# Extreme Events and Optimal Monetary Policy*

Jinill Kim[†] Francisco Ruge-Murcia[‡]

This revision: January 2018

**Abstract**

This paper studies the implication of extreme shocks for monetary policy. The analysis is based on a small-scale New Keynesian model with sticky prices and wages where shocks are drawn from asymmetric Generalized Extreme Value distributions. A nonlinear perturbation solution of the model is estimated by the simulated method of moments. Under the Ramsey policy, the central bank responds nonlinearly and asymmetrically to shocks. The trade-off between targeting a gross inflation rate above 1 (or a net inflation rate above 0) as insurance against extreme shocks and targeting an average gross inflation at unity to avoid adjustment costs is unambiguously decided in favour of strict price stability.

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

*This paper was previously circulated under the title "Extreme Events and the Fed." The authors benefitted from comments by Luigi Boccola, Marcelle Chauvet, the editor (J. Fernández-Villaverde), and anonymous referees. Ruge-Murcia acknowledges the support from the Social Sciences and Humanities Research Council (SSHRC), and from the Bank of Canada through its Fellowship Program.

[†]Department of Economics, Korea University, Korea. E-mail: jinillkim@korea.ac.kr
[‡]Department of Economics, McGill University, Canada. E-mail: francisco.ruge-murcia@mcgill.ca

# 1. Introduction

Economies are occasionally subjected to extreme shocks that can have profound and long lasting effects—think, for example, of the oil shocks in the 1970s or the financial shocks associated with the Great Recession. Thus, it is important to design policy by taking into account the fact that extreme events can happen sometimes. This paper studies the positive and normative implications of extreme shocks for monetary policy using a small-scale New Keynesian model with sticky prices and sticky—and more downwardly-rigid—wages (see Kim and Ruge-Murcia, 2009). Crucially, the model relaxes the usual assumption that shocks are normally distributed and assumes instead that they are drawn from asymmetric distributions with an arbitrarily long tail. Methodologically, we use tools from extreme value theory, which is a branch of statistics concerned with extreme deviations from the median of probability distributions. This theory was developed primarily in meteorology and engineering, where designers are interested in protecting structures against infrequent—but potentially damaging—events like earthquakes and hurricanes.[1]

Previous research on the positive analysis of monetary policy typically works under the dual assumptions that the propagation mechanism is linear and that 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, and a second-order approximation is consistent with any two-parameter distribution. Since the normal distribution satisfies this two-degrees-of-freedom specification, the normal distribution is also widely used in normative analysis. This strategy leads to tractable models, but, as we argue below, it is unsatisfactory for understanding policy responses to extreme events.

Instead, the shock innovations in our model are assumed to be drawn from generalized extreme value (GEV) distributions. This distribution is widely used in extreme value theory to model the maxima (or minima) of a sequence of random variables.[2] The distribution has three independent parameters that determine its first, second, and third moments. To be consistent with considering three moments of the distribution, we approximate the model dynamics using a third-order perturbation, and so our approximate solution is nonlinear. The nonlinear model is estimated by the simulated method of moments (SMM). In order to disentangle the relative contribution of asym-

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

[2] In the context of financial markets, this distribution could be motivated, for example, by Stein (2014), who argues that the most optimistic investors drive asset prices, and by Adrian and Duarte (2017), who model financial intermediaries subject to occasionally binding value-at-risk constraints. More generally, since many economic shocks—strikes, weather, political uncertainty, changes in commodity prices, etc.—feature long tails, the most constructive interpretation is to think of the GEV distribution as a way to capture a wide array of potentially large disturbances that are summarized here using a parsimonious number of structural shocks.

metric shocks and nonlinearity to our results, we also estimate a nonlinear version of the model with normal innovations. Results show that the data prefer a specification where monetary policy innovations are drawn from a positively skewed distribution, and productivity and preference innovations are drawn from a negatively skewed distribution. This conclusion is based on structural estimates from the model and also supported by reduced-form estimates from the raw data.

Using the estimated parameters, we examine the normative implications of the model under the Ramsey policy. We find that the benevolent monetary authority responds asymmetrically to shocks and the change in the nominal interest rate is generally larger than that under the Taylor policy. In addition to investigating the optimal monetary policy response to large shocks, this paper derives specific policy prescriptions concerning optimal inflation targets. This issue is important because in light of the recent Global Financial Crisis, Williams (2009, 2014), Blanchard et al. (2010), and Ball (2014) propose increasing inflation targets in order to provide a larger buffer zone from the zero lower bound on interest rates. In one of the few contributions to the literature on optimal policy in an environment with extreme shocks, Svensson (1993) notes the tension between: (i) acting prudently and incorporating systematically the possibility of extreme shocks into policy (e.g., by raising the inflation target) and (ii) taking a wait-and-see approach. Under the wait-and-see approach, the monetary authority acts only if and when an extreme shock occurs and adjusts the policy variables appropriately to counteract its effects. Our model incorporates such a trade-off and uses quantitative analysis to compare these two strategies using a well-defined welfare metric. We show that the solution to the trade-off is solved unambiguously in favour of the wait-and-see approach. The reason is simply that while prudence calls for an optimal gross inflation target above 1 as an insurance against extreme shocks that would require costly nominal wage cuts, such target involves price-and-wage adjustment costs that must paid in every period. Thus, under the Ramsey policy the optimal gross inflation rate is virtually indifferent from 1 (i.e., strict price stability).

The New Keynesian model used to study extreme shocks is based on our previous work (Kim and Ruge-Murcia, 2009). We use this model for two reasons. First, this model is highly nonlinear because of the asymmetry in wage adjustment costs. Nonlinearity is a key ingredient in evaluating the economic implications of extreme shocks because it can give rise to prudent behavior. Second, this model has a well-defined cost of deflation in the form of very costly nominal wage cuts.[3] However, this project makes a distinct contribution from Kim and Ruge-Murcia (2009). Our previous contribution attempts to evaluate Tobin's argument that inflation "greases the wheels" of the labor market. That is, that inflation eases the adjustment of the labor market after an adverse

[3]As an alternative, one could consider the costs associated with the zero lower bound on nominal interest rates. We do not pursue this strategy here, but refer the reader to the extensive literature on this topic (see, e.g., Coibion et al., 2012, and the references therein).

shock by speeding the decline of real wages. Instead, this paper examines the recent argument that in anticipation of extreme shocks, the monetary authority should increase inflation targets (see the literature cited above). By relaxing the usual assumption that shocks are normally distributed and using a third-order perturbation method to solve the model, we can quantify the skewness risk faced by the authority in this environment and its implications for monetary policy.

Previous literature concerned with the implications of asymmetric shocks includes work by Barro (2006), Andreasen (2012), and Gourio (2012) on rare disasters, and by Ferreira (2016), Zeke (2016) and Ruge-Murcia (2017) on skewness risk. Disasters are low probability events where output (or consumption) drops by at least 15 percent from peak to trough as, for example, during the Great Depression. This paper complements literature on rare disasters by showing that even in the relatively calm, postwar U.S., agents face the possibility of large decreases in consumption primarily associated with the business cycle. This paper contributes to the literature on skewness risk by evaluating its positive and normative implications for monetary policy and inflation.

The paper is organized as follows. Section 2 presents a small-scale New Keynesian model of an economy occasionally subject to extreme shocks. Section 3 discusses the estimation method, describes the data, and reports estimates of three versions of the model: a benchmark version with GEV innovations and asymmetric costs, and two alternative versions with normal innovations and symmetric costs. This section also examines the positive implications of the model for the moments of key macroeconomic variables and studies the responses of the economy to large shocks using impulse-response analysis. Sections 4 studies optimal monetary policy under the Ramsey policy. Finally, section 5 concludes and discusses some limitations of our analysis.

## 2. An Economy Subject to Extreme Shocks

The agents in this economy consist of firms that produce differentiated goods, households with idiosyncratic job skills, and a monetary authority. This section describes their behavior and the resulting equilibrium.

### 2.1 Firms

Firm $i \in [0,1]$ hires heterogeneous labour supplied by households and combines it as

$$n_{i,t} = \left( \int_0^1 (n_{i,t}^h)^{1/\omega} dh \right)^{\omega}, \tag{1}$$

where $h \in [0,1]$ is an index for households and $\omega > 1$ is a parameter that determines the elasticity of substitution between labour types. This labour aggregate is employed to produce output using the technology

$$y_{i,t} = z_t n_{i,t}^{1-\alpha}, \tag{2}$$

where $y_{i,t}$ is output, $\alpha \in (0,1)$ is a parameter and $z_t$ is a productivity shock. The price of the labour input is

$$W_{i,t} = \left( \int_0^1 (W_t^h)^{1/(1-\omega)} dh \right)^{1-\omega}, \tag{3}$$

where $W_t^h$ is the nominal wage of household $h$.

The productivity shock follows the process

$$\ln(z_t) = \varphi \ln(z_{t-1}) + \epsilon_t, \tag{4}$$

where $\varphi \in (-1,1)$ and $\epsilon_t$ is an innovation assumed to be independent and identically distributed (i.i.d.) with mean zero and skewness different from zero. By allowing for non-zero skewness, this specification relaxes the standard assumption that shocks are symmetrically distributed around the mean, and, hence, a positive realization is as likely as a negative realization of the same magnitude. Frequently, the assumption of symmetry is not explicit but rather the result of assuming that shocks are drawn from normal distributions. Instead, in our economy, innovations are drawn from an asymmetric distribution. Since agents face the possibility of extreme realizations from the long tail of the distribution, they are subject to skewness risk. In the empirical part of the paper, we assume that innovations are drawn from a generalized extreme value (GEV) distribution.

Goods market frictions induce a convex cost whenever nominal prices are adjusted. This cost is represented using the linex function (Varian, 1974)

$$\Gamma_t^i = \Gamma(P_{i,t}/P_{i,t-1}) = \gamma \left( \frac{\exp\left(-\eta\left(P_{i,t}/P_{i,t-1} - 1\right)\right) + \eta\left(P_{i,t}/P_{i,t-1} - 1\right) - 1}{\eta^2} \right), \tag{5}$$

where $\gamma \in (0,\infty)$ and $\eta \in (-\infty,\infty)$ are parameters. This model of price rigidity generalizes the one in Rotemberg (1982) by allowing adjustment costs to be asymmetric. Asymmetric price adjustment costs are consistent with the empirical evidence on price changes reported by Peltzman (2000) for individual goods in a Chicago supermarket chain and for components of the producer price index. Zbaracki et al. (2004) find that price adjustment costs in a manufacturing firm—interpreted broadly to include physical and managerial costs—are convex and increasing in the size of the adjustment. They also find that the managerial time and effort involved in price increases is different than for decreases.

Under the function (5), the adjustment cost depends on both the sign and magnitude of the price change, with $\eta > 0$ corresponding to the case where a nominal price increase involves a smaller frictional cost than a price decrease of the same magnitude. The converse is true in the case where $\eta < 0$. In the special case where $\eta$ approaches zero, (5) nests the quadratic function in Rotemberg (1982). Hence, it is straightforward to compare statistically the model with asymmetric costs and the restricted version with quadratic costs.

The firm maximizes

$$E_s \sum_{t=s}^{\infty} \beta^{t-s} (\Lambda_t/\Lambda_s) \left( \left(1 - \Gamma_t^i\right) (P_{i,t}/P_t) c_{i,t} - \int_0^1 (W_t^h/P_t) n_t^h dh \right), \tag{6}$$

where $E_s$ is the expectation conditional on information available at time $s$, $\beta \in (0, 1)$ is the discount factor, $\Lambda_t$ is the marginal utility of consumption, $c_{i,t}$ is total consumption demand for good $i$, $\left(1 - \Gamma_t^i\right) (P_{i,t}/P_t) c_{i,t}$ is real revenue net of adjustment costs, $n_t^h$ is hours worked by household $h$, $\int_0^1 (W_t^h/P_t) n_t^h dh$ is the real wage bill, and $P_t$ is the aggregate price index, which is defined as

$$P_t = \left( \int_0^1 (P_{i,t})^{1/(1-\upsilon)} di \right)^{1/(1-\upsilon)}. \tag{7}$$

The maximization is subject to a downward-sloping consumption demand function (see (14), below), the technology (2), and the condition that supply must meet demand for good $i$ at the posted price. The optimal demand for labour $h$ is

$$n_t^h = \left( \frac{W_t^h}{W_t} \right)^{-\omega/(\omega-1)} n_{i,t}, \tag{8}$$

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

## 2.2 Households

Household $h$ maximizes

$$E_s \sum_{t=s}^{\infty} \beta^{t-s} \left( \log(c_t^h) - \frac{(n_t^h)^{1+\chi}}{1+\chi} \right) u_t, \tag{9}$$

where $c_t^h$ is consumption, $\chi$ is a positive parameters, and $u_t$ is a preference shock.[4] The weight of the disutility of labour is set to 1 in (9), but this normalization is inconsequential because this weight only scales the number of hours worked in steady state and does not affect the dynamics of the model. Consumption is an aggregate of the differentiated goods produced by firms,

$$c_t^h = \left( \int_0^1 (c_{i,t}^h)^{1/\upsilon} di \right)^{\upsilon}, \tag{10}$$

[4]The assumption of logarithmic consumption preferences is based on preliminary results that show that the curvature parameter in a more general CRRA function is not statistically different from 1. Moreover, for empirically plausible values, this parameter has a limited effect on the quantitative implications of the model.

where $\upsilon > 1$ is a parameter that determines the elasticity of substitution between goods. The preference shock follows the process

$$\ln(u_t) = \varrho \ln(u_{t-1}) + \zeta_t, \tag{11}$$

where $\varrho \in (-1, 1)$ and $\zeta_t$ is an innovation assumed to be i.i.d. with mean zero, skewness different from zero, and independent of the productivity innovation, $\epsilon_t$.

Labour market frictions induce a convex cost whenever nominal wages are adjusted. This cost is represented using the function

$$\Phi_t^n = \Phi(W_t^n/W_{t-1}^n) = \phi\left(\frac{\exp\left(-\psi\left(W_t^n/W_{t-1}^n - 1\right)\right) + \psi\left(W_t^n/W_{t-1}^n - 1\right) - 1}{\psi^2}\right), \tag{12}$$

where $\phi \in (0, \infty)$ and $\psi \in (-\infty, \infty)$. In the case where $\psi > 0$, a nominal wage decrease involves a larger frictional cost than a wage increase of the same magnitude, and wages are, therefore, more downwardly than upwardly rigid. When $\psi \to 0$ the cost function is the quadratic function and wage increases and decreases of the same magnitude are equally costly. When $\psi \to \infty$ the cost function takes the shape of an "L" and wages are completely flexible upwards and inflexible downwards.

Downward wage rigidity is discussed by Keynes (1936, ch. 21) and is consistent with the observation that the cross-sectional distribution of individual wages is positively skewed with a peak at zero and very few nominal wage cuts. For example, see Akerlof et al. (1996), and Card and Hyslop (1997) for the United States; Fehr and Goette (2005) for Switzerland; Kuroda and Yamamoto (2003) for Japan; and Castellanos et al. (2004) for Mexico. Recent literature examines the implications of downward nominal wage rigidity for monetary policy (Kim and Ruge-Murcia 2009, 2011), business cycle asymmetries (Abbritti and Fahr 2013), and currency pegs (Schmitt-Grohé and Uribe 2016). Kim and Ruge-Murcia (2009) provide statistical evidence in favour of downward nominal wage rigidity in the form of a positive and statistically significant coefficient of the asymmetry parameter $\psi$ in (12).

The household is subject to the budget constraint

$$c_t^h + \frac{B_t^h - I_{t-1}B_{t-1}^h}{P_t} = \left(1 - \Phi_t^h\right)\left(\frac{W_t^h n_t^h}{P_t}\right) + \frac{D_t^h}{P_t}, \tag{13}$$

where $B_t^h$ is a one-period nominal bond, $I_t$ is the gross nominal interest rate, and $D_t^h$ are dividends. In addition to this budget constraint and a no-Ponzi-game condition, utility maximization is subject to the demand for labour $h$ by firms (see (8)). The optimal consumption of good $i$ satisfies

$$c_{i,t}^h = \left(\frac{P_{i,t}}{P_t}\right)^{-\upsilon/(\upsilon-1)} c_t^h, \tag{14}$$

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

### 2.3 Monetary Policy

The monetary authority (or "the Fed") sets the interest rate following the Taylor-type rule

$$\ln(I_t/I) = \rho_1 \ln(I_{t-1}/I) + \rho_2 \ln(\Pi_t/\Pi) + \rho_3 \ln(n_t/n) + \xi_t, \tag{15}$$

where $\rho_1 \in (-1, 1)$, $\rho_2$ and $\rho_3$ are parameters; variables without time subscript denote steady-state values; and $\xi_t$ is a monetary shock that represents factors that affect the nominal interest rate beyond the control of the Fed. We assume that $\xi_t$ is i.i.d. with mean zero, skewness different from zero, and independent of the innovations to productivity ($\epsilon_t$) and the preference shock ($\zeta_t$).

### 2.4 The GEV Distribution

Under the Fisher-Tippett theorem (Fisher and Tippett 1928), the maxima of a sample of i.i.d. random variables converge in distribution to one of three possible distributions: the Gumbel, the Fréchet, and the Weibull distributions. Jenkinson (1955) shows that these distributions can be represented in a unified way using a generalized extreme value (GEV) distribution. The probability density function (PDF) of the GEV distribution is

$$f(x) = (1/\kappa)\tau(x)^{\varsigma+1} \exp(-\tau(x)), \tag{16}$$

with $\tau(x) = ((1 + (x - \mu)\varsigma/\kappa))^{-1/\varsigma}$ when $\varsigma \neq 0$ and $\tau(x) = \exp(-(x - \mu)/\kappa)$ when $\varsigma = 0$. In this function, $\mu$ is the location parameter, $\kappa$ is the scale parameter, and $\varsigma$ is the 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 skewness is positive. In the case where the shape parameter is negative, the 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 paper because, as we will see below, the U.S. data prefer specifications where the skewness of the innovations is relatively large.[5] There are values of the shape parameter for which some moments of the distribution do not exist—for example, the mean is not defined when this parameter is larger than or equal to 1—but this turns out to be not empirically relevant here. For additional details on the GEV distribution and examples of practical applications, see Coles (2001) and Embrechts et al. (2011)

[5] In preliminary work, we considered using the Skew-normal distribution, whose skewness is bounded between −1 and 1. However, parameter estimates 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.

### 2.5 Equilibrium

In the symmetric equilibrium, all firms are identical and all households are identical. This means that all firms charge the same price, demand the same quantity of labour, and produce the same quantity of output; all households supply the same amount of labour and receive the same wages; and net bond holdings are zero.

Equilibrium in the goods market implies the aggregate resource constraint

$$c_t = y_t - (y_t \Gamma_t + w_t n_t \Phi_t) \,, \tag{17}$$

where $y_t$ is aggregate output and $w_t = W_t/P_t$ is the real wage. In the special 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 (17)). When prices and wages are constant, there are no deadweight losses ($\Gamma_t = \Phi_t = 0$). This result indeed holds in the cases where: (i) there is no uncertainty or (ii) certainty-equivalent applies. However, in the more relevant case where the social welfare function is concave in inflation and there is uncertainty, optimal gross inflation may be different from 1 as a result of precautionary behaviour by the planner.

### 2.6 Model Solution

The model is solved using a perturbation method that approximates the policy functions using a third-order polynomial in the state variables and moments of the innovations. Jin and Judd (2002) explain in detail this method and establish the conditions under which the approximate solution exists. The solution is nonlinear by construction because it contains linear, quadratic, and cubic terms in the state variables. The solution also features a risk adjustment factor that depends on both the variance and the skewness of the innovations.[6]

## 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 real wage, and the nominal interest rate from 1964Q2

[6] It would be ideal to check the accuracy of our third-order perturbation based on Euler equation errors. However, given the size of our model, we rely instead on Caldara et al. (2012)—who report that second- and third-order perturbations perform well relative to projection solutions in terms of Euler equation accuracy—and on Aruoba et al. (2017)—who report that perturbation solutions match the nonlinear inflation and wage dynamics in terms of posterior predictive check.

to 2015Q4. The sample starts in 1964 because aggregate data on wages and hours worked are not available prior to that year. The sample ends with the latest available observation at the time the data was collected. The raw data were taken from the website 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 the estimate of civilian non-institutional population produced by the Bureau of Labor Statistics (BLS). Hours worked are measured by average weekly hours of production and non-supervisory employees in manufacturing. The real wage is hourly compensation in the non-farm business sector divided by the CPI. Real per-capita consumption, hours worked, and the real wage are quadratically detrended in order to make these series consistent with model, where there is no long-run growth. The rate of price inflation is the percentage change in the CPI expressed as a gross quarterly rate. The nominal interest rate is the effective federal funds rate.[7] The original interest rate series, which is quoted as a net annual rate, is transformed into a gross quarterly rate. Except for the nominal interest rate, all data are seasonally adjusted at the source.

## 3.2 Estimation Method

The model is estimated by the simulated method of moments (SMM). Defining $\theta \in \Theta$ to be a $q \times 1$ vector of structural parameters, the SMM estimator, $\widehat{\theta}$, is the value that solves

$$\min_{\{\theta\}} \left( (1/T) \sum_{t=1}^{T} m_t - (1/\lambda T) \sum_{\iota=1}^{\lambda T} m_\iota(\theta) \right)' \mathbf{W} \left( (1/T) \sum_{t=1}^{T} m_t - (1/\lambda T) \sum_{\iota=1}^{\lambda T} m_\iota(\theta) \right), \qquad (18)$$

where $\mathbf{W}$ is a $p \times p$ weighting matrix, $T$ is the sample size, $\lambda$ is a positive integer, $m_t$ is a $p \times 1$ vector of empirical observations on variables whose moments are of interest to us, and $m_\iota(\theta)$ is a synthetic counterpart of $m_t$ with elements obtained from the stochastic simulation of the model. Note that the SMM estimator minimizes the weighted distance between the unconditional moments predicted by the model and those computed from the data, where the moments predicted by the model are computed on the basis of artificial data simulated from the model. Duffie and Singleton (1993) show that under general regularity conditions SMM delivers consistent and asymptotically normal parameter estimates.

In this application, the weighting matrix is the diagonal of the inverse of the matrix with the long-run variance of the moments. This weighting matrix makes the objective function scale

[7] The target for nominal federal funds rate was virtually at its lower bound, and it did not change between late 2008 and late 2015. We abstract from this issue in our baseline estimation; but in preliminary work, we estimated the model using data until 2008 only, and estimates were similar to those reported here.

free and gives a larger weight to the moments that are more precisely estimated.[8] The long-run variance of the moments is computed using the Newey-West estimator with a Bartlett kernel and bandwidth given by the integer of $4(T/100)^{2/9}$, where $T = 208$ is the sample size. The number of simulated observations is 100 times larger than the sample size (i.e., $\lambda = 100$).[9] In order to attenuate the effect of starting values on the results, the simulated sample contains 100 additional "training" observations that are discarded for the purpose of computing the moments. The dynamic simulations of the nonlinear model are based on the pruned version of the solution using scheme proposed by Andreasen et al. (2017).

The estimated parameters are the curvature of labour in the utility function ($\chi$), the parameters of the adjustment cost functions for prices ($\gamma$ and $\eta$) and wages ($\phi$ and $\psi$), the monetary policy rule ($\rho_1$, $\rho_2$, and $\rho_3$), and the parameters of the distributions of productivity, preference, and monetary shocks. During the estimation procedure the discount factor ($\beta$) is fixed to 0.995, which is close to the the mean of the inverse ex-post real interest rate in the sample period. The steady-state (gross) inflation target ($\Pi$) in the monetary policy rule is set to 1.[10] The elasticity 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 labour in total income is approximately this value. Finally, the elasticities of substitution between goods and between labour types are respectively fixed to values standard in the literature, namely, $\upsilon = 1.1$ and $\omega = 1.4$.

The moments used to estimate these parameters are the variances, covariances, autocovariances and skewness of consumption, hours worked, price inflation, the real wage, and the nominal interest rate: 25 moments in total. In addition to the benchmark nonlinear model with GEV innovations and asymmetric adjustment costs, we estimate two alternative models: a model with asymmetric adjustment costs but normal innovations, and a model with GEV innovations but quadratic adjustment costs. Comparing these two specifications with the benchmark model allows us to evaluate the relative contribution of relaxing the assumptions of symmetric (normal) shock innovations and symmetric (quadratic) adjustment costs.[11]

[8] For a comparison of the efficiency of the SMM estimator under this and other weighting matrices see Ruge-Murcia (2012, sec. 4.3)

[9] Using a relative large value for $\lambda$ is important to accurately estimate the skewness implied by the model, but it has the drawback that it makes computationally costly the use of bootstrap methods to compute standard errors. For this reason, we report asymptotic standard errors in section 3.3 with the warning that Monte-Carlo results in Ruge-Murcia (2012) show that they may not always be a good approximation to the actual variability of SMM estimates in small samples.

[10] This choice is driven primarily by computational convenience. The steady state of the model can be computed analytically when $\Pi = 1$, but it must be computed numerically when $\Pi \neq 1$. Since the model has to be solved in each iteration of the optimization routine that minimizes (18), setting $\Pi = 1$ reduces substantially computational costs. However, sensitivity analysis shows that similar values of $\Pi$ deliver basically the same results.

[11] Also, in an online appendix, we report results from the estimation of the benchmark model using Hodrick-Prescott filtered data and show that parameter estimates are moderately robust to the method used to detrend the raw data.

### 3.3 Parameter Estimates

Table 1 reports SMM estimates and asymptotic standard errors for the different versions of the model. Let us focus first on the estimates obtained under the assumption that innovations follow GEV distributions and adjustment costs are asymmetric. The curvature parameter of labor in the utility function is relatively high (about 6.2), but it is imprecisely estimated. The adjustment cost parameters for wages and prices are quantitatively similar, but the former is somewhat larger than the latter suggesting that wages may be more rigid than prices. However, notice that the asymmetry parameter for wages is large, positive, and statistically significant different from zero (at the 10% level). The one-sided test of the hypothesis that this parameter is zero against the alternative that it is positive can be rejected at the 5% level. Thus, wage rigidity is primarily downward rigidity, rather than upward rigidity. In contrast, the asymmetry parameter for prices is negative and statistically significant at the 5% level meaning that prices are more upwardly than downwardly rigid. The fact that both asymmetry parameters are different from zero means that we can reject the hypothesis that adjustment costs are symmetric (quadratic as in Rotemberg, 1982) against the alternative that they are asymmetric (linex as in equations (5) and (12)).

Figure 1 reports the estimated adjustment cost functions for wages and prices (thick lines) and compares them with quadratic cost functions (thin lines).[12] The horizontal axis is inflation at the annual rate and the vertical axis is cost as percent of output. From equation (17), notice that wage and price adjustment costs as a proportion of output are, respectively, $w_t n_t \Phi_t / y_t$ and $\Gamma_t$. (For the purpose of constructing figure 1, we fixed $w_t$, $n_t$, and $y_t$ to their steady-state values.) This figure supports the following conclusions. First, the adjustment costs predicted by the model are not implausibly high in the neighborhood of price stability where gross inflation equals 1, but they can increase quickly outside this region. Second, downward nominal wage rigidity is quantitatively and economically important so that, for example, an annual gross inflation rate of 2% induces wage adjustment costs equal to 0.07% of output, but a deflation rate of 2% induces adjustment costs equal to 0.34% of output. Finally, despite its statistical significance, the asymmetry in price rigidity is quantitatively modest. For instance inflation and deflation rates of 2% respectively induce price adjustment costs of 0.19% and 0.12% of output.

Estimates of the scale and shape parameters imply that productivity innovations are negatively skewed with skewness equal to $-0.59$. The result that the shape parameter is not statistically different from zero means that, among extreme value distributions, the one that best describes productivity innovations is the Gumbel distribution. Figure 2 plots the estimated probability

[12] The quadratic functions have the same values for $\phi$ and $\gamma$ as the estimated linex functions (180.31 and 118.31, respectively), but with asymmetry parameters, $\psi$ and $\eta$, set to zero.

density function (PDF) (thick line) and compares it with the PDF of a normal distribution with the same standard deviation (thin line). Note that the PDF of the GEV distribution has more probability mass in the left tail and less mass in the right tail, than the normal distribution. Thus, extreme negative productivity shocks can occasionally happen, but large positive ones are unlikely.

Additional evidence on the skewness of productivity innovations is reported in the last columns of table 1 (under the heading "Single Equation") and in figure 3. For this analysis, we use the series on total factor productivity (TFP) series constructed by John Fernald (Fernald, 2014) and available from the website of the Federal Reserve Bank of San Francisco (www.frbsf.org). The series, which is an annual growth rate at the quarterly frequency, was first converted into a quarterly growth rate. Then, in order to construct a measure of productivity consistent with the model—where productivity is stationary in levels—we time aggregated the data to obtain a productivity index in levels, as in $\ln(\hat{z}_t) = \ln(\hat{z}_{t-1}) + \Delta \ln\left(z_t^f\right)$, where $\Delta \ln\left(z_t^f\right)$ is Fernald's measure of productivity. Finally, we detrended the index by projecting $\ln(\hat{z}_t)$ on a constant and a quadratic trend using an ordinary least squares (OLS) regression. The residuals from this regression are empirical counterparts of the productivity shock, $\ln(z_t)$, in the model. Using this measure of $\ln(z_t)$, we estimated the autogression (4) by OLS. The residuals of this regression are empirical counterparts of the productivity innovation, $\epsilon_t$, in the model. We estimated the parameters of the GEV distribution of these residuals by the method of maximum likelihood. The last column of table 1 reports single-equation estimates of the time series process of productivity (4) and of the parameters of the GEV distribution of its innovations.

Table 1 shows that the autoregressive coefficient from the single-equation estimation is quantitatively close to the SMM estimates from the full model. Estimates of the distribution parameters are also similar to those of the full model, but since the shape parameter is statistically different from zero, single-equation estimates suggest that productivity innovations follow a Weibull, rather than a Gumbel, distribution. The left panel in figure 3 plots the histogram of the residuals of the regression (4) and shows that they are negatively skewed with a skewness of $-0.10$, which is smaller than, but still consistent with, the estimate of $-0.59$ obtained from the full model.

An alternative way to illustrate the departure of productivity innovations from normality is the quantile-quantile (q-q) plot in figure 4.[13] The left panel in figure 4 plots the quantiles of the residuals of the regression (4) versus the theoretical quantiles of a standard normal distribution. If the residuals were normally distributed the plot would be close to linear. As we can see in figure 4, the lower quantiles of the empirical distribution of the residuals features larger realizations (in

[13] Quantile-quantile plots are widely used in extreme value analysis to illustrate the thick tails associated with extreme realizations.

absolute value) than are consistent with a normal distribution. These observations account for the negative skewness of productivity innovations and as shown in the figure, they are appear to be primarily associated with recessions.

The preference shock is very persistent and estimates of the scale and shape parameters imply that its innovations are negatively skewed with skewness equal to $-0.92$. In this case, the shape parameter is statistically different from zero, which means that innovations to the preference shock follow a Weibull distribution. The middle panel of figure 2 plots the estimated PDF of preference innovations and shows that the distribution has more probability mass in the left tail and less mass in the right tail, than a normal distribution with the same standard deviation, as for the productivity innovations. Hence, large positive innovations are unlikely but large negative ones can occasionally happen.[14]

Regarding monetary policy, table 1 shows that the smoothing parameter in the Taylor rule is moderately large and statistically significant. The coefficients of inflation and output are positive, but only the former is statistically significant. The long-run response to inflation is 1.39 while the long-run response to output is 0.37. The estimated scale and shape parameters imply that monetary policy innovations are positively skewed—in contrast to the other two innovations—with skewness equal to 0.87. The right panel of figure 2 plots the estimated PDF of monetary policy innovations and shows that the 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. This means that extreme positive monetary shocks can happen sometimes, but large negative ones are uncommon.

The last column of table 1 reports single-equation estimates of the Taylor rule (15) and of the parameters of the GEV distribution of its innovations. The coefficients of the Taylor rule were estimated by OLS, and the parameters of the GEV distribution of its residuals were estimated by maximum likelihood. Table 1 shows that the coefficients from the single-equation estimation of the Taylor rule are similar to the SMM estimates from the full model. The estimate of the scale parameter is also similar to that of the full model, but the estimate of the shape parameter is negative and statistically different from zero. Thus, single-equation estimates indicate that monetary policy shocks follow a Weibull distribution, rather than the Gumbel distribution implied by estimates from the full model. However, in both cases the predicted skewness of monetary policy shocks is positive.[15] Figure 3 plots the histogram of the residuals of the regression (15) and shows that they are positively skewed with a skewness of 1.06, which is quantitatively larger than, but

[14] No single-equation estimates are reported for this shock because we could not construct an empirical counterpart as we did for productivity (Solow residual) and monetary policy (residuals of an estimated Taylor rule).

[15] Recall that for the GEV distribution, a positive shape parameter is sufficient for positive skewness, but a negative shape parameter may imply either positive or negative skewness depending on the relative magnitudes of the shape and scale parameters.

still consistent with, the estimate of 0.87 obtained from the full model.

The right panel in figure 4 shows the quantile-quantile plot for the residuals of the Taylor rule (15), which compares the quantiles of these residuals with the theoretical quantiles of a normal distribution. The "extreme event" in this case is a large monetary policy innovation associated with the Volcker's disinflation policy. This observation and the other ones in the upper quantiles of the empirical distribution are larger than would be consistent with a normal distribution and they account for the positive skewness reported in figure 3. Overall, the quantile-quantile plots in figure 4 suggest that the few, extreme observations in the data are important to identify the parameters of the innovation distributions.

## 3.4 Alternative Models

Let us consider the SMM estimates based on the first alternative model with asymmetric adjustment costs but normal innovations. Two middle columns of table 1 show that these estimates are similar to those reported for the model with GEV innovations with two exceptions: the wage adjustment cost parameter is one order of magnitude smaller, and the price asymmetry parameter is smaller and not statistically different from zero. Hence, for this model one cannot reject the null hypothesis that price adjustment costs are quadratic.

The lower panel of table 1 compares the goodness of fit of all models. The panel reports the value of the objective function at the minimum, the root mean squared error (RMSE), and the mean absolute error (MAE). Notice all three statistics are larger for the version with normal innovations than for the version with GEV innovations. The main reason for this result is that the model with GEV innovations accounts better for the skewness in the data than the model with normal innovations. Column 1 of table 2 reports estimates of the skewness of six key U.S. macroeconomic series, namely consumption, hours worked, the real wage, wage inflation, price inflation, and the nominal interest rate. Consumption, hours worked, the real wage, and wage inflation are negatively skewed, primarily as a result of large negative observations associated with recessions, while price inflation and the nominal interest rate are positively skewed. The skewness predicted by the model is computed from an artificial sample with $20,000$ observations simulated using the parameter values reported in table 1. For the model with normal innovations column 3 shows that, despite the fact that innovations are symmetric, this model can generate some skewness as a result of the nonlinearity of the model. However, in some cases (e.g., consumption and hours worked), the predicted skewness is much lower than in the data, and in the case of wage inflation the predicted skewness is of sign opposite to the data. For the model with GEV innovations, column 2 shows that the combination of nonlinearity and asymmetric innovations deliver skewness that is quantitatively similar to the data. Notice, however, that also in this case the predicted skewness for wage inflation is of sign opposite to the data.

Let us now turn to the SMM estimates based on the other alternative model with GEV innovations but quadratic adjustment costs. Table 1 shows that the point estimate of the wage adjustment cost parameter is extremely large, but not statistically significant. Still, this model would imply implausibly large wage adjustments costs as a fraction of output: for example, an inflation rate of 2% would entail adjustment costs of the order of 50% of output. The point estimate of the price adjustment cost parameter is small and not statistically significant. The parameters of the shock processes are roughly similar to those of the model with GEV innovations and asymmetric adjustment costs, but shape parameters tend to be quantitatively large (in absolute value). As we can see in table 2, this implies large negative skewness for productivity and preference innovations and large positive skewness for monetary policy innovations. Large skewness accounts for the relative success of this model in matching the skewness observed in the data (for example, this model predicts skewness in wage inflation of the same sign as in the data). However, the skewness of innovations under this model are one order magnitude larger than under the model with asymmetric adjustment costs and, in the case of productivity and monetary policy innovations, much larger than in the data. Also notice that the value of the objective function at the minimum, the RMSE, and the MAE are larger for the model with quadratic adjustment costs than for the model with asymmetric adjustment costs. The quantitative differences in the measures of fit across model is driven by the skewness (especially that of consumption), which is better accounted for by the models with GEV innovations.

Table 2 also reports results of Jarque-Bera tests that the data follows a normal distribution. The hypothesis is rejected at the 5% significance level for all U.S. time series, including our measures of productivity and monetary policy innovations. The hypothesis is also rejected for all series simulated under the three versions of the model, including the version with normal innovations. The reason for the latter result is simply that normal shocks propagated through a nonlinear model would produce non-normal series. The result that the hypothesis is rejected (cannot be rejected) when innovations are GEV (normal) is, of course, obtained by construction.

### 3.5 Impulse Responses

Using our benchmark estimates in the first column of table 1, we study the response of the economy to productivity, preference, and monetary shocks using impulse-response analysis. We compute the responses to shock innovations in the 5th and 95th percentiles, which are assumed to occur when all variables are equal to the unconditional mean of their ergodic distribution. The size (in absolute value) of these innovations is not same for an asymmetric distribution like the GEV, but the point here is that the likelihood of the two realizations is the same. As we will see in figures 5 through 7, responses are qualitatively similar to those reported in earlier New Keynesian literature, except for the fact that in this model they are asymmetric, with shocks of a given sign having larger effects than the equally likely shock of the opposite sign.

Figure 5 plots the responses to productivity shocks for the model with GEV innovations. The vertical axis is the percentage deviation from the mean of the ergodic distribution and the horizontal axis is quarters. The positive shock in the 95th percentile of the distribution induces an increase in consumption, which is persistent as a result of intertemporal smoothing. Hours worked increase on impact and, following a hump, return to their unconditional mean from below. Price inflation and the nominal interest rate decrease; in the case of the interest rate because the inflation coefficient in the Taylor rule is quantitatively larger than that of output. Finally, the wage inflation increases on impact, goes below the mean of its ergodic distribution for a brief period, and then increases again returning to its unconditional mean from above. Since prices are more flexible than wages, the shock induces an increase in the real wage. Due to the strong wage rigidity, the response of wage inflation is relatively muted. Qualitatively, the effects of the negative shock in the 5th percentile are the opposite to those just described. However, the key observation in figure 4 is that the effects of the negative shock are larger than those of the equally likely positive shock. This result is partly due to the fact that the size of these two innovations is different for the asymmetric GEV distribution: the negative innovation takes productivity $-2.16$ percent below the steady state, while the positive innovation takes it 1.69 percent above. This relatively modest difference in shock size is amplified by the nonlinearity of the model leading to the asymmetric responses reported in figure 5.

Figure 6 plots the responses to monetary shocks. The positive shock raises the interest rate and induces a decrease in consumption and hours worked and an increase in price inflation and wage inflation. Since the response of wage inflation is much larger than that of price inflation, the real wage increases. The negative shock has the converse effects, but, as before, the key result is the asymmetry in the responses to monetary shocks: the positive shock induces larger responses than the equally likely negative shock. The fact that GEV distribution of monetary shocks has a large positive skewness means that the size of the two innovations is substantially different: the positive innovation takes the interest rate 0.81 percent above the steady state, while the negative innovation takes it $-0.54$ percent below. Hence, both the shock asymmetry and the model nonlinearity account for the asymmetric responses reported in figure 6.

Finally, figure 7 plots the responses to preference shocks. The positive shock increases the utility level of consumption and, hence, consumption increases. The associated increase in output, induces an increase in hours worked. Recall that, by construction, this shock does not alter the marginal rate of substitution between leisure and consumption. Both price and wage inflation increase and in roughly the same magnitude so that response of the real wage to the positive shock is relatively muted. The negative shock induces the converse effects, but effects are asymmetric. For example, the negative shock induces a decrease in consumption of $-0.59$ percent, while the positive shock induces an increase of 0.36 percent. The largest asymmetry concerns the real wage: the negative shock induces an increase 0.15 percent, while the positive shock induces a decrease of $-0.07$ percent, which is half the size of the response to the negative shock.

In summary, results in this section show that negative productivity and preference shocks have larger effects than equally likely positive shocks, while positive monetary policy shocks have larger effects than equally likely negative shocks. These results are due the asymmetry of the shock innovations and their amplification through the nonlinear propagation mechanism of the model. In turn, these asymmetric effects induce negative skewness in consumption, hours worked, and the real wage, and positive skewness in price inflation, wage inflation, and the nominal interest rate, which are roughly in line with the U.S. data.

# 4. The Ramsey Policy

Consider a monetary authority that follows the Ramsey policy of maximizing the households' welfare by choosing $\{c_t, n_t, W_t, I_t, \Omega_t, \Pi_t\}_{t=s}^{\infty}$ to maximize

$$E_s \sum_{t=s}^{\infty} \beta^{t-s} \left( \log(c_t^h) - \frac{(n_t^h)^{1+\chi}}{1+\chi} \right) u_t,$$

subject to the resource constraint and the first-order conditions of firms and households, and taking the previous values for wages, goods prices, and shadow prices as given. It is assumed that the monetary authority can commit to the implementation of the optimal policy and that it discounts future utility at the same rate as households. The model is solved using a third-order perturbation with parameter estimates equal to those reported in the first column of table 1.

## 4.1 Decision Rules

Figures 8 and 9 plot the decision rules that solve the model when policy is implemented by the Ramsey planner. Figure 8 (resp. figure 9) plots the rule as a function of the productivity shock (resp. preference shock) holding all the other state variables at their values in the deterministic steady state. In these figures, the vertical axis is the percentage deviation from the deterministic steady state and the horizontal axis is the size of the shock normalized by their standard deviation.

The thick line is the nonlinear decision rule implied by our third-order perturbation, while the thin line is the linear policy function implied by a first-order perturbation.

Figure 8 shows that the nonlinear decision rules for consumption, hours worked, wage inflation, and the real wage are upward-sloping and concave functions of the productivity shock. For price inflation and the nominal interest rate the decision rules are downward-sloping and convex functions of the productivity shock. The departure from the linear rule is specially pronounced in the cases of consumption, hours worked, and the nominal interest rate. The key result from figure 8 is that, for all variables, the policy rules imply larger adjustment in the variables when the economy is hit by a negative, than when it is hit by a positive, productivity shock. That is, the response of the Ramsey planner to productivity shocks is asymmetric. Up to the extent that productivity shocks are negatively skewed—and, hence, there are more large realizations from the left than from the right side of the distribution—the optimal policy calls for occasional, large interest rate adjustments in response to extreme productivity shocks

Figure 9 shows that under the Ramsey policy, the planner adjusts the nominal interest rate in response to the preference shock and undoes its effects on consumption, hours worked, the real wage, price inflation, and wage inflation. For this reason, the coefficients of this shock in their decision rules are zero and the plots in figure 9 are horizontal lines that cross the vertical axis at zero. In contrast, the decision rule for the nominal interest rate is upward-sloping and linear. That is, the coefficients of the higher-order terms in the nonlinear decision rule are zero and, thus, the rule in the nonlinear model is just a shifted version of the rule in the linear model. The linearity of this decision rule implies that the optimal interest rate response to preference shocks is symmetric.

### 4.2 Impulse Responses

Figure 10 plots the optimal responses to productivity shocks under the Ramsey policy. As before, the shocks are innovations in the 5th and 95th percentiles of the GEV distribution and take place when all variables are equal to the unconditional mean of their ergodic distribution. The positive productivity shock induces a persistent increase in consumption and hours worked as agents take advantage of their temporarily high productivity. Wage inflation increases and price inflation decreases leading to an unambiguous increase in the real wage. Qualitatively these responses are similar to those obtained under the Taylor rule policy. The negative shock induces the converse effects, but their magnitudes are larger than for the positive shock in all cases. As discussed in section 3.5, the size of the productivity innovation at the 5th percentile is only somewhat larger than that at the 95th percentile, and, thus, the asymmetry in the responses reported in figure 10 is due primarily to the nonlinearity of the model.

Regarding the nominal interest rate, notice that the positive productivity shock calls for an interest rate decrease of basically the same magnitude as under the Taylor policy (see figure 4). Instead, the negative productivity shock calls for an interest rate increase of a larger magnitude under the Ramsey than under the Taylor policy. Hence, the asymmetry of the interest rate response to productivity shocks is more pronounced under the Ramsey than under the Taylor policy and, in particular, involves stronger responses to large negative shocks.

Figure 11 plots the optimal responses to the preference shock under the Ramsey policy. As anticipated from the decision rules in figure 9, the interest rate response under the Ramsey policy undoes the effects of this shock on the other variables, and they remain at the mean of their ergodic distribution after the shock. The Ramsey policy calls for an increase in the nominal interest rate in response to the positive preference shock, and a decrease in response to the negative preference shock. However, since the decision rule is linear, the asymmetry in the responses in figure 11 is primarily the result of the asymmetry in the distribution of the innovations to the preference shock.

### 4.3 Optimal Inflation

We measure the optimal inflation rate by the mean of the ergodic distribution of inflation under the Ramsey policy. For the parameter estimates reported in table 1, mean annual gross inflation computed from a simulation of $20,000$ observation is basically 1 (or, more precisely, 0.99998). This means that optimal expected inflation is essentially the same as the inflation rate in the deterministic steady state. This result is remarkable because the model is nonlinear and, consequently, it does not feature certainty equivalence. Hence, one would expect different average inflation rates in the stochastic and deterministic steady states of the model. In particular, a prudent Ramsey planner—who faces skewness risk in the form of possibly large negative shocks from the left tail of the productivity distribution, which may require costly downward nominal wage adjustments—should target an average rate of price inflation above unity. However, this conjecture is not confirmed in our model because the Ramsey planner actually needs to trade off the benefits of acting prudently with the costs of systematically incurring price and wage adjustment costs when gross inflation is above 1. A similar result is reported by Coibion et al. (2012), which shows in a calibrated model that for costly, but infrequent, episodes at the zero lower bound (ZLB) on interest rates, the optimal inflation rate is low.[16]

In one of the few contributions to the literature on optimal policy in an environment with

[16] In previous work (Kim and Ruge-Murcia, 2009) we found that downward nominal wage rigidity dictates an optimal level of inflation for the U.S. economy of about 0.35% per year. We note that, while the second-order perturbation used to solve the model in Kim and Ruge-Murcia (2009) implies a small but positive inflation target, the third-order perturbation used to solve the model in this paper implies a target of practically zero (net) inflation. This observation suggests that the third-order terms in the approximate policy rule are quantitatively and economically important.

extreme shocks, Svensson (1993) notes the tension between: (i) acting prudently and incorporating systematically the possibility of extreme shocks into policy and (ii) taking a wait-and-see approach. An example of the first strategy is increasing the inflation target, as proposed by Williams (2009, 2014), Blanchard et al. (2010), and Ball (2014) some of whom advocate inflation targets of 4% per year, as opposed to the 2% per year currently used by the Federal Reserve and some other central banks. In our model, the trade-off between the two options in Svensson (1993) arises for the following reason. On the one hand, prudence induces the policy market to target a gross inflation rate above 1 in order to avoid costly nominal wage cuts. On the other hand, price and wage adjustment costs induce the policy maker to target a gross inflation rate equal to 1 and, instead, to aggressively adjust the policy variable(s) when an extreme, negative shock occurs. As we can see, the quantitative welfare analysis of these two strategies in our model unambiguously favours strict price stability and a wait-and-see approach.

## 5. Conclusion

This paper uses tools from extreme value theory to study the positive and normative implication of extreme events for monetary policy. Our New Keynesian model incorporates a trade-off between (i) acting prudently and systematically incorporating the possibility of extreme shocks into policy (e.g., by targeting a gross inflation target above 1) and (ii) taking a wait-and-see approach whereby the central banker targets a gross inflation rate close to 1 but adjusts policy variables aggressively when a extreme negative shocks hits the economy. We evaluate the welfare implication of these two approaches and find that for our estimated model, this trade-off is solved unambiguously in favour of the wait-and-see approach. The intuition is simple: the cost of price and wage adjustments required under the prudent policy in every period overrides the potential benefit of the prudent policy in the expectation of large, but infrequent, extreme negative shock. As a result, the optimal gross inflation rate under the Ramsey policy is virtually indifferent from 1 (that is, strict price stability).

In interpreting our findings it is important to keep in mind that the mechanism through which extreme shocks affect welfare in our model is by increasing the likelihood of deflation and, hence, costly nominal wage cuts. An alternative strategy that we did not pursue here is to consider the cost associated with the zero lower bound on nominal interest rates. Since the ZLB and downward nominal wage rigidity may lead to similar nonlinear dynamics, it is possible that they are modeling substitutes (e.g., see Coibion et al., 2012) and results may be robust to explicitly modeling the zero lower bound. We plan to examine this conjecture in future research.

**Table 1: SMM Estimates**

| | Model | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | GEV with Asymmetric Costs | | Normal with Asymmetric Costs | | GEV with Quadratic Costs | | Single Equation | |
| | Estimate | s.e | Estimate | s.e | Estimate | s.e. | Estimate | s.e |
| Labor parameter | 6.205 | 4.124 | 1.172 | 1.783 | 0.104 | 0.769 | | |
| Nominal rigidity | | | | | | | | |
| Wage adjustment cost | 180.31$^{\dagger}$ | 101.13 | 33.85 | 33.54 | 75534.78 | 56975.31 | | |
| Wage asymmetry | 450.64$^{\dagger}$ | 234.51 | 602.48* | 172.46 | | | | |
| Price adjustment cost | 118.31* | 53.49 | 115.25* | 30.04 | 10.94 | 10.93 | | |
| Price asymmetry | −126.20* | 39.25 | 27.46 | 40.52 | | | | |
| Productivity shock | | | | | | | | |
| Autoregressive coeff. | 0.898* | 0.018 | 0.898* | 0.029 | 0.907* | 0.034 | 0.934* | 0.021 |
| Scale$\times$10 | 0.137* | 0.037 | 0.145* | 0.027 | 0.132* | 0.029 | 0.076* | 0.004 |
| Shape | −0.480 | 0.699 | | | −0.757* | 0.259 | −0.285* | 0.034 |
| Preference shock | | | | | | | | |
| Autoregressive coeff. | 0.994* | 0.002 | 0.995* | 0.002 | 0.999* | 0.001 | | |
| Scale$\times$10 | 0.851* | 0.209 | 0.797* | 0.186 | 0.206 | 3.196 | | |
| Shape | −0.609$^{\dagger}$ | 0.349 | | | −2.077 | 10.299 | | |
| Monetary policy rule | | | | | | | | |
| Smoothing | 0.638* | 0.197 | 0.823* | 0.051 | 0.441* | 0.115 | 0.921* | 0.020 |
| Inflation | 0.502* | 0.239 | 0.256* | 0.082 | 0.707* | 0.124 | 0.123* | 0.022 |
| Output | 0.133 | 0.106 | 0.075* | 0.033 | 0.097* | 0.042 | 0.052* | 0.010 |
| Scale$\times10^2$ | 0.287 | 0.205 | 0.168$^{\dagger}$ | 0.091 | 0.194$^{\dagger}$ | 0.105 | 0.196* | 0.009 |
| Shape | −0.049 | 0.405 | | | 0.317* | 0.101 | −0.117* | 0.014 |
| Objective function$\times10^2$ | 0.901 | | 2.154 | | 6.865 | | | |
| RMSE | 0.770 | | 4.330 | | 2.467 | | | |
| MAE | 0.308 | | 1.104 | | 0.813 | | | |

*Note:* s.e. denotes asymptotic standard errors. The superscripts $*$ and $\dagger$ denote statistical significance at the 5% and 10% levels. For the normal distribution, the scale parameter is the standard deviation. No single-equation estimates are reported for the preference shock because we could not construct an empirical counterpart for this shock as we did for productivity (Solow residual) and monetary policy (residuals of an estimated Taylor rule).

**Table 2: Skewness and Jarque-Bera Tests**

| Variable | U.S. Data | Model: GEV with Asymmetric Costs | Normal with Asymmetric Costs | GEV with Quadratic Costs |
|---|---|---|---|---|
| Skewness | (1) | (2) | (3) | (4) |
| Consumption | −0.707 | −0.586 | −0.138 | −0.372 |
| Hours worked | −0.711 | −0.592 | −0.564 | −0.565 |
| Real wage | −0.334 | −0.356 | −0.309 | −0.392 |
| Wage inflation | −0.525 | 1.096 | 1.833 | −0.325 |
| Price inflation | 0.679 | 0.920 | 0.723 | 0.770 |
| Nominal interest rate | 0.587 | 0.458 | 0.210 | 0.440 |
| Productivity innovations | −0.104 | −0.593 | −0.015 | −1.348 |
| Monetary policy innovations | 1.059 | 0.854 | 0.002 | 5.054 |
| Preference innovations | – | −0.865 | −0.018 | −6.583 |
| Jarque-Bera test | | | | |
| Consumption | 0.003 | $< 0.001$ | $< 0.001$ | $< 0.001$ |
| Hours worked | 0.002 | $< 0.001$ | $< 0.001$ | $< 0.001$ |
| Real wage | 0.084 | $< 0.001$ | $< 0.001$ | $< 0.001$ |
| Wage inflation | $< 0.001$ | $< 0.001$ | $< 0.001$ | $< 0.001$ |
| Price inflation | $< 0.001$ | $< 0.001$ | $< 0.001$ | $< 0.001$ |
| Nominal interest rate | 0.006 | $< 0.001$ | $< 0.001$ | $< 0.001$ |
| Productivity innovations | 0.022 | $< 0.001$ | 0.463 | $< 0.001$ |
| Monetary policy innovations | $< 0.001$ | $< 0.001$ | 0.300 | $< 0.001$ |
| Preference innovations | – | $< 0.001$ | 0.439 | $< 0.001$ |

*Note:* The table reports the unconditional skewness and the $p$-value of jarqof the actual US series and of artificial data simulated from the model using the parameters reported in table 1 .

## References


[1] Abbritti, M., Fahr, S., 2013. "Downward Wage Rigidity and Business Cycle Asymmetries." Journal of Monetary Economics 60, 871–86.

[2] Adrian, T. and Duarte, F, 2017. "Financial Vulnerability and Monetary Policy." Federal Reserve Bank of New York Staff Report No. 804.

[3] Akerlof, G. A., Dickens, W. T., Perry, G. L., 1996. "The Macroeconomics of Low Inflation." Brookings Papers on Economic Activity 1996(1), 1–76.

[4] 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.

[5] Andreasen, M. M., Fernández-Villaverde, J., Rubio-Ramírez, J. F., 2013. "The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications," Review of Economic Studies, forthcoming.

[6] Aruoba, B., Bocola, L., Schorfheide, F., 2017. "Assessing DSGE Model Nonlinearities." Journal of Economic Dynamics and Control 83, 34-54

[7] Ball, L., 2014. "The Case for a Long-Run Inflation Target of Four Percent." IMF Working Paper 14/92.

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

[9] Caldara, D., Fernandez-Villaverde, J., Rubio-Ramirez, J. and Yao, W., 2012. "Computing DSGE Models with Recursive Preferences and Stochastic Volatility." Review of Economic Dynamics 15, 188–206.

[10] Card, D., Hyslop, D., 1997. "Does Inflation Greases the Wheels of the Labor Market?" In: Romer, C., Romer, D. (eds.), Reducing Inflation: Motivation and Strategy. University of Chicago Press: Chicago, pp. 71–122.

[11] Castellanos, S. G., Garcia-Verdú, R., Kaplan D. S., 2004. "Nominal Wage Rigidity in Mexico: Evidence from Social Security Records." Journal of Development Economics 75, 507–33.

[12] Coibion, O., Gorodnichenko, Y., Wieland, J., 2012. "The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise Their Inflation Targets in Light of the Zero Lower Bound?" Review of Economic Studies 79, 1371–406.

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

[14] Duffie, D., Singleton, K. J., 1993. "Simulated Moments Estimation of Markov Models of Asset Prices." Econometrica 61, 929–52.

[15] Embrechts, P., Klüppelberg, C., Mikosch, T., 2011. Modelling Extremal Events for Insurance and Finance. Springer: London.

[16] Fehr, E., Goette, L., 2005. "Robustness and Real Consequences of Nominal Wage Rigidity." Journal of Monetary Economics 52, 779–804.

[17] Fernald, J. G., 2014, "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity." Federal Reserve Bank of San Francisco Working Paper 2012-19.

[18] Ferreira, T., 2016, "Asymmetric Risk Shocks and the Credit Channel," Federal Reserve Board, Mimeo.

[19] 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, 180–90.

[20] 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, 158–71.

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

[22] Jin, H., Judd, K. L., 2002. "Perturbation Methods for General Dynamic Stochastic Models." Hoover Institution, mimeo.

[23] Keynes, J. M., 1936. The General Theory of Employment, Interest and Money. Macmillan Cambridge University Press: Cambridge.

[24] Kim, J., Ruge-Murcia, F., 2009. "How Much Inflation is Necessary to Grease the Wheels?" Journal of Monetary Economics 56, 365–77.

[25] —— 2011. "Monetary Policy when Wages are Downwardly Rigid: Friedman meets Tobin." Journal of Economic Dynamics and Control 35, 2064–77.

[26] Kuroda, S., Yamamoto, I., 2003. "Are Japanese Nominal Wages Downwardly Rigid?: Examination of Nominal Wage Change Distributions." Monetary and Economic Studies 21, 1–30.

[27] Peltzman. S., 2000. "Price Rise Faster than They Fall." Journal of Political Economy 108, 466–502.

[28] Rotemberg, J. J., 1982. "Sticky Prices in the United States." Journal of Political Economy 90, 1187–211.

[29] 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, 914–38.

[30] Ruge-Murcia, F., 2017, "Skewness Risk and Bond Prices," Journal of Applied Econometrics 32, 379-400.

[31] Schmitt-Grohe, S., Uribe, M., 2016. "Downward Nominal Wage Rigidity, Currency Pegs, and Involuntary Unemployment," Journal of Political Economy, 124, 1466-1514.

[32] Stein, J., 2014. "Challenges for Monetary Policy Communication." At the Money Marketeers of New York University, New York.

[33] Svensson, L. E. O., 2003. "Optimal Policy with Low-Probability Extreme Events." NBER Working Paper No. 10196.

[34] Varian, H., 1974. "A Bayesian Approach to Real Estate Assessment." In: Feinberg, S. E., Zellner, A. (eds.), Studies in Bayesian Economics in Honour of L. J. Savage, North-Holland: Amsterdam.

[35] Williams, J. C., 2009. "Heeding Daedalus: Optimal Inflation and the Zero Bound." Brookings Papers on Economic Activity 2, pp. 1–37.

[36] Williams, J. C., 2014. "Monetary Policy in a Low R-star World." Federal Reserve Bank of San Francisco Economic Letter 2016-23.

[37] Zbaracki, M. J., Ritson, M., Levy, D., Dutta,S., Bergen, M., 2004. "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets." Review of Economics and Statistics 86, 514–33.

[38] Zeke, D., 2016, "Financial Frictions, Volatily, and Skewness," UCLA, mimeo.

Figure 1: Estimated Price and Wage Adjustment Costs

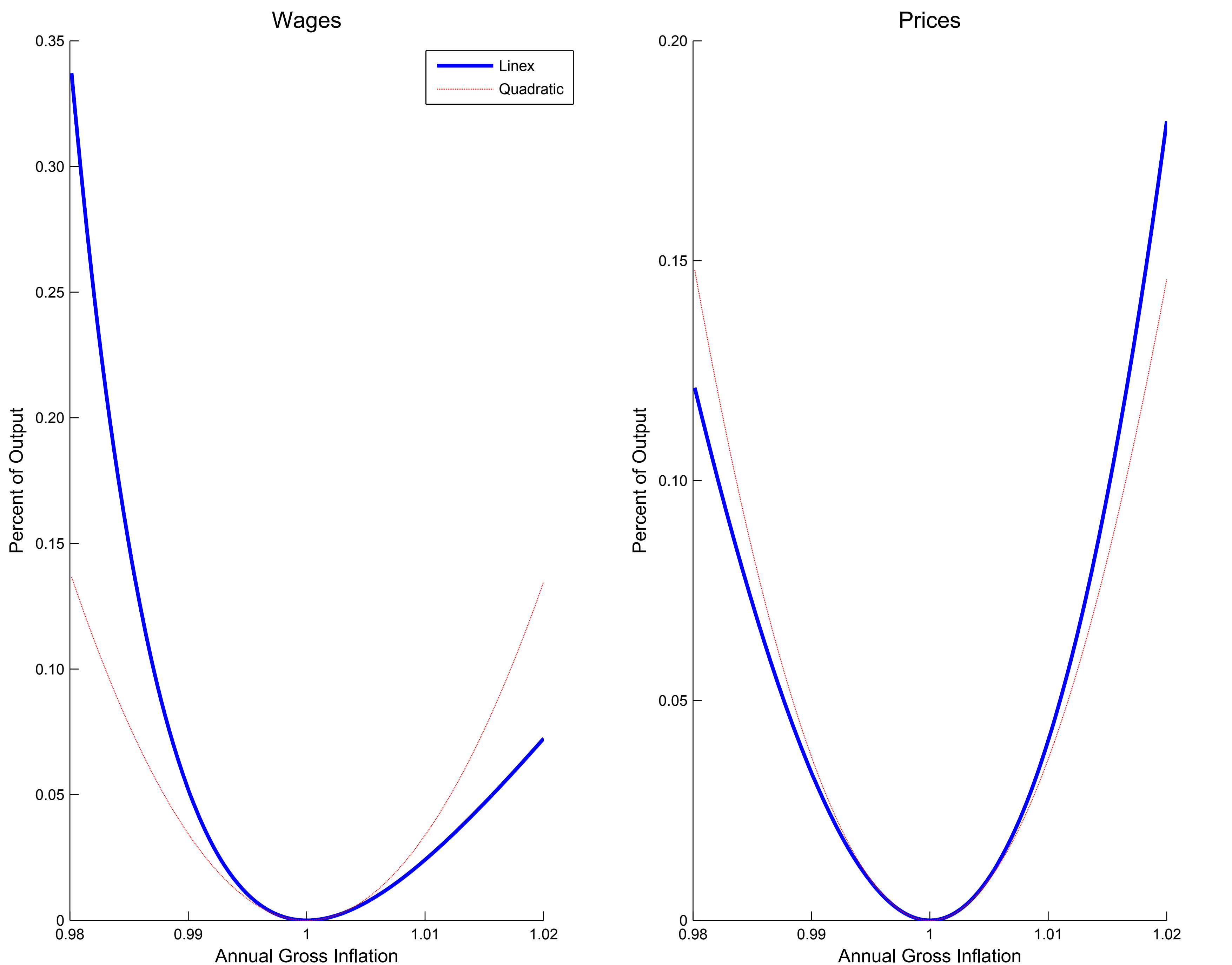
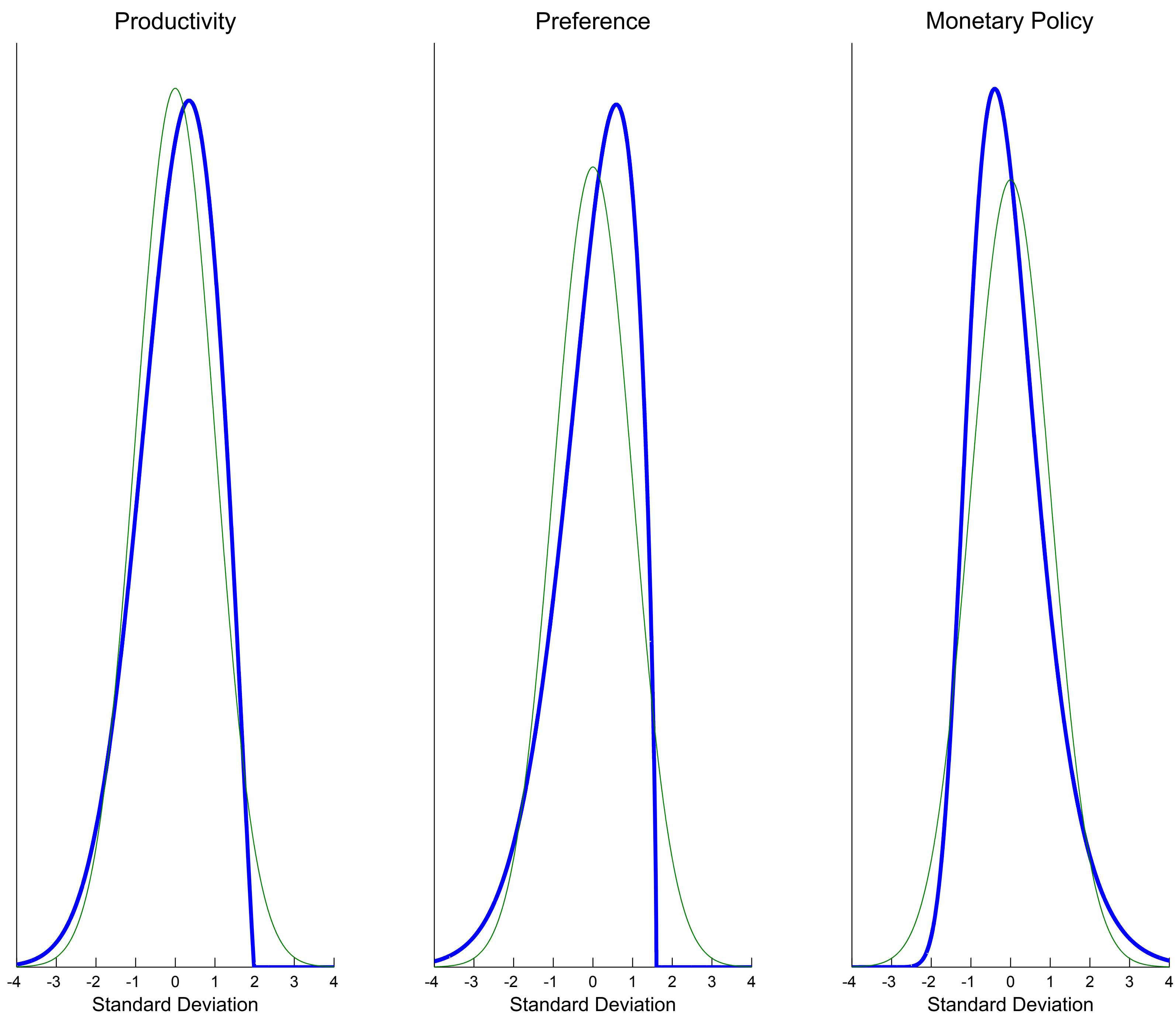


Figure 2: Estimated Probability Density Functions



# Figure 3: Histogram of Residuals Computed from the U.S. Data

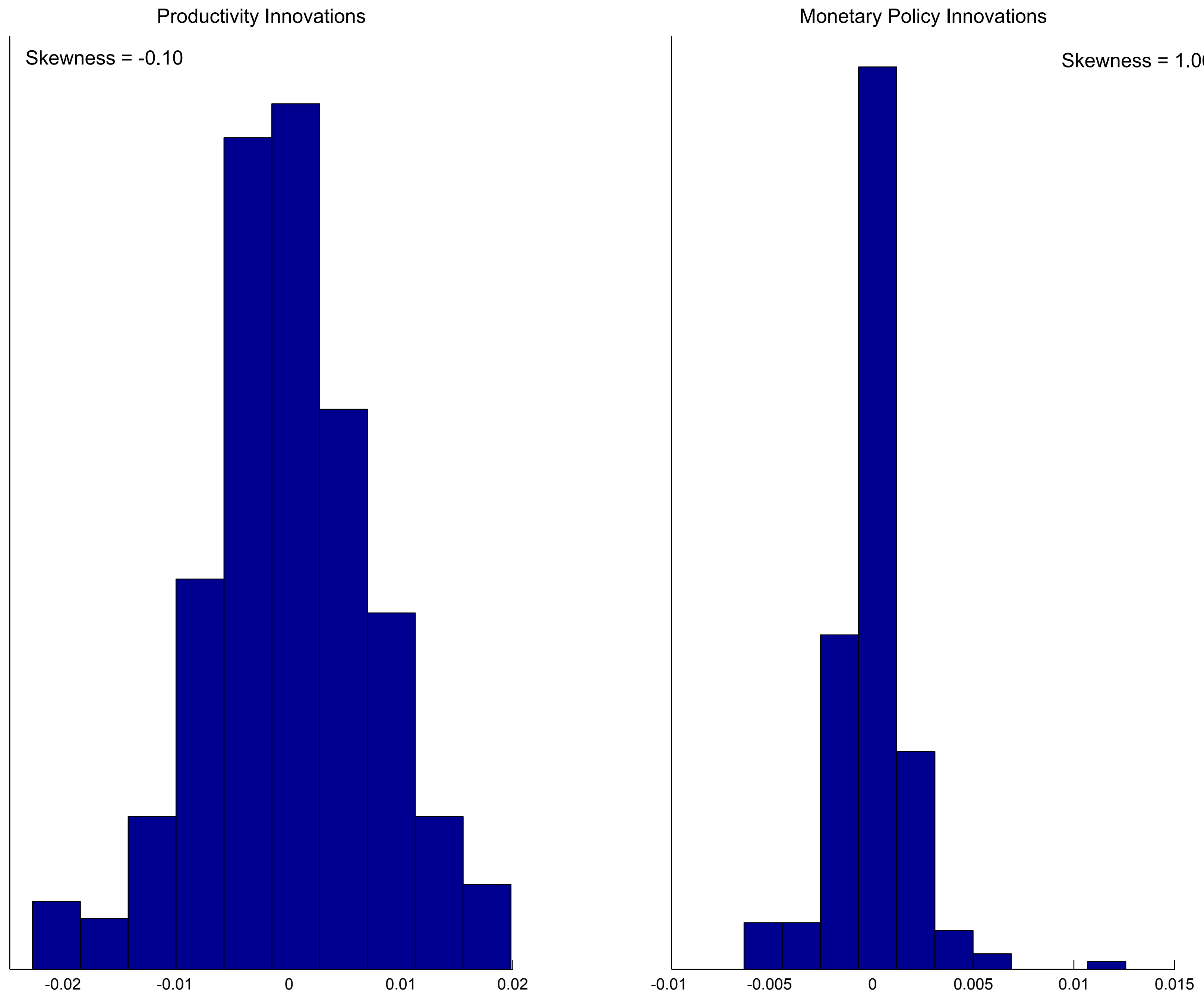


Figure 4: Quantile-quantile Plot of U.S. Residuals versus Normal Distribution

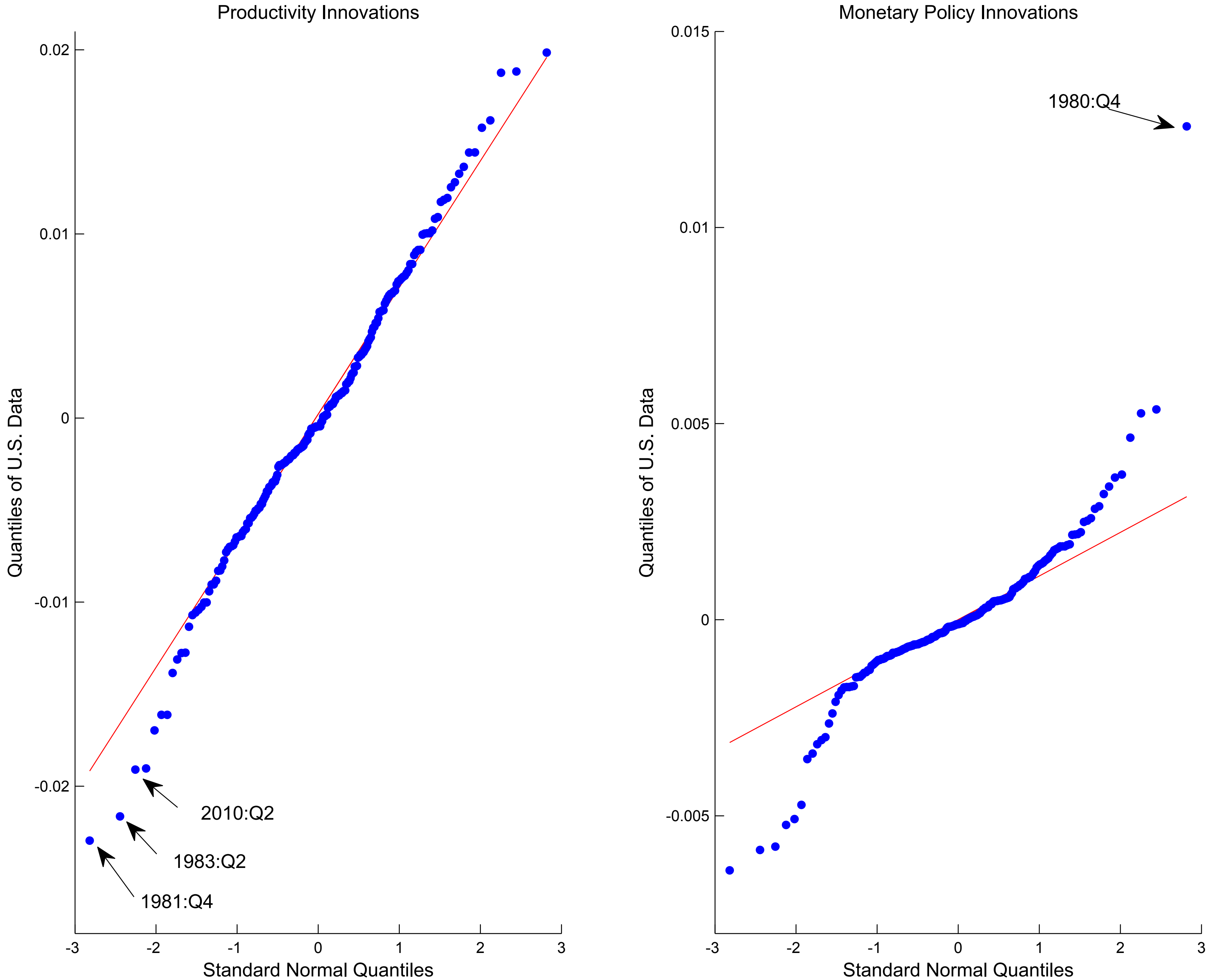
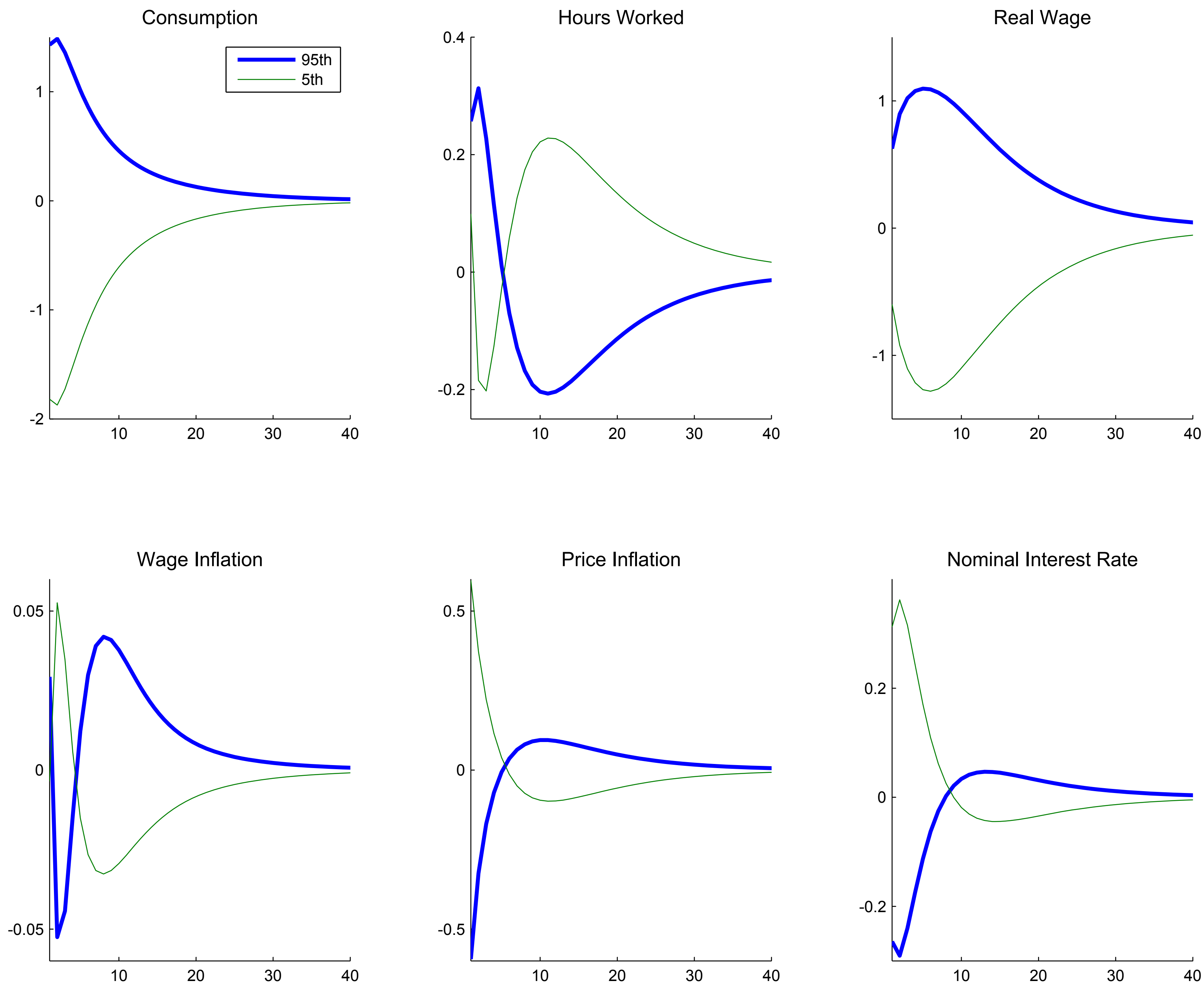
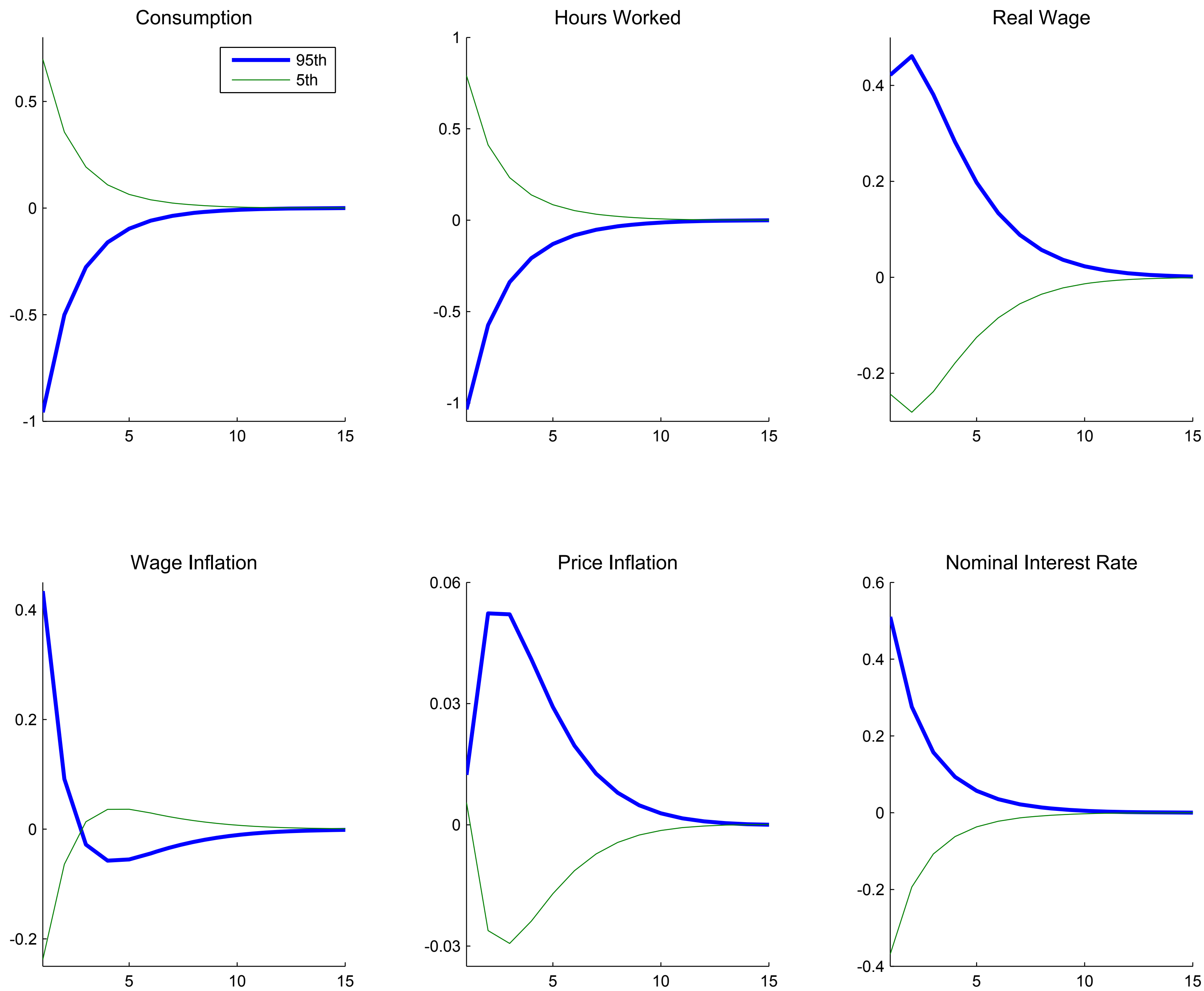


Figure 5: Responses to Productivity Shock under Taylor Policy



# Figure 6: Responses to Monetary Shock under Taylor Policy



# Figure 7: Responses to Preference Shock under Taylor Policy

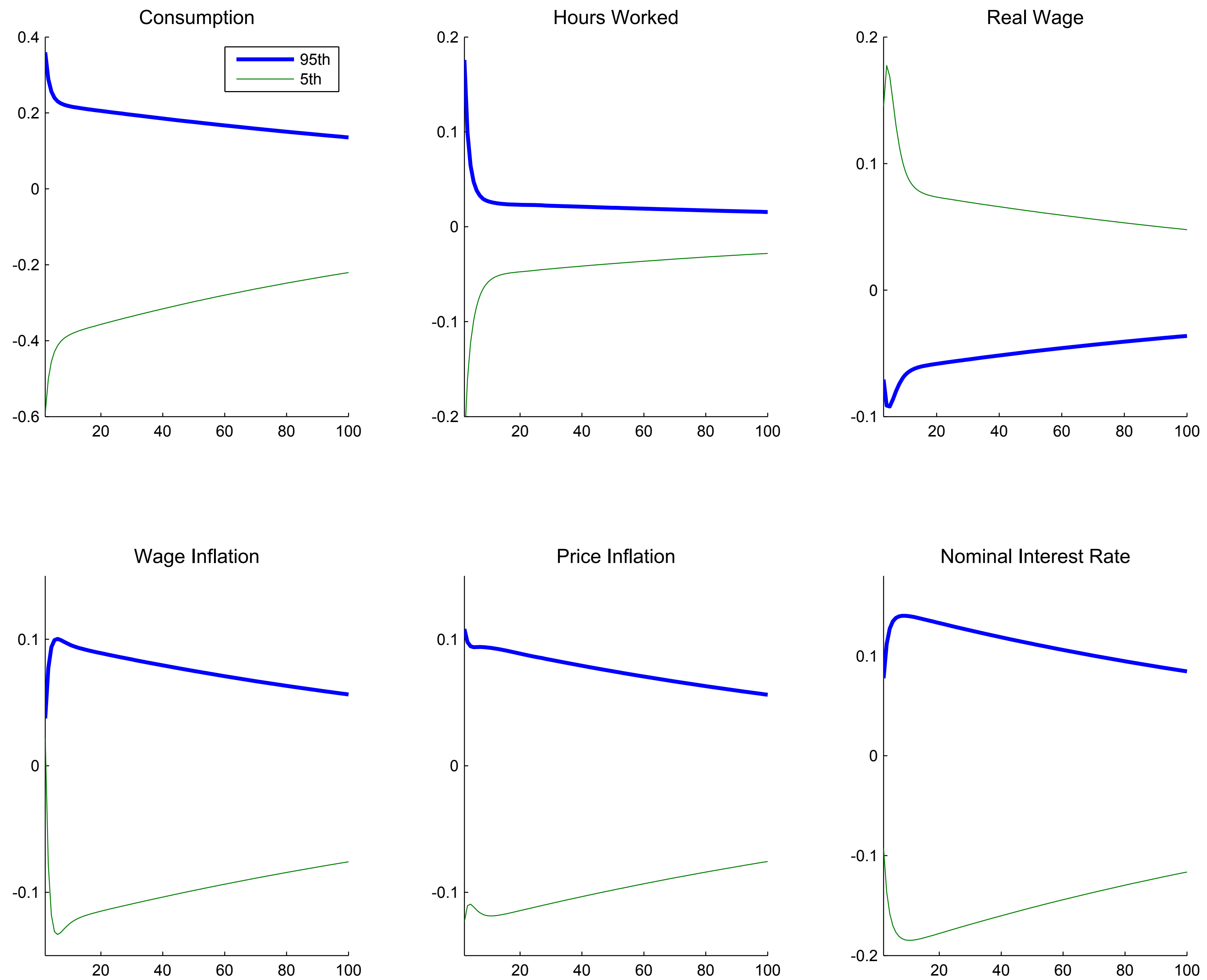


Figure 8: Decision Rules of the Ramsey Planner. Productivity Shock

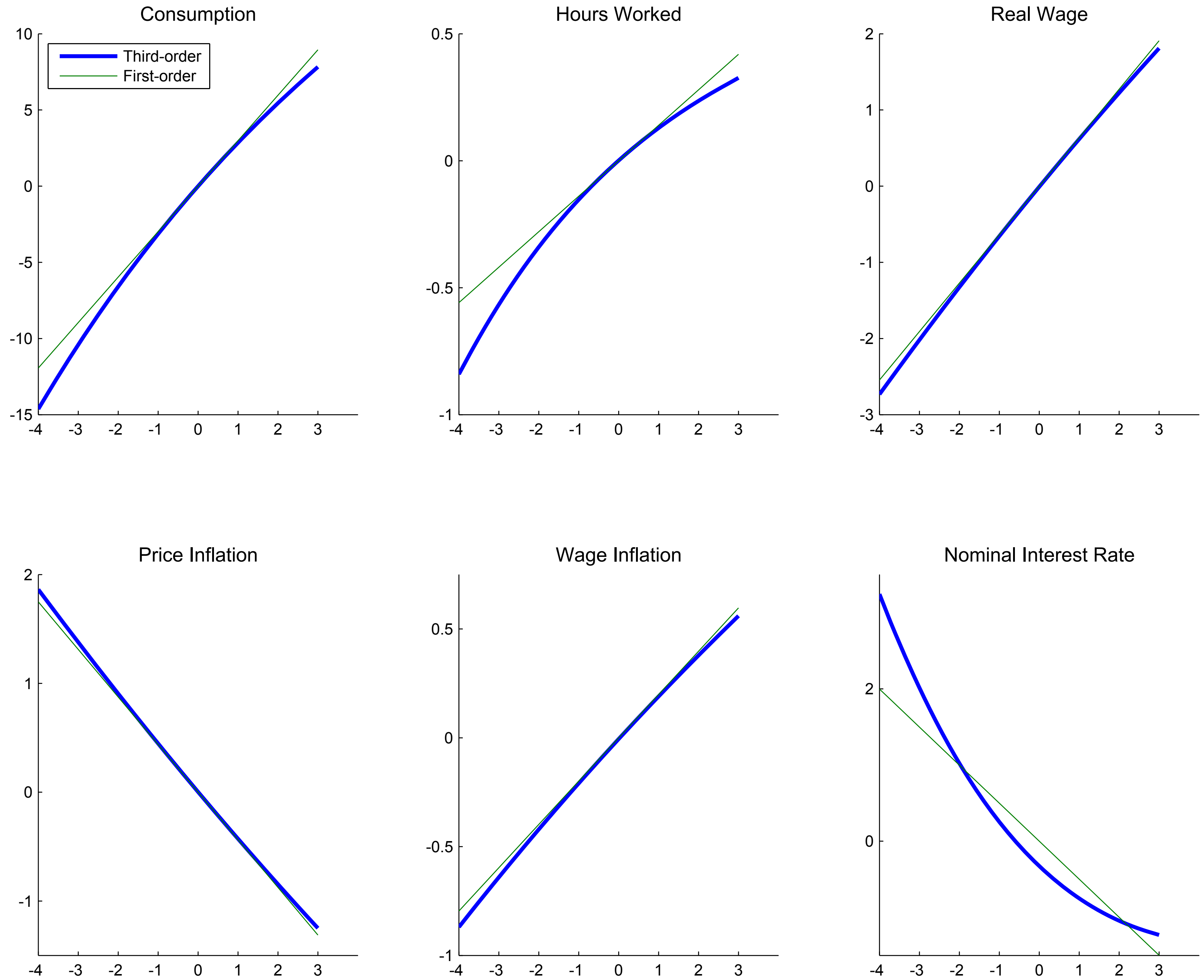
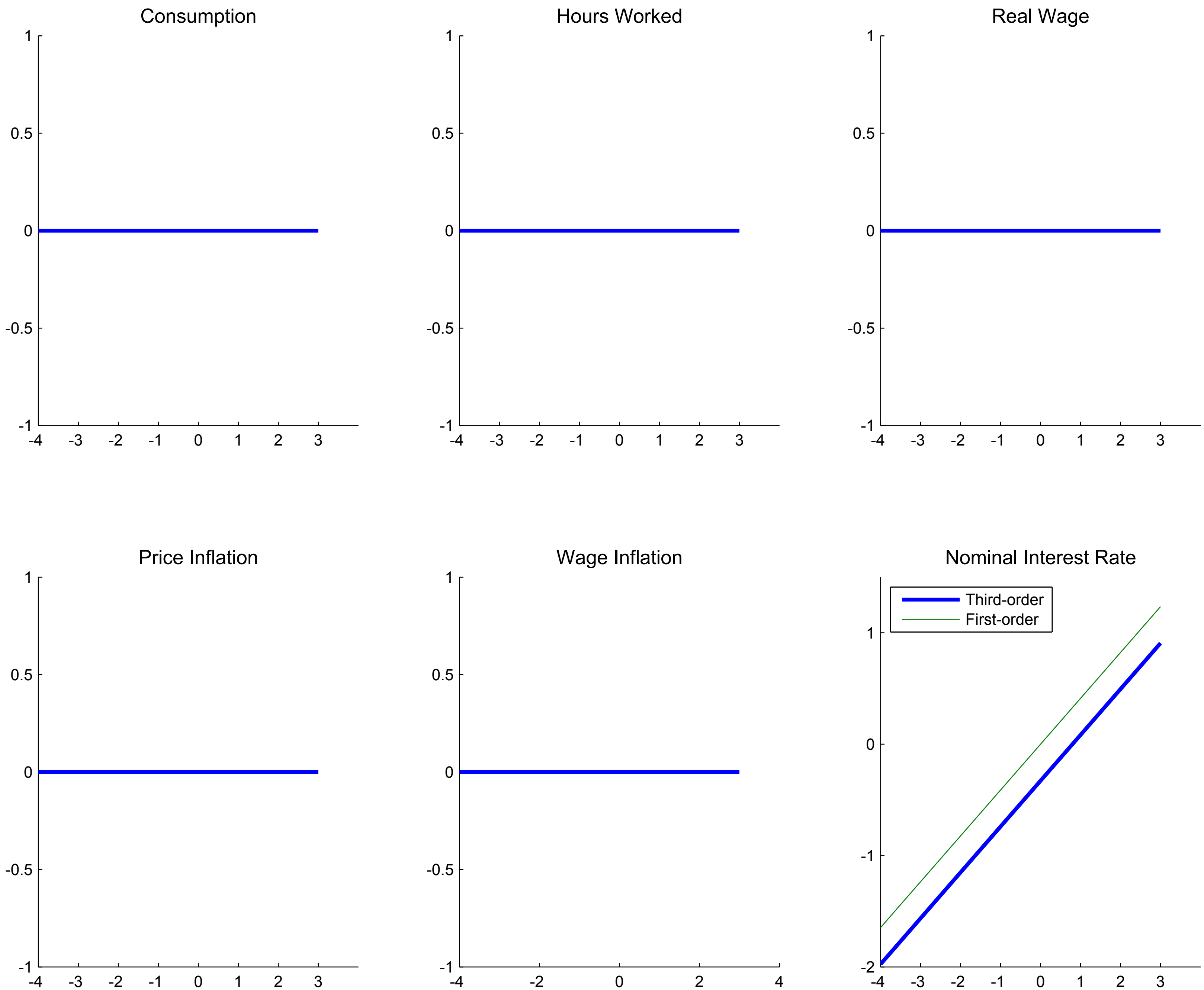
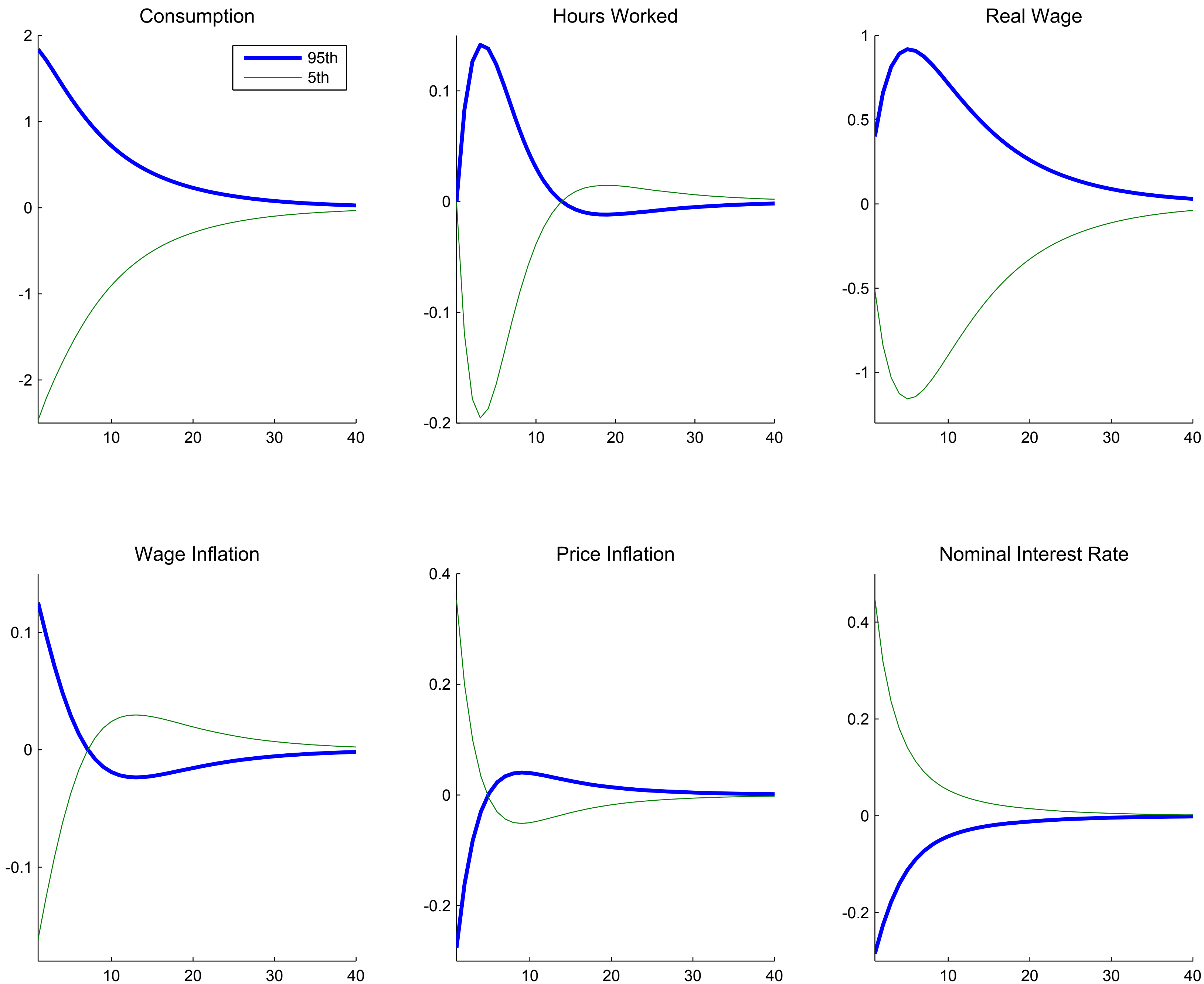


Figure 9: Decision Rules of the Ramsey Planner. Preference Shock



# Figure 10: Responses to Productivity Shock under Ramsey Policy



# Figure 11: Responses to Preference Shock under Ramsey Policy