

Extreme Events and Optimal Monetary Policy*

Jinill Kim[†]

Francisco Ruge-Murcia[‡]

First draft: August 2014

This revision: May 2016

Abstract

This paper studies the positive and normative 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 (GEV) distributions. A nonlinear perturbation of the model is estimated by the simulated method of moments. Under both the Taylor and Ramsey policies, the central bank responds non-linearly and asymmetrically to shocks. The trade-off between targeting a gross inflation rate above 1 as insurance against extreme shocks, and strict price stability is unambiguously solved in favour of the latter.

JEL Classification: E4, E5.

Key Words: 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”. We benefitted from comments by Luigi Boccia. Ruge-Murcia acknowledges the support from the Social Sciences and Humanities Research Council of Canada, and Kim acknowledges the support from Korea Research Foundation.

[†]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 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 wages. In particular, wages are downwardly rigid as in 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 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 primarily developed in meteorology and in engineering, where designers are interested in protecting structures against infrequent—but potentially damaging—events like earthquakes or 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-degree-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 to understand policy responses to extreme events.

Instead, the innovations to productivity and monetary policy 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. 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 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 innovations. Results show that the data prefer a specification where monetary policy innovations are drawn from a positively skewed distribution and productivity

¹Key contributions in extreme value theory are 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).

innovations are drawn from a negatively skewed distribution. This conclusion is based on structural estimates from the model and on reduced-form estimates from the raw data.

Using the estimated parameters, we pursue the positive and normative implications of the model. We find that under the Taylor rule policy, the monetary authority reacts more strongly to large negative than to large positive productivity shocks. Conversely, the monetary authority reacts more strongly to large positive than to large negative monetary shocks. Under the Ramsey policy, the benevolent monetary authority also responds asymmetrically to productivity shocks and the change in the nominal interest rate is generally larger than 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, Blanchard, Dell’Ariccia, and Mauro (2010) 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. Our model incorporates such a trade-off and uses quantitative analysis to compare these two strategies using a well-defined welfare metric. In our model, the solution to this trade-off is solved unambiguously in favor 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 be paid in every period. Thus, under both the Ramsey policy and a strict inflation-targeting policy, the optimal gross inflation rate is virtually indifferent from 1.

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 two versions of the model: one version with GEV innovations and a benchmark version with normal innovations. Section 4 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 5 and 6 respectively study optimal monetary policy under the Ramsey policy and under a strict-inflation-targeting policy. Finally, Section 7 concludes.

2. An Economy Subject to Extreme Shocks

The agents in this economy are 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 labor supplied by households and combines it as

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

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

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

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

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

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

The productivity shock follows the process

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

where $\zeta \in (-1, 1)$ and ϵ_t is an innovation assumed to be independent and identically distributed (i.i.d.) with mean zero and skewness different from zero. By allowing 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 this 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 ϵ_t is drawn from an asymmetric 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 due to Varian (1974),

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

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 $\eta < 0$. Finally, note that (5) nests the quadratic function in Rotemberg (1982) as the special case where η approaches zero. Hence, it is straightforward to compare statistically the model with asymmetric costs and the restricted version with quadratic costs.

The firm maximizes

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

where E_s is the expectation conditional on information available at time s , $\beta \in (0, 1)$ is the discount factor, Λ_t is the marginal utility of consumption, $c_{i,t}$ is total consumption demand for good i , $(1 - \Gamma_t^i) (P_{i,t} / P_t) c_{i,t}$ 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

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

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 labor h is

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

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

2.2 Households

Household h maximizes

$$E_s \sum_{t=s}^{\infty} \beta^{t-s} U(c_t^h, n_t^h), \quad (9)$$

where $U(\cdot)$ is an instantaneous utility function and c_t^h is consumption. Consumption is an aggregate of the differentiated goods produced by firms,

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

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

$$U(c_t^h, n_t^h) = \frac{(c_t^h)^{1-\varrho}}{1-\varrho} - \frac{(n_t^h)^{1+\chi}}{1+\chi}, \quad (11)$$

where ϱ and χ are non-negative parameters. The weight of the disutility of labor in this specification is set to 1, 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.

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

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

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 $\psi \rightarrow \infty$, the cost function takes the shape of an “L” so that wages are completely flexible upwards and completely inflexible downwards, as in Benigno and Ricci (2011).

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 McLaughlin (1994), Akerlof, Dickens, and Perry (1996), and Card and Hyslop (1997) for the United States, Fehr and Goette (2005) for Switzerland, Kuroda and Yamamoto (2003) for Japan, and Castellanos, Garcia-Verdú, and Kaplan, (2004) for Mexico. Recent literature examines the implications of downward nominal wage rigidity for monetary policy (Kim and Ruge-Murcia, 2009 and 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 favor of downward nominal wage rigidity in the form of a positive and statistically significant coefficient of the asymmetry parameter ψ in (12).

The household is subject to the budget constraint

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

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 labor h by firms (see (8)). The optimal consumption of good i satisfies

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

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

2.3 Monetary Policy

The monetary authority (or “the Fed” for short) sets the interest rate following the Taylor-type rule

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

where $\rho_1 \in (-1, 1)$, ρ_2 and ρ_3 are constant parameters, variables without time subscript denote steady-state values, and ξ_t is a monetary shock that represents factors that affect the nominal interest rate beyond the control of the Fed.² We assume that ξ_t is i.i.d. with mean zero, skewness different from zero, and independent of the productivity innovation, ϵ_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 either of three possible distributions. These distributions are 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

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

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, μ is the location parameter, κ is the scale parameter, and ς is the shape parameter.

²In preliminary work, we considered a more general specification where the interest rate also responds directly to the productivity shock. However, its coefficient was quantitatively small and not statistically different from zero, while the other parameter estimates were similar to those reported here.

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 project because, as we will see below, the U.S. data prefer specifications where the skewness of the innovations is relatively large.³ 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, see Coles (2001), de Haan and Ana Ferreira (2006), and Embrechts, Klüppelberg, and Mikosch (2011).

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 labor, and produce the same quantity of output, that all households supply the same amount of labor and receive the same wages, and that net bond holdings are zero.

Equilibrium in the goods market implies the aggregate resource constraint,

$$c_t = y_t - (y_t \Gamma_t + w_t n_t \Phi_t), \quad (17)$$

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 (17)). It follows that optimal gross inflation should be 1, meaning that prices and wages are constant and, consequently, 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; but 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 behavior by the planner.

³In 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.

2.6 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.

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 between 1964Q2 to 2015Q4. 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 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). Civilian non-institutional population is defined as persons older than 15 years of age who are not inmates of institutions or on active duty in the Armed Forces. 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.⁴ 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.

⁴The 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.

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, $\hat{\theta}$, is the value that solves

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

where

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

$\mathbf{W}$ is a $q \times q$ weighting matrix, T is the sample size, λ 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. 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 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,

$$\sqrt{T}(\hat{\theta} - \theta) \rightarrow N(\mathbf{0}, (1 + 1/\lambda)(\mathbf{J}' \mathbf{W}^{-1} \mathbf{J})^{-1} \mathbf{J}' \mathbf{W}^{-1} \Sigma \mathbf{W}^{-1} \mathbf{J} (\mathbf{J}' \mathbf{W}^{-1} \mathbf{J})^{-1}), \quad (19)$$

where

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

and $\mathbf{J} = E(\partial m_{\iota}(\theta)/\partial \theta)$ is a finite Jacobian matrix of dimension $p \times q$ and full column rank.

In this application, the weighting matrix is the diagonal of the inverse of the matrix with the long-run variance of the moments. That is, $\mathbf{W}$ has diagonal entries equal to those of Σ^{-1} and non-diagonal entries equal to zero. This weighting matrix is attractive because it makes the objective function scale-free and gives a larger weight to the moments that are more precisely estimated. Σ is computed using the Newey-West estimator with a Barlett kernel and bandwidth given by the integer of $4(T/100)^{2/9}$ where $T = 208$ is the sample size. The number of simulated observations is twenty times larger than the sample size (i.e., $\lambda = 20$). In order to attenuate the effect of starting values on the results, the simulated sample contains 100 additional “training” observations that are discarded for the purpose of computing the moments. The dynamic simulations of the non-linear model are based on the pruned version of the solution. We use here the pruning scheme proposed by Andreasen, Fernández-Villaverde, and Rubio-Ramírez (2013), but using the unpruned solution delivers similar results as those reported.

The estimated parameters are the parameters of the adjustment cost functions for prices (γ and η) and wages (ϕ and ψ), the parameters of the monetary policy rule (ρ_1 , ρ_2 , and ρ_3), and the parameters of the distributions of productivity and monetary shocks. The moments used to estimate these parameters are the variances, covariances, autocovariances and skewness of consumption, hours worked, price inflation, wage inflation, and the nominal interest rate: 25 moments in total. During the estimation procedure the discount factor (β) is fixed to 0.995, which is close to the mean of the inverse ex-post real interest rate in the sample period. The parameters of the utility function are set to $\varrho = 1$, meaning that the utility of consumption is logarithmic, and $\chi = 0$, meaning that the disutility of labor is linear.⁵ The steady-state (gross) inflation target (Π) in the monetary policy rule is set to 1. 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 $v = 1.1$ and $\omega = 1.4$, respectively. This value for v is standard in the literature. Sensitivity analysis with respect to ω indicates that results are robust to using similarly plausible values. Finally, note that—in addition to the nonlinear model with GEV innovations—we also estimate another version of the nonlinear model where innovations are normally distributed.

The local identification of the model parameters requires that $\text{rank}(E(\partial m_i(\theta)/\partial \theta)) = q$, where 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 $\hat{\theta}$ for both versions of the model.

3.3 Parameter Estimates

Estimates of the parameters and 95% confidence intervals for the two versions of the model are reported in table 1. The confidence intervals were computed using a parametric bootstrap with 199 replications.⁶ Estimates of the price and wage rigidity parameters (ϕ and γ) are similar under both distributions and they are statistically different from zero in all cases. The adjustment cost parameter for wages (ϕ) is larger than the one for prices (γ), which suggests that wages are more rigid than prices. However, since the asymmetry parameter for wages (ψ) is large, positive, and statistically significant, wage rigidity is primarily downward rigidity, rather than upward rigidity. The large estimate of ψ means that the wage cost function has the shape of a smoothed “L” with the kink at a gross annual inflation rate of 0.998. Thus, wage increases and mild wage decreases of up

⁵The functional form of the utility function is based on preliminary estimates where $\varrho = 0.96$ (0.27) and $\chi = 2.21$ (4.38), with the figures in parenthesis denoting standard errors. Note that ϱ is not statistically different from 1 and χ is not statistically different from 0.

⁶We use bootstrap, rather than asymptotic, standard errors because Monte-Carlo results in Ruge-Murcia (2012) show that the latter are not always a good approximation to the actual variability of SMM estimates in small samples.

0.2% per year are essentially costless, but wage decreases larger than 0.2% per year are prohibitively costly. In contrast, the asymmetry parameter for prices (η) is two orders of magnitude smaller than for wages and it is not statistically different from zero at the 5% level. The fact that η is not different from zero means that we cannot reject the hypothesis that price adjustment costs are symmetric (quadratic as in Rotemberg, 1982) against the alternative that they are asymmetric (linear as in equation (5)).

For the model with GEV innovations, the scale and shape parameters imply that productivity innovations are negatively skewed with skewness equal to -0.58 . The result that the shape parameter is negative and statistically different from zero means that among extreme value distributions the one that best describes productivity innovations is the Weibull distribution. Figure 1 plots the estimated probability density function (PDF) (thick line) and compares it with the PDF of a normal distribution with the same standard deviation (thin line). Note that PDF of the GEV (Weibull) 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 column of table 1 and in figure 2. 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. 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 it 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 $\ln(\hat{z}_t) = \ln(\hat{z}_{t-1}) + \Delta \ln(z_t^f)$, where $\Delta \ln(z_t^f)$ 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 $\ln(z_t)$ in the model. Using this measure of $\ln(z_t)$, estimates of (4) were computed by OLS and estimates of the parameters of the GEV distribution of the residuals were computed by the method of maximum likelihood.

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 and support the conclusion that productivity innovations follow a Weibull distribution in that the shape parameter is negative and statistically different from zero. Figure 2 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 quantitatively smaller than, but still consistent with, the estimate of -0.58 obtained from the full model.

Regarding monetary policy, table 1 shows that under both versions of the model the smoothing parameter in the Taylor rule is moderately large and the coefficients of inflation and output are positive and statistically significant. The long-run response to inflation for the versions with GEV and normal innovations are 2.25 and 2.13, respectively, while the long-run response to output are 1.21 and 1.01, respectively. Hence, the systematic part of the reaction function is basically the same in both models. This means that, as far as monetary policy is concerned, the difference in the results across model is primarily due to the difference in innovation distributions. For the version with GEV innovations, the estimated scale and shape parameters imply that monetary policy innovations are positively skewed, with skewness equal to 3.18. The result that the shape parameter is positive and statistically different from zero means that among extreme value distributions the one that best describes monetary policy shocks is the Fréchet distribution. Figure 1 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 unlikely.

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 Fréchet distribution implied by estimates from the full model. However, in both cases the predicted skewness of monetary policy shocks is positive. (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.) Figure 2 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 smaller than, but still consistent with, the estimate of 3.18 obtained from the full model.

Figure 3 reports the fit for the two versions of the model by comparing actual and predicted moments. In the panels, the horizontal axis are the moments computed from U.S. data while the

vertical axis and dots are the moments predicted by the model. The straight line is the 45 degree line. If a model were to predict moments that fit perfectly those of the data, all dots would be on this line. We can see in this figure that the model with GEV innovations fit the data better than the model with normal innovations. This impression is statistically confirmed by two measures of fit, namely the root mean squared error (RMSE) and mean absolute error (MAE), which are also reported in the figure. Most of the difference comes from the fact that the model with normal innovations does not match well the unconditional skewness of the data. For instance, the model with normal innovations predicts much lower skewness for consumption and real wages than in the data and the corresponding dots are, therefore, relatively far from the 45 degree line. Overall, these results suggest that the nonlinear model with GEV innovations can account better for the non-Gaussian features of the data than the nonlinear model with normal innovations. Additional evidence to support this conclusion is reported below in section 4.1.

4. Implications of Extreme Events

This section examines the positive implications of extreme events for macroeconomic variables and monetary policy. Normative implications are derived in section 5 below.

4.1 Skewness

Table 2 reports estimates of the skewness of six key macroeconomic series, namely consumption, hours worked, the real wage, wage inflation, price inflation, and the nominal interest rate. The table also reports p -values of the Jarque-Bera test of the hypothesis that the data follow a normal distribution. Notice that 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 hypothesis that the data follow a normal distribution is rejected at the 5% level in all cases.

We now examine whether the economic model can account for this non-Gaussian feature of the data. To that end, we simulate an artificial sample of 4000 observations under each version model, compute the unconditional skewness of each series, and test the null-hypothesis of normality using the Jarque-Bera test. Consider first the model with normal innovations. Table 2 shows that, despite the fact that innovations are symmetric, this model can produce 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 other cases (real wage and wage inflation) the predicted skewness of opposite sign to that in the data.

Consider now the model with GEV innovations. Table 2 shows that the combination of non-linearity and asymmetric innovations deliver skewness that is quantitatively similar to that in the data. Notice that, except for consumption under the model with normal innovations, the hypothesis that the data are drawn from a normal distribution can be rejected at the 5% level.

4.2 Means

Table 3 reports the mean deviation from the deterministic steady state for consumption, hours worked, the real wage, wage inflation, price inflation, and the nominal interest rate. Since this mean is zero for linear models, the values reported in table 3 are a measure of the departure from certainty equivalence in our nonlinear model. For the model with GEV innovations the mean is calculated using the parameters reported in the first column of table 1. In addition, in order to quantify the contribution of skewness to the departure from certainty equivalence, table 3 reports the mean for a version of the model where innovations are normally distributed with standard deviation equal to that of the GEV distribution and all other parameters are those in the first column of table 1.

As expected, the mean of consumption is below the deterministic steady state because a prudent agent in a stochastic economy would consume less and save more than its counterpart in a certainty-equivalent economy. Comparing the two columns in table 3 shows that shock asymmetry induces a further reduction in mean consumption from -0.014% below the steady-state value to -0.06% below. The reason is simply that in the latter case agents are subject to an additional source of risk: skewness risk or the possibility of large draws from the long tails of the asymmetric distributions of productivity and monetary shocks, both of which reduce consumption (see below). The contribution of skewness risk to the departure from certainty equivalence is also substantial for hours worked and the real wage. The interest rate is below its deterministic steady state because as agents try to save, in the aggregate they will push bond prices up and yields down.

The gross price and wage inflation rates are above their deterministic steady value of 1 and imply annual net inflation rates of about 0.08% . This result is also driven by prudence as agents would like to avoid the large costs associated with downward nominal wage adjustments. However, as we discuss below, the price and wage adjustment costs that agents must systematically pay when gross inflation is above 1 moderate this prudence motive and, hence, the mean reported in table 3 is relatively small.

4.3 Asymmetric Responses

We study the response of the economy to productivity and monetary shocks using impulse-response analysis. 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 5th and 95th percentiles. 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. Regarding timing, we assume that shocks occur when all variables are equal to the unconditional mean of their ergodic distribution. As we will see below, 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 4 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 that is persistent as a result of intertemporal smoothing. Hours worked decrease on impact and, following a hump, return to their unconditional mean from below. Price inflation and the nominal interest rate decrease, in the latter case because the inflation coefficient in the Taylor rule is quantitatively larger than that of output. Finally, the nominal wage 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 price decrease induces an increase in the real wage. Note that due to the strong wage rigidity, the response of wage inflation is muted and wages decreases are above the -0.2% below which adjustment costs are exceedingly high. Qualitatively, the effects of the negative shock in the 5th percentile are the opposite to those just described. The key observation in figure 4 is that the effects of the negative shock are much larger those of the equally-likely positive shock. This result is partly due the fact that the size of these two innovations is different for the asymmetric GEV distribution: the negative innovation takes productivity -2.88 percentage points below the steady state, while the positive innovation takes it 2.30 percentage points above. However, since this difference is relatively small for the two percentiles considered, the asymmetric responses are primarily due to the nonlinearity of the model.

Figure 5 plots the responses to a monetary shock. The positive shock raises the interest rate and induces a decrease in consumption, hours, and price inflation, and an increase in wage inflation and the real wage. In this case, the quantitative response of price and wage inflation is similar, but since they move in different directions, the increase in the real wage is unambiguous and relatively

large. The negative shock has the converse effects, but, as before, the key feature of this figure is the asymmetry in the responses to monetary shocks: the positive shock induces much larger responses than the equally-likely negative shock. Compared with figure 4, however, the fact that GEV distribution of monetary shocks has a large skewness (3.18) means that the size of the two innovations considered is substantially different: the positive innovation takes the interest rate 0.74 percentage points above the steady state, while the negative innovation takes it -0.42 percentage points below. Thus, in this case both the shock asymmetry and the model nonlinear account for the asymmetric responses reported in figure 5.

5. 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

$$E_s \sum_{t=s}^{\infty} \beta^{t-s} U(c_t, n_t),$$

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. 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. We focus on the case where productivity innovations are drawn from a GEV distribution.

5.1 Decision Rules

Figure 6 plots the decision rules that solve the model when policy is implemented by the Ramsey planner. In this figure, the horizontal axis is the size of the productivity shock normalized by its standard deviation, and the vertical axis is the percentage deviation from the deterministic steady state. 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 approximation.

For consumption and hours worked the nonlinear decision rules are concave and imply larger changes in these variables than the linear rule when the economy is hit by a large negative productivity shock. In contrast, for price inflation, wage inflation, and the real wage, the nonlinear rules imply smaller changes than the linear rule when the economy is hit by a large negative productivity shock. However, the departure from linearity is limited.

For the nominal interest rate the nonlinear policy rule is convex and there is a large departure from the linear rule, especially in the case of large negative shocks. In particular, the nonlinear rule implies larger interest rate adjustment for negative than for positive productivity shocks, and very large adjustments in response to large negative productivity shocks.

5.2 Impulse Responses

Figure 7 plots the responses of the economy to productivity shocks. 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 in hours worked as agents take advantage of their temporarily high productivity. Wage inflation increases and price inflation decreases leading to an increase in the real wage. The nominal interest rate decreases. Compared with the response under the Taylor-rule policy (see figure 4), the decrease in the interest rate is much larger: -0.5% under the Ramsey policy compared with -0.3% under the Taylor-rule policy.

The negative shock induces the converse effects, but their magnitudes are larger than for the positive shock. As discussed in section 4.3, the size of the productivity innovation at the 5th percentile only somewhat larger than that at the 95th percentile and, thus, the asymmetry in the responses reported in figure 7 is primarily due to the nonlinearity of the model

5.3 Optimal Inflation

We measure of the optimal inflation rate by the mean of the ergodic distribution of inflation under the Ramsey policy. For the parameter estimates reported in the first column of table 1, mean annual gross inflation computed from a simulation of 4000 observation is 1.000002 with a 95% confidence interval spanning $[0.99999, 1.00004]$. Since this confidence interval includes 1, the hypothesis that optimal gross inflation is 1 cannot be rejected at the 5% significance level. This means that optimal expected inflation is statistically 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 realized 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, Gorodnichenko, and Wieland (2012) who find in a calibrated model that for costly, but infrequent, episodes at the zero lower bound on interest rates, the optimal inflation rate is low.

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, and ii) taking a wait-and-see approach. An example of the first strategy is increasing the inflation target, as recently proposed by Blanchard, Dell’Ariccia, and Mauro (2010) who 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 because, on one hand, prudence induces the policy maker 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 of the wait-and-see approach.

6. The Optimal Inflation Target under Strict Targeting

This section computes the rate of inflation that would maximize unconditional welfare when the monetary authority follows a policy of strict inflation targeting. Figure 8 plots unconditional welfare, expressed in consumption equivalents, for different values of the inflation target. First, note that downward nominal wage rigidity implies that deflation entails a substantial welfare loss. Second, welfare is approximately the same for gross inflation targets between 1 (strict price stability) and 1.02, which is the value targeted by the Federal Reserve and other central banks. Finally, the optimal gross inflation target is 1.0001, meaning a net rate of 0.04% per year. This is basically the same optimal inflation rate as under the Ramsey policy.

This result is interesting because under strict inflation targeting the policy maker has limited knowledge and less flexibility with respect to shocks compared with the Ramsey planner. This would suggest a larger buffer above zero net inflation compared with the Ramsey policy in order to avoid paying the cost associated with nominal wage cuts, should an extreme negative productivity shock hit the economy. However, as under the Ramsey policy, the inflation targeter trade-off between prudence and systematic price and wage adjustments costs results in the policy maker adopting a policy of strict price stability, understood to be inflation very close to zero, where these adjustment costs are small.

7. 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 incorporating systematically 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/if a extreme negative shocks hits the economy. We evaluate the welfare implication of these two approaches and find for our estimated model, this trade-off is solved unambiguously in favor of the wait-and-see approach. The intuition is simple: the cost of price and wage adjustment required under the prudent policy in every period override the potential benefits of targeting a gross inflation inflation in the expectation of large, but infrequent, extreme negative shock. As a result, under both the Ramsey policy and a strict-inflation-targeting policy, the optimal gross inflation rate is virtually indifferent from 1.

Table 1: Parameter Estimates

	Notation	Distribution		Single Equation
		GEV	Normal	
		(1)	(2)	(3)
Nominal Rigidity				
Wage adjustment cost	ϕ	383.48 (79.96, 6209.83)	394.66 (78.76, 5748.11)	
Wage asymmetry $\times 10^{-2}$	ψ	237.75 (42.40, 8706.97)	275.36 (47.52, 7651.46)	
Price adjustment cost	γ	143.43 (73.71, 195.17)	130.65 (54.46, 249.55)	
Price asymmetry	η	-559.59 (-821.04, 4.13)	-684.60 (-2941.68, 201.89)	
Productivity				
Autoregressive coeff.	ζ	0.849 (0.766, 0.867)	0.918 (0.844, 0.958)	0.934 (0.893, 0.976)
Scale $\times 10$	κ_ϵ	0.171 (0.121, 0.225)	0.102 (0.054, 0.157)	0.076 (0.069, 0.084)
Shape	ς_ϵ	-0.480 (-0.601, -0.130)		-0.285 (-0.350, -0.219)
Monetary Policy				
Smoothing	ρ_1	0.870 (0.782, 0.999)	0.908 (0.786, 0.990)	0.921 (0.882, 0.960)
Inflation	ρ_2	0.293 (0.185, 0.466)	0.196 (0.134, 0.311)	0.123 (0.080, 0.166)
Output	ρ_3	0.157 (0.095, 0.278)	0.093 (0.061, 0.191)	0.052 (0.033, 0.071)
Scale $\times 10^2$	κ_ξ	0.233 (0.097, 0.480)	0.229 (0.054, 0.424)	0.196 (0.180, 0.214)
Shape	ς_ξ	0.185 (-0.405, 0.310)		-0.117 (-0.144, -0.090)

Note: The table reports SMM estimates of the model parameters under each distribution. The figures in parenthesis are 95% confidence intervals computed using a parametric bootstrap with 199 replications. In the case of the normal distribution, the scale parameter is the standard deviation.

Table 2: Skewness and Jarque-Bera Testss

	U.S.	Distribution	
	Data	GEV	Normal
	(1)	(2)	(3)
Skewness			
Consumption	-0.707	-0.323	-0.044
Hours	-0.711	-0.707	-0.363
Real Wage	-0.334	-0.296	0.204
Wage inflation	-0.525	-0.795	2.597
Price inflation	0.679	0.541	0.937
Nominal interest rate	0.587	1.619	1.147
Jarque-Bera Tests			
Consumption	0.004	< 0.001	0.075
Hours	0.002	< 0.001	< 0.001
Real wage	0.046	< 0.001	< 0.001
Wage inflation	< 0.001	< 0.001	< 0.001
Price inflation	< 0.001	< 0.001	< 0.001
Nominal interest rate	0.006	< 0.001	< 0.001

Note: The table reports the unconditional skewness of actual U.S. series and of artificial data simulated from the model, and the p -values of the Jarque-Bera test of the hypothesis that the data follows a normal distribution.

Table 3: Mean Deviation from Deterministic Steady State

	Distribution	
	Normal with same	
	GEV	SD as GEV
	(1)	(2)
Consumption	−0.060	−0.014
Hours	−0.091	−0.021
Real Wage	−0.537	−1.039
Wage inflation	0.019	−0.036
Price inflation	0.019	−0.036
Nominal interest rate	−0.067	−0.106

Note: The table reports the mean deviation from the deterministic steady state value of each variable expressed in percent. SD is standard deviation.

References

- [1] Abbritti, M., Fahr, S., 2013. Downward Wage Rigidity and Business Cycle Asymmetries. *Journal of Monetary Economics* 60, pp. 871-886
- [2] Akerlof, G. A., Dickens, W. T., Perry, G. L., 1996. The Macroeconomics of Low Inflation. *Brookings Papers on Economic Activity*, pp. 1-76.
- [3] 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. University of Pennsylvania, unpublished manuscript.
- [4] Benigno, P., Ricci, L. A., 2011. The Inflation-Unemployment Trade-Off at Low Inflation. *American Economic Review* 101, pp. 1436-1466.
- [5] Blanchard, O., Dell’Ariccia, G., Mauro, P., 2010. Rethinking Macroeconomic Policy. IMF Staff Position Note SPN/10/03.
- [6] 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.
- [7] 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, pp.507-533.
- [8] 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, pp. 1371-1406.
- [9] Coles, S., 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer: London.
- [10] de Haan, L., Ferreira, A., 2006. *Extreme Value Theory: An Introduction*. Springer: London.
- [11] Duffie, D., Singleton, K. J., 1993. Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica* 61, pp. 929-952.
- [12] Embrechts, P., Klüppelberg, C., Mikosch, T., 2011. *Modelling Extremal Events for Insurance and Finance*. Springer: London.
- [13] Fehr, E., Goette, L., 2005. Robustness and Real Consequences of Nominal Wage Rigidity. *Journal of Monetary Economics* 52, pp. 779-804.

- [14] Fernald, J. G., 2014, A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. Federal Reserve Bank of San Francisco Working Paper 2012-19.
- [15] Fisher, R.A., Tippett, L. H. C., 1928. On the Estimation of the Frequency Distributions of the Largest or the Smallest Member of a Sample. *Proceedings of the Cambridge Philosophical Society* 24, pp. 180-190.
- [16] Gallant, A. R., Rossi, P. E., Tauchen, G., 1993. Nonlinear Dynamic Structures. *Econometrica* 61, pp. 871-908.
- [17] Gnedenko, B. V., 1943. Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics* 44, pp. 423-453.
- [18] 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, 158-171.
- [19] Jin, H., Judd, K. L., 2002. Perturbation Methods for General Dynamic Stochastic Models. Hoover Institution, Mimeo.
- [20] Keynes, J. M., 1936. *The General Theory of Employment, Interest and Money*. Macmillan Cambridge University Press: Cambridge.
- [21] Kim, J., Ruge-Murcia, F., 2009. How Much Inflation is Necessary to Grease the Wheels?. *Journal of Monetary Economics* 56, pp. 365-377.
- [22] Kim, J., Ruge-Murcia, F., 2011. Monetary Policy when Wages are Downwardly Rigid: Friedman meets Tobin. *Journal of Economic Dynamics and Control* 35, pp. 2064-2077.
- [23] Koop, G., Pesaran, M. H., Potter, S., 1996. Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics* 74, pp. 119-147.
- [24] Kuroda, S., Yamamoto, I., 2003. Are Japanese Nominal Wages Downwardly Rigid? (Part I): Examination of Nominal Wage Change Distributions. *Monetary and Economic Studies* 21 (August), pp. 1-30.
- [25] Lee, B.-S., Ingram, B. F., 1991. Simulation Estimation of Time-Series Models. *Journal of Econometrics* 47, pp. 195-205.
- [26] McLaughlin, K. J., 1994. Rigid Wages? *Journal of Monetary Economics* 34, pp. 383-414.

- [27] Peltzman. S., 2000. Price Rise Faster than They Fall. *Journal of Political Economy* 108, pp. 466-502.
- [28] Rotemberg, J. J., 1982. Sticky Prices in the United States. *Journal of Political Economy* 90, pp. 1187-1211.
- [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, pp. 914-938.
- [30] Schmitt-Grohe, S., Uribe, M., 2016. Downward Nominal Wage Rigidity, Currency Pegs, and Involuntary Unemployment, *Journal of Political Economy*, forthcoming.
- [31] Svensson, L. E. O., 2003. Optimal Policy with Low-Probability Extreme Events. NBER Working Paper No. 10196.
- [32] 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.
- [33] 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, pp. 514-533

Figure 1: Estimated Probability Density Functions of Innovations

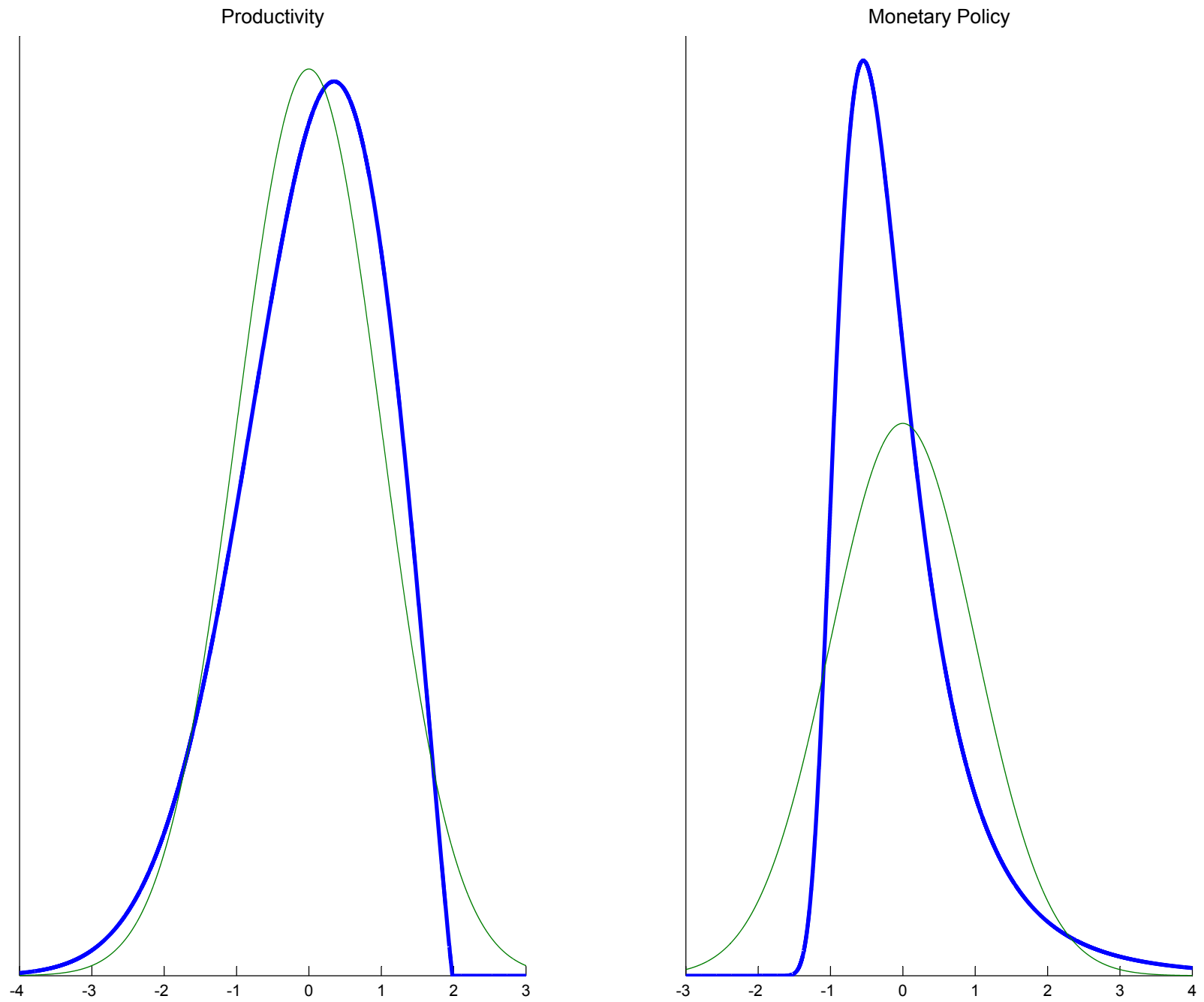


Figure 2: Histogram of Residuals Computed from the U.S. Data

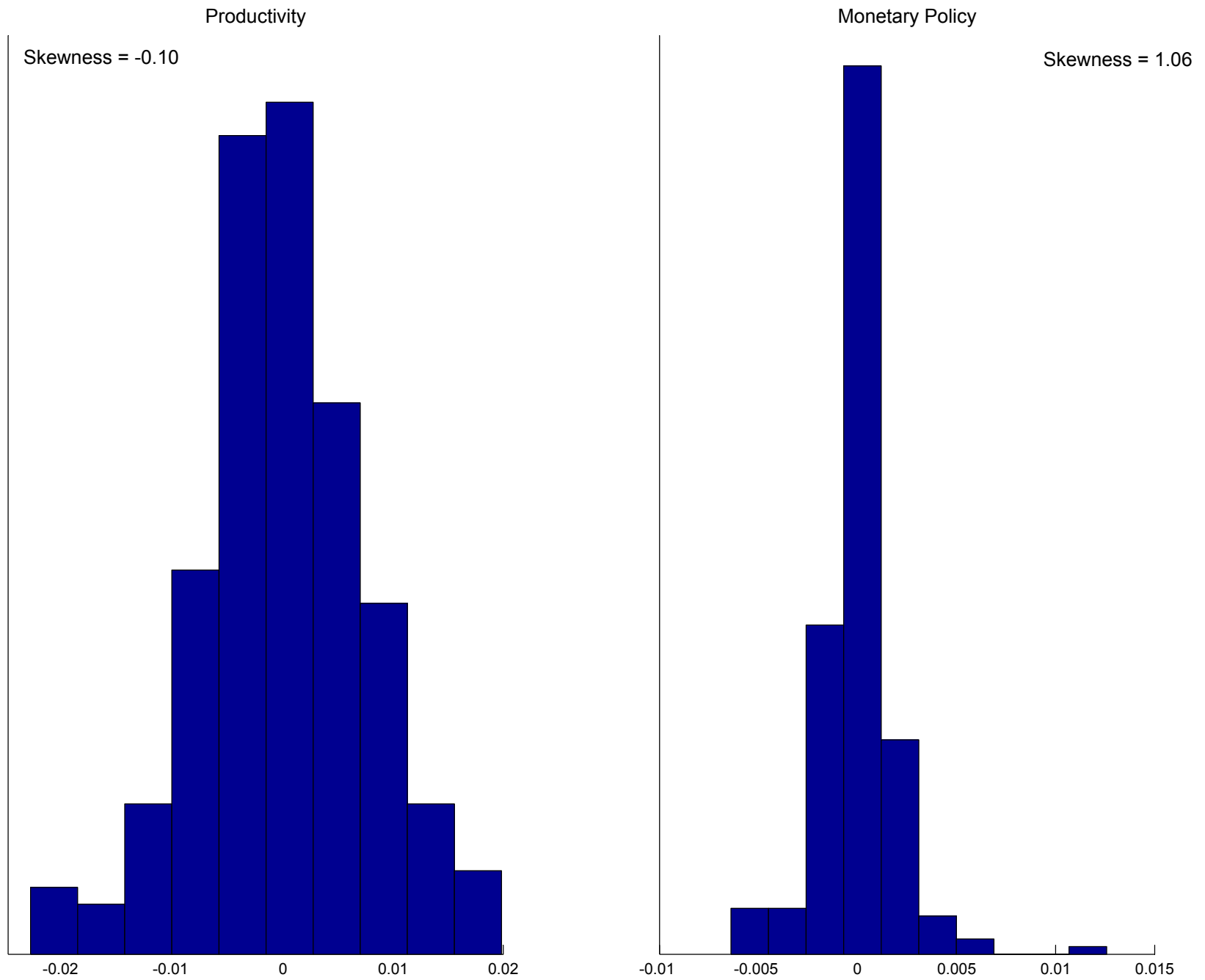


Figure 3: Model Fit

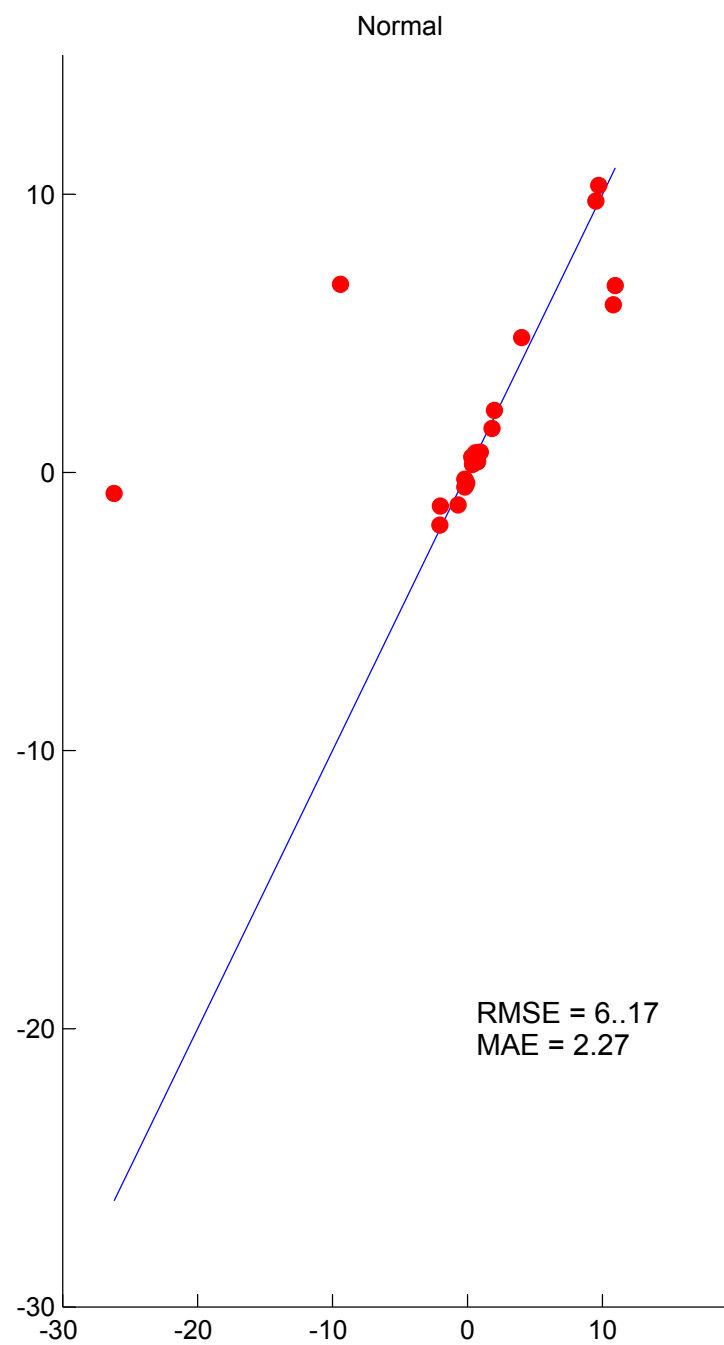
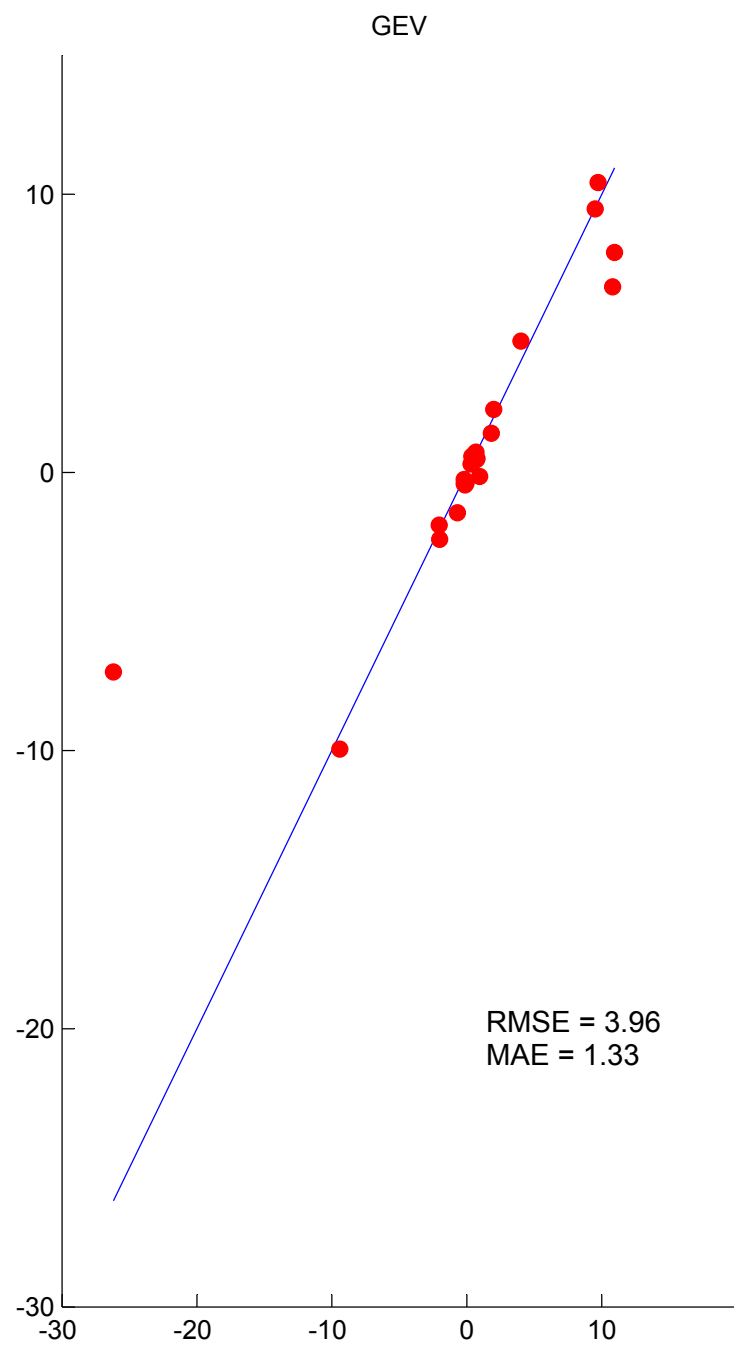


Figure 4: Responses to a Productivity Shock under Taylor Rule Policy

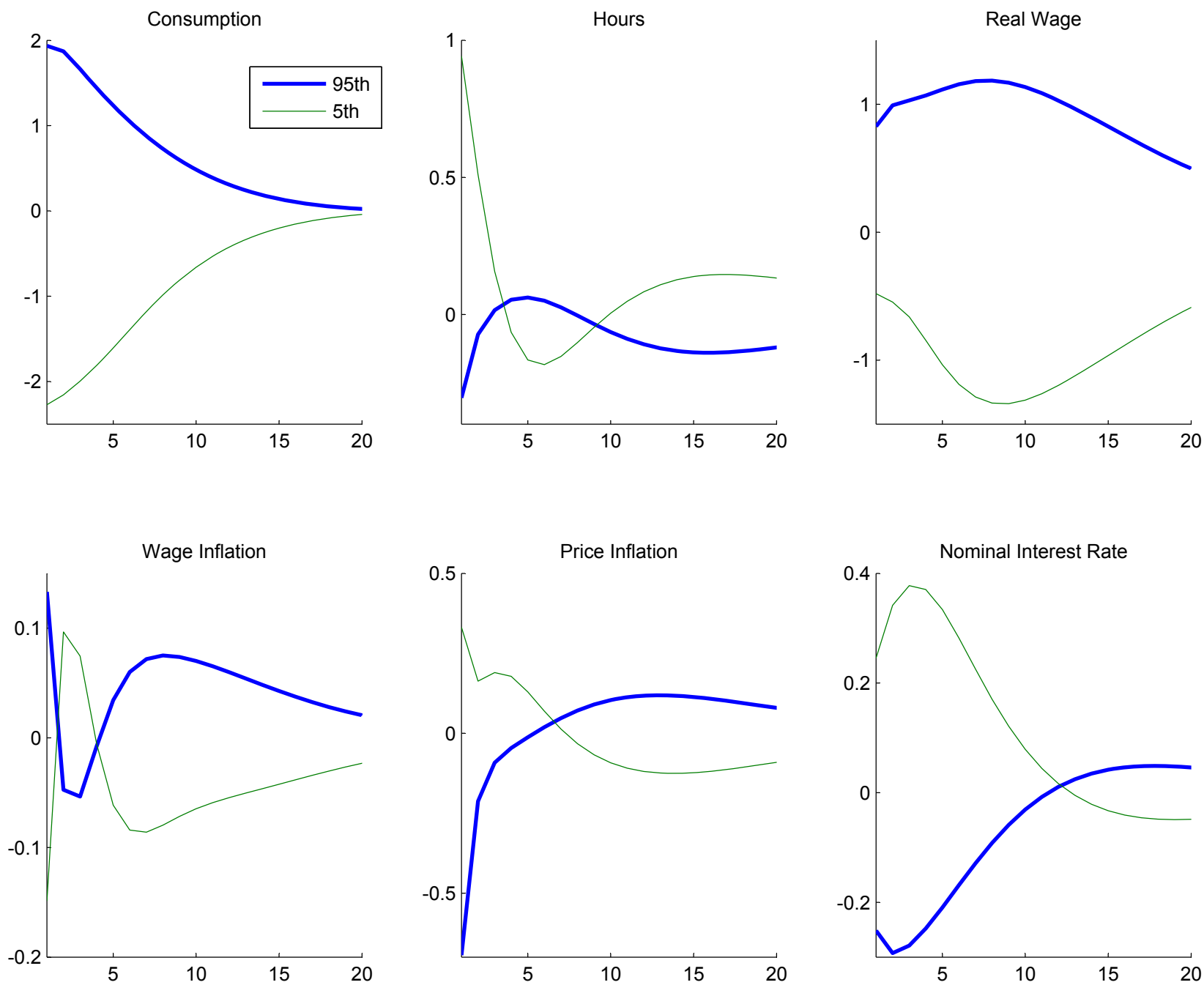


Figure 5: Responses to a Monetary Shock under Taylor Rule Policy

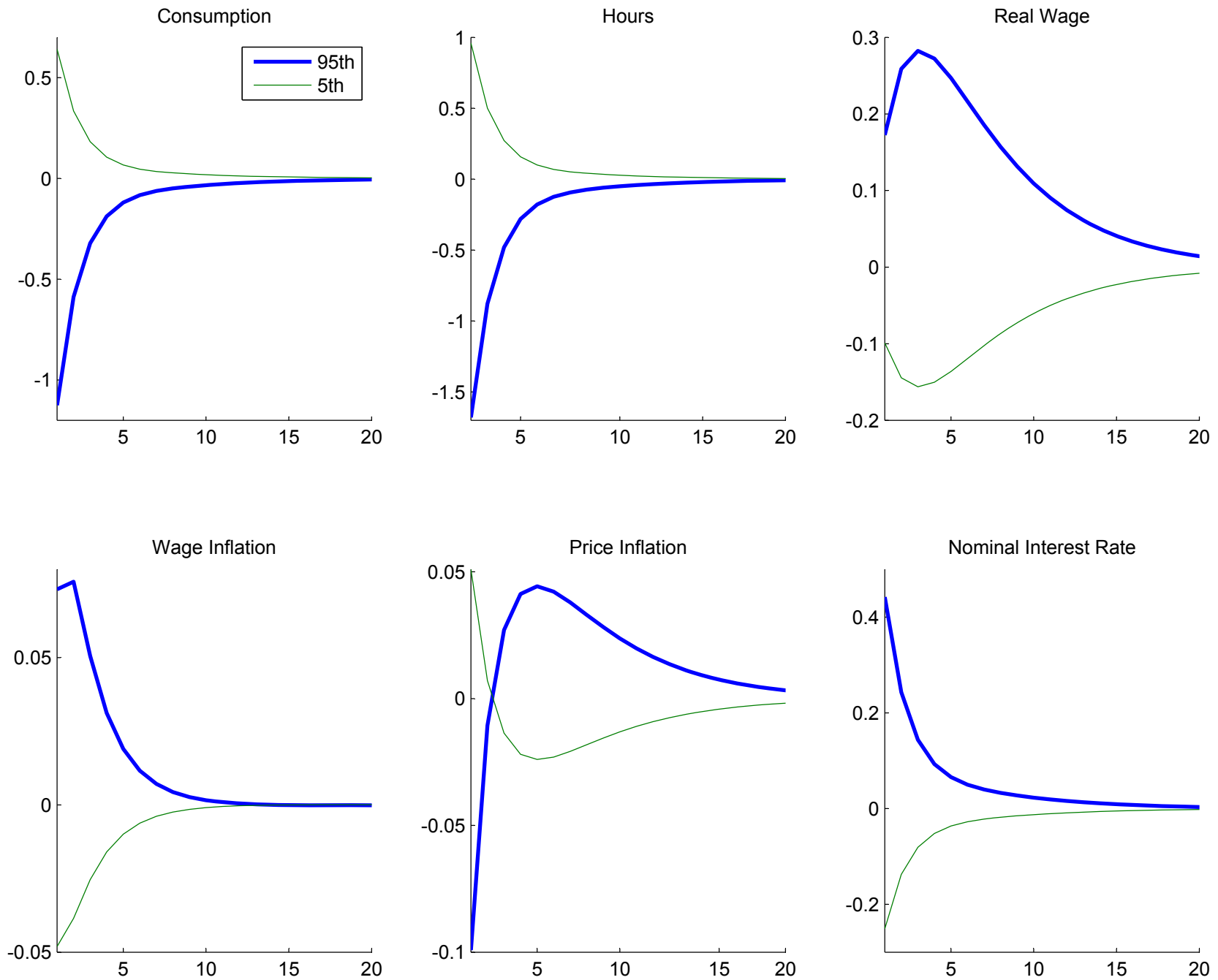


Figure 6: Decision Rules of the Ramsey Planner

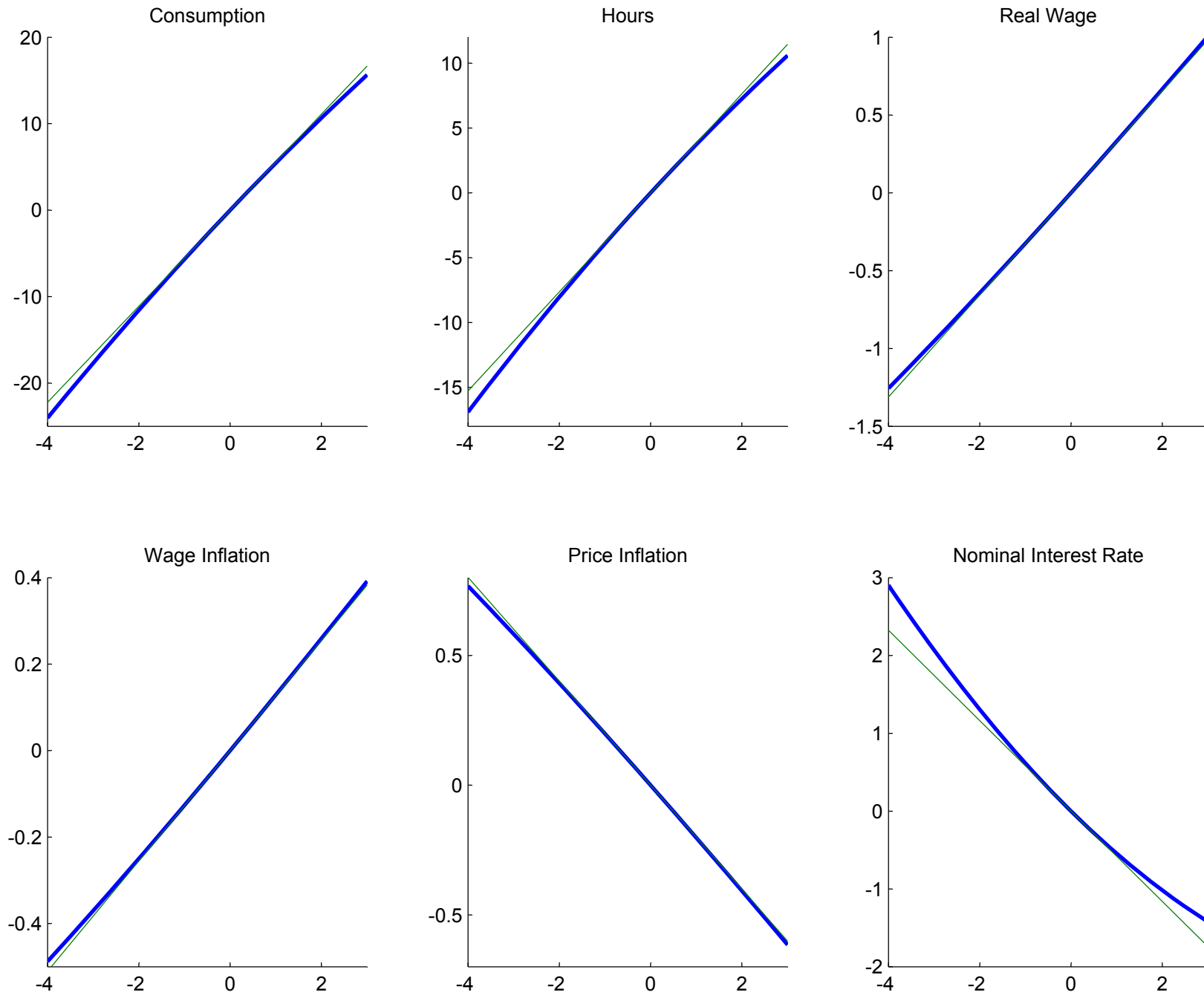


Figure 7: Responses to a Productivity Shock under Ramsey Policy

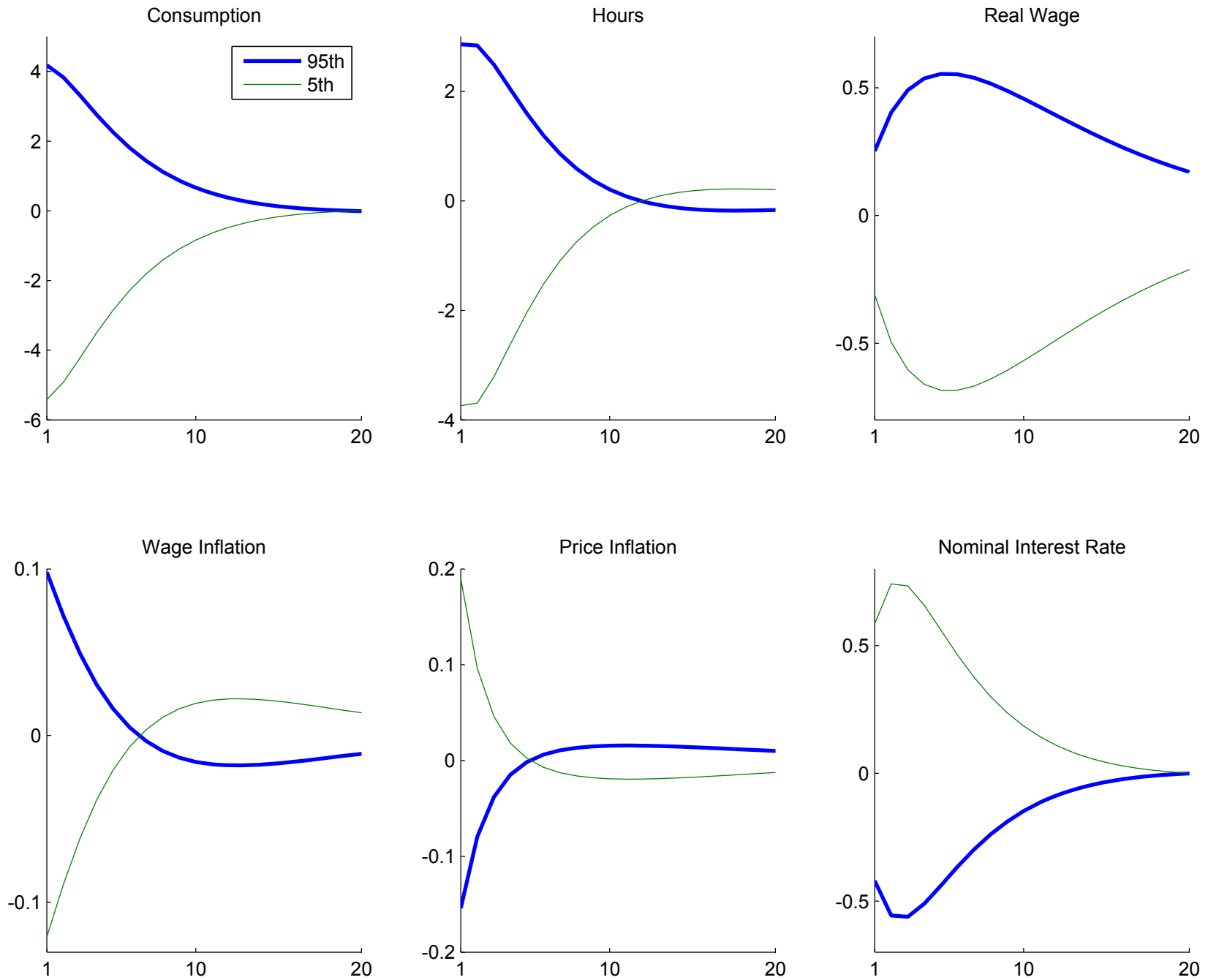


Figure 8: Unconditional Welfare

