and reports estimates for the three estimated specifications. Section 4 examines the positive implications of the model using impulse-response analysis. Section 5 studies optimal monetary policy in an environment where extreme events can happen. Finally, Section 6 concludes.

2 The Model

The agents in this economy are 1) infinitely-lived households with idiosyncratic job skills, 2) firms that produce each a differentiated good, and 3) a monetary authority. This section describes the behavior of these agents and the resulting equilibrium.

2.1 Households

Household $h \in [0, 1]$ maximizes

$$E_\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau} U(c_t^h, n_t^h),$$

where E_τ is the expectation conditional on information available at time τ, $\beta \in (0, 1)$ is the discount factor, $U(\cdot)$ is an instantaneous utility function, c_t^h is consumption, and n_t^h is hours worked. Consumption is an aggregate of differentiated goods indexed by $j \in [0,1]$,

$$c_t^h = \left(\int_0^1 (c_{j,t}^h)^{1/\mu} dj \right)^{\mu}, \tag{1}$$

where $\mu > 1$ is a parameter that determines the elasticity of substitution between goods. The utility function is

$$U(c_t^h, n_t^h) = \ln(c_t^h) - \psi n_t^h / z_t, \tag{2}$$

where $\psi > 0$ is a weight and z_t is a labor supply shock that shifts the disutility of labor. (The stochastic process of this and the other shocks in the model are specified in section 2.4). The linear specification for hours worked in (2) is based on Hansen (1985).[2]

Nominal wages are assumed to be rigid as a result of labor-market frictions that induce a cost whenever there is an adjustment. This cost is modeled using the convex function

$$\Phi_t^h = \Phi(W_t^h / W_{t-1}^h) = \left(\frac{\phi}{2} \right) \left(\frac{W_t^h}{W_{t-1}^h} - 1 \right)^2, \tag{3}$$

where W_t^h is the nominal wage and $\phi \geq 0$ is a parameter. In the especial case where $\phi = 0$, nominal wages are flexible.

In every period the household is subject to the budget constraint

$$c_t^h + \frac{Q_t A_t^h - A_{t-1}^h}{P_t} + \frac{B_t^h - i_{t-1} B_{t-1}^h}{P_t} = \left(1 - \Phi_t^h\right) \left(\frac{W_t^h n_t^h}{P_t} \right) + \frac{D_t^h}{P_t}, \tag{4}$$

where A_t is a complete set of Arrow-Debreu securities, Q_t is a vector of prices, B_t is a one-period nominal bond, i_t is the gross nominal interest rate, D_t are dividends, and P_t is an aggregate price index. The index is defined as

$$P_t = \left(\int_0^1 (P_{j,t})^{1/(1-\mu)} dj \right)^{1/(1-\mu)}, \tag{5}$$

with $P_{j,t}$ denoting the nominal price of good j. In addition to this budget constraint and a no-Ponzi-game condition, utility maximization is subject to the demand for labor h by firms (see the firms' problem below).

First-order conditions include a wage Phillips curve that equates the marginal costs and benefits of increasing W_t^h,

$$\begin{aligned} & \Lambda_t^h n_t^h \left(\left(\frac{\varsigma}{\varsigma - 1} \right) \left(1 - \Phi_t^h\right) + \Omega_t^h \left(\Phi_t^h\right)' \right) \\ = \; & \Lambda_t^h n_t^h \left(1 - \Phi_t^h\right) + \left(\frac{\varsigma}{\varsigma - 1} \right) \left(\frac{n_t^h / z_t}{W_t^h / P_t} \right) + \beta E_t \left(\frac{\Lambda_{t+1}^h}{\Pi_{t+1}} \left(\Omega_t^h\right)^2 n_{t+1}^h \left(\Phi_{t+1}^h\right)' \right), \end{aligned} \tag{6}$$

[2] In preliminary work, we considered a more general formulation of the utility function but estimates were quantitatively close to the log-linear specification in the text. Section 5 examines the implications of the model under alternative parameterizations of the utility function.

where $\Lambda_t^h = 1/c_t^h$ is the marginal utility of consumption, $\varsigma/(\varsigma-1)$ is the elasticity of substitution between labor types, $\Pi_t = P_t/P_{t-1}$ is gross price inflation, $\Omega_t^h = W_t^h/W_{t-1}^h$ is gross wage inflation for the labor of type h, and $\left(\Phi_t^h\right)'$ denotes the derivative of the cost function with respect to its argument. The costs in the left-hand side of (6) are the wage adjustment cost and the decrease in labor income as firms substitute away from the more expensive labor. The benefits in the right-hand side are the increase in labor income per hour worked, the increase in leisure time, and the reduction in the future expected wage adjustment cost.

The optimal consumption of good j satisfies

$$c_{j,t}^h = \left(\frac{P_{j,t}}{P_t}\right)^{-\mu/(\mu-1)} c_t^h, \tag{7}$$

which is decreasing in the relative price with elasticity $-\mu/(\mu-1)$.

2.2 Firms

Firm $j \in [0,1]$ produces output, $y_{j,t}$, using the technology

$$y_{j,t} = x_t n_{j,t}^{1-\alpha}, \tag{8}$$

where $n_{j,t}$ is labor input, $\alpha \in (0,1)$ is a production parameter and x_t is a productivity shock. Labor input is an aggregate of heterogeneous labor supplied by households,

$$n_{j,t} = \left(\int_0^1 (n_{j,t}^h)^{1/\varsigma} dh\right)^{\varsigma}, \tag{9}$$

where $\varsigma > 1$ is a parameter that determines the elasticity of substitution between labor types. The price of the labor input is

$$W_{j,t} = \left(\int_0^1 (W_t^h)^{1/(1-\varsigma)} dh\right)^{1-\varsigma}. \tag{10}$$

Nominal price adjustments entail a convex cost that is modeled using the function

$$\Gamma_t^j = \Gamma(P_{j,t}/P_{j,t-1}) = \left(\frac{\gamma}{2}\right)\left(\frac{P_{j,t}}{P_{j,t-1}} - 1\right)^2, \tag{11}$$

where $\gamma \geq 0$ is a parameter. In the case where $\gamma = 0$, prices are flexible. This model of price rigidity is due to Rotemberg (1982).

The firm maximizes

$$E_\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau}(\Lambda_t/\Lambda_\tau)\left(\left(1-\Gamma_t^j\right)(P_{j,t}/P_t)c_{j,t} - \int_0^1 (W_t^h/P_t)n_t^h dh\right),$$

where the term in parenthesis is profits, Λ_t is the aggregate counterpart of Λ_t^h, and $c_{j,t} = \int_0^1 c_{j,t}^h dh$ is total consumption demand for good j. The maximization is subject to the downward-sloping consumption demand function (7), the technology (8), and the condition that supply must meet demand for good j at the posted price.

First-order conditions include the price Phillips curve,

$$\begin{aligned} & \Lambda_t c_{j,t} \left(\left(\frac{\mu}{\mu - 1} \right) \left(1 - \Gamma_t^j \right) + \Pi_t^j \left(\Gamma_t^j \right)' \right) \\ = \; & \Lambda_t c_{j,t} \left(1 - \Gamma_t^j \right) + \left(\frac{\mu}{\mu - 1} \right) \left(\frac{\lambda_t \Lambda_t y_{j,t}}{P_{j,t}/P_t} \right) + \beta E_t \left(\frac{\Lambda_{t+1}}{\Pi_{t+1}} \left(\Pi_t^j \right)^2 c_{j,t+1} \left(\Gamma_{t+1}^j \right)' \right), \end{aligned} \tag{12}$$

where λ_t is the real marginal cost and Π_t^j is the gross inflation rate of good j. This condition equates the marginal costs and benefits of increasing $P_{j,t}$. The costs are the decrease in sales, which is proportional to the elasticity of substitution between goods, and the price adjustment cost. The benefits are the increase in revenue for each unit sold, the decrease in the marginal cost, and the reduction in the future expected price adjustment cost.

The optimal demand for (total) labor satisfies the condition that the marginal productivity of labor equals its real cost,

$$(1 - \alpha) x_t n_{j,t}^{-\alpha} = W_{j,t}/P_{j,t}.$$

The optimal demand for labor h is

$$n_t^h = \left(\frac{W_t^h}{W_t} \right)^{-\varsigma/(\varsigma - 1)} n_{j,t},$$

where $-\varsigma/(\varsigma - 1)$ is the elasticity of demand of labor h with respect to its relative wage.

2.3 The Fed

The monetary authority (let us call it the Fed, for short) sets the interest rate following the Taylor-type rule

$$\ln(i_t/i) = \eta_1 \ln(i_{t-1}/i) + \eta_2 \ln(\Pi_t/\Pi) + \eta_3 \ln(n_t/n) + e_t, \tag{13}$$

where $\eta_1 \in (-1, 1)$, η_2 and η_3 are constant parameters, variables without time subscript denote steady-state values, and e_t is a monetary shock that represents factors beyond the control of the Fed that also affect the nominal interest rate.

2.4 Shocks

Define the 3×1 vector $\xi_t = [\ln(z_t)\ \ln(x_t)\ \ln(e_t)]'$ with the current realization of the labor supply, productivity, and monetary shocks, the 3×1 vector $\varepsilon_t = [\varepsilon_{z,t}\ \varepsilon_{x,t}\ \varepsilon_{e,t}]'$ with the innovations to these shocks, and the 3×3 matrix

$$\rho = \begin{bmatrix} \rho_z & 0 & 0 \\ 0 & \rho_x & 0 \\ 0 & 0 & \rho_e \end{bmatrix}, \tag{14}$$

with all eigenvalues inside the unit circle. Then, the process of the structural shocks of the model is

$$\xi_t = \rho \xi_{t-1} + \varepsilon_t. \tag{15}$$

We assume that the vector of innovations, ε_t, is independent and identically distributed (i.i.d.) with mean zero, diagonal variance-covariance matrix, and its elements are drawn from a generalized extreme value (GEV) distribution.

The GEV distribution due to Jenkinson (1955) is the most widely used distribution in extreme value analysis. The reason is that under the Fisher-Tippett theorem (Fisher and Tippett, 1928), the maxima of a sample of i.i.d. random variables converge in distribution to either of three possible distributions, namely, the Gumbel, Fréchet, and Weibull distributions, but, as shown by Jenkinson, all three can be represented in a unified way using the GEV distribution. The distribution is described by three parameters: a location, a scale and a shape parameter. Depending on whether the shape parameter is zero, larger than zero, or smaller than zero, the GEV distribution corresponds to either the Gumbel, the Fréchet, or the Weibull distribution, respectively.

The shape parameter also determines the thickness of the long tail and the skewness of the distribution. In the case where the shape parameter is non-negative (the GEV is either the Gumbel or the Fréchet distribution), skewness is positive. In the case where the shape parameter is negative (the GEV is the Weibull distribution), skewness can be negative or positive depending on the relative magnitudes of the shape and scale parameters. The fact that the GEV distribution allows for both positive and negative skewness of a potentially large magnitude is particularly attractive for this project because, as we will see below, the U.S. data prefer specifications where the skewness of the innovations is relatively large.[3] Finally, 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 this parameter is

[3]In preliminary work, we considered using the skew normal distribution, whose skewness is bounded between -1 and 1. However, parameter estimates consistently hit the boundary of the parameter space because, in fact, matching the unconditional skewness of the data with our model requires innovations with skewness larger than 1 in absolute value.

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. For additional details on the GEV distribution, see Embrechts, Klüppelberg, and Mikosch (1997) and Coles (2001).

2.5 Equilibrium

In the symmetric equilibrium, all firms are identical and all households are identical. That is, all firms charge the same price, demand the same input quantities, and produce the same output quantity, while all households supply the same amount of labor and receive the same wages. As a result, net holdings of bonds and Arrow-Debreu securities are zero. Substituting the profits of the (now) representative firm into the budget constraint of the (now) representative household delivers the aggregate resource constraint,

$$c_t = y_t - (y_t \Gamma_t + w_t n_t \Phi_t), \tag{16}$$

where y_t is aggregate output and $w_t = W_t/P_t$ is the real wage. In the especial case where prices and wages are flexible, $c_t = y_t$, meaning that all output produced is available for private consumption. Instead, when prices and wages are rigid, part of the output is lost to frictional costs (the term is parenthesis in (16)).

2.6 Solution

Since the model does not have an exact solution, we use a perturbation method to approximate the policy functions around the deterministic steady state using a third-order polynomial and to characterize the local dynamics (Jin and Judd, 2002). A third-order perturbation is of the minimum order necessary to capture the effect of skewness in the policy functions and, as we will see below, allows for rich non-linear dynamics. The solution method is implemented using the MATLAB codes described in Ruge-Murcia (2012), which extend those originally written by Schmitt-Grohé and Uribe (2004) for a second-order perturbation.

In order to appreciate the non-linearity of the solution and the contribution of skewness to the policy functions, it will helpful to explicitly write the third-order perturbation. For a generic variable k in the model, the policy function that solves the dynamic model takes the general form $f(s_t, \sigma)$ where s_t is the vector of state variables and σ is a perturbation parameter that represents the departure from certainty. Then, the third-order approximation of $f(s_t, \sigma)$ around the deterministic steady state ($s_t = s$ and $\sigma = 0$) can be written in tensor notation as

$$[f(s_t, \sigma)]^k = [f(s, 0)]^k + [f_s(s, 0)]^k_a [(s_t - s)]^a \tag{17}$$

$$+(1/2)[f_{ss}(s,0)]^k_{ab}[(s_t-s)]^a[(s_t-s)]^b$$
$$+(1/6)[f_{sss}(s,0)]^k_{abc}[(s_t-s)]^a[(s_t-s)]^b[(s_t-s)]^c$$
$$+(1/2)[f_{\sigma\sigma}(s,0)]^k[\sigma][\sigma]$$
$$+(1/2)[f_{s\sigma\sigma}(s,0)]^k_a[(s_t-s)]^a[\sigma][\sigma]$$
$$+(1/6)[f_{\sigma\sigma\sigma}(s,0)]^k[\sigma][\sigma][\sigma],$$

where a, b, and c are indices, and we have used the results $[f_{s\sigma}(s,0)]^k_a = [f_{\sigma s}(s,0)]^k_a = 0$ (Schmitt-Grohé and Uribe, 2004, p. 763), $[f_{ss\sigma}(s,0)]^k_{ab} = [f_{\sigma ss}(s,0)]^k_{ab} = [f_{s\sigma s}(s,0)]^k_{ab} = 0$ (Ruge-Murcia, 2012, p. 936) and $[f_{\sigma\sigma s}(s,0)]^k_a = [f_{\sigma s\sigma}(s,0)]^k_a = [f_{s\sigma\sigma}(s,0)]^k_a$ by Clairaut's theorem. In this notation, elements like $[f_s(s,0)]^k_a$ and $[f_{ss}(s,0)]^k_{ab}$ are coefficients that depend nonlinearly on structural parameters.

As one would expect to see in a third-order expansion, the policy function (17) includes linear, quadratic, and cubic terms in the state variables. Obviously, the latter two terms are nonlinear and, thus, the overall solution expressing the relation between endogenous and state variables in the model is nonlinear. As we will see below this has important implications for the dynamics of the model and the effect of shocks.

The policy function also includes linear, quadratic, and cubic terms in the perturbation parameter, but the coefficient of the linear term (not shown) is zero. This is just another way of saying that the linear solution (a first-order perturbation) features certainty equivalence. The quadratic term is proportional to the variance, and the cubic term is proportional to the skewness, of the innovations. 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. Thus, the contribution of skewness to the solution is a constant that shifts the policy function above and beyond the shift induced by the variance. Depending on the magnitude of the skewness and its coefficient, this contribution may be sizable.

The policy function also includes cross-terms between the state variables and the perturbation parameter, but, except for the time-varying term in the variance, the other terms have coefficients equal to zero.

3 Estimation

3.1 Data

The data used to estimate the model are quarterly observations of real per-capita consumption, hours worked, the price inflation rate, the wage inflation rate, and the nominal interest rate between 1964Q2 to 2012Q4. The sample starts in 1964 because aggregate data on wages and hours worked are not available prior to that year, and ends with the latest available observation at the time the data was collected. The raw data were taken from the Web Site of the Federal Reserve Bank of St. Louis (*www.stlouisfed.org*). Real consumption is measured by personal consumption expenditures on nondurable goods and services divided by the consumer price index (CPI). The measure of population used to convert this variable into per-capita terms is an estimate produced by the Bureau of Labor Statistics (BLS). Hours worked are measured by average weekly hours of production and nonsupervisory employees in manufacturing. The rates of price and wage inflation are measured by the percentage change in the CPI and the average hourly earnings for private industries, respectively. The nominal interest rate is the effective federal funds rate. Except for the nominal interest rate, all data are seasonally adjusted at the source.

3.2 Method

The model is estimated by the simulated method of moments (SMM). Ruge-Murcia (2012) explains in detail the application of SMM to the estimation of non-linear dynamic models and provides Monte-Carlo evidence on its small-sample properties. Defining $\theta \in \Theta$ to be a $q \times 1$ vector of structural parameters, the SMM estimator, $\widehat{\theta}$, is the value that solves

$$\min_{\{\theta\}} \mathbf{M}(\theta)'\mathbf{W}\mathbf{M}(\theta), \tag{18}$$

where

$$\mathbf{M}(\theta) = (1/T)\sum_{t=1}^{T} m_t - (1/\kappa T)\sum_{\iota=1}^{\kappa T} m_\iota(\theta),$$

$\mathbf{W}$ is a $q \times q$ weighting matrix, T is the sample size, κ is a positive integer, m_t is a $p \times 1$ vector of empirical observations on variables whose moments are of our interest, and $m_\iota(\theta)$ is a synthetic counterpart of m_t with elements obtained from the stochastic simulation of the model. Intuitively, 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. In particular,

$$\sqrt{T}(\widehat{\theta}-\theta) \to N(\mathbf{0}, (1+1/\kappa)(\mathbf{J}'\mathbf{W}^{-1}\mathbf{J})^{-1}\mathbf{J}'\mathbf{W}^{-1}\mathbf{S}\mathbf{W}^{-1}\mathbf{J}(\mathbf{J}'\mathbf{W}^{-1}\mathbf{J})^{-1}), \tag{19}$$

where

$$\mathbf{S} = \lim_{T\to\infty} Var\left((1/\sqrt{T})\sum_{t=1}^{T}\mathbf{m}_t\right), \tag{20}$$

and $\mathbf{J} = E(\partial m_\iota(\theta)/\partial\theta)$ is a finite Jacobian matrix of dimension $p \times q$ and full column rank.[4] In this application, the weighting matrix is the identity matrix and $\mathbf{S}$ is computed using the Newey-West estimator with a Barlett kernel and bandwidth given by the integer of $4(T/100)^{2/9}$ where $T = 195$ is the sample size. The number of simulated observations is ten times larger than the sample size, that is $\kappa = 10$. Thus, the simulated sample has $10 \times 195 = 1950$ observations, which is sufficient to accurately estimate the higher-order moments. The dynamic simulations of the non-linear model are based on the pruned version of the solution, as suggested by Kim, Kim, Schaumburg and Sims (2008).

The estimated parameters are the coefficients that determine price and wage rigidity (ϕ and γ) and the parameters of the productivity, labor supply, and monetary shocks. The moments used to estimate these parameters are the variances, covariances, autocovariances and skewness of the five data series. That is, a total of 25 moments. During the estimation procedure the discount factor (β) is fixed to 0.995, which is the mean of the inverse ex-post real interest rate in the sample period. The steady-state (gross) inflation target (Π) in the monetary policy rule is set to 1. The disutility weight (ψ) is set to 1, but this is just an inconsequential normalization because this parameter only scales the number of hours worked in steady state and does not affect either the model dynamics or the moments used to estimate the model. The curvature parameter of the production function $(1-\alpha)$ is set to 2/3 based on data from the National Income and Product Accounts (NIPA) that show that the share of labor in total income is approximately this value. Finally, the elasticities of substitution between goods and between labor types are fixed to $\mu = 1.1$ and $\varsigma = 1.4$, respectively. This value for μ is standard in the literature. Sensitivity analysis with respect to ς indicates that results are robust to using similarly plausible values.

In addition to the nonlinear model with GEV innovations, we estimate two benchmark versions of the model. The first one is a nonlinear model with normal innovations. This

[4] SMM is generally less efficient than GMM as a result of the term $(1+1/\kappa)$ that captures the uncertainty associated with computing the moments by simulation. However, this term decreases geometrically with κ so that, for example, for $\kappa = 10$, SMM asymptotic standard errors are only 5% larger than those of GMM. On the other hand, for medium scale models computing the moments by simulation can be more computationally efficient than computing them analytically because the latter requires time-consuming matrix inversions.

model allows us to examine the effects of departing from the usual Gaussian assumption. The second one is a linear model with GEV innovations. In this certainty-equivalent model agents ignore the variance and skewness of shocks when making their economic decisions. Thus, this model allows us examine the effect of departing from certainty-equivalence and to quantify the contribution of nonlinearity to our results.

3.3 Identification

Although it is difficult to verify that parameters are globally identified, local identification simply requires

$$rank\left\{E\left(\frac{\partial \mathbf{m}_\iota(\theta)}{\partial \theta}\right)\right\} = q, \tag{21}$$

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

3.4 Estimates

Estimates of the parameters under each of the three models are reported in table 1. Standard errors are computed using the k-step bootstrap proposed by Davidson and MacKinnon (1999) and Andrews (2002).[5] In this preliminary version of the paper, the number of replications is 19 and the number of steps is 5. In current work, we are examining the robustness of our results to using a larger number of replications and steps.

For the two nonlinear models, estimates of the price and wage rigidity parameters are similar across distributions and suggest that wages are substantially more rigid than prices. Quantitatively, these estimates are in line with earlier literature. In contrast, for the linear model, the estimate of the wage rigidity parameter is implausibly large and, in order to match the volatility of wage inflation, this model requires a much larger standard deviation for the labor supply shock than the nonlinear models.

Productivity and labor supply shocks are very persistent and, for the models with GEV innovations, the scale and shape parameters imply that their innovations are negatively skewed. Figure 1 plots the cumulative distribution function (CDF) of productivity innovations estimated under each model (thick line). For the two models with GEV innovations, the figure also plots, as a comparison, the CDF of a normal distribution with the same standard deviation as the GEV (thin line). Notice that CDFs of the GEV distribution have more probability mass in the left tail, and less mass in the right tail, than the normal distribution.

[5] The use of the k-step bootstrap was suggested to us by Silvia Gonçalves.

Thus, large negative productivity surprises can occasionally happen, but large positive ones are unlikely.

Figure 2 plots the estimated CDF of the innovations to the labor supply shock. As in figure 1, the GEV distributions have more mass in the left tail, and less mass in the right tail, than the normal distribution. However, for the linear version of the model, the left tail is much thicker than that of the nonlinear version, and a normal distribution with the same standard deviation features a much larger second moment (this is apparent from the thin line in the right panel of figure 2). Also, notice in table 1 that the estimated standard deviation and skewness of the innovations are, respectively, 16 and 3 times larger in the linear than in the nonlinear model. This observation suggests that whether the propagation mechanism of the model is linear or nonlinear has important quantitative implications for the estimates of parameters of this shock. Finally, although the skewness of the labor supply shock is large, its contribution to the model dynamics is relatively small, except for the fact that its asymmetry is crucial to capture the positive skewness of wage inflation.[6]

The smoothing parameter in the Taylor rule in table 1 is moderately large, and the coefficients of inflation and output are both positive, as expected. For the models with GEV innovations, the estimated scale and shape parameters imply that monetary shocks are positively skewed. In particular, the nonlinear version of the model implies a skewness of 1.09, which is quantitatively close to the 0.79 obtained when we estimated the Taylor rule (equation (13)) alone by maximum likelihood using data on the interest rate, inflation and hours worked.[7] Figure 3 plots the estimated CDF of this shock and shows that its GEV distribution has more probability mass in the right tail, and less mass in the left tail, than a normal distribution with the same standard deviation. That is, the monetary shock, which represents factors that affect the interest rate outside the control of the monetary authority, is not symmetrically distributed around zero. Instead, because the shock is positively skewed, large positive monetary shocks can happen sometimes while large negative ones are uncommon.

The key empirical result in table 1 is that the skewness of all three innovation distributions is quantitatively large and statistically significant. Thus, the null hypothesis that innovations are drawn from a symmetric distribution (whether normal or Student's t) can be safely rejected. (*Note:* To be verified by a larger bootstrap.) The asymmetry in the shocks means that agents in this economy face skewness risk and that the monetary authority will be called

[6]This observation is based on (unreported) sensitivity analysis where we estimated our model without a labor supply shock and found that the fit of the model worsened, primarily because the model could not generate quantitatively large skewness in wage inflation.

[7]When we used the rate of employment instead of hours worked as measure of output, the estimated skewness was about 0.50. Thus, the positive skewness of monetary policy shocks appears to be robust to the estimation method and output measure.

to set policy during and following extreme events, whether large drops in productivity or large interest rate increases due to causes outside its control.

4 Implications

In this section, we examine the positive implications of extreme events for monetary policy.

4.1 Skewness and Kurtosis

Figure 4 plots the histograms of five key macroeconomic series in the U.S. The most obvious observation from this figure is that the data look quite different from a normal distribution (continuous line): none of the series is symmetric around the mean and all of them feature extreme realizations at one tail of the distribution. Consumption and hours worked are negatively skewed, with observations in the left tail corresponding to periods of recession, while price inflation, wage inflation, and the nominal interest rates are positively skewed. We compute the skewness and kurtosis of these series and report them in table 2. The skewness are quantitatively large and the kurtosis are larger than 3 in all cases, with 3 being the kurtosis of the normal distribution. In order to evaluate statistically these departures from normality, we carry out the test proposed by Jarque and Bera (1987) and report the results in table 3. The Jarque-Bera test is a goodness-of-fit test that evaluates the hypothesis that the data follow a normal distribution and it is based on sample estimates of the skewness and excess kurtosis, both of which are should be zero if the data are normal. Given the plots in figure 5, it is not surprising that the p-values are all below 0.05 for all series and the hypotheses can be rejected.

We now examine to what extent the different versions of the model can account for the non-Gaussian features of the U.S. data. To that end, we simulate an artificial sample of 2000 observations under each model and perform the same empirical analysis that we carried out above for each sample. Consider first the nonlinear model with normal innovations. Table 2 shows that this model predicts skewness close to zero for all variables, and kurtosis close to 3 for most variables but less than 3 for consumption and the nominal interest rate. The hypothesis that the data follows a normal distribution cannot be rejected for hours worked, price inflation and wage inflation. It is interesting to note that the hypothesis is rejected for consumption and the nominal interest rate, but this is so primarily because their kurtosis is well below 3 and, thus, their excess kurtosis is negative. In summary, these results indicate that despite its nonlinearity, this version of the model cannot account for the departures from normality in the data.

In contrast, the linear and nonlinear models with GEV innovations predict skewness and kurtosis that are quantitatively large and of the same sign as the actual data. In addition, the hypothesis that date follows a normal distribution can be rejected using the Jarque-Bera test (see table 3). Thus, both versions of the model—linear and nonlinear—can account for the non-Gaussian features of the U.S., although, as pointed out above, the empirical estimates of the parameters are not as plausible in the case of the linear model.

In light of the results reported above—the statistical and economic significance of the skewness and the implausible estimates of the linear model—we focus from now on the nonlinear model with GEV innovations.

4.2 Impulse-Responses

This section uses impulse-response analysis to study the effects of extreme shocks on our model economy. Since the model is nonlinear, the effects of a shock depend on its sign, size, and timing (see Gallant, Rossi and Tauchen, 1993, and Koop, Pesaran, and Potter, 1996). Regarding sign and size, we compute the responses to shock innovations in the 1st and 99th percentiles. We focus on these percentiles because we are concerned here with extreme realizations. Naturally, the size (in absolute value) of these innovations is not same for an asymmetric distribution like the GEV, but the point is that the likelihood of the two realizations is the same. Regarding timing, we assume that shocks occur when the system is at the stochastic steady state (where all variables are equal to the unconditional mean of their ergodic distribution). Responses are reported in figures 5 through 7 (thick lines), with the vertical axis denoting percentage deviation from the stochastic steady state and the horizontal axis denoting periods. As a comparison, we plot the responses for a normal distribution with the same standard deviation as the GEV (thin lines). Since the model is nonlinear, the effects of normal shocks need not be symmetric. Thus, these responses allow us to evaluate the contribution of nonlinearity (as opposed to shock asymmetry) to our results.

Figure 5 plots the responses to productivity shocks. A positive shock in the 99th percentile of the distribution induces an increase in consumption and hours worked, with the effect on consumption being very persistent as a result of intertemporal smoothing. Price inflation and the nominal interest rate decrease, in the latter case because the inflation coefficient in the Taylor rule is quantitatively much larger than that of employment. Finally, the nominal wage increases on impact but decreases thereafter. Since prices are more flexible than wages and they decrease by a larger proportion there is an unambiguous increase in the real wage. Qualitatively, the effects of the negative shock in the 1st percentile are the opposite to those just described. However, note that since the magnitude of the (negative) shock is quantitatively much larger than the magnitude of the equally-likely positive shock, its effects are much larger. This asymmetric is particular obvious when we compare these responses with those obtained using a normal distribution with the same standard deviation as the GEV distribution.

Figure 6 plots the responses to labor supply shocks. A negative shock in the 1st percentile of the distribution, increase the marginal disutility of labor and lead to a reduction in hours worked and, hence, consumption. Price and wage inflation increase but in this case the effect on wages is large enough that real wages actually increase. With price inflation raising, the nominal interest rate increases as well. Except for its effect of real wages, this shock qualitatively resembles a negative productivity shock. Notice, however, that the positive shock in the 99th percentile is quantitatively small as a result of the sharp asymmetric of the innovation distribution. Thus, the effects of this equally-likely shock are negligible.

Finally, consider figure 7 that plots the responses to a monetary shock. A positive shock that raises the interest rate induces a decrease in consumption, hours, price inflation, and wage inflation. Since prices drop by more than wage, the real wage increases. Again, an important feature of this figure is the asymmetry in the responses to monetary shocks.

4.3 How Does the Fed React to Extreme Events?

In this model, the Fed follows a linear rule in inflation and employment. However, since the latter variables are themselves nonlinear functions of the state variables, the Fed reacts nonlinearly to shocks. The extent of the nonlinearity can seen in the policy functions plotted in figure 8. In this figure, the think line is the nonlinear policy function implied by our third-order perturbation, while the thin line is the linear policy function implied by a first-order approximation. Consider first the reaction to productivity and labor supply shocks. Under both the first- and third-order perturbations, the Fed cuts the interest rate for positive shocks and increases it for negative shocks. The intuition simply is that negative shocks lead to an increase in inflation, and thus under a Taylor rule, the Fed must raise the interest rate. However, the response is larger under the nonlinear than under the linear solution, with the opposing being true in the case of positive shocks. In the case of monetary shocks, notice that the Fed's reaction is the same under solutions because the Taylor rule is log-linear in the monetary shock and, thus, both solution methods coincide.

5 Optimal Policy

Consider now case where the monetary authority follows a Ramsey policy of maximizing the households' welfare. That is, the monetary authority chooses $\{c_t,\, n_t,\, W_t,\, i_t,\, \Omega_t,\, \Pi_t\}_{t=\tau}^{\infty}$ to maximize

$$E_\tau \sum_{t=\tau}^{\infty} \delta^{t-\tau} U(c_t, n_t),$$

subject to the resource constraint and the first-order conditions of firms and households, and taking as given previous values for wages, goods prices, and shadow prices. The monetary authority can commit to the implementation of the optimal policy and may evaluate future utilities using a discount factor $\delta \in (0,1)$ that may differ from the one used by households.

5.1 Impulse-Responses under the Optimal Policy

Figures 9 and 10 respectively plot the responses to productivity and labor supply shocks. As before, we consider innovations in the 1st and 99th percentiles for the GEV distribution and a normal shock with the same standard deviation as the GEV, as a comparison.

Figure 9 plots the responses to productivity shocks. Qualitatively, responses are similar to those in figure 5 under the Taylor policy. That is, a positive leads to an increase in consumption, hours worked, wage inflation, and real wages, and a decrease in the interest rate and price inflation. Notice, however, that the response of hours worked under the Ramsey policy is much larger than under the Taylor policy. As before, a key feature of this responses is their asymmetry. Finally, figure 10 plots the responses to labor supply shocks. Although the qualitative responses are similar to those in figure 6 under the Taylor policy, there are some quantitative differences. Most notably, the effects on price and wage inflation are much smaller under the Ramsey policy. The asymmetry of these responses is especially large because a positive shock in the 99th percentile is a relatively small shock, while the equally-likely negative shock in the 1st percentile is very large.

5.2 Comparison with the Taylor Policy

In this section, we compare monetary policy under the Ramsey and Taylor policies. To that end, we compare the policy functions under both monetary regimes. Figure 11 plots the policy function under Ramsey for the third- and first-order perturbations. Comparing figures 11 and 8, notice that the Ramsey planner reacts much more aggressive to negative productivity shocks than the Taylor central banker, and react relatively less to labour supply shocks. A key difference, however, is that under the Ramsey policy the interest rate is cut for negative realizations of the labor supply shocks but it is raised under the Taylor rule.

Another important difference between both policies is the effect of uncertainty on the average rate of inflation. Under the Taylor policy, the mean of the ergodic distribution is 0.41 points (at an annual rate) above its deterministic steady state, while under the Ramsey policy, it is -0.003 below. Concerning the interest rate, the mean of the ergodic distributions are respectively, 0.38 and -0.03 points from its value in the deterministic steady state.

6 Conclusions

This paper studies the effect of extreme shocks for monetary policy under a realistic Taylor-type policy and under an ideal Ramsey policy. It is shown that extreme shocks and non-linearities in the propagation mechanism are important to account for the non-Gaussian features of the data.

Table 1

Parameter Estimates

	Model					
	Nonlinear GEV		Nonlinear Normal		Linear GEV	
Parameter	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.
Nominal Rigidities						
Wages	230.66	0.0003	282.28	0.0019	9932.8	0.0001
Prices	14.119	0.0206	45.644	0.0716	31.302	0.0433
Productivity Shock						
Autoregresive coeff.	0.9582	0.0207	0.8481	0.0235	0.9328	0.0251
Scale ($\times 10^{-2}$)	0.8994	0.2272	–	–	1.1735	0.2260
Shape	−1.2044	0.0554	–	–	−1.1889	0.2121
Standard deviation ($\times 10^{-2}$)	0.9990	0.2304	1.5016	0.1434	1.2915	0.2227
Skewness	−2.6551	0.1820	0	–	−2.6022	0.7508
Labor Supply Shock						
Autoregresive coeff.	0.9955	0.0111	0.9682	0.0151	0.9964	0.0024
Scale ($\times 10^{-4}$)	0.4183	0.6152	–	–	0.7147	0.6435
Shape	−3.7550	0.0621	–	–	−4.8812	0.1817
Standard deviation ($\times 10^{-2}$)	0.1321	0.1064	0.7581	0.4006	2.1074	1.2433
Skewness	−45.501	3.2520	0	–	−165.63	26.156
Monetary Policy						
Smoothing	0.8443	0.0597	0.8618	0.0626	0.6933	0.0432
Inflation	0.3836	0.0769	0.3847	0.0834	0.3838	0.0588
Output	0.1425	0.0371	0.1371	0.0475	0.0631	0.0497
Scale ($\times 10^{-2}$)	0.4199	0.0979	–	–	0.2977	0.0816
Shape ($\times 10^{-1}$)	−0.9169	1.8867	–	–	0.7753	0.8508
Standard deviation ($\times 10^{-2}$)	0.5322	0.1331	0.4434	0.1141	0.4277	0.1363
Skewness	1.0858	1.8810	0	–	1.6978	1.1941

Note: s.e. denotes standard errors computed using a k-step bootstrap with 19 replications. For the GEV distributions, the standard deviation and skewness are those implied by the scale and shape parameters. The superscript $*$ denotes statistical significance at the 5% level.

Table 2
Unconditional Moments

		Model		
	U.S. Data	Nonlinear GEV	Nonlinear Normal	Linear GEV
Skewness				
Consumption	−0.874	−0.566	0.014	−0.725
Hours	−0.580	−0.625	0.078	−0.538
Price inflation	0.656	1.150	0.094	0.987
Wage inflation	1.023	0.899	−0.066	0.991
Nominal interest rate	0.641	0.703	0.044	0.659
Kurtosis				
Consumption	3.581	3.419	2.720	3.796
Hours	3.190	3.574	3.092	3.720
Price inflation	6.051	5.701	3.114	4.715
Wage inflation	3.996	4.312	3.062	4.227
Nominal interest rate	3.767	3.972	2.663	3.695

Note: The table reports unconditional moments of actual U.S. series and of artificial data simulated from the four versions of the model considered. The sample size of the artificial data is 2000 observations.

Table 3

Jarque-Bera Test

Series	U.S. Data	Model: Nonlinear GEV	Model: Nonlinear Normal	Model: Linear GEV
Jarque-Bera				
Consumption	0.001	0.001	0.037	0.001
Hours	0.011	0.001	0.249	0.001
Price inflation	0.001	0.001	0.130	0.001
Wage inflation	0.001	0.001	0.402	0.001
Nominal interest rate	0.003	0.001	0.008	0.001

Note: The table reports p-values of the Jarque-Bera test of the hypothesis that the data follows a normal distribution. For the models, the test is based on 2000 simulated observations.

References

[1] Andrews, D. W. K., 2002. Higher-Order Improvements of a Computationally Attractive k-step Bootstrap for Extremum Estimators. Econometrica 70, pp. 119-162.

[2] Blanchard, O., Dell'Ariccia, G., Mauro, P., 2010. Rethinking Macroeconomic Policy. IMF Staff Position Note SPN/10/03.

[3] Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer: London.

[4] Davidson, R, J. G. MacKinnon, 1999. Bootstrap Testing in Nonlinear Models. International Economic Review 40, pp. 487-508.

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

[6] Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events. Springer: London.

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

[8] Jarque, C. M., Bera, A. K., 1987. A Test for Normality of Observations and Regression Residuals. International Statistical Review 55, pp. 163–172.

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

[10] Hansen, G. D., 1985. Indivisible Labor and the Business Cycle. Journal of Monetary Economics 16, 309-327.

[11] 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.

[12] Kim, J., Ruge-Murcia, F., 2009. How Much Inflation is Necessary to Grease the Wheels?. Journal of Monetary Economics 56, pp. 365-377.

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

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

[15] Rotemberg, J. J., 1982. Sticky Prices in the United States. Journal of Political Economy 90, 1187-1211.

[16] 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.

[17] 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.

Figure 1: Estimated Cumulative Distribution Function of Productivity Shock

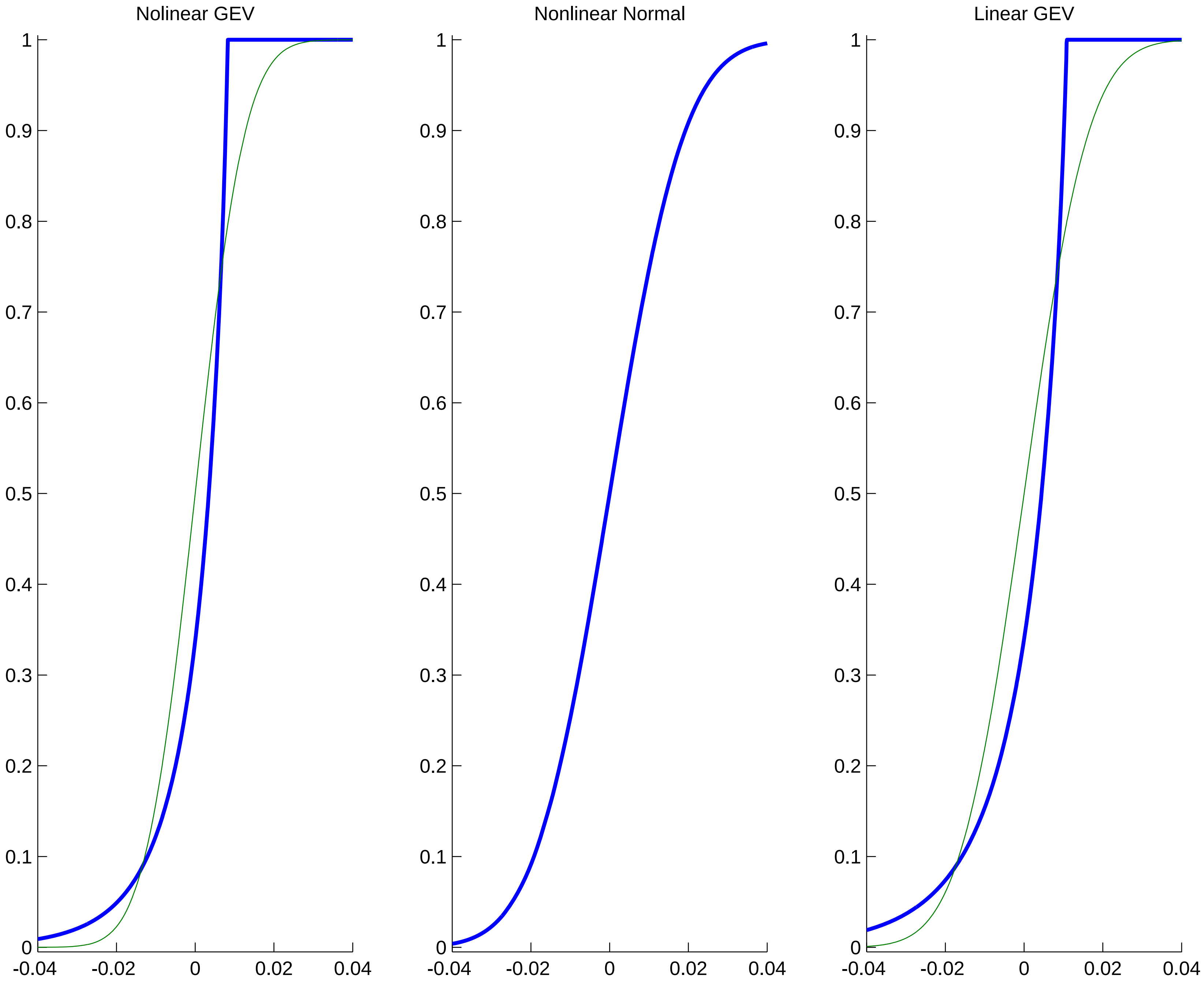

Figure 2: Estimated Cumulative Distribution Function of Labor Supply Shock

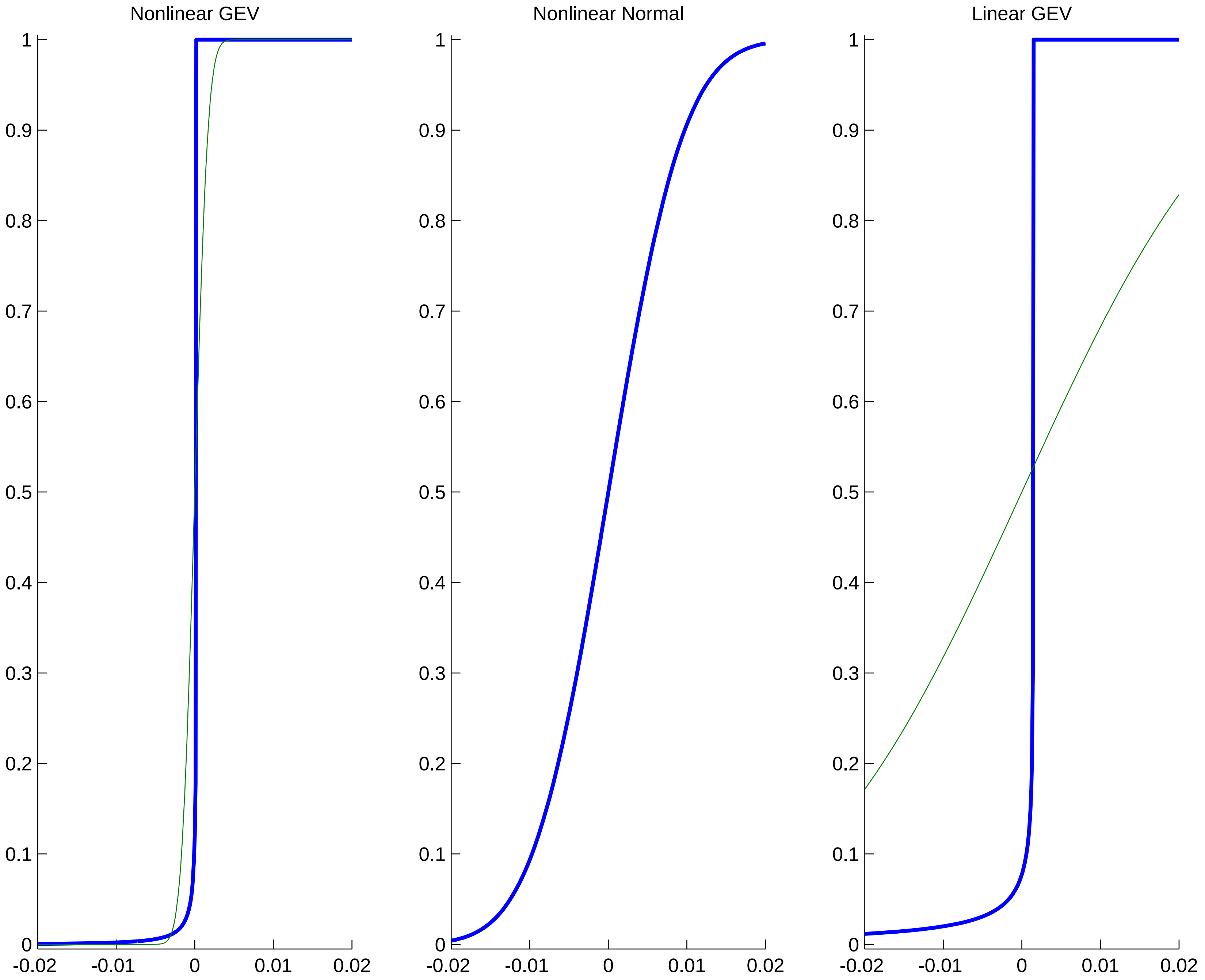

Figure 3: Estimated Cumulative Distribution Function of Monetary Policy Shock

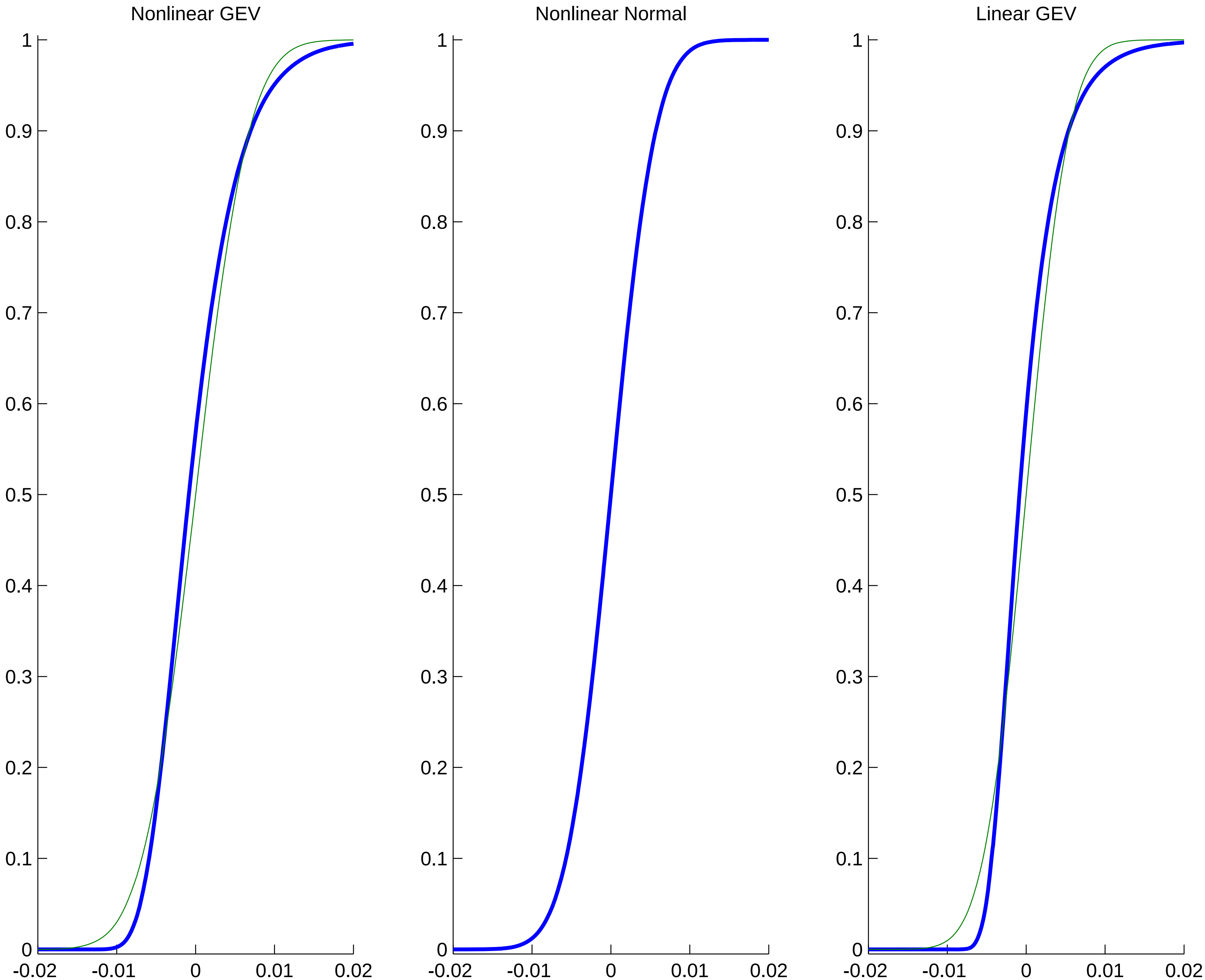

Figure 4: Asymmetry of U.S. Macroeconomic Data

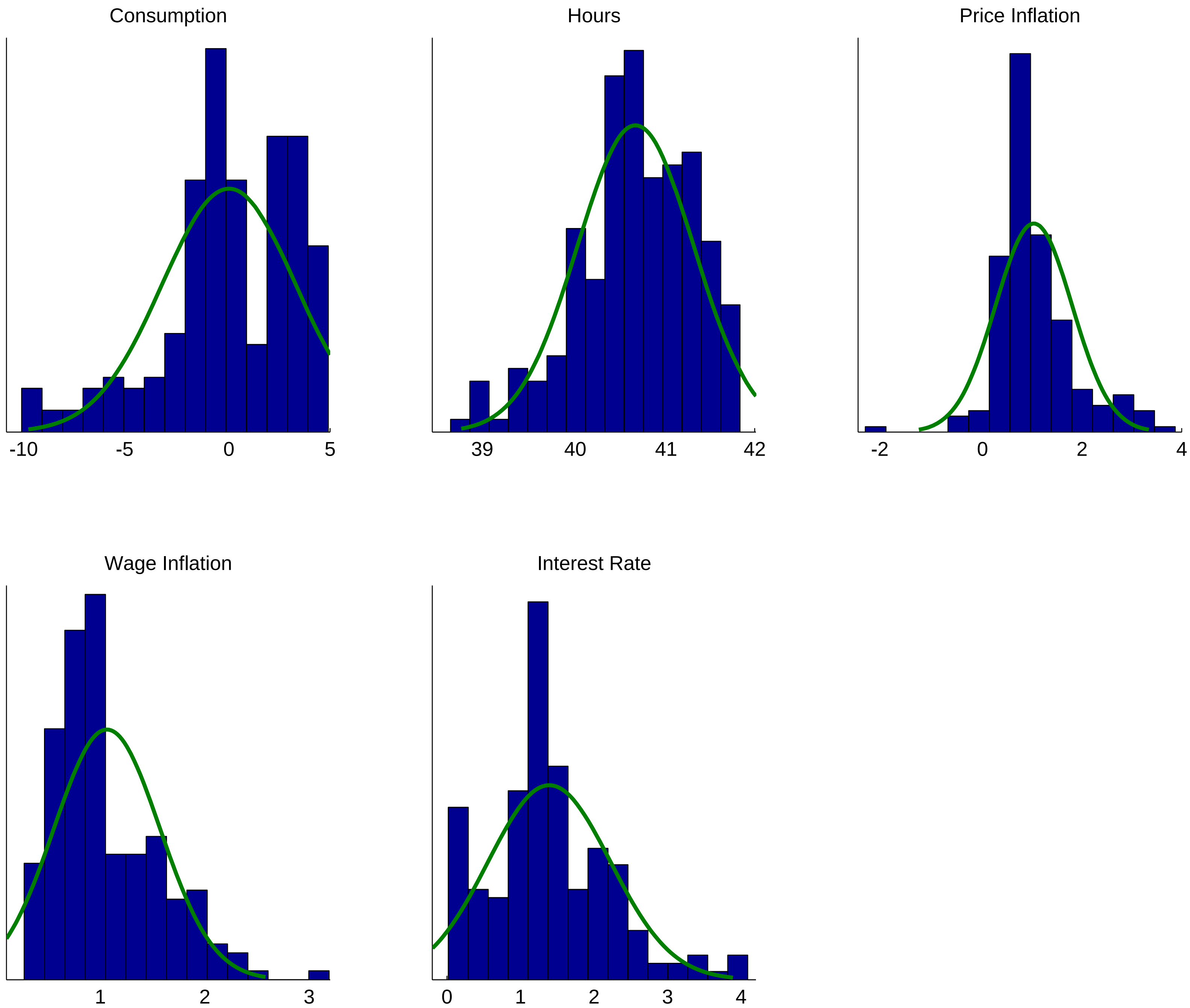

Figure 5: Responses to Extreme Productivity Shocks under Taylor Rule Policy

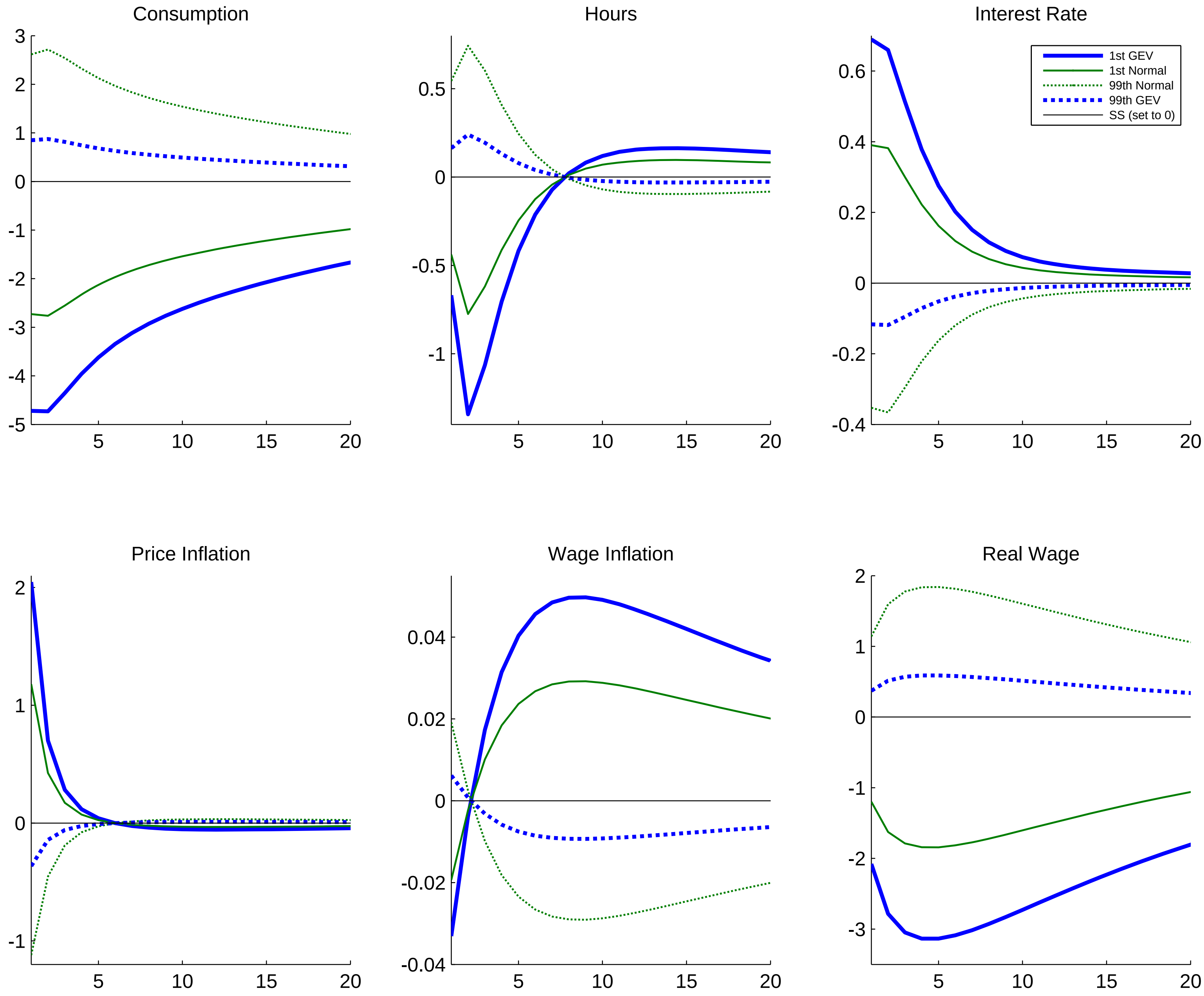

Figure 6: Responses to Extreme Labor Supply Shocks under Taylor Rule Policy

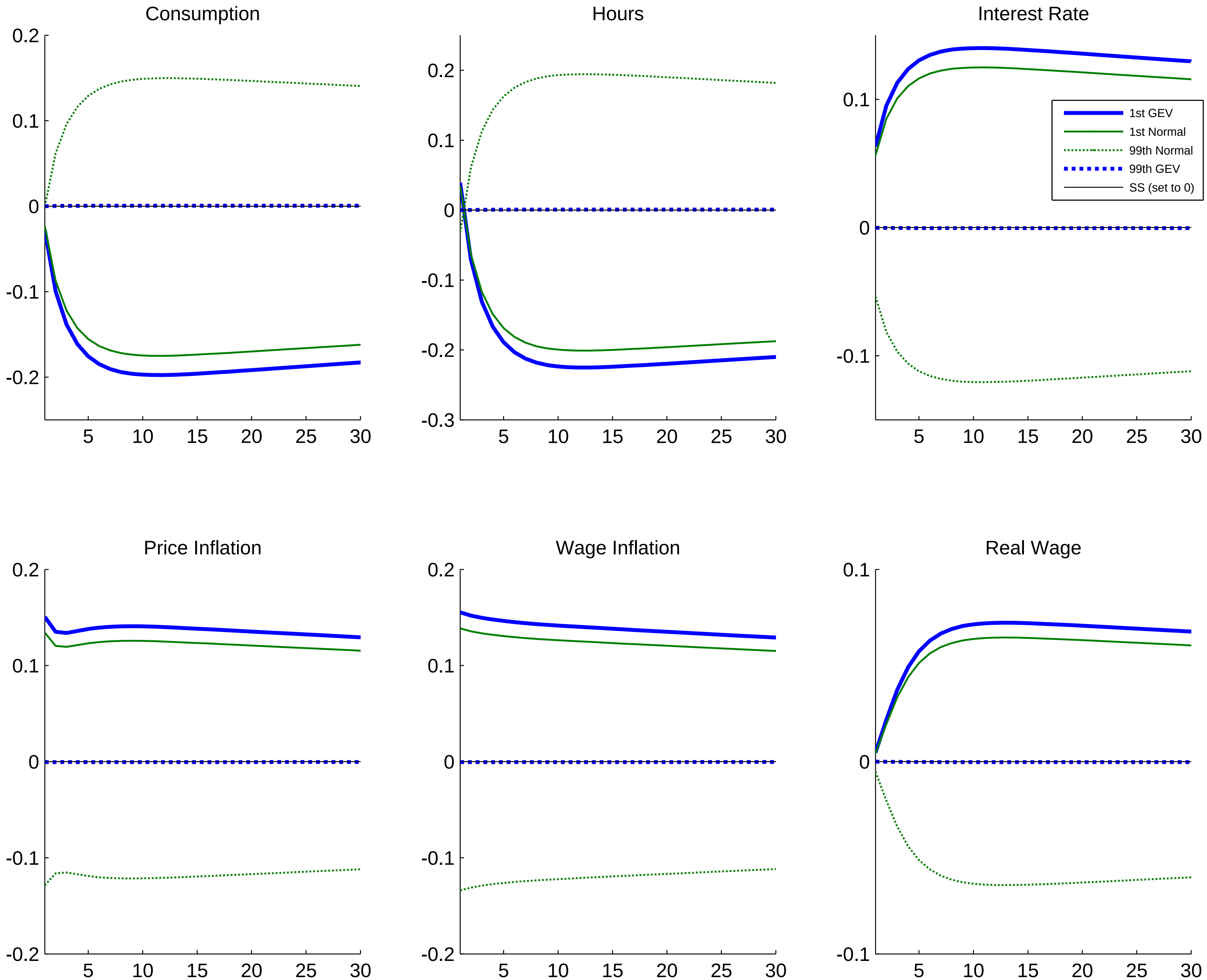

Figure 7: Responses to Extreme Monetary Shocks under Taylor Rule Policy

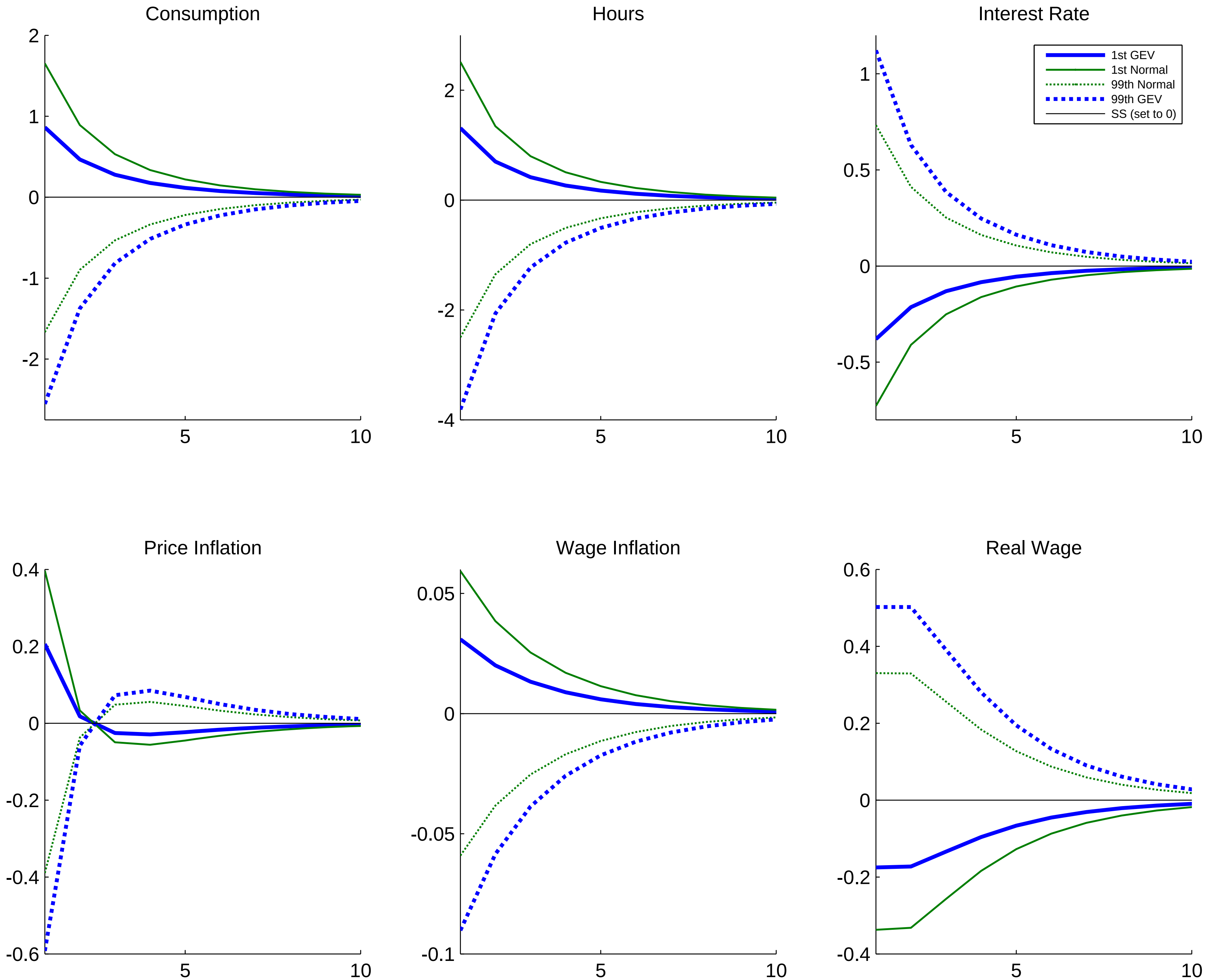

Figure 8: Interest Rate Policy Function

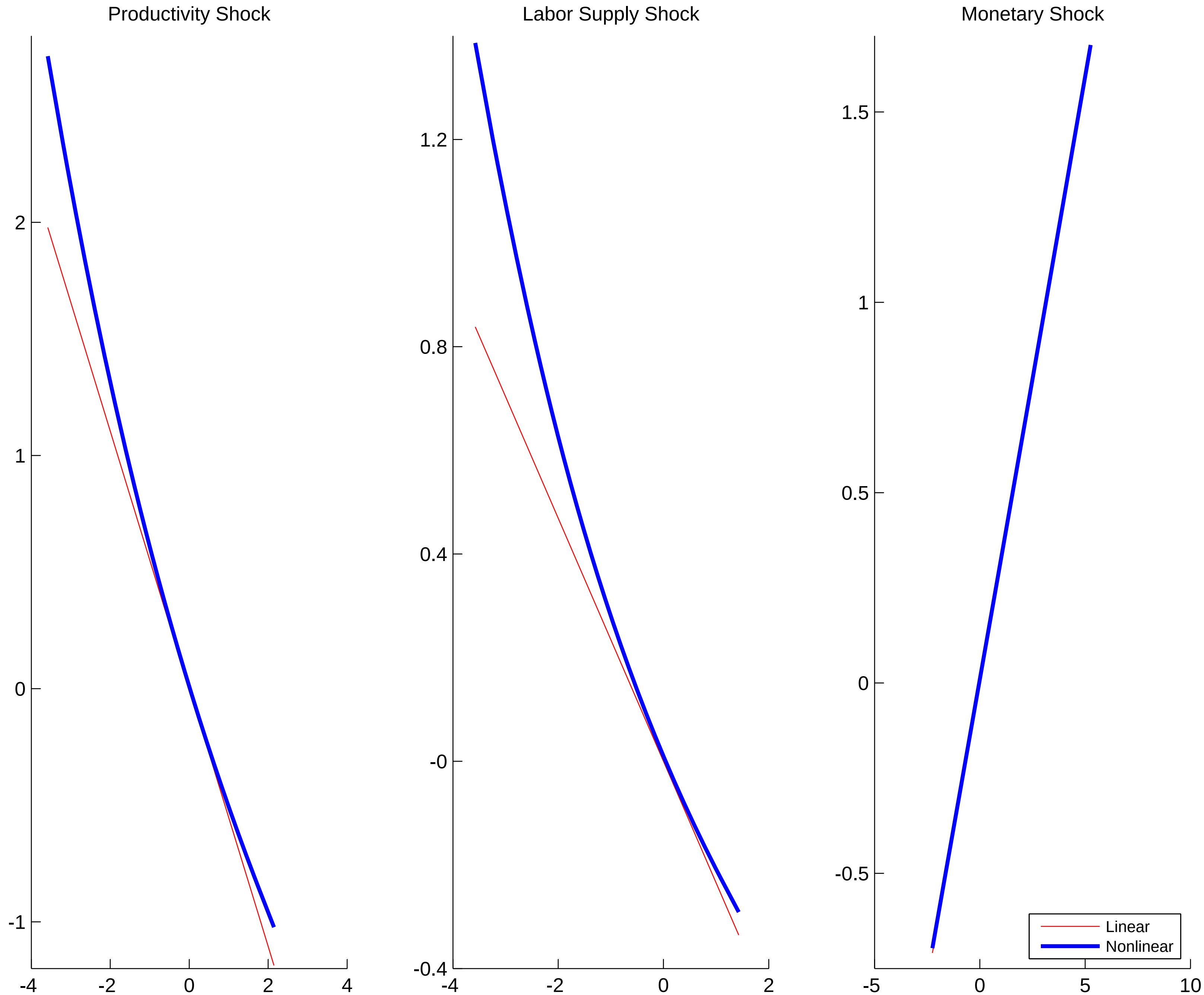

Figure 9: Optimal Responses to Extreme Productivity Shocks

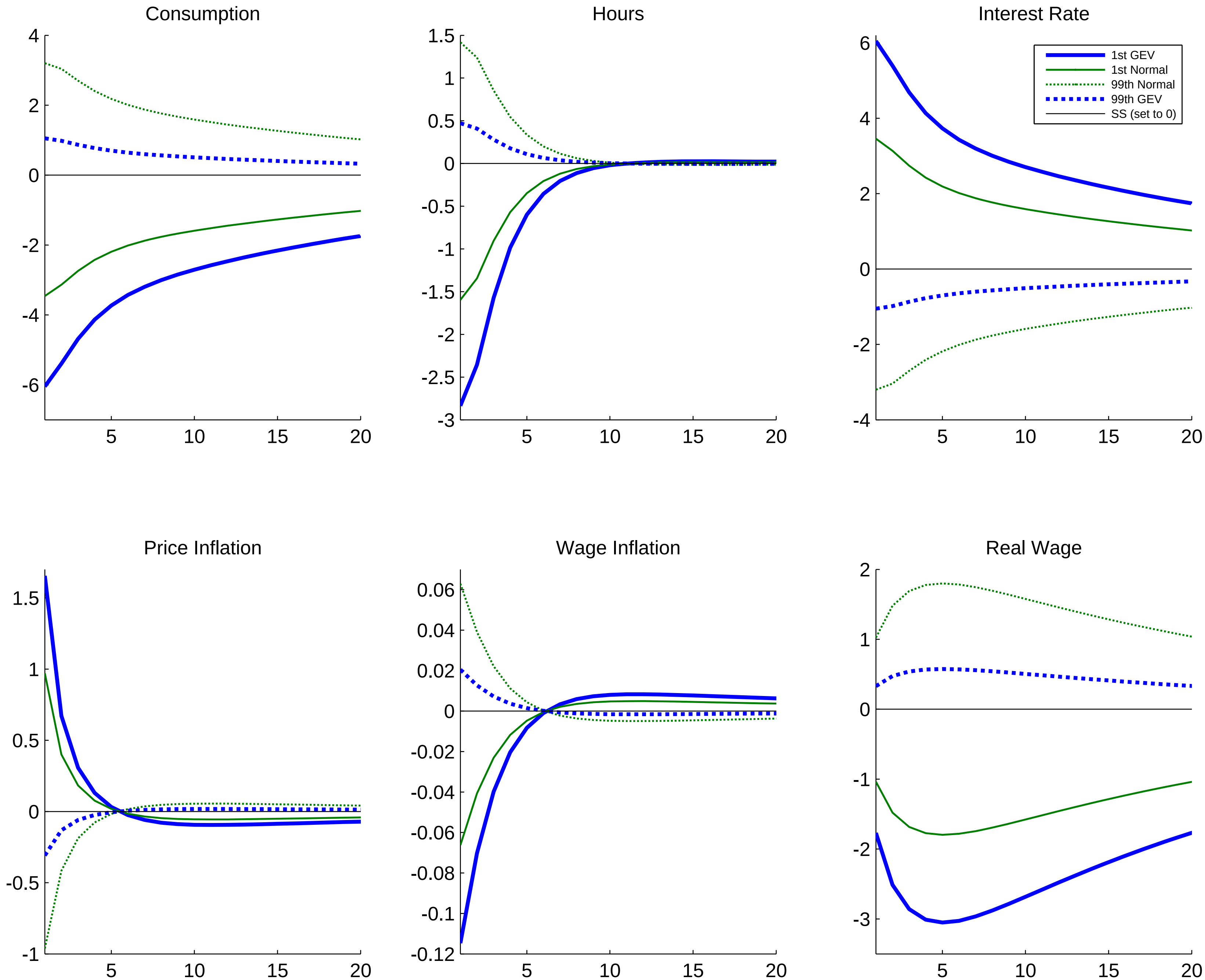

Figure 10: Optimal Responses to Extreme Labor Supply Shocks

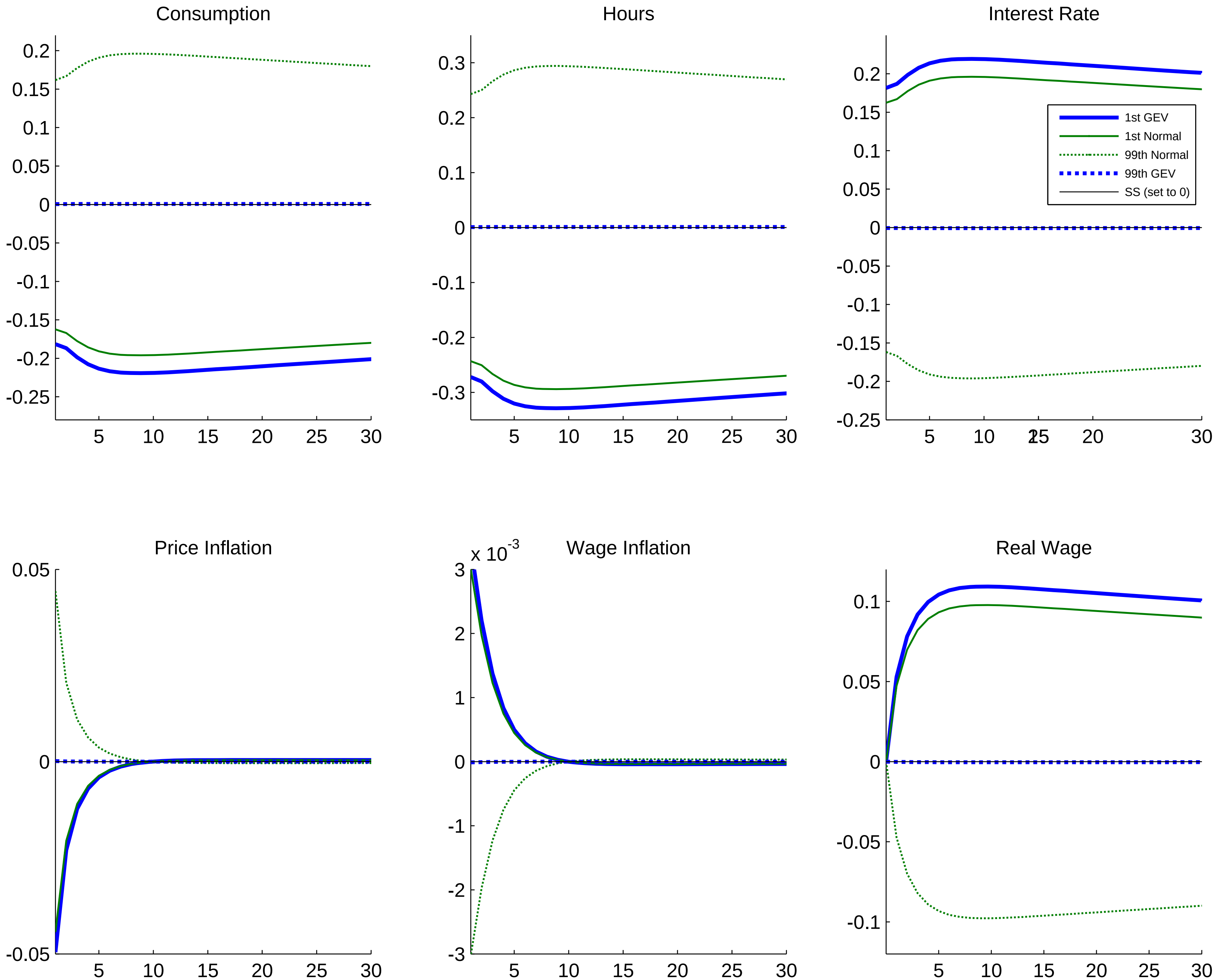

Figure 11: Optimal Interest Rate Policy Function