Assessing DSGE Model Nonlinearities

S. Borağan Aruoba[*]   Luigi Bocola   Frank Schorfheide

University of Maryland   *University of*   *University of*

and Visiting Scholar   *Pennsylvania*   *Pennsylvania, NBER,*

Federal Reserve Bank   *and Visiting Scholar*

of Philadelphia   *Federal Reserve Bank*

of Philadelphia

November 26, 2013

Abstract

We develop a new class of nonlinear time-series models to identify nonlinearities in the data and to evaluate nonlinear DSGE models. U.S. output growth and the federal funds rate display nonlinear conditional mean dynamics, while inflation and nominal wage growth feature conditional heteroskedasticity. We estimate a DSGE model with asymmetric wage/price adjustment costs and use predictive checks to assess its ability to account for nonlinearities. While it is able to match the nonlinear inflation and wage dynamics, thanks to the estimated downward wage/price rigidities, these do not spill over to output growth or the interest rate. (JEL C11, C32, C52, E32)

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are now widely used for empirical research in macroeconomics, as well as for forecasting and quantitative policy analysis in central banks. In these models, decision rules of economic agents are derived from assumptions about agents' preferences and production technologies utilizing some fundamental principles such as optimization, rational expectations, and competitive equilibrium. In practice, this means that the functional forms and parameters of equations that describe the behavior of economic agents are tightly restricted by the equilibrium conditions. Consequently, a careful evaluation of the DSGE model-implied restrictions is an important aspect of empirical research.

Until recently, much of the research that estimates DSGE models used first-order approximations to the equilibrium decision rules. This made linear models such as vector autoregressions (VARs) appropriate for evaluating the restrictions of the DSGE model. With the advance of methods to estimate DSGE models using higher-order approximations, as developed in Fernández-Villaverde and Rubio-Ramírez (2007), an important avenue of research has opened. The end of the Great Moderation also makes nonlinear models all the more relevant for empirical macroeconomics.[1] While there is a burgeoning literature on both the methods to solve nonlinear DSGE models and their applications, there does not seem to be an obvious nonlinear time-series model in use to evaluate these DSGE models.

The objective of this paper is to develop a class of time-series models that mimic

---

[1]There are in principle two types of nonlinearities that can appear in a nonlinear DSGE model. First are (approximately) smooth nonlinearities, where decision rules display curvature and possibly asymmetries such as those that are generated by asymmetric loss or cost functions. Second are kinks in decision rules such as those that are generated by the zero lower bound on nominal interest rates. This paper is about the former. While the latter is also crucial and we intend to extend our work to address this type of nonlinearities, solving and estimating DSGE models with kinks in decision rules is very difficult. See, for example, Gust, Lopez-Salido, and Smith (2012), Aruoba and Schorfheide (2013), Bocola (2013), and references therein.

nonlinearities of DSGE models and to use these models as a benchmark for the evaluation of a nonlinear DSGE model. Motivated by the popular second-order perturbation approximations of DSGE model dynamics, we consider autoregressive models that involve quadratic terms of lagged endogenous variables as well as interactions between current-period innovations and lagged endogenous variables, which generate conditional heteroskedasticity. These time-series models are derived from a perturbation solution to a nonlinear difference equation and have a recursively linear structure that makes it straightforward to characterize stability properties and derive moments. While multivariate extensions are possible, we focus in this paper on univariate specifications, which we refer to as QAR(p,q) models, where "Q" stands for quadratic.[2] In the empirical work, we use $p = q = 1$.

After documenting some of the theoretical properties of the QAR models, the first step of the empirical analysis is to fit QAR(1,1) models to growth of real gross domestic product (GDP), inflation, nominal wage growth, and interest rate data for the U.S. We start our sample in 1960 but consider various subsamples, using 1983 (the end of the Volcker era and the start of the Great Moderation) and 2007 (the end of the Great Moderation and the start of the Great Recession) as additional start and end points. We find three sets of important nonlinearities across the variables and samples we consider. First, GDP growth displays pronounced nonlinearities in the post-1983 samples with sharp output losses during recessions and relatively slow recoveries. Second, for inflation and wage growth the long samples that start in 1960 and extend beyond the 1990s exhibit high volatility in times of high inflation and wage growth, which is mainly driven by the observations in the 1970s. Finally, QAR estimates for interest rates imply an asymmetric behavior by the Federal Reserve, cutting interest rates faster but not increasing them as fast in the post-1983 era.

The second step of the empirical analysis consists of the estimation of a DSGE model. In our application we focus on the estimation and evaluation of a New Keyne-

---

[2]The abbreviation QAR has previously been used for quantile autoregressions; see Koenker and Xiao (2006).

sian DSGE model with asymmetric price and wage adjustment costs, building on work by Kim and Ruge-Murcia (2009). This model can generate downward nominal wage and price rigidity and is interesting for several reasons. First, it is well known that in the absence of the zero-lower-bound (ZLB) constraint on nominal interest rates, unrealistically large shocks, or high degrees of risk aversion, New Keynesian DSGE models do not generate significant nonlinearities (see, for instance, An (2007)). However, once one allows for asymmetric adjustment costs, agents' decision rules can become strongly nonlinear. Thus, ex ante, to the extent that there are nonlinearities in the data, the model may be able to deliver some of these.

Second, downward rigidity is a well-documented feature of nominal wage changes at the micro level, e.g., Gottschalk (2005), Barattieri, Basu, and Gottschalk (2010), and Daly, Hobijn, and Lucking (2012). Third, a number of papers have incorporated downward nominal wage rigidity into DSGE models to study its macroeconomic effects. For instance, Kim and Ruge-Murcia (2009) study optimal monetary policy in the presence of downward nominal wage rigidity. Schmitt-Grohe and Uribe (2013) use downward nominal wage rigidity to generate large output losses and a jobless recovery in a deflation (or liquidity-trap) equilibrium of a New Keynesian model with ZLB constraint. Thus, a careful evaluation of the nonlinearities that this mechanism generates is important.

In estimating the DSGE model, we use the same data set as in the estimation of the univariate QAR models and consider two samples, one long and one short, both of which end in 2007 to avoid using data where the ZLB starts to bind. By and large, the parameter estimates for the DSGE models are consistent with estimates that have been reported elsewhere in the literature. In particular, our estimates indicate asymmetries in the adjustment costs for both prices and nominal wages that make increases less costly than decreases.

The final and most important step of the analysis is to conduct a posterior predictive check of the DSGE model that compares coefficient estimates obtained from

data simulated from an estimated DSGE model to coefficient estimates obtained from actual data. The predictive check amounts to assessing how far the QAR estimates obtained from the actual data lie in the tails of the predictive distribution. The general conclusion is that the DSGE model does not generate very strong nonlinearities except for inflation and nominal wage growth, both of which show conditional heteroskedasticity. This means that the asymmetric adjustment costs in prices and wages are able to deliver asymmetric behavior in inflation and nominal wage growth in line with the data, but this asymmetry does not spill over to real GDP growth or to the policy instrument of the Federal Reserve.

Our work is related to several branches of the literature. There exists a large body of work on nonlinear time-series models.[3] However, none of these model classes seem to be directly usable for our purposes since either the nonlinearities do not match the nonlinearities of DSGE models solved with higher-order perturbation methods, or the models have undesirable instability properties.

The proposed QAR family is most closely related to generalized autoregressions (GAR) discussed in Mittnik (1990) in the sense that the conditional mean of the dependent variable y_t is a polynomial function of its lags. Our QAR models also involve interactions between lagged dependent variables y_{t-j} and innovations u_t, which is a defining property of bilinear models and linear autoregressive conditional heteroskedasticity (LARCH) models, e.g., those in Giraitis, Robinson, and Surgailis (2000). However, rather than simply augmenting a linear autoregressive model by quadratic terms and interactions between lagged endogenous variables and innovations, we derive its structure from a second-order perturbation approximation to the solution of a nonlinear difference equation along the lines of Holmes (1995). To the extent that both conditional mean and variance depend on quadratic functions of

---

[3]These include regime-switching models, e.g., Hamilton (1989) and Sims and Zha (2006), time-varying coefficient models, e.g., Cogley and Sargent (2002) and Primiceri (2005), threshold and smooth-transition autoregressive models, e.g., Tong and Lim (1980) and Teräsvirta (1994), and bilinear models, e.g., Granger and Andersen (1978) and Rao (1981).

the innovations u_t our model is also related to the class of (G)ARCH-M models, e.g., those in Engle, Lilien, and Robins (1987) and Grier and Perry (1996). Finally, the QAR model can be viewed as a set of tight restrictions on the coefficients of a Volterra (1930) representation of a nonlinear time series.

There exists an abundant literature that develops methods to evaluate DSGE models based on comparisons with more flexible and densely parameterized time-series models. However, much of the existing econometric work is based on linearized DSGE models. A natural benchmark for the evaluation of such models is provided by vector autoregressions (VARs) that relax the cross-coefficient restrictions. In fact, there exists an extensive literature that develops and applies methods to evaluate DSGE models based on comparisons with VARs, e.g., Cogley and Nason (1994), Schorfheide (2000), Christiano, Eichenbaum, and Evans (2005), Del Negro, Schorfheide, Smets, and Wouters (2007), and Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007).

In this paper we use so-called posterior predictive checks to evaluate a prototypical DSGE model. A general discussion of the role of predictive checks in Bayesian analysis can be found in Lancaster (2004) and Geweke (2005). Canova (1994) is the first paper that uses predictive checks to assess implications of a DSGE model. While Canova (1994)'s checks were based on the prior predictive distribution, we use posterior predictive checks in this paper as, for instance, in Chang, Doh, and Schorfheide (2007). Finally, Abbritti and Fahr (2013) use a model with asymmetric wage adjustment costs and search and matching frictions to investigate the ability of the model to deliver nonlinearities, focusing on skewness and turning-point statistics.

The remainder of the paper is organized as follows. In Section 2 we review the structure of second-order perturbation approximations of DSGE models. The QAR model is developed in Section 3, where we discuss some of its theoretical properties as well as Bayesian inference. Estimates of the QAR model for U.S. data are presented in Section 4. The DSGE model with asymmetric price and wage adjustment costs

is introduced in Section 5. The estimation and evaluation of the DSGE model is presented in Section 6 and Section 7 concludes. An online appendix contains detailed derivations of the properties of the QAR model, as well as details about the Markov chain Monte Carlo (MCMC) methods employed in this paper.

2 DSGE Model Nonlinearities

Most estimated nonlinear DSGE models are solved with perturbation methods because they can be efficiently applied to models with a large state space. A DSGE model solved by second-order perturbation can be generically written as

$$
\begin{aligned}
c_{i,t} &= \psi_{1i}(\theta) + \psi_{2ij}(\theta)x_{j,t} + \psi_{3ij}(\theta)z_{j,t} \\
&\quad + \psi_{4ijk}(\theta)x_{j,t}x_{k,t} + \psi_{5ijk}(\theta)x_{j,t}z_{k,t} + \psi_{6ijk}(\theta)z_{j,t}z_{k,t} \\
x_{i,t+1} &= \zeta_{1i}(\theta) + \zeta_{2ij}(\theta)x_{j,t} + \zeta_{2ij}(\theta)z_{j,t} \\
&\quad + \zeta_{4ijk}(\theta)x_{j,t}x_{k,t} + \zeta_{5ijk}(\theta)x_{j,t}z_{k,t} + \zeta_{6ijk}(\theta)z_{j,t}z_{k,t} \\
z_{i,t+1} &= \xi_{2ij}(\theta)z_{j,t} + \xi_{3i}(\theta)\epsilon_{i,t+1},
\end{aligned}
\tag{1}
$$

where θ denotes the parameters of the model and the DSGE model variables are grouped into control variables $c_{i,t}$, e.g., consumption; endogenous state variables $x_{i,t}$, e.g., the capital stock; and exogenous state variables $z_{i,t}$, e.g., total factor productivity. The notation $a_{ijk}x_{j,t}x_{k,t}$ in (1) is shorthand for $\sum_{j=1}^{n}\sum_{k=1}^{n}a_{ijk}x_{j,t}x_{k,t}$. Since not all of the control and state variables are observable, it is common to augment the system by a measurement equation of the form

$$
y_{i,t} = A_{1i}(\theta) + A_{2ij}(\theta)c_{j,t} + A_{3ij}(\theta)x_{j,t} + A_{4ij}(\theta)z_{j,t} + e_{i,t},
\tag{2}
$$

where the $e_{i,t}$'s are measurement errors. Typically, the vector of observables y_t is composed of a subset of the state and control variables such that the A matrices are simple selection matrices.

Nonlinear features may arise endogenously or exogenously. Curvature in utility functions, adjustment cost function, and production functions can generate nonlinear

decision rules of households and firms endogenously. An example of an exogenous nonlinearity is stochastic volatility in the exogenous shocks that generate business cycle fluctuations. In (1) the endogenous nonlinearity is captured by the quadratic functions of x_t and z_t that appear in the law of motion of the control variables $c_{i,t}$ and the endogenous state variables $x_{i,t+1}$. The representation assumes there are no exogenous nonlinearities as the exogenous states $x_{i,t}^{exo}$ evolve according to a linear autoregressive process.

The objective of this paper is to propose an econometric method to assess the empirical adequacy of the nonlinear terms in the DSGE model solution (1). The DSGE model generates cross-coefficient restrictions between the first-order terms and the higher-order terms that may or may not be correctly specified. In principle, one could try to estimate two versions of the state-space model given by (1) and (2): a restricted version that imposes the functional relationship between the low-dimensional DSGE model parameter vector θ and the state-space coefficients $\psi(\cdot)$, $\zeta(\cdot)$, and $\xi(\cdot)$ and an unrestricted version in which the ψ's, ζ's, and ξ's are freely estimated. The discrepancy between the restricted and unrestricted estimates provides a measure of empirical adequacy. However, due to the large number of parameters and some inherent identification problems, the unrestricted estimation of the state-space system (1) and (2) is difficult to implement. In fact, even the literature that evaluates linearized DSGE models has by and large abstained from trying to estimate unrestricted state-space representations.

A more common approach in the literature on linearized DSGE models is to compare properties of the DSGE model to properties of an unrestricted VAR. This comparison can take many different forms, e.g., assessing the discrepancy between unrestricted VAR coefficient estimates and the DSGE-model-implied VAR approximation as in Smith (1993), or the comparison of VAR and DSGE model impulse responses as in Cogley and Nason (1994) or Christiano, Eichenbaum, and Evans (2005). Since our goal is to identify nonlinearities, a linear VAR would not be of any use. Instead, we compare parameter estimates of nonlinear autoregressive time-series

models obtained from actual U.S. data and data simulated from a nonlinear DSGE model. If the DSGE model is well specified, then the estimates of the auxiliary models ought to be very similar. This comparison is formalized as a Bayesian posterior predictive check. We proceed by providing a detailed description of the auxiliary time-series model that is used for the DSGE model evaluation.

3   Quadratic Autoregressive Models

The most popular (and empirically successful) nonlinear time-series models are those capturing time variation in the coefficients of linear time-series models, e.g., Markov-switching models, time-varying coefficient models, GARCH models, and stochastic volatility models. However, none of these models provides a good characterization of the nonlinearity generated endogenously by the DSGE model solution in (1). For this reason we develop a new class of nonlinear autoregressive time-series models that are more closely tied to the DSGE model solution in (1).

We introduce the specification of a first-order quadratic autoregressive (QAR) model in Section 3.1. We subsequently characterize some of its important properties in Section 3.2 and describe the implementation of posterior inference in Section 3.3. Section 3.4 provides generalizations of the basic specification and discusses the relationship of our QAR models to other nonlinear time-series models.

3.1   Specification of the QAR(1,1) Model

The starting point is a perturbation approximation of the solution for a nonlinear difference equation of the form

$$y_t = f(y_{t-1}, \omega u_t), \quad u_t \overset{iid}{\sim} N(0,1). \tag{3}$$

We assume that the process characterized by (3) has a unique deterministic steady state that solves the equation $y_* = f(y_*, 0)$. Following the literature on perturbation

methods, e.g., Holmes (1995) and Lombardo (2011), we construct an approximate solution of the form

$$y_t^* = y_t^{(0)} + \omega y_t^{(1)} + \omega^2 y_t^{(2)}. \tag{4}$$

It turns out that this solution is second-order accurate in the sense that

$$y_t = y_t^* + \mathcal{O}_p(\omega^3) \tag{5}$$

as $\omega \longrightarrow 0$.

To obtain y_t^*, we take a second-order Taylor expansion of the function f around $y_t = y_*$ and $\omega = 0$:

$$
\begin{aligned}
y_t - y_* &= f_y(y_{t-1} - y_*) + f_u \omega u_t \\
&\quad + \frac{1}{2} f_{y,y}(y_{t-1} - y_*)^2 + f_{y,u}(y_{t-1} - y_*)\omega u_t \\
&\quad + \frac{1}{2} f_{u,u}(\omega u_t)^2 + \text{higher-order terms},
\end{aligned}
\tag{6}
$$

where $f_{x,y}$ denotes the (x, y)'th derivative of f evaluated at the point $(y_t = y_*, \omega = 0)$. Substituting (4) into (6) and neglecting terms of order $\mathcal{O}_p(\omega^3)$, one obtains:

$$
\begin{aligned}
& y_t^{(0)} - y_* + \omega y_t^{(1)} + \omega^2 y_t^{(2)} \\
&= f_y\left(y_{t-1}^{(0)} - y_* + \omega y_{t-1}^{(1)} + \omega^2 y_{t-1}^{(2)}\right) + f_u \omega u_t \\
&\quad + \frac{1}{2} f_{y,y}\left(y_{t-1}^{(0)} - y_* + \omega y_{t-1}^{(1)} + \omega^2 y_{t-1}^{(2)}\right)^2 \\
&\quad + \frac{1}{2} f_{y,u}\left(y_t^{(0)} - y_* + \omega y_{t-1}^{(1)} + \omega^2 y_{t-1}^{(1)}\right)\omega u_t + \frac{1}{2} f_{u,u}\omega^2 u_t^2 + \mathcal{O}_p(\omega^3).
\end{aligned}
\tag{7}
$$

We set $y_t^{(0)} = y_{t-1}^{(0)} = y_*$ and then match terms of the same ω-order on the left-hand side and the right-hand side of (7) to obtain the laws of motion for $y_t^{(1)}$ and $y_t^{(2)}$:

$$
\begin{aligned}
y_t^{(1)} &= f_y y_{t-1}^{(1)} + f_u u_t, \\
y_t^{(2)} &= f_y y_{t-1}^{(2)} + \frac{1}{2} f_{y,y}\left(y_{t-1}^{(1)}\right)^2 + \frac{1}{2} f_{y,u} y_{t-1}^{(1)} u_t + \frac{1}{2} f_{u,u} u_t^2.
\end{aligned}
$$

Notice that $y_t^{(1)}$ follows an AR(1) process and that conditional on $y_t^{(1)}$ the dynamics of $y_t^{(2)}$ are also linear. Substituting the expressions for $y_t^{(1)}$ and $y_t^{(2)}$ into (4) and

collecting terms, we obtain that a second-order accurate perturbation approximation of the nonlinear difference equation (3) is given by the system:

$$y_t = y_* + f_y(y_{t-1} - y_*) + \frac{1}{2}f_{y,y}\left(\omega y_{t-1}^{(1)}\right)^2 + \left(f_u + \frac{1}{2}f_{y,u}y_{t-1}^{(1)}\right)\omega u_t + \frac{1}{2}f_{u,u}\omega^2 u_t^2$$

$$y_t^{(1)} = f_y y_{t-1}^{(1)} + f_u u_t. \tag{8}$$

We undertake a few additional modifications. We define $s_t \equiv \omega y_t^{(1)}$ and introduce the parameters

$$\phi_0 = y_*, \quad \phi_1 = f_y, \quad \tilde{\phi}_2 = \frac{1}{2}f_{y,y}, \quad \tilde{\gamma} = \frac{1}{2\omega}f_{y,u}, \quad \sigma = f_u\omega.$$

Moreover, we drop the term $\frac{1}{2}f_{u,u}u_t^2$ to obtain a conditional Normal distribution of y_t. Overall, this leads to the nonlinear state-space model:

$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \tilde{\phi}_2\tilde{s}_{t-1}^2 + (1 + \tilde{\gamma}\tilde{s}_{t-1})\sigma u_t \tag{9}$$

$$\tilde{s}_t = \phi_1\tilde{s}_{t-1} + \sigma u_t, \quad u_t \overset{iid}{\sim} N(0,1).$$

To complete the specification of the time-series model we assume that the initial values in period $t = -T_*$ have distribution F_{-T_*} and that the innovations u_t are normally distributed:

$$(y_{-T_*}, \tilde{s}_{-T_*}) \sim F_{-T_*}, \quad u_t \overset{iid}{\sim} N(0,1). \tag{10}$$

We refer to (9) as the QAR(1,1) model. The first number indicates the number of lags in the conditional mean function, and the second number indicates the number of lags that interact with the innovation u_t.

It is convenient to reparameterize the QAR(1,1) model as follows. Define ϕ_2, γ, and s_t such that

$$\phi_2 = \tilde{\phi}_2\frac{\sigma^2}{1 - \phi_1^2}, \quad \gamma = \frac{\sigma}{\sqrt{1 - \phi_1^2}}\tilde{\gamma}, \quad \text{and} \quad s_t = \frac{\sqrt{1 - \phi_1^2}}{\sigma}\tilde{s}_t. \tag{11}$$

Under the reparameterization, the coefficients ϕ_2 and γ interact with standardized versions of s_{t-1}^2 and s_{t-1}, respectively. Thus, (9) becomes

$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2 s_{t-1}^2 + (1 + \gamma s_{t-1})\sigma u_t \tag{12}$$

$$s_t = \phi_1 s_{t-1} + \sqrt{1 - \phi_1^2}u_t, \quad u_t \overset{iid}{\sim} N(0,1).$$

3.2 Important Properties of the QAR(1,1) Model

In order to appreciate two of the important implications of the recursively linear structure of the QAR(1,1) model given by (12), consider the alternative specification (omitting the constant term and the volatility dynamics) $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-1}^2 + u_t$, $0 < \phi_1 < 1$ and $\phi_2 > 0$. It is straightforward to verify that this specification has two steady states: $y_*^{(1)} = 0$ and $y_*^{(2)} = (1 - \phi_1)/\phi_2$. The second steady state arises as an artefact of the quadratic representation even if the underlying nonlinear model (3) has only a single steady state. Moreover, from writing $\Delta y_t = (-1 + \phi_1 + \phi_2 y_{t-1})y_{t-1} + u_t$ notice that the system becomes explosive if a large shock has pushed y_{t-1} above $y_*^{(2)}$. This explosiveness can arise regardless of the value of ϕ_1.

The multiplicity of steady states and the undesirable explosive dynamics have been pointed out in the context of second-order perturbation solutions of DSGE models by Kim, Kim, Schaumburg, and Sims (2008), who proposed an ex-post modification of quadratic autoregressive equations to ensure that unwanted higher-order terms do not propagate forward and generate explosive behavior not present in the underlying nonlinear model. This modification is called "pruning" in the literature.[4] Our derivation of the QAR model in Section 3.1 automatically generates a recursively linear structure with a unique steady state and nonexplosive dynamics for suitably restricted values of ϕ_1. If the marginal distribution of s_{-T_*} is $N(0, 1)$, then the process s_t, $t \geq -T_*$, is strictly stationary under the restriction $|\phi_1| < 1$. In turn, the vector process $z_t = [s_{t-1}, s_{t-1}^2, u_t]'$ is strictly stationary and we can rewrite the law of motion

---

[4]Lombardo (2011) constructs a pruned perturbation solution of a DSGE model directly rather than by ex-post adjustment. Lan and Meyer-Gohde (2013) solve DSGE models by constructing approximate second-order Volterra series expansions for the model variables, which also eliminates unwanted higher-order terms. Andreasen, Fernández-Villaverde, and Rubio-Ramírez (2013) derive the moments of observables from a general state-space representation for pruned DSGE models to facilitate moment-based estimation.

of y_t in (9) as

$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + g(z_t) = \phi_0 + \sum_{j=0}^{\infty} \phi_1^j g(z_{t-j}). \tag{13}$$

This representation highlights that y_t is a stationary process. Since $g(z_t)$ is a nonlinear function of u_t and its history, the process is, however, not linear in u_t anymore. In fact, under the assumption that y_t was initialized in the infinite past $(T_* \longrightarrow -\infty)$, we obtain the following representation:

$$y_t = \phi_0 + \sigma \sum_{j=0}^{\infty} \phi_1^j u_{t-j} + \sigma \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \left(\tilde{\gamma}\mathcal{I}\{l > j\}\phi_1^{l-j} + \tilde{\phi}_2 \sum_{k=0}^{\min\{j,l\}} \phi_1^{j+l-k} \right) u_{t-j}u_{t-l}. \tag{14}$$

(14) is a discrete-time Volterra series expansion, in which the Volterra kernels of order one and two are tightly restricted and the kernels of order larger than two are equal to zero.[5] The recursively linear structure also facilitates the computation of higher-order moments of y_t. Further details are provided in the appendix.

For the purpose of the empirical analysis the most important characteristic of (9) is that the model is able to generate nonlinear dynamics that are akin to the nonlinear dynamics of DSGE models solved with perturbation methods. In particular, impulse responses defined as

$$IRF_t(h) = \mathbb{E}_t[y_{t+h}|u_t = 1] - \mathbb{E}_t[y_{t+h}]$$

are state dependent. For instance, for $h = 1$ we obtain

$$IRF_t(0) = \sigma(1 + \gamma s_{t-1}), \quad IRF_t(1) = \sigma\left(\phi_1(1 + \gamma s_{t-1}) + 2\phi_1\phi_2\sqrt{1 - \phi_1^2}s_{t-1}\right). \tag{15}$$

Moreover, the model generates conditional heteroskedasticity. The conditional variance of y_t is given by

$$\mathbb{V}_{t-1}[y_t] = (1 + \gamma s_{t-1})^2\sigma^2. \tag{16}$$

---

[5]The infinite sequences of coefficients on terms $\{u_{t-j}\}_{j\geq 0}$, $\{u_{t-j}u_{t-l}\}_{j\geq 0,l\geq 0}$, $\{u_{t-j}u_{t-l}u_{t-k}\}_{j\geq 0,l\geq 0,k\geq 0}$, etc. are called Volterra kernels.

3.3 Posterior Inference for the QAR(1,1) Model

We estimate the QAR(1,1) model using Bayesian methods. Starting point is a joint distribution of data, parameters, and initial states:

$$p(Y_{0:T}, \theta, s_0) = p(Y_{1:T}|y_0, s_0, \theta)p(y_0, s_0|\theta)p(\theta),$$

where $p(Y_{1:T}|y_0, s_0, \theta)$ is a likelihood function that conditions on the initial values of y_0 and s_0, $p(y_0, s_0|\theta)$ characterizes the distribution of the initial values, $p(\theta)$ is the prior density of the QAR(1,1) parameters, and $\theta = [\phi_0, \phi_1, \phi_2, \gamma, \sigma^2]'$. Since for large values of $|s_{t-1}|$ the term $(1 + \gamma s_{t-1})$ in (12) may become close to zero or switch signs, we replace it by

$$\left((1 - \vartheta) \exp\left[\frac{\gamma}{1 - \vartheta} s_{t-1} \right] + \vartheta \right), \tag{17}$$

where $1 + \gamma s_{t-1}$ is the first-order Taylor series expansion of (17). The exponential transformation guarantees non-negativity of the time-varying standard deviation, and the constant ϑ provides some regularization ensuring that the shock standard deviation is strictly greater than $\sigma \exp(\vartheta)$ in all states of the world.

It is convenient to factorize the likelihood function into conditional densities as follows:

$$p(Y_{1:T}|y_0, s_0, \theta) = \prod_{t=1}^{T} p(y_t|y_{0:t-1}, s_0, \theta).$$

Given s_0 and θ, it is straightforward to evaluate the likelihood function iteratively. Conditional on s_{t-1} the distribution of y_t is normal. The equation for y_t in (12) can be solved for u_t to determine s_t, which completes iteration t. In addition to the likelihood function, we need to specify an initial distribution $p(y_0, s_0|\theta)$. We assume that the system was in its steady state in period $t = -T_*$, that is, $y_{-T_*} = \phi_0$ and $s_{-T_*} = 0$. Based on iterating the original system (9) forward we compute a mean and variance for (y_0, s_0) and assume that the initial values are normally distributed. Further details of this initialization are provided in the appendix. Since the dimension of θ is small, we use a single-block random walk Metropolis (RWM) algorithm to generate draws from the posterior of θ.

3.4 Further Discussion

The QAR(1,1) model in (8) has a straightforward generalization in which we include additional lag terms:

$$
\begin{aligned}
y_t &= \phi_0 + \sum_{l=1}^{p} \phi_{1,l}(y_{t-l} - \phi_0) + \sum_{l=1}^{p}\sum_{m=l}^{p} \tilde{\phi}_{2,lm} s_{t-l} s_{t-m} + \left(1 + \sum_{l=1}^{q} \tilde{\gamma}_l s_{t-l}\right) \sigma u_t \quad (18) \\
\tilde{s}_t &= \sum_{l=1}^{p} \phi_{1,l} s_{t-l} + \sigma u_t.
\end{aligned}
$$

We refer to (18) as the QAR(p,q) model. As in the standard AR(p) model, the stationarity of y_t is governed by the roots of the lag polynomial $1 - \sum_{l=1}^{p} \phi_{1,l} z^l$. The quadratic terms generate an additional $p(p+1)/2$ coefficients in the conditional mean equation for y_t. Since the number of coefficients grows at rate p^2, a shrinkage estimation method is required even for moderate values of p, in order to cope with the dimensionality problem. The QAR model can also be extended to the vector case, which is an extension that we are pursuing in ongoing research. The empirical analysis presented in Section 6 is based on the QAR(1,1) specification.

The QAR model is closely related but not identical to some of the existing nonlinear time-series models. For $\gamma = 0$, the QAR(1,1) can be viewed as a pruned version of the generalized autoregressive model (GAR) discussed in Mittnik (1990), which augments the standard AR model by higher-order polynomials of the lagged variables. The conditional heteroskedasticity in (9) has a linear autoregressive structure. For $\phi_2 = 0$, our model is a special case of the LARCH model studied in Giraitis, Robinson, and Surgailis (2000). Since the conditional variance of y_t can get arbitrarily close to zero, likelihood-based estimation of LARCH models is intrinsically difficult. We circumvent these difficulties by introducing exponential transformation in (17).

Grier and Perry (1996, 2000) have estimated GARCH-M models on macroeconomic time series. GARCH-M models provide a generalization of the ARCH-M mod-

els proposed by Engle, Lilien, and Robins (1987) and can be written as

$$y_t \;=\; \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2(\sigma_t^2 - \sigma^2) + \sigma_t u_t$$
$$\sigma_t^2 - \sigma^2 \;=\; \gamma_1(u_{t-1}^2 - \sigma^2) + \gamma_2(\sigma_{t-1}^2 - \sigma^2).$$

Under suitable parameter restrictions, y_t can be expressed as a nonlinear function of u_t and its lags. As in the case of the QAR model, y_t depends on the sequence $\{u_{t-j}^2\}$. In addition, the term $\sigma_t u_t$ introduces interactions between u_t and u_{t-j}^2, $j > 1$. However, coefficients on terms of the form $u_{t-j} u_{t-l}$, $j \neq l$ are restricted to be zero. From our perspective, the biggest drawback of the GARCH-M model is that nonlinear conditional mean dynamics are tied to the volatility dynamics: In the absence of conditional heteroskedasticity, the dynamics of y_t are linear. The QAR model is much less restrictive in this regard: y_t can be conditionally homoskedastic ($\gamma = 0$) but at the same time have nonlinear conditional mean dynamics, that is, $\phi_2 \neq 0$.

4 QAR Empirics

We begin the empirical analysis by fitting the QAR(1,1) model to per capita output growth, nominal wage growth, GDP deflator inflation, and federal funds rate data.[6] The choice of data is motivated by the DSGE model that is being evaluated subsequently. The DSGE model features potentially asymmetric wage and price adjustment costs, and we will assess whether the nonlinearities generated by this DSGE

---

[6]All series are quarterly and obtained from the FRED database of the Federal Reserve Bank of St. Louis. Output growth is the log difference of real GDP (GDPC96). We compute log differences of civilian noninstitutional population (CNP16OV) and remove a one-sided eight-quarter moving average to smooth population growth. The smoothed population growth series is used to obtain per capita GDP growth rates. Inflation is the log difference of the GDP deflator (GDPDEF). Nominal wage growth is the log difference of compensation per hour in the nonfarm business sector (COMPNFB). For interest rates, we use quarterly averages of monthly effective federal funds rates (FEDFUNDS).

Table 1: Estimation Samples and Pre-Samples

Sample	Estimation Sample	Pre-Sample for Prior
1	1960:Q1 - 1983:Q4	1955:Q1 - 1959:Q4
2	1960:Q1 - 2007:Q4	1955:Q1 - 1959:Q4
3	1960:Q1 - 2012:Q4	1955:Q1 - 1959:Q4
4	1984:Q1 - 2007:Q4	1955:Q1 - 1983:Q4
5	1984:Q1 - 2012:Q4	1955:Q1 - 1983:Q4

model are consistent with the nonlinearities in U.S. data. We report parameter estimates for the QAR model in Section 4.1 and explore the properties of the estimated models in Section 4.2.

4.1 Estimation of QAR(1,1) Model on U.S. Data

We estimate QAR(1,1) models for output growth, inflation, nominal wage growth, and interest rates using five different sample periods, which are summarized in Table 1. The longest sample spans the period from 1960:Q1 to 2012:Q4. This sample includes the high-inflation episode of the 1970s, the subsequent disinflation period, and the Great Recession of 2008-09. We then split this sample after 1983:Q4 into a pre-Great-Moderation sample that ranges from 1960:Q1 to 1983:Q4 and a post-Great-Moderation sample from 1984:Q1 to 2012:Q4. Since the 2008-09 recession involves large negative GDP growth rates that may be viewed as outliers, and the federal funds rate has been at or near the lower bound of 0% since 2008, we consider two additional samples that exclude the Great Recession data and end in 2007:Q4.

To specify the prior distribution for the QAR parameters we use normal distributions for ϕ_0, ϕ_2, and γ. The autoregressive coefficient ϕ_1 is *a priori* also normally distributed, but the normal distribution is truncated to ensure stationarity of the QAR model. Finally, the prior distribution of σ is of the inverted gamma form. We

use pre-sample information to parameterize the priors. The pre-sample periods for our five estimation samples are provided in the last column of Table 1. Throughout the estimation the tuning constant ϑ in (17) is fixed at $\vartheta = 0.1$. The prior distributions for ϕ_1, the first-order autoregressive coefficient, are centered at the pre-sample first-order autocorrelations of the four time series. The inverse Gamma distribution of σ is centered at the residual standard deviation associated with the pre-sample estimation of an AR(1) model. Finally, the prior mean of ϕ_0 is specified such that the implied $\mathbb{E}[y_t]$ of the QAR(1,1) model corresponds to the pre-sample mean of the respective time series. The priors for ϕ_2 and γ are centered at zero and have a standard deviation of 0.1. Further details are provided in the appendix.

Figure 1 summarizes the posterior distributions of ϕ_2 and γ. Detailed estimation results for the remaining QAR(1,1) parameters are tabulated in the appendix. The ϕ_2 posteriors for GDP growth using the three samples starting in 1960 are essentially centered at zero with the 90% credible interval covering both positive and negative values. The γ posterior medians for the same samples are slightly negative, around -0.05, but the 90% credible sets also cover positive values, providing only some mild evidence for conditional variance dynamics. For the two post-Great-Moderation samples, the ϕ_2 estimates drop to about -0.1 and the credible set now excludes zero. The strongest evidence for nonlinearity in GDP growth is present in the 1984-2012 sample, which includes large negative growth rates of output during the Great Recession, in the form of $\phi_2 < 0$ and $\gamma < 0$. Nonlinearities in wages and inflation are reflected in positive estimates of γ. These nonlinearities are most pronounced for the 1960-2007 and the 1960-2012 samples. For the federal funds rate we obtain estimates of ϕ_2 near zero and estimates for γ of about 0.4 for samples that include the pre-1984 observations. For samples that start after the Great Moderation, the pattern is reversed: The estimates of ϕ_2 are around -0.2 and the estimates of γ are close to zero. We will discuss the interpretation of these estimates in Section 4.2.

Figure 2 depicts log marginal likelihood differentials for the QAR(1,1) versus a linear autoregressive AR(1) model. The AR(1) models are estimated by setting $\phi_2 =$

Figure 1: Posterior Medians and Credible Intervals for QAR Parameters

Parameter ϕ_2

Parameter γ
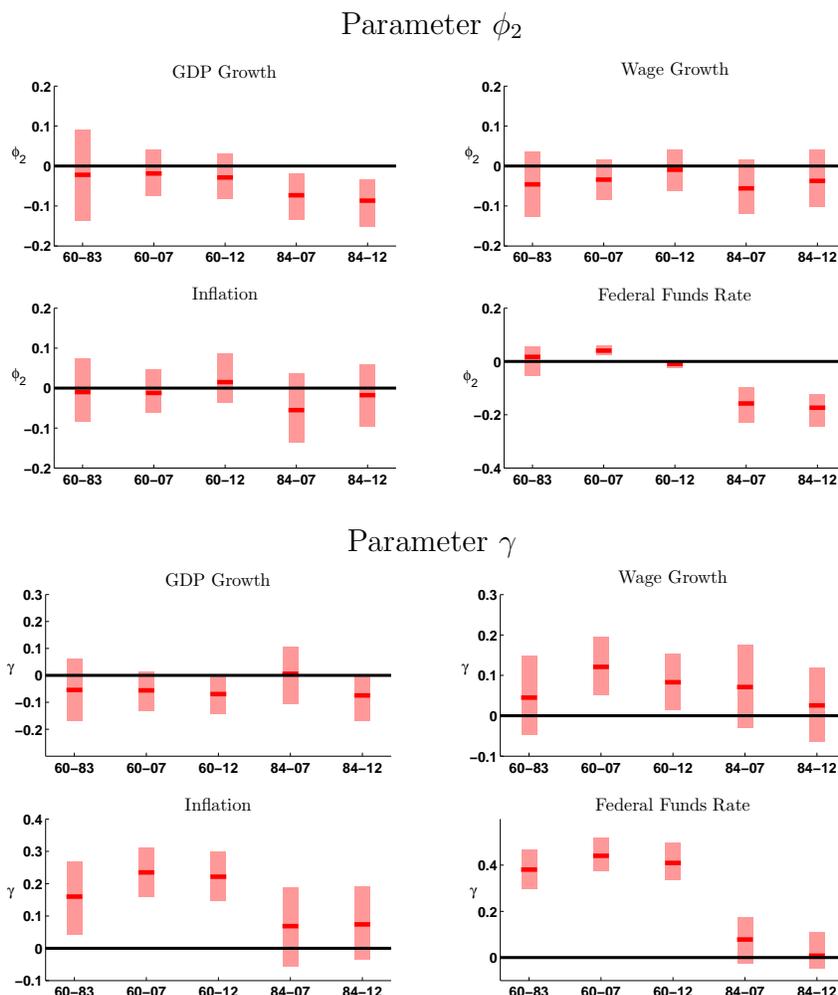
Notes: The solid bars indicate posterior medians and the shaded boxes delimit 90% equal-tail-probability credible intervals.

$\gamma = 0$ and using the same priors for ϕ_0, ϕ_1, and σ as in the estimation of the QAR(1,1) model. A positive value indicates evidence in favor of the nonlinear QAR(1,1). Under equal prior probabilities, the difference in log marginal data density between the two models has the interpretation of log posterior odds. By and large, the marginal likelihood differentials favor the QAR(1,1) specification. The evidence in favor of the nonlinear specification is strongest for the federal funds rate. Marginal likelihood differentials range from 20 to 60. For output growth, there is substantial evidence in favor of the QAR model for the post-Great-Moderation samples, whereas for inflation

Figure 2: Log Marginal Data Density Differentials: QAR(1,1) versus AR(1)
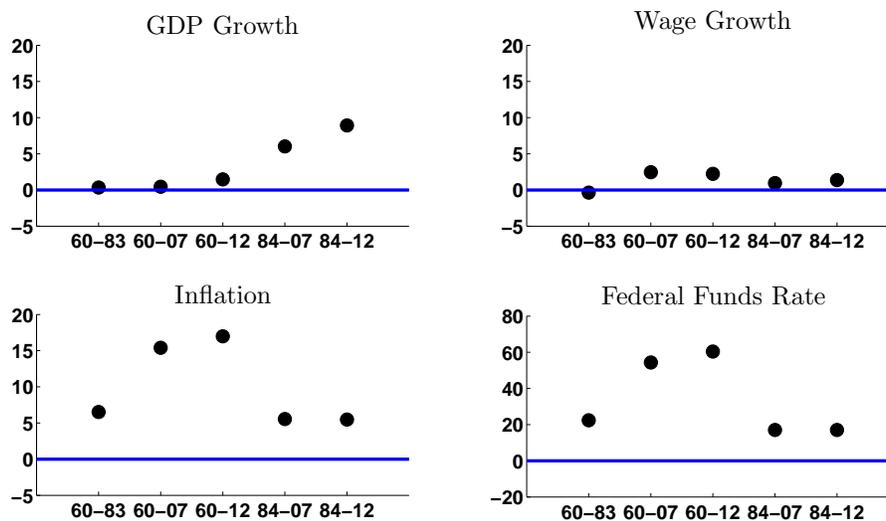
Notes: The figure depicts log marginal data density differentials. A positive number provides evidence in favor of the QAR(1,1) specification.

large positive log marginal likelihood differentials are obtained for the 1960-2007 and the 1960-2012 samples. For wage growth, the evidence in favor of the nonlinear specification is less strong: Log marginal likelihood differentials are around 2.

4.2 Properties of the Estimated QAR Models

In this section we discuss what the nonlinearities we identified in the previous section mean for each variable. For ease of exposition, we focus on the subsample that "maximizes" the nonlinearities for each variable, which roughly corresponds to picking the subsample that has the largest marginal data density differential between the AR(1) and the QAR(1,1) models.

GDP Growth. Our results show that the posterior medians of ϕ_2 and γ for GDP growth are less than zero. The largest estimates (in absolute terms) are obtained for the 1984-2012 sample. As (16) shows, with a negative γ, the periods of below-mean growth (likely to be recessions) are also periods where volatility is higher, which is a

Figure 3: Impulse Responses of GDP Growth (in Absolute Terms)
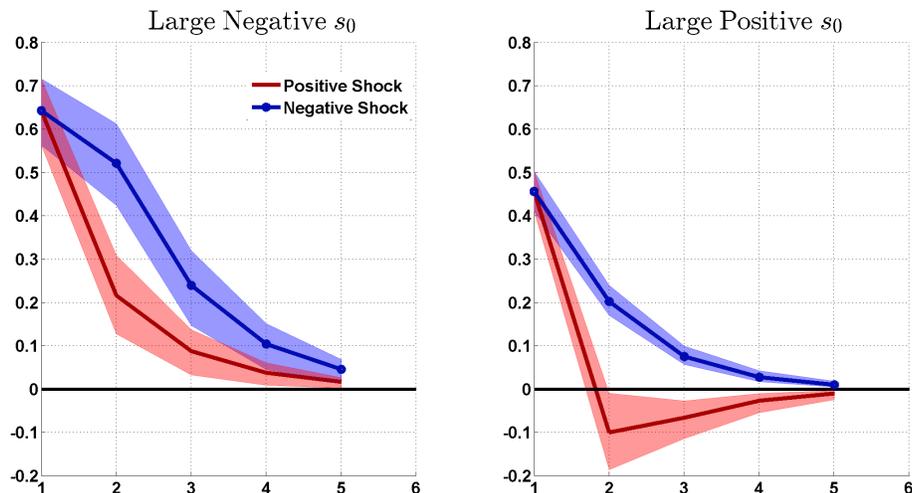
Notes: 1984-2012 sample. Solid and dashed lines correspond to median impulse responses to one-standard-deviation shocks and shaded bands represent 60% credible intervals (equal tail probability). To initialize the latent state s_0 we compute two-quarter moving averages based on the states associated with the estimated QAR model and calculate the minimum and the maximum of the smoothed series. For the left panel (large negative s_0) the initialization is based on the minimum and in the right panel (large positive s_0) it is based on the maximum.

well-known business cycle fact. A negative ϕ_2 along with a negative γ also implies that the response to shocks is a function of the initial state s_0. Using the formulas in (15), Figure 3 depicts the absolute responses of GDP growth to a negative and a positive one-standard-deviation shock. In the left panel, we assume that the initial state s_0 takes on large negative values, whereas the responses in the right panel condition on large positive s_0's.[7] This figure highlights, that regardless of the initial state, negative shocks are more persistent than positive shocks. Moreover, both shocks are more persistent in recessions. Combining these results, we deduce that multiple positive shocks are necessary to recover from a recession, while a small number of negative shocks can generate a recession.

The impulse-response findings are consistent with the time-series plot of GDP growth, which is provided in Figure 4. In this figure, and all the other time-series plots

---

[7]To obtain the s_0 for a given draw, we compute a two-period moving average to smooth the s_t series and use the minimum and the maximum values for this smoothed series.

Figure 4: Annualized GDP Growth
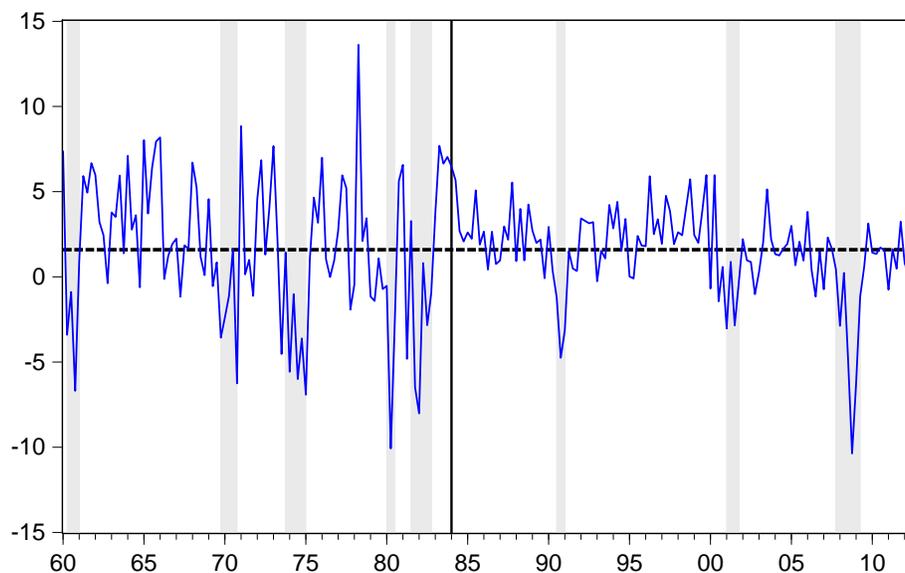
Notes: Shaded areas indicate NBER recessions and the dashed horizontal line represents the sample mean of the series.

below, shaded areas indicate NBER recessions and the solid vertical line indicates the year 1984, which is the starting point of two of the five estimation samples. The unconditional mean of the variable is shown as a horizontal dashed line. In the post-1983 sample, the most extreme observations are all during recessions, confirming the effect of $\gamma < 0$. Looking at the quarters just prior to and just after NBER recessions, we see that the declines in GDP growth are always very sharp, but the recoveries, defined as getting back to and staying at pre-recession levels, take much longer.

It is easy to see why the nonlinearities identified in the samples starting in 1984 are not as pronounced in the samples that start in 1960. First, prior to 1984 there are more episodes of large positive GDP growth rates. These are, in absolute terms, as large as the negative growth rates observed between 1960 and 2012. Thus, recessions are not necessarily periods of higher volatility. Second, the recoveries from recessions are as sharp as the entries, not displaying the clear asymmetry in the later sample. These findings explain why a linear AR(1) is a good description of GDP growth

pre-1984.[8]

Inflation and Wage Growth. The nonlinearities in the inflation dynamics are most pronounced in the 1960-2007 sample with $\gamma > 0$ and $\phi_2 = 0$. Once again referring to the conditional variance formula in (16), we conclude that periods of above-mean inflation are associated with high volatility. In fact, Figure 5 shows that the period from 1970 to 1980 has high and volatile inflation. A similar conclusion can be reached in the post-1983 sample, but to a lesser degree. Figure 6 shows that nominal wage growth displays properties similar to inflation. In the 1960-2007 sample, which is also the relevant one for nominal wage growth, volatility tends to be high when the level is high. Because nominal wage growth is more volatile than inflation, and there are many large negative observations, the estimate of γ is smaller for the former series.

Federal Funds Rate. Figure 7 shows the plot of the federal funds rate. Based on the QAR(1,1) estimation results, there are two samples with strong nonlinearities. In the 1960-2007 sample, we find a positive γ. As was the case for inflation and nominal wage growth, this is due to the observations from late the 1960s to the mid-1980s, which are typically above the unconditional mean; thus, volatility is higher when the level is higher. For the 1984-2012 sample we find $\phi_2 < 0$ and $\gamma = 0$. In this period the extreme observations are equally likely to be positive or negative and thus $\gamma = 0$ is reasonable. $\phi_2 < 0$ implies that interest rates fall faster than they rise. This seems to be consistent with the Federal Reserve's operating procedures in the post-1983 sample, and it can have two separate explanations. First, the Federal Reserve may have an asymmetric policy rule, in which reactions to deviations from targets may

---

[8]Qualitatively, our results for GDP growth are in line with findings by Brunner (1997), who estimated three nonlinear models for real gross national product. Based on a sample from 1947 to 1990 the author obtained strong evidence of countercyclical volatility, that is, recessions are periods of high volatility. Moreover, Brunner (1997) detects nonlinear conditional mean dynamics: According to the impulse responses, the effects of a negative shock accumulate faster than those of a positive shock, in line with our findings. Similarly, McKay and Reis (2008) find that the brevity and violence of contractions and expansions are about equal in a sample that encompasses our longest sample, once again in line with our results.

Figure 5: Annualized GDP Deflator Inflation
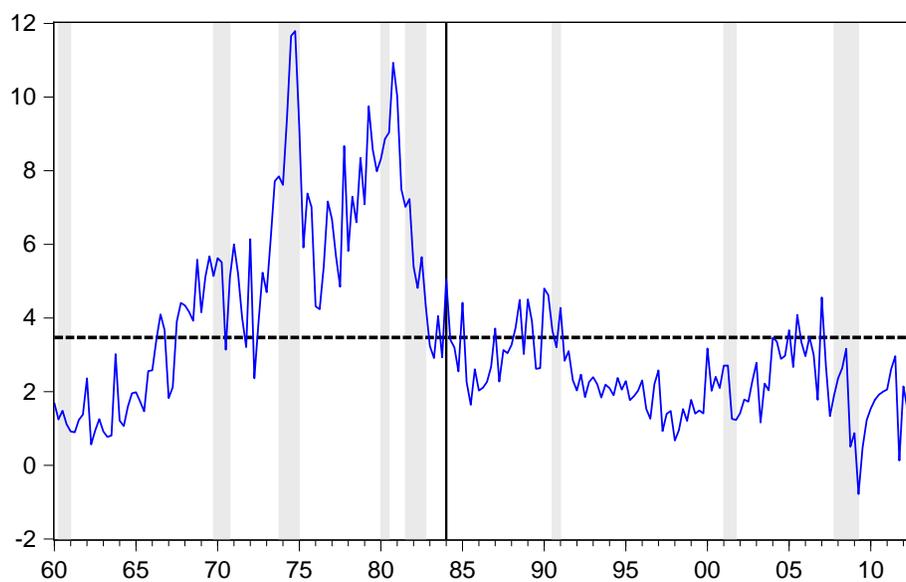
Notes: Shaded areas indicate NBER recessions and the dashed horizontal line represents the sample mean of the series.

Figure 6: Annualized Nominal Wage Growth
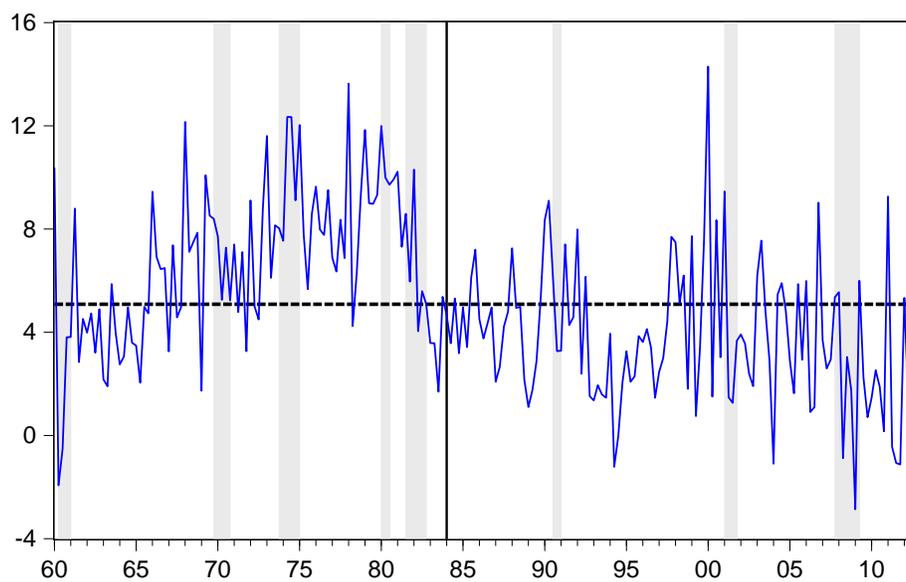
Notes: Shaded areas indicate NBER recessions and the dashed horizontal line represents the sample mean of the series.

Figure 7: Federal Funds Rate
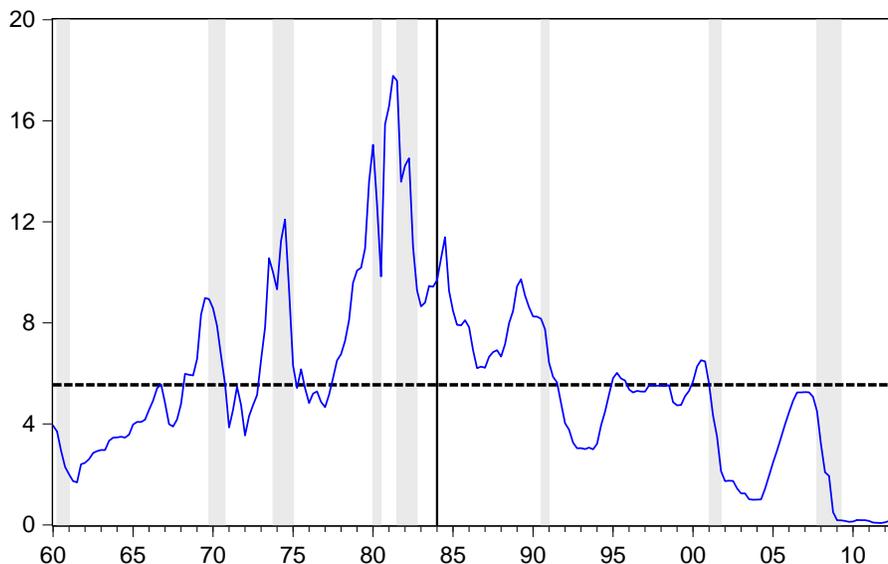
Notes: Shaded areas indicate NBER recessions and the dashed horizontal line represents the sample mean of the series.

depend on the sign of the deviation. This can happen, for example, if the Federal Reserve is risk averse and wants to avoid recessions: When GDP growth falls, the central bank is willing to cut the policy rate quickly, but when GDP growth starts to improve, it is reluctant to increase the policy rate immediately. Second, the variables that the Federal Reserve tracks may have asymmetries themselves. Given our finding that $\phi_2 < 0$ for GDP in this sample, the second explanation is certainly reasonable. There is some evidence for the first explanation as well. For example, Dolado, Maria-Dolores, and Ruge-Murcia (2004) and Cukierman and Muscatelli (2008) estimate a nonlinear Taylor rule using GMM and find that U.S. monetary policy is better characterized by a nonlinear policy rule after 1983, especially with respect to the reaction to output gap deviations.

To sum up, the estimation of QAR(1,1) models provides evidence of interesting and substantial nonlinearities in the U.S. macroeconomic time series. For the two samples that start in 1960 and extend beyond the 1990s the nonlinearities are re-

flected in the run-up in inflation in the 1970s, with spillovers to nominal wage growth and the federal funds rate. In the shorter post-1983 samples, there are two important nonlinearities: the asymmetries in GDP growth, which are particularly pronounced if the 2008-09 recession is included in the sample, and the federal funds rate. In the remainder of this paper we examine whether a DSGE model with asymmetric adjustment costs for prices and wages can possibly generate the nonlinearities documented in this section.

5 A DSGE Model with Asymmetric Price and Wage Adjustment Costs

By now there exists a large empirical literature on the estimation of New Keynesian DSGE models, including small-scale models such as the one studied in Lubik and Schorfheide (2004) and Rabanal and Rubio-Ramírez (2005), as well as variants of the Smets and Wouters (2007) model. It turns out that, in the absence of zero-lower-bound constraints on nominal interest rates, high degrees of risk aversion, large shocks, or exogenous nonlinearities such as stochastic volatility, these models do not generate strong nonlinearities, in the sense that first-order and higher-order perturbation approximations deliver very similar decision rules. In order to generate stronger nonlinearities that can be captured in higher-order perturbation approximations, we consider a model with potentially asymmetric price and wage adjustment costs that builds on that of Kim and Ruge-Murcia (2009).

The model economy consists of a final good producing firm, a continuum of intermediate goods producing firms, a representative household, and a monetary and a fiscal authority. The model replaces Rotemberg-style quadratic adjustment cost functions by linex adjustment cost functions, which can capture downward (as well as upward) nominal price and wage rigidities. Our model abstracts from capital accumulation. In the subsequent empirical analysis we examine whether the asym-

metric adjustment costs can generate the observed nonlinearities in inflation and wage growth and whether the effects of asymmetric adjustment costs translate into nonlinearities in GDP growth and the federal funds rate. In a nutshell, asymmetric price adjustments should lead to asymmetric quantity adjustments. To the extent that the central bank sets interest rates in response to inflation and output movements, nonlinearities in the target variables may translate into nonlinearities in the interest rate itself.

Final Good Production. The perfectly competitive, representative, final good producing firm combines a continuum of intermediate goods indexed by $j \in [0,1]$ using the technology

$$Y_t = \left(\int_0^1 Y_t(j)^{1-\lambda_{p,t}} dj \right)^{\frac{1}{1-\lambda_{p,t}}}. \tag{19}$$

Here, $1/\lambda_{p,t} > 1$ represents the elasticity of demand for each intermediate good. The firm takes input prices $P_t(j)$ and output prices P_t as given. Profit maximization implies that the demand for intermediate goods is

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-1/\lambda_{p,t}} Y_t. \tag{20}$$

The relationship between intermediate goods prices and the price of the final good is

$$P_t = \left(\int_0^1 P_t(j)^{\frac{\lambda_{p,t}-1}{\lambda_{p,t}}} dj \right)^{\frac{\lambda_{p,t}}{\lambda_{p,t}-1}}. \tag{21}$$

Intermediate Goods. Intermediate good j is produced by a monopolist who has access to the following production technology:

$$Y_t(j) = A_t H_t(j), \tag{22}$$

where A_t is an exogenous productivity process that is common to all firms. Intermediate goods producers buy labor services $H_t(j)$ at a nominal price of W_t. Moreover, they face nominal rigidities in terms of price adjustment costs. These adjustment costs, expressed as a fraction of the firm's revenues, are defined by the linex function

$$\Phi_p(x) = \varphi_p \left(\frac{\exp\left(-\psi_p\left(x-\pi\right)\right) + \psi_p\left(x-\pi\right) - 1}{\psi_p^2} \right), \tag{23}$$

where we let $x = P_t(j)/P_{t-1}(j)$ and π is the steady-state inflation rate associated with the final good. The parameter ϕ_p governs the overall degree of price stickiness, and ψ_p controls the asymmetry of the adjustment costs. Taking as given nominal wages, final goods prices, the demand schedule for intermediate products, and technological constraints, firm j chooses its labor inputs $H_t(j)$ and the price $P_t(j)$ to maximize the present value of future profits:

$$
\mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s Q_{t+s|t} \left(\frac{P_{t+s}(j)}{P_{t+s}} \left(1 - \Phi_p \left(\frac{P_{t+s}(j)}{P_{t+s-1}(j)} \right) \right) Y_{t+s}(j) - \frac{1}{P_{t+s}} W_{t+s} H_{t+s}(j) \right) \right].
$$
(24)

Here, $Q_{t+s|t}$ is the time t value of a unit of the consumption good in period $t+s$ to the household, which is treated as exogenous by the firm.

Labor Packers. Labor services used by intermediate goods producers are supplied by a perfectly competitive labor packer. The labor packer aggregates the imperfectly substitutable labor services of households according to the following technology:

$$
H_t = \left(\int_0^1 H_t(k)^{1-\lambda_w} dk \right)^{\frac{1}{1-\lambda_w}}.
$$
(25)

The labor packer chooses demand for each type of labor in order to maximize his profits, taking as given input prices $W_t(k)$ and output prices W_t. Optimal labor demand is then given by

$$
H_t(k) = \left(\frac{W_t(k)}{W_t} \right)^{-\frac{1}{\lambda_w}} H_t.
$$
(26)

Perfect competition implies that labor cost W_t and nominal wages paid to workers are related as follows:

$$
W_t = \left(\int_0^1 W_t(k)^{\frac{\lambda_w-1}{\lambda_w}} dk \right)^{\frac{\lambda_w}{\lambda_w-1}}.
$$
(27)

Households. Each household consists of a continuum of family members indexed by k. The family members provide perfect insurance to each other, which equates their marginal utility in each state of the world. A household member of type k derives utility from consumption $C_t(k)$ relative to a habit stock. We assume that the habit stock is given by the level of technology A_t. This assumption ensures that the

economy evolves along a balanced growth path even if the utility function is additively separable in consumption, real money balances, and leisure. The household member derives disutility from hours worked $H_t(k)$ and maximizes

$$\mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s \left(\frac{(C_{t+s}(k)/A_{t+s})^{1-\tau} - 1}{1-\tau} - \chi_H \frac{H_{t+s}^{1+1/\nu}(k)}{1+1/\nu} \right) \right], \tag{28}$$

where β is the discount factor, $1/\tau$ is the intertemporal elasticity of substitution, and χ_H is a scale factor that determines the steady-state hours worked. Moreover, ν is the Frisch labor supply elasticity.

The household member is a monopolist in the supply of labor services. As a monopolist, he chooses the nominal wage and labor, taking as given the demand from the labor packer. We assume that labor market frictions induce a cost in the adjustment of nominal wages. Adjustment costs are paid as a fraction of labor income and they have the same linex structure assumed for prices:

$$\Phi_w(x) = \varphi_w \left(\frac{\exp\left(-\psi_w \left(x - \gamma\pi\right)\right) + \psi_w \left(x - \gamma\pi\right) - 1}{\psi_w^2} \right), \tag{29}$$

where $x = W_t(k)/W_{t-1}(k)$, $\gamma\pi$ is the growth rate of nominal wages, and γ is the average growth rate of technology defined below. In addition to his labor choices, the household member faces a standard consumption/saving trade-off. He has access to a domestic bond market where nominal government bonds $B_t(k)$ are traded that pay (gross) interest R_t. Furthermore, he receives aggregate residual real profits $D_t(k)$ from the firms and has to pay lump-sum taxes T_t. Thus, the household's budget constraint is of the form

$$
\begin{aligned}
& P_t C_t(k) + B_t(k) + T_t \\
& = W_t(k) H_t(k) \left(1 - \Phi_w \left(\frac{W_t(k)}{W_{t-1}(k)} \right) \right) + R_{t-1} B_{t-1}(k) + P_t D_t(k) + P_t SC_t,
\end{aligned}
$$

where $SC_t(k)$ is the net cash inflow that household k receives from trading a full set of state-contingent securities. We denote the Lagrange multiplier associated with the budget constraint by λ_t. The usual transversality condition on asset accumulation applies, which rules out Ponzi schemes.

Monetary and Fiscal Policies. Monetary policy is described by an interest rate feedback rule of the form

$$R_t = R_t^{*\,1-\rho_R} R_{t-1}^{\rho_R} e^{\epsilon_{R,t}}, \tag{30}$$

where $\epsilon_{R,t}$ is a monetary policy shock and R_t^* is the (nominal) target rate. Our specification of R_t^* implies that the central bank reacts to inflation and deviations of output growth from its equilibrium steady state γ:

$$R_t^* = r\pi^* \left(\frac{\pi_t}{\pi^*}\right)^{\psi_1} \left(\frac{Y_t}{\gamma Y_{t-1}}\right)^{\psi_2}. \tag{31}$$

Here, r is the steady-state real interest rate, π_t is the gross inflation rate defined as $\pi_t = P_t/P_{t-1}$, and π^* is the target inflation rate, which in equilibrium coincides with the steady-state inflation rate. The fiscal authority consumes a fraction ζ_t of aggregate output Y_t, where $\zeta_t \in [0, 1]$ follows an exogenous process. The government levies a lump-sum tax (subsidy) to finance any shortfalls in government revenues (or to rebate any surplus).

The model economy is perturbed by four exogenous processes. Aggregate productivity evolves according to

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln z_t, \quad \text{where} \quad \ln z_t = \rho_z \ln z_{t-1} + \epsilon_{z,t}. \tag{32}$$

Thus, on average, technology grows at the rate γ, and z_t captures exogenous fluctuations of the technology growth rate. Define $g_t = 1/(1 - \zeta_t)$. We assume that

$$\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \epsilon_{g,t}. \tag{33}$$

The inverse-demand elasticity for intermediate goods evolves according to a first-order autoregressive processes in logs:

$$\ln \lambda_{p,t} = (1 - \rho_p) \ln \lambda_p + \rho_p \ln \lambda_{p,t-1} + \epsilon_{p,t}. \tag{34}$$

Finally, the monetary policy shock $\epsilon_{R,t}$ is assumed to be serially uncorrelated. The five innovations are independent of each other at all leads and lags and are normally distributed with mean zero and standard deviations σ_z, σ_g, σ_p, and σ_R, respectively.

6 Estimation and Evaluation of the DSGE Model

The estimation and evaluation of the DSGE model proceeds in three steps. In Section 6.1, the DSGE model is estimated for two samples: 1960-2007 and 1984-2007, using the same series that were studied in Section 4. In Section 6.2, we use posterior predictive checks that are based on posterior mean estimates of QAR(1,1) parameters to assess whether the nonlinearities captured in the second-order approximated DSGE model are commensurate with the nonlinearities captured by the QAR(1,1) model. Finally, in Section 6.3 we assess the effect of adjustment cost asymmetries on the model's ability to generate nonlinear inflation and wage growth dynamics.

6.1 DSGE Model Estimation on U.S. Data

The second step in the empirical analysis consists of estimating the DSGE model based on the same data used to estimate the QAR(1,1) models in Section 4. The marginal prior distributions for the DSGE model parameters are summarized in Table 2. We use pre-sample evidence to quantify *a priori* beliefs about the average growth rate of the economy, as well as average inflation and real interest rates. We use the same priors for both samples. The prior mean for τ implies a risk-aversion coefficient of 2. Our prior for the Frisch labor supply elasticity covers some of the low values estimated in the microeconometrics literature, as well as a value of 2 advocated in the real-business-cycle literature based on steady-state considerations. The prior for the price-adjustment-cost parameter φ_p is specified indirectly through a prior for the slope $\kappa(\varphi_p)$ of the New Keynesian Phillips curve. This prior encompasses values that imply an essentially flat as well as a fairly steep Phillips curve. The prior for the wage rigidity is directly specified on φ_w and spans values ranging from 0 to 30. The priors for the asymmetry parameters ψ_p and ψ_w are centered at zero (symmetric adjustment costs) and have a large variance, meaning that the asymmetries could potentially be strong. We do not restrict the signs of ψ_p and ψ_w, i.e., we allow *a priori* for both downward and upward price and wage rigidities. The priors for the monetary

Table 2: Prior for Structural Parameters of DSGE Model

Parameter	Distribution	Para (1)	Para (2)
$400\left(\frac{1}{\beta}-1\right)$	Gamma	2.00	1.00
π^A	Gamma	3.00	1.00
γ^A	Gamma	2.00	1.50
$1/g$	Fixed	0.85	
τ	Gamma	2.00	1.00
ν	Gamma	0.50	1.00
$\kappa(\varphi_p)$	Gamma	0.30	0.20
φ_w	Gamma	15.0	7.50
ψ_w	Uniform	-200	200
ψ_p	Uniform	-300	300
ψ_1	Gamma	1.50	0.50
ψ_2	Gamma	0.20	0.10
ρ_r	Beta	0.50	0.20
ρ_g	Beta	0.80	0.10
ρ_z	Beta	0.20	0.10
ρ_p	Beta	0.60	0.20
$100\sigma_r$	InvGamma	0.20	2.00
$100\sigma_g$	InvGamma	0.75	2.00
$100\sigma_z$	Beta	0.75	2.00
$100\sigma_p$	Beta	0.75	2.00

Notes: Para (1) and Para (2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; and s and ν for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu,s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region.

policy rule coefficients are centered at 1.5 (reaction to inflation), 0.2 (output growth), and 0.5 (interest rate smoothing). Finally, we use priors for the parameters associated with the exogenous shock processes, then generate *a priori* reasonable magnitudes for the persistence and volatility of the observables.

The DSGE model presented in Section 5 is solved using a second-order approximation, which leads to a nonlinear state-space representation. We use the particle filter developed in Fernández-Villaverde and Rubio-Ramírez (2007) to evaluate the likelihood function of the DSGE model. To facilitate the likelihood evaluation with

Table 3: Posterior Estimates for DSGE Model Parameters

Parameter	1960:Q1-2007:Q4		1984:Q1-2007:Q4	
	Mean	90% Interval	Mean	90% Interval
$400\left(\frac{1}{\beta}-1\right)$	0.47	[0.08, 1.04]	1.88	[0.47, 3.01]
π^A	3.19	[2.57, 3.84]	3.34	[2.44, 4.32]
γ^A	2.04	[1.57, 2.77]	1.98	[1.59, 2.36]
τ	4.83	[2.75, 7.28]	4.10	[2.35, 6.06]
ν	0.37	[0.21, 0.52]	0.10	[0.05, 0.17]
$\kappa(\varphi_p)$	0.02	[0.01, 0.04]	0.21	[0.12, 0.35]
φ_w	18.7	[8.47, 38.1]	11.7	[5.34, 20.2]
ψ_w	67.4	[33.2, 99.5]	59.4	[21.7, 90.9]
ψ_p	150	[130, 175]	165	[130, 192]
ψ_1	1.77	[1.51, 2.12]	2.57	[1.93, 3.26]
ψ_2	1.41	[0.97, 1.85]	0.79	[0.42, 1.18]
ρ_r	0.81	[0.23, 0.72]	0.73	[0.64, 0.80]
ρ_g	0.95	[0.92, 0.98]	0.96	[0.94, 0.98]
ρ_z	0.48	[0.23, 0.72]	0.07	[0.01, 0.20]
ρ_p	0.89	[0.86, 0.94]	0.90	[0.76, 0.98]
$100\sigma_r$	0.17	[0.14, 0.21]	0.17	[0.12, 0.23]
$100\sigma_g$	0.88	[0.58, 1.29]	0.83	[0.49, 1.30]
$100\sigma_z$	0.44	[0.31, 0.62]	0.47	[0.38, 0.56]
$100\sigma_p$	2.62	[0.46, 7.23]	6.54	[4.56, 9.37]

Notes: For 90% credible interval we are reporting the 5th and 95th percentile of the posterior distribution.

the particle filter, the measurement equation contains mean-zero *iid* Gaussian measurement errors. The measurement error variances are set equal to 10% of the sample variances of GDP growth, inflation, interest rates, and nominal wage growth. Posterior inference is implemented with a single-block RWM algorithm, described in detail in An and Schorfheide (2007). Theoretical convergence properties of so-called particle MCMC approaches are established in Andrieu, Doucet, and Holenstein (2010).

Posterior summary statistics for the DSGE model parameters are reported in Table 3. The most interesting and important estimates are those of the asymmetry parameters in the price and wage adjustment cost function. The wage and price

rigidity estimates differ substantially across subsamples. For instance, the estimated slope of the New Keynesian Phillips curve is 0.02 for the 1960-2007 sample, whereas it increases to 0.2 for the post-1983 sample. Likewise, the estimated wage rigidity is larger over the long sample. The positive estimates of ψ_p and ψ_w imply that it is more expensive to lower prices and wages than to raise them and that the asymmetry in prices is more pronounced than in wages. The asymmetry of the adjustment costs is more pronounced for prices ($\hat{\psi}_p$ equals 150 and 165, respectively) than for wages ($\hat{\psi}_w$ equals 67 and 59, respectively).

Compared to the estimates reported by Kim and Ruge-Murcia (2009) and Abbritti and Fahr (2013), who report estimates of $\hat{\psi}_w = 3,844$ and $\hat{\psi}_w = 24,700$, respectively, our estimates of the ψ_w's are considerably smaller.[9] In our experience, such large values of ψ_w lead to a clear deterioration of the model's ability to track U.S. data. Moreover, the second-order solution of the DSGE model relies on a third-order approximation of the linex cost function that becomes very inaccurate for large values of ψ. In particular, we found that for values of ψ_w above 500 the adjustment costs for large positive wage changes (that lie in the support of the ergodic distribution) would become negative due to the polynomial approximation of the linex function.

We estimate the risk-aversion parameter τ to be fairly large, around 4, and the Frisch labor supply elasticity to be fairly low, ranging from 0.1 to 0.4. The estimates of ν are in line with those reported in Ríos-Rull, Schorfheide, Fuentes-Albero, Kryshko, and Santaeulalia-Llopis (2012). The policy rule coefficient estimates are similar to the ones reported elsewhere in the DSGE model literature. The coefficient ψ_1 on inflation

---

[9]Kim and Ruge-Murcia (2009) estimated their DSGE model simulated method of moments (SMM). While they also used consumption and hours-worked data in their estimation, the SMM objective function includes only second moments. The authors find that the covariance of consumption and hours worked, respectively, with wage growth playing a crucial role in their estimation. Abbritti and Fahr (2013) use a calibration approach to parameterize their model. Given their preferred calibration of the exogenous technology, discount factor, and monetary policy shocks, they find that a very large value of ψ_w is needed to match the volatility and skewness of wage growth observed in the data.

is larger for the post-1983 sample, which is consistent with the view that following the Volcker disinflation the Federal Reserve has responded more aggressively to inflation movements. The government spending shock, which should be viewed as a generic demand shock, is the most persistent among the serially correlated exogenous shocks: ρ_g is approximately 0.95. The estimated autocorrelation ρ_z of technology growth shock, which generates most of the serial correlation in output growth rates, drops from 0.48 for the long sample to 0.07 for the post-1983 sample.

6.2 Posterior Predictive Checks

We proceed by examining whether QAR(1,1) parameter estimates obtained from data that are simulated from the estimated DSGE model are similar to the estimates reported in Section 4 computed from actual data. This comparison is formalized through a posterior predictive check. The role of posterior predictive checks in Bayesian analysis is discussed in the textbooks by Lancaster (2004) and Geweke (2005) and reviewed in the context of the evaluation of DSGE models in Del Negro and Schorfheide (2011). The posterior predictive check is implemented with the following algorithm.

Posterior Predictive Check. Let $\theta^{(i)}$ denote the i'th draw from the posterior distribution of the DSGE model parameter θ.

1. For $i = 1$ to n:

   (i) Conditional on $\theta^{(i)}$ simulate a pre-sample of length T_0 and an estimation sample of size T from the DSGE model. The second-order approximated DSGE model is simulated using the pruning algorithm described in Kim, Kim, Schaumburg, and Sims (2008). A Gaussian *iid* measurement error is added to the simulated data. The measurement error variance is identical to the one imposed during the estimation of the DSGE model. Denote the simulated data by $Y_{-T_0+1:T}^{(i)}$.

(ii) Based on the simulated trajectory $Y^{(i)}_{-T_0+1:T}$ estimate the QAR(1,1) model as described in Section 4.1. The prior for the QAR(1,1) parameters is elicited from the pre-sample $Y^{(i)}_{-T_0+1:0}$ and the posterior is based on $Y^{(i)}_{1:T}$. Denote the posterior median estimates of the QAR parameters by $\mathcal{S}\big(Y^{(i)}_{-T_0+1:T}\big)$.

2. The empirical distribution of $\big\{\mathcal{S}\big(Y^{(i)}_{-T_0+1:T}\big)\big\}^n_{i=1}$ approximates the posterior predictive distribution of $\mathcal{S}|Y_{-T_0+1:T}$. Examine how far the actual value $\mathcal{S}(Y_{1:T})$, computed from U.S. data, lies in the tail of its predictive distribution. $\square$

The predictive check is carried out for each QAR(1,1) parameter estimate separately. The results are summarized in Figure 8. The top panel corresponds to the 1960-2007 sample, whereas the bottom panel contains the results from the 1984-2007 sample. The red dots signify the posterior median estimates obtained from U.S. data and correspond to the horizontal bars in Figure 1. The blue rectangles delimit the 90% credible intervals associated with the posterior predictive distributions, and the solid horizontal bars indicate the medians of the predictive distributions. The length of the credible intervals reflects both parameter uncertainty, i.e., the fact that each trajectory $Y^{(i)}_{-T_0+1:T}$ is generated from a different parameter draw $\theta^{(i)}$, and sampling uncertainty, meaning that if one were to hold the parameters θ fixed, the variability in the simulated finite-sample trajectories generates variability in posterior mean estimates. Because the posterior variance of the DSGE model parameters is fairly small, these intervals mostly capture sampling variability. Accordingly, they tend to be larger in the bottom panel (short sample) than in the top panel (long sample).

By and large, the QAR parameter estimates for output growth, wage growth, and inflation from model-generated data are very similar to those obtained from actual data, in the sense that most actual estimates do not fall far in the tails of their respective posterior predictive distributions. Only interest rates exhibit large discrepancies between actual and model-implied estimates of the QAR(1,1) parameters.

Overall, the estimated DSGE model does not generate very strong nonlinearities. Posterior predictive distributions for $\hat{\phi}_2$ and $\hat{\gamma}$ typically cover both positive and nega-

Figure 8: Predictive Checks Based on QAR(1,1)Estimates

1960-2007 Sample

1984-2007 Sample
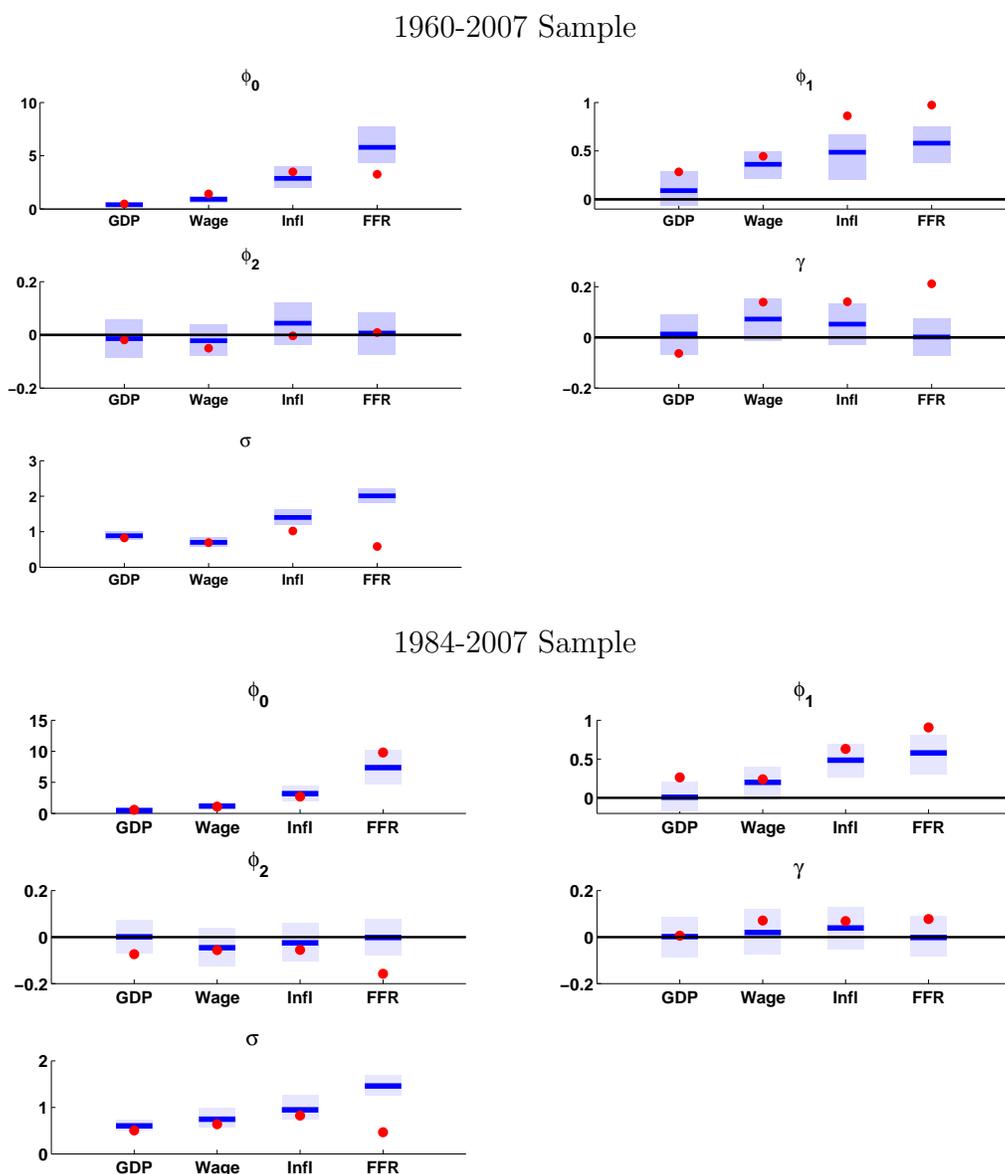
Notes: Dots correspond to posterior median estimates from U.S. data. Solid horizontal lines indicate medians of posterior predictive distributions for parameter estimates and the boxes indicate 90% credible intervals associated with the posterior predictive distributions.

tive values. The only exceptions are the predictive distributions for wage growth and inflation $\hat{\gamma}$ conditional on the 1960-2007 sample, which imply that $\hat{\gamma}$ is positive. Recall from Table 3 that, for this sample, we estimate sizable adjustment costs ($\hat{\kappa} = 0.02$ and

$\hat{\varphi}_w = 18.7$). Moreover, the asymmetry parameter estimates are substantially larger than zero: $\hat{\psi}_p = 150$ and $\hat{\psi}_w = 67.4$. The model-implied positive estimates of γ imply that high inflation and wage-growth rates are associated with high levels of volatility, which describes the experience of the U.S. economy in the 1970s and early 1980s. However, the nonlinear inflation and wage dynamics do not generate any spillovers to nonlinearities in GDP growth or the interest rate. Figure 8 indicates that the predictive distributions for the corresponding $\hat{\phi}_2$ and $\hat{\gamma}$ are centered at zero. For the 1984-2007 sample the overall magnitude of the estimated adjustment costs is smaller, which flattens the adjustment cost functions, makes the asymmetries less important for equilibrium dynamics, and shifts the predictive distribution for the inflation and wage growth $\hat{\gamma}$'s toward zero.

Two types of nonlinearities present in the data are not predicted by the estimated DSGE model. First, for the short sample $\hat{\phi}_2$ for GDP growth is negative, because the post-1983 sample exhibits a pronounced drop in output growth during the recessions but does not feature positive growth rates of similar magnitudes in early parts of expansions. Second, the interest rate exhibits strong nonlinearities in the data, i.e., a large positive $\hat{\gamma}$ in the 1960-2007 sample and a large negative $\hat{\phi}_2$ in the 1984-2007 sample, which the DSGE model is unable to reproduce.

To sum up, of the nonlinearities we identified in Section 4, the only ones the DSGE model seems able to deliver are the conditional heteroskedasticity in inflation and nominal wage growth. It is able to do so by relying on the asymmetric adjustment costs that penalize downward adjustments more than upward adjustments. However, while ex-ante reasonable, these asymmetries in prices do not spill over to quantities. Moreover, since the interest rate feedback rule in the model does not feature any asymmetries, which would result from the central bank having an asymmetric loss function, and since there are no asymmetries in GDP growth in the model, the policy instrument does not display the asymmetry we identified in the data.

6.3   The Role of Asymmetric Adjustment Costs

To further study the role of asymmetric adjustment costs in generating nonlinear wage and inflation dynamics, we repeat the predictive checks based on $\hat{\phi}_2$ and $\hat{\gamma}$ for alternative choices of ψ_p and ψ_w. We focus on the 1960-2007 sample because the nonlinearities are more pronounced there than in the post-1983 sample. For each draw $\theta^{(i)}$ from the posterior distribution of the DSGE model parameters, we replace $\psi_p^{(i)}$ and $\psi_w^{(i)}$ with alternative values $\bar{\psi}_p$ and $\bar{\psi}_w$. In particular, we consider an elimination of the asymmetries, i.e., $\bar{\psi}_p = \bar{\psi}_w = 0$ and an increase to $\bar{\psi}_p = \bar{\psi}_w = 300$. The results are plotted in Figure 9. A decrease of the asymmetry in the adjustment costs moves the predictive distributions of $\hat{\phi}_2$ and $\hat{\gamma}$ toward zero, whereas an increase shifts them further away from zero. Relative to the overall width of the predictive intervals the location shifts are fairly small. This highlights that a precise measurement of nonlinearities is very difficult using quarterly observations.

For nominal wage growth, the increase in the asymmetry parameters essentially eliminates the gap between the median of the posterior predictive distributions for $\hat{\phi}_2$ and $\hat{\gamma}$ and the estimates obtained from actual data, which are -0.05 and 0.14, respectively. For inflation, the medians of the predictive distributions for $\hat{\phi}_2$ and $\hat{\gamma}$ shift slightly upward, toward 0.05 and 0.06, respectively. This implies that the actual value of $\hat{\phi}_2$ lies further in the tail of the predictive distribution if ψ_w is increased, whereas the actual value of $\hat{\gamma}$ is less far in the tails. While an increase of ψ_w improves the outcome of the predictive check constructed from the QAR parameter estimates for nominal wage growth, if judged from the overall posterior distribution, the increased asymmetries lead to a deterioration of fit in other dimensions of the model, which is why the posterior estimates for ψ_p and ψ_w are only about 150 and 68, respectively.

Figure 9: Effect of Adjustment Costs on Nonlinearities
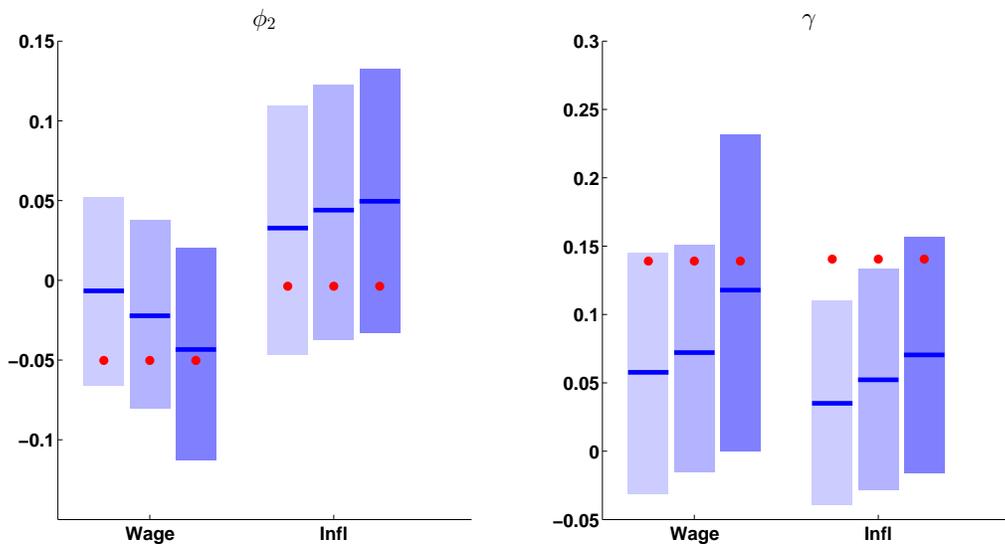
Notes: 1960-2007 sample. Box plots of posterior predictive distribution for ϕ_2 and γ estimates for different parameter values of the adjustment cost functions. No asymmetric costs is $\psi_p = \psi_w = 0$ (light blue); high asymmetric costs is $\psi_p = \psi_w = 300$ (dark blue). Large dots correspond to posterior median estimates based on U.S. data.

7 Conclusion

Building on the specification of generalized autoregressive models, bilinear models, and LARCH models, this paper uses a perturbation approximation of a nonlinear difference equation to obtain a new class of nonlinear time-series models that can be used to assess nonlinear DSGE models. We use these univariate QAR(1,1) models to identify nonlinearities in the U.S. data and to construct predictive checks to assess a DSGE models ability to capture nonlinearities that are present in the data. The QAR(1,1) estimates obtained from U.S. data highlight nonlinearities in output growth, inflation, nominal wage growth, and interest rate dynamics. Output growth displays sharp declines and slow recoveries in the post-1983 sample. Inflation and nominal wage growth both display conditional heteroskedasticity in the 1960-2007 sample. Finally, downward adjustments in the federal funds rate seem to be easier than upward adjustments in the post-1983 sample.

Among the nonlinearities identified through estimation of the QAR models, the only ones our estimated DSGE model seems able to capture are the conditional heteroskedasticity in inflation and nominal wage growth. The model does so by relying on the asymmetric adjustment costs that penalize downward adjustments more than upward adjustments. The model is not able to generate the apparent nonlinearities in output growth and the federal funds rate.

The tools developed in this paper can be used to identify nonlinearities in any time series: Doing this for other key series, such as labor market variables in the U.S., and for key variables in other countries will be a useful exercise. The predictive checks simply require a simulation from the model and can be applied to any model, whether or not it is estimated, and should be a part of the toolbox for researchers working with DSGE models. Finally, we leave to future research multivariate extensions of the QAR model, where the main challenge is to cope with the dimensionality of the model.

References

ABBRITTI, M., AND S. FAHR (2013): "Downward Wage Rigidity and Business Cycle Asymmetries," *Journal of Monetary Economics*, forthcoming.

AN, S. (2007): "Bayesian Estimation of DSGE Models: Lessons from Second-Order Approximations," *Manuscript, Singapore Management University*, Available at http://www.mysmu.edu/faculty/sungbae/.

AN, S., AND F. SCHORFHEIDE (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4), 113–172.

ANDREASEN, M. M., J. FERNÁNDEZ-VILLAVERDE, AND J. F. RUBIO-RAMÍREZ (2013): "The Pruned State-Space System for Nonlinear DSGE Models: Theory and Empircal Applications," *Manuscript, University of Pennsylvania*, (4).

ANDRIEU, C., A. DOUCET, AND R. HOLENSTEIN (2010): "Particle Markov Chain Monte Carlo Methods," *Journal of the Royal Statistical Society Series B*, 72(3), 269–342.

ARUOBA, S. B., AND F. SCHORFHEIDE (2013): "Macroeconomic Dynamics Near the ZLB: A Tale of Two Equilibria," *NBER Working Paper*, 19248.

BARATTIERI, A., S. BASU, AND P. GOTTSCHALK (2010): "Some Evidence on the Importance of Sticky Wages," *NBER Working Paper*.

BOCOLA, L. (2013): "The Pass-Through of Sovereign Risk," *Manuscript, University of Pennsylvania*.

BRUNNER, A. D. (1997): "On the Dynamic Properties of Asymmetric Models of Real GNP," *Review of Economics and Statistics*, 79(2), 321–326.

CANOVA, F. (1994): "Statistical Inference in Calibrated Models," *Journal of Applied Econometrics*, 9, S123–144.

CHANG, Y., T. DOH, AND F. SCHORFHEIDE (2007): "Non-stationary Hours in a DSGE Model," *Journal of Money, Credit, and Banking*, 39(6), 1357–1373.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

COGLEY, T., AND J. M. NASON (1994): "Testing the Implications of Long-Run Neutrality for Monetary Business Cycle Models," *Journal of Applied Econometrics*, 9, S37–70.

COGLEY, T., AND T. J. SARGENT (2002): "Evolving Post-World War II U.S. Inflation Dynamics," in *NBER Macroeconomics Annual 2001*, ed. by B. S. Bernanke, and K. Rogoff, vol. 16, pp. 331–88. MIT Press, Cambridge.

CUKIERMAN, A., AND A. MUSCATELLI (2008): "Nonlinear Taylor Rules and Asymmetric Preferences in Central Banking: Evidence from the United Kingdom and the United States," *The B.E. Journal of Macroeconomics*, 8(1), 1–31.

DALY, M., B. HOBIJN, AND B. LUCKING (2012): "Why Has Wage Growth Stayed Strong?," *Federal Reserve Bank of San Francisco Economic Letters*, 10.

DEL NEGRO, M., AND F. SCHORFHEIDE (2011): "Bayesian Macroeconometrics," in *The Oxford Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. van Dijk, pp. 293–389. Oxford University Press.

DEL NEGRO, M., F. SCHORFHEIDE, F. SMETS, AND R. WOUTERS (2007): "On the Fit of New Keynesian Models," *Journal of Business and Economic Statistics*, 25(2), 123–162.

DOLADO, J., R. MARIA-DOLORES, AND F. J. RUGE-MURCIA (2004): "Nonlinear Monetary Policy Rules: Some New Evidence for the U.S," *Studies in Nonlinear Dynamics & Econometrics*, 8(3), 1–34.

ENGLE, R. F., D. M. LILIEN, AND R. P. ROBINS (1987): "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica*, 55(2), 391–407.

FERNÁNDEZ-VILLAVERDE, J., AND J. F. RUBIO-RAMÍREZ (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74(4), 1059–1087.

FERNÁNDEZ-VILLAVERDE, J., J. F. RUBIO-RAMÍREZ, T. J. SARGENT, AND M. W. WATSON (2007): "ABCs (and Ds) of Understanding VARs," *American Economic Review*, 97(3), 1021–1026.

GEWEKE, J. (2005): *Contemporary Bayesian Econometrics and Statistics*. John Wiley & Sons, Hoboken, New Jersey.

GIRAITIS, L., P. ROBINSON, AND D. SURGAILIS (2000): "A Model for Long Memory Conditional Heteroskedasticity," *Annals of Applied Probability*, 10(3), 1002–1024.

GOTTSCHALK, P. (2005): "Downward Nominal-Wage Flexibility: Real or Measurement Error," *Review of Economics and Statistics*, 87(3), 556–568.

GRANGER, C. W., AND A. P. ANDERSEN (1978): *Introduction to Bilinear Time Series Models*. Vandenhoeck & Ruprecht, Göttingen.

GRIER, K. B., AND M. J. PERRY (1996): "Inflation, Inflation Uncertainty, and Relative Price Dispersion: Evidence from Bivariate GARCH-M Models," *Journal of Monetary Economics*, 38, 391–405.

——— (2000): "The Effects of Real and Nominal Uncertainty on Inflation and Output Growth: Some GARCH-M Evidence," *Journal of Applied Econometrics*, 15, 45–58.

GUST, C., D. LOPEZ-SALIDO, AND M. E. SMITH (2012): "The Empirical Implications of the Interest-Rate Lower Bound," *Manuscript, Federal Reserve Board*.

HAMILTON, J. D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–384.

HOLMES, M. (1995): *Introduction to Perturbation Methods*. Cambridge University Press.

KIM, J., S. KIM, E. SCHAUMBURG, AND C. A. SIMS (2008): "Calculating and Using Second-Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models," *Journal of Economic Dynamics and Control*, 32, 3397–3414.

KIM, J., AND F. RUGE-MURCIA (2009): "How Much Inflation Is Necessary to Grease the Wheels?," *Journal of Monetary Economics*, 56, 365–377.

KOENKER, R., AND Z. XIAO (2006): "Quantile Autoregression," *Journal of the American Statistical Association*, 101(475), 980–990.

LAN, H., AND A. MEYER-GOHDE (2013): "Solving DSGE Models with a Nonlinear Moving Average," *Journal of Economic Dynamics and Control*, forthcoming.

LANCASTER, T. (2004): *An Introduction to Modern Bayesian Econometrics*. Blackwell Publishing.

LOMBARDO, G. (2011): "On Approximating DSGE Models by Series Expansions," Available at https://sites.google.com/site/giovannilombardohomepage/.

LUBIK, T. A., AND F. SCHORFHEIDE (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94(1), 190–217.

MCKAY, A., AND R. REIS (2008): "The Brevity and Violence of Contractions and Expansions," *Journal of Monetary Economics*, 55, 738–751.

MITTNIK, S. (1990): "Modeling Nonlinear Processes with Generalized Autoregressions," *Applied Mathematical Letters*, 3(4), 71–74.

PRIMICERI, G. E. (2005): "Time Varying VARs and Monetary Policy," *Review of Economic Studies*, 72(3), 821–852.

RABANAL, P., AND J. F. RUBIO-RAMÍREZ (2005): "Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach," *Journal of Monetary Economics*, 52(6), 1151–1166.

RAO, S. T. (1981): "On the Theory of Bilinear Time Series Models," *Journal of the Royal Statistical Society Series B*, 43(2), 244–255.

RÍOS-RULL, J.-V., F. SCHORFHEIDE, C. FUENTES-ALBERO, M. KRYSHKO, AND R. SANTAEULALIA-LLOPIS (2012): "Methods versus Substance: Measuring the Effects of Technology Shocks," *Journal of Monetary Economics*, 59(8), 826–845.

SCHMITT-GROHE, S., AND M. URIBE (2013): "The Making of a Great Contraction with a Liquidity Trap and a Jobless Recovery," *Manuscript, Columbia University*.

SCHORFHEIDE, F. (2000): "Loss Function-Based Evaluation of DSGE Model," *Journal of Applied Econometrics*, 15(6), 645–670.

SIMS, C. A., AND T. ZHA (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96(1), 54–81.

SMETS, F., AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

SMITH, A. A. (1993): "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8, S63–S84.

TERÄSVIRTA, T. (1994): "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89(425), 208–218.

TONG, H., AND K. LIM (1980): "Threshold Autoregression, Limit Cycles and Cyclical Data," *Journal of the Royal Statistical Society Series B*, 42(3), 245–292.

VOLTERRA, V. (1930): *Theory of Functionals and of Integral and Integro-Differential Equations.* Dover Publications.

Online Appendix: Assessing DSGE Model Nonlinearities

S. Borağan Aruoba, Luigi Bocola, and Frank Schorfheide

This Version: November 26, 2013

A   QAR(1,1) Model

This section shows how to derive important moments for the QAR(1,1) model given by

$$y_t = \phi_1 y_{t-1} + \phi_2 s_{t-1}^2 + (1 + \gamma s_{t-1})\sigma u_t, \quad u_t \sim iidN(0,1) \tag{A.1}$$

$$s_t = \phi_1 s_{t-1} + \sigma u_t, \quad |\phi_1| < 1 \tag{A.2}$$

by exploiting the recursively linear structure of the model. The model corresponds to (9) in the main text. To simplify the presentation, we dropped the tildes for ϕ_2, γ, and s.

A.1   Moments

We now derive the time-invariant mean and autocovariances for y_t, assuming the process is stationary and was initialized in the infinite past. Due to the recursively linear structure of the model, we begin with the derivation of the moments of s_t.

Moments of s_t. The process s_t in (A.2) is linear and has a moving average representation of the form

$$s_t = \sigma \sum_{j=0}^{\infty} \phi_1^j u_{t-j}.$$

The mean and the autocovariances of s_t are given by

$$\mathbb{E}[s_t] = 0, \quad \mu_{s^2} = \mathbb{E}[s_t^2] = \frac{\sigma^2}{1 - \phi_1^2}, \quad \mathbb{E}[s_t s_{t-h}] = \phi_1^h \mu_{s^2}.$$

Since the innovations u_t are *iid* standard normal variates, we obtain the following third and fourth moments:

$$\mathbb{E}[s_t^3] = \sum_{j=0}^{\infty} \phi_1^{3j} \mathbb{E}[u_{t-j}^3] = 0, \quad \mathbb{E}[s_t^4] = \sum_{j=0}^{\infty} \phi_1^{4j} \mathbb{E}[u_{t-j}^4] = \frac{3\sigma^4}{1-\phi_1^4}.$$

Mean of y_t. Taking expectations on both sides of (A.1) we obtain

$$\mathbb{E}[y_t] = \phi_1 \mathbb{E}[y_{t-1}] + \phi_2 \mu_{s^2} + (1 + \gamma\mathbb{E}[s_{t-1}])\sigma\mathbb{E}[u_t] = \phi_1\mathbb{E}[y_t] + \frac{\phi_2\sigma^2}{1-\phi_1^2}.$$

Here we used the expression for μ_{s^2} obtained previously as well as the fact that u_t and s_{t-1} are independent. In turn,

$$\mu_y = \mathbb{E}[y_t] = \frac{\phi_2\sigma^2}{(1-\phi_1)(1-\phi_1^2)}. \tag{A.3}$$

Variance of y_t. Consider the centered second moment of y_t:

$$
\begin{aligned}
\mathbb{V}[y_t] &= \mathbb{E}\left[(\phi_1(y_{t-1}-\mu_y) + \phi_2(s_{t-1}^2 - \mu_{s^2}) + \sigma(1+\gamma s_{t-1})u_t)^2\right] \\
&= \mathbb{E}\Big[\phi_1^2(y_{t-1}-\mu_y)^2 + \phi_2^2(s_{t-1}^2-\mu_{s^2})^2 + \sigma^2(1+\gamma s_{t-1})^2 u_t^2 \\
&\quad 2\phi_1\phi_2(y_{t-1}-\mu_y)(s_{t-1}^2-\mu_{s^2}) + 2\phi_2\sigma(s_{t-1}^2-\mu_{s^2})(1+\gamma s_{t-1})u_t \\
&\quad +2\phi_1\sigma(1+\gamma s_{t-1})(y_{t-1}-\mu_y)u_t\Big] \\
&= \phi_1^2\mathbb{E}[(y_{t-1}-\mu_y)^2] + \phi_2^2\mathbb{E}[(s_{t-1}^2-\mu_{s^2})^2] + \sigma^2(1+\gamma^2\mu_{s^2}) \\
&\quad +2\phi_1\phi_2\mathbb{E}[(y_{t-1}-\mu_y)(s_{t-1}^2-\mu_{s^2})].
\end{aligned}
$$

The time-invariant solution is

$$\mathbb{V}[y_t] = \frac{1}{1-\phi_1^2}\left[\phi_2^2\mathbb{V}[s_t^2] + \sigma^2(1+\gamma^2\mathbb{E}[s_t^2]) + 2\phi_1\phi_2 Cov[y_t, s_t^2]\right],$$

where

$$
\begin{aligned}
Cov[y_t, s_t^2] &= \mathbb{E}\Big[(\phi_1(y_{t-1}-\mu_y) + \phi_2(s_{t-1}^2-\mu_{s^2}) + (1+\gamma s_{t-1})\sigma u_t) \\
&\quad \times(\phi_1^2(s_{t-1}^2-\mu_{s^2}) + 2\phi_1\sigma s_{t-1}u_t + \sigma^2(u_t^2-1))\Big] \\
&= \phi_1^3\mathbb{E}[(y_{t-1}-\mu_y)(s_{t-1}^2-\mu_{s^2})] + \phi_1^2\phi_2\mathbb{E}[(s_{t-1}^2-\mu_{s^2})^2] \\
&\quad +2\phi_1\gamma\sigma^2\mu_{s^2},
\end{aligned}
$$

which implies

$$Cov[y_t, s_t^2] = \frac{1}{1 - \phi_1^3}\left[\phi_1^2\phi_2\mathbb{V}[s_t^2] + 2\phi_1\gamma\sigma^2\mathbb{E}[s_t^2]\right].$$

Interestingly,

$$
\begin{aligned}
Cov[y_t, s_t] &= \mathbb{E}\left[(\phi_1(y_{t-1} - \mu_y) + \phi_2(s_{t-1}^2 - \mu_{s^2}) + (1 + \gamma s_{t-1})\sigma u_t)(\phi_1 s_{t-1} + \sigma u_t)\right] \\
&= \phi_1^2 Cov[y_{t-1}, s_{t-1}] + \sigma^2.
\end{aligned}
$$

All other terms drop out because $\mathbb{E}[u_t] = \mathbb{E}[s_t] = \mathbb{E}[s_t^3] = 0$. Thus, solving for the time-invariant solution leads to the "first order" variance expression

$$Cov[y_t, s_t] = \mathbb{E}[s_t^2] = \frac{\sigma^2}{1 - \phi_1^2}.$$

Autocovariances of y_t. Consider $\mathbb{E}[(y_t - \mu_y)(y_{t-1} - \mu_y)]$:

$$
\begin{aligned}
Cov[y_t, y_{t-1}] &= \mathbb{E}\left[(\phi_1(y_{t-1} - \mu_y) + \phi_2(s_{t-1}^2 - \mu_{s^2}) + (1 + \gamma s_{t-1})\sigma u_t)(y_{t-1} - \mu_y)\right] \\
&= \phi_1\mathbb{V}[y_{t-1}] + \phi_2 Cov[y_{t-1}, s_{t-1}^2].
\end{aligned}
$$

In general, higher-order autocovariances can be computed recursively:

$$
\begin{aligned}
Cov[y_t, y_{t-h}] &= \mathbb{E}\left[(\phi_1(y_{t-1} - \mu_y) + \phi_2(s_{t-1}^2 - \mu_{s^2}) + (1 + \gamma s_{t-1})\sigma u_t)(y_{t-h} - \mu_y)\right] \\
&= \phi_1 Cov[y_{t-1}, y_{t-h}] + \phi_2 Cov[y_{t-h}, s_{t-1}^2].
\end{aligned}
$$

The term $Cov[y_{t-h}, s_{t-1}^2]$ can also be calculated recursively:

$$
\begin{aligned}
Cov[y_{t-h}, s_{t-1}^2] &= \mathbb{E}\left[(y_{t-h} - \mu_y)(\phi_1^2(s_{t-2} - \mathbb{E}[s_{t-2}^2]) + 2\phi_1 s_{t-2}\sigma u_{t-1} + \sigma(u_{t-1})^2 - 1)\right] \\
&= \phi_1^2 Cov[y_{t-h}, s_{t-2}^2].
\end{aligned}
$$

A.2 Initialization and Identification

In order to compute the likelihood function recursively, it is necessary to initialize s_0. We write the joint distribution of observables, initial state, and parameters as

$$p(Y_{0:T}, \theta, s_0) = p(Y_{1:T}|y_0, s_0, \theta)p(y_0, s_0|\theta)p(\theta)$$

and use MCMC methods to generate draws from the posterior

$$p(\theta, s_0|Y_{0:T}) \propto p(Y_{1:T}|y_0, s_0, \theta)p(y_0, s_0|\theta)p(\theta).$$

We will approximate the distribution of (y_0, s_0) using a normal distribution

$$\begin{bmatrix} y_0 \\ s_0 \end{bmatrix}\Big|\theta \sim N\left(\begin{bmatrix} \mu_y \\ \mu_s \end{bmatrix}, \begin{bmatrix} \Sigma_{yy} & \Sigma_{ys} \\ \Sigma_{sy} & \Sigma_{ss} \end{bmatrix}\right). \tag{A.4}$$

The moments of this normal distribution are calculated as follows. We will assume that the system was in its steady state in period $t = -T_*$, i.e., $s_{-T_*} = 0$ and $y_{-T_*} = \phi_0$. In principle, T_* could be infinite, but this will create some problems if $\phi_1 = 1$. In order to simplify the time subscripts a bit, we shift the time index by T_* periods. Starting from $s_0 = 0$ and $y_0 = \phi_0$, we will calculate the first and second moments of y_t, s_t, and s_t^2 recursively, starting with

$$\mathbb{E}[s_0] = 0, \ \mathbb{E}[y_0] = \phi_0, \ \mathbb{V}[s_0] = 0, \ \mathbb{V}[y_0] = 0, \tag{A.5}$$

$$\text{Cov}[y_0, s_0] = 0, \ \text{Cov}[y_0, s_0^2], \ \mathbb{V}[s_0^2] = 0.$$

The process for s_t is linear autoregressive of order one and we obtain

$$\mathbb{E}[s_t] = \phi_1\mathbb{E}[s_{t-1}], \quad \mathbb{V}[s_t] = \phi_1^2\mathbb{V}[s_{t-1}] + \sigma^2. \tag{A.6}$$

Since the innovations ϵ_t are *iid* standard normal variates, we see that the third moment is zero:

$$\mathbb{E}[s_t^3] = \sum_{j=0}^{t-1} \phi_1^{3j}\mathbb{E}[\epsilon_{t-j}^3] = 0.$$

Now consider

$$\begin{aligned} \mathbb{V}[s_t^2] &= \mathbb{E}[(s_t^2 - \mathbb{V}[s_t])^2] & \text{(A.7)} \\ &= \mathbb{E}[(\phi_1^2(s_{t-1}^2 - \mathbb{V}[s_{t-1}]) + 2\phi_1 s_{t-1}\sigma\epsilon_t + \sigma^2(\epsilon_t^2 - 1))^2] \\ &= \phi_1^4\mathbb{V}[s_{t-1}^2] + 4\phi_1^2\sigma^2\mathbb{V}[s_{t-1}] + 2\sigma^4. \end{aligned}$$

A formula for the mean of y_t is obtained by taking expectations of the observation equation:

$$\mathbb{E}[y_t] = \phi_0(1 - \phi_1) + \phi_1\mathbb{E}[y_{t-1}] + \phi_2\mathbb{V}[s_{t-1}]. \tag{A.8}$$

The covariance between y_t and s_t is given by

$$
\begin{aligned}
\mathrm{Cov}[y_t, s_t] &= \mathbb{E}[(y_t - \mathbb{E}[y_t])s_t] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(A.9)}\\
&= \mathbb{E}\big[\big(\phi_1(y_{t-1} - \mathbb{E}[y_{t-1}]) + \phi_2(s_{t-1}^2 - \mathbb{E}[s_{t-1}^2]) + (1 + \gamma s_{t-1})\sigma\epsilon_t\big)\big(\phi_1 s_{t-1} + \sigma\epsilon_t\big)\big]\\
&= \phi_1^2 \mathrm{Cov}[y_{t-1}, s_{t-1}] + \sigma^2.
\end{aligned}
$$

All other terms drop out because the first and third moments of s_{t-1} and ϵ_t are equal to zero. The covariance between y_t and s_t^2 is given by

$$
\begin{aligned}
\mathrm{Cov}[y_t, s_t^2] &= \mathbb{E}[(y_t - \mathbb{E}[y_t])(s_t^2 - \mathbb{V}[s_t])] \qquad\qquad\qquad\qquad\qquad\text{(A.10)}\\
&= \mathbb{E}\big[\big(\phi_1(y_{t-1} - \mathbb{E}[y_{t-1}]) + \phi_2(s_{t-1}^2 - \mathbb{V}[s_{t-1}]) + (1 + \gamma s_{t-1})\sigma\epsilon_t\big)\\
&\qquad \times(\phi_1^2(s_{t-1}^2 - \mathbb{V}[s_{t-1}]) + 2\phi_1\sigma s_{t-1}\epsilon_t + \sigma^2(\epsilon_t^2 - 1))\big]\\
&= \phi_1^3 \mathrm{Cov}[y_{t-1}, s_{t-1}^2] + \phi_1^2\phi_2\mathbb{V}[s_{t-1}^2] + 2\phi_1\gamma\sigma^2\mathbb{E}[s_{t-1}^2].
\end{aligned}
$$

The variance of y_t can be computed as follows:

$$
\begin{aligned}
\mathbb{V}[y_t] &= \mathbb{E}\big[(\phi_1(y_{t-1} - \mathbb{E}[y_{t-1}] + \phi_2(s_{t-1}^2 - \mathbb{V}[s_{t-1}]) + \sigma(1 + \gamma s_{t-1})\epsilon_t)^2\big] \text{ (A.11)}\\
&= \phi_1^2\mathbb{V}[y_{t-1}] + \phi_2^2\mathbb{V}[s_{t-1}^2] + \sigma^2(1 + \gamma^2\mathbb{V}[s_{t-1}])\\
&\quad + 2\phi_1\phi_2\mathrm{Cov}[y_{t-1}, s_{t-1}^2].
\end{aligned}
$$

We can iterate Equations (A.6) to (A.11) forward for T_* periods to obtain the moments for the initial distribution of (y_0, s_0) in (A.4).

Note, that for $\gamma = \phi_2 = 0$, s_0 and y_0 become perfectly correlated conditional on θ since for a linear model $y_0 = s_0 + \phi_0$. This may affect our posterior sampler when we include s_0 into the parameter vector. To avoid the singularity we add a small constant to the covariance matrix of (y_0, s_0).

A.3 MCMC Implementation

The RWM algorithm mentioned in Section 3.3 is used to implement the posterior inference. Using a preliminary covariance for the proposal distribution in the RWM

algorithm that is constructed from the prior variance of the QAR parameters, we generate an initial 100,000 draws from the posterior. Based on the last 50,000 draws, we compute a covariance matrix that replaces the preliminary covariance matrix of the proposal distribution. We then continue the chain, generating an additional 60,000 draws and retaining the last 50,000 to construct summary statistics for the posterior.

A.4 Detailed Estimation Results

Table A-1: Prior Distribution for QAR(1,1) Model, Samples Starting in 1960

	GDP Growth	Wage Growth	Inflation	Fed Funds Rate
ϕ_0	$N(0.48, 2)$	$N(1.18, 2)$	$N(2.38, 2)$	$N(2.50, 2)$
ϕ_1	$N^\dagger(0.36, 0.5)$	$N^\dagger(-0.02, 0.5)$	$N^\dagger(0.00, 0.5)$	$N^\dagger(0.66, 0.5)$
σ	$IG(1.42, 4)$	$IG(0.82, 4)$	$IG(1.87, 4)$	$IG(0.58, 4)$
ϕ_2	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$
γ	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$

Notes: (†) The prior for ϕ_1 is truncated to ensure stationarity. The IG distribution is parameterized such that $p_{IG}(\sigma|\nu, s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$.

Table A-2: Prior Distribution for QAR(1,1) Model, Samples Starting in 1984

	GDP Growth	Wage Growth	Inflation	Fed Funds Rate
ϕ_0	$N(0.43, 2)$	$N(1.58, 2)$	$N(4.38, 2)$	$N(6.08, 2)$
ϕ_1	$N^\dagger(0.28, 0.5)$	$N^\dagger(0.34, 0.5)$	$N^\dagger(0.85, 0.5)$	$N^\dagger(0.94, 0.5)$
σ	$IG(1.33, 4)$	$IG(0.88, 4)$	$IG(1.83, 4)$	$IG(1.45, 4)$
ϕ_2	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$
γ	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$	$N(0, 0.1)$

Notes: (†) The prior for ϕ_1 is truncated to ensure stationarity. The IG distribution is parameterized such that $p_{IG}(\sigma|\nu, s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$.

Table A-3: Posterior Estimates for QAR(1,1) Model, 1960:Q1 to 1983:Q4

Data	ϕ_0	ϕ_1	ϕ_2	γ	σ	s_0
GDP	0.42	0.28	-0.02	-0.05	1.16	1.42
	[0.11 , 0.69]	[0.11 , 0.46]	[-0.14 , 0.09]	[-0.17 , 0.06]	[0.91 , 1.53]	[1.02 , 1.85]
WAGE	1.75	0.41	-0.05	0.04	0.52	0.89
	[1.49 , 1.98]	[0.23 , 0.58]	[-0.13 , 0.04]	[-0.05 , 0.15]	[0.40 , 0.68]	[0.63 , 1.15]
INFL	4.24	0.87	-0.01	0.16	1.52	-1.97
	[2.28 , 5.84]	[0.80 , 0.95]	[-0.08 , 0.07]	[0.04 , 0.27]	[1.08 , 2.12]	[-4.68 , 0.79]
FFR	4.84	0.92	0.02	0.38	0.62	-1.56
	[0.86 , 6.75]	[0.88 , 0.96]	[-0.05 , 0.05]	[0.30 , 0.47]	[0.41 , 1.00]	[-4.21 , 0.14]

Notes: We report posterior means and 90% equal-tail-probability credible sets in brackets.

Table A-4: Posterior Estimates for QAR(1,1) Model, 1960:Q1 to 2007:Q4

Data	ϕ_0	ϕ_1	ϕ_2	γ	σ	s_0
GDP	0.48	0.29	-0.02	-0.06	0.69	1.37
	[0.33 , 0.63]	[0.16 , 0.41]	[-0.07 , 0.04]	[-0.13 , 0.01]	[0.58 , 0.82]	[1.19 , 1.56]
WAGE	1.41	0.44	-0.03	0.12	0.48	1.22
	[1.25 , 1.59]	[0.33 , 0.55]	[-0.09 , 0.02]	[0.05 , 0.20]	[0.40 , 0.57]	[1.00 , 1.42]
INFL	3.51	0.85	-0.01	0.23	1.06	-1.31
	[2.74 , 4.47]	[0.79 , 0.91]	[-0.06 , 0.05]	[0.16 , 0.31]	[0.81 , 1.38]	[-2.90 , 0.31]
FFR	2.96	0.96	0.04	0.44	0.28	-0.74
	[2.16 , 4.16]	[0.95 , 0.97]	[0.02 , 0.06]	[0.37 , 0.52]	[0.22 , 0.42]	[-1.27 , 0.45]

Notes: We report posterior means and 90% equal-tail-probability credible sets in brackets.

Table A-5: Posterior Estimates for QAR(1,1) Model, 1960:Q1 to 2012:Q4

Data	ϕ_0	ϕ_1	ϕ_2	γ	σ	s_0
GDP	0.45	0.33	-0.03	-0.07	0.68	1.41
	[0.28 , 0.60]	[0.22 , 0.44]	[-0.08 , 0.03]	[-0.14 , 0.00]	[0.58 , 0.81]	[1.19 , 1.61]
WAGE	1.29	0.43	-0.01	0.08	0.54	1.31
	[1.12 , 1.46]	[0.32 , 0.53]	[-0.06 , 0.04]	[0.01 , 0.15]	[0.46 , 0.63]	[1.11 , 1.50]
INFL	3.23	0.84	0.02	0.22	1.09	-1.26
	[2.55 , 4.16]	[0.78 , 0.90]	[-0.04 , 0.09]	[0.15 , 0.30]	[0.87 , 1.36]	[-2.82 , 0.22]
FFR	3.54	0.96	-0.01	0.41	0.22	0.43
	[2.29 , 5.06]	[0.94 , 0.97]	[-0.02 , 0.00]	[0.33 , 0.50]	[0.13 , 0.37]	[-0.94 , 1.47]

Notes: We report posterior means and 90% equal-tail-probability credible sets in brackets.

Table A-6: Posterior Estimates for QAR(1,1) Model, 1984:Q1 to 2007:Q4

Data	ϕ_0	ϕ_1	ϕ_2	γ	σ	s_0
GDP	0.57	0.26	-0.07	0.01	0.25	1.06
	[0.44, 0.70]	[0.10 , 0.44]	[-0.13 , -0.02]	[-0.10 , 0.11]	[0.20 , 0.32]	[0.91,1.21]
WAGE	1.09	0.24	-0.06	0.07	0.41	0.10
	[0.93,1.21]	[0.06,0.42]	[-0.12,0.02]	[-0.03,0.17]	[0.32,0.53]	[-0.09,0.29]
INFL	2.72	0.63	-0.06	0.07	0.68	2.42
	[2.30,3.13]	[0.48,0.78]	[-0.14,0.04]	[-0.06,0.19]	[0.52,0.89]	[1.76,2.93]
FFR	9.80	0.91	-0.16	0.08	0.22	0.79
	[8.68,11.56]	[0.87,0.93]	[-.23,-.10]	[-0.03,0.17]	[0.15,0.32]	[-0.26,1.64]

Notes: We report posterior means and 90% equal-tail-probability credible sets in brackets.

Table A-7: Posterior Estimates for QAR(1,1) Model, 1984:Q1 to 2012:Q4

Data	ϕ_0	ϕ_1	ϕ_2	γ	σ	s_0
GDP	0.53	0.36	-0.09	-0.07	0.28	1.09
	[0.38 , 0.66]	[0.22 , 0.52]	[-0.15 , -0.03]	[-0.17 , -0.00]	[0.23 , 0.35]	[0.87 , 1.28]
WAGE	0.98	0.18	-0.04	0.03	0.48	0.20
	[0.83 , 1.14]	[0.02 , 0.36]	[-0.10 , 0.04]	[-0.06 , 0.12]	[0.38 , 0.60]	[0.03 , 0.37]
INFL	2.51	0.63	-0.02	0.07	0.76	2.54
	[2.12 , 2.93]	[0.48 , 0.77]	[-0.10 , 0.06]	[-0.03 , 0.19]	[0.61 , 0.97]	[1.80 , 3.00]
FFR	10.00	0.92	-0.17	0.01	0.19	1.00
	[8.72 , 11.43]	[0.90 , 0.94]	[-0.25 , -0.12]	[-0.05 , 0.11]	[0.15 , 0.29]	[0.05 , 1.40]

Notes: We report posterior means and 90% equal-tail-probability credible sets in brackets.

B   The DSGE Model

B.1   First-Order Conditions

Intermediate Goods Producers. Taking as given nominal wages, final good prices, the demand schedule for intermediate products, and technological constraints, firm j chooses its labor inputs $H_t(j)$ and the price $P_t(j)$ to maximize the present value of future profits. After using the production function to substitute our $Y_t(j)$ from the present value of future profits in (24) (see main text) we can write the objective function of the firm as

$$\mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s Q_{t+s|t}\left(\frac{P_{t+s}(j)}{P_{t+s}}\left(1-\Phi_p\left(\frac{P_{t+s}(j)}{P_{t+s-1}(j)}\right)\right)A_{t+s}H_{t+s}(j)-\frac{1}{P_{t+s}}W_{t+s}H_{t+s}(j)\right)\right].$$
(A.12)

This objective function is maximized with respect to $H_t(j)$ and $P_t(j)$ subject to

$$A_{t+s}H_{t+s}(j)=\left(\frac{P_t(j)}{P_t}\right)^{-1/\lambda_{p,t}}Y_{t+s}.$$

We use $\mu_{t+s}\beta^s Q_{t+s|t}$ to denote the Lagrange multiplier associated with this constraint. Setting $Q_{t|t}=1$, the first-order condition with respect to $P_t(j)$ is given by

$$
\begin{aligned}
0 \;=\; & \frac{1}{P_t}\left(1-\Phi_p\left(\frac{P_t(j)}{P_{t-1}(j)}\right)\right)A_t H_t(j)-\frac{P_t(j)}{P_t P_{t-1}(j)}\Phi_p'\left(\frac{P_t(j)}{P_{t-1}(j)}\right)A_t H_t(j) \qquad \text{(A.13)} \\
& -\frac{\mu_t}{\lambda_{p,t}P_t}\left(\frac{P_t(j)}{P_t}\right)^{-1/\lambda_{p,t}-1}Y_t+\beta\mathbb{E}_t\left[Q_{t+1|t}\frac{P_{t+1}^2(j)}{P_{t+1}P_t^2(j)}\Phi_p'\left(\frac{P_{t+1}(j)}{P_t(j)}\right)A_{t+1}H_{t+1}(j)\right].
\end{aligned}
$$

Taking first-order conditions with respect to $H_t(j)$ yields

$$\frac{W_t}{P_t}=\frac{P_t(j)}{P_t}\left(1-\Phi_p\left(\frac{P_t(j)}{P_{t-1}(j)}\right)\right)A_t-\mu_t A_t.$$
(A.14)

Households. The first-order condition with respect to consumption is given by

$$P_t\lambda_t=\left(\frac{C_t(k)}{A_t}\right)^{-\tau}\frac{1}{A_t}.$$
(A.15)

We define

$$Q_{t+1|t}=\frac{\lambda_{t+1}P_{t+1}}{\lambda_t P_t}.$$
(A.16)

Using this definition, the first-order condition for bond holdings becomes

$$1 = \beta \mathbb{E}_t \left[Q_{t+1|t} \frac{R_t}{\pi_{t+1}} \right]. \tag{A.17}$$

Member k is a monopolistic competitor with respect to his wage choice. Taking into account the demand for labor of type k the relevant portion of the utility function for the wage decision is

$$\mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s \left(\cdots - \chi_H \frac{1}{1+1/\nu} \left(\frac{W_{t+s}(k)}{W_{t+s}} \right)^{-(1+1/\nu)/\lambda_w} H_t^{1+1/\nu} \right) \right].$$

The relevant portion of the budget constraint after substituting $H_{t+s}(k)$ by the labor demand schedule is

$$\cdots = W_{t+s}(k) \left(\frac{W_{t+s}(k)}{W_{t+s}} \right)^{-1/\lambda_w} H_{t+s} \left(1 - \Phi_w \left(\frac{W_{t+s}(k)}{W_{t+s-1}(k)} \right) \right) + \cdots,$$

where the demand for aggregated labor services H_{t+s} is taken as given. Taking first-order conditions with respect to $W_t(k)$ yields

$$
\begin{aligned}
0 &= \frac{\chi_H}{\lambda_w W_t} \left(\frac{W_t(k)}{W_t} \right)^{-\frac{1+1/\nu}{\lambda_w}-1} H_t^{1+1/\nu} + \lambda_t \left(\frac{W_t(k)}{W_t} \right)^{-1/\lambda_w} H_t \left(1 - \Phi_w \left(\frac{W_t(k)}{W_{t-1}(k)} \right) \right) non \text{(A.18)} \\
&\quad - \frac{\lambda_t}{\lambda_w} \frac{W_t(k)}{W_t} \left(\frac{W_t(k)}{W_t} \right)^{-1/\lambda_w-1} H_t \left(1 - \Phi_w \left(\frac{W_t(k)}{W_{t-1}(k)} \right) \right) \tag{A.19} \\
&\quad - \lambda_t \frac{W_t(k)}{W_{t-1}(k)} \left(\frac{W_t(k)}{W_t} \right)^{-1/\lambda_w} H_t \Phi_w' \left(\frac{W_t(k)}{W_{t-1}(k)} \right) \\
&\quad + \beta \mathbb{E}_t \left[\lambda_{t+1} \frac{W_{t+1}^2(k)}{W_t^2(k)} \left(\frac{W_{t+1}(k)}{W_{t+1}} \right)^{-1/\lambda_w} H_{t+1} \Phi_w' \left(\frac{W_{t+1}(k)}{W_t(k)} \right) \right].
\end{aligned}
$$

B.2 Equilibrium Relationships

We consider the symmetric equilibrium in which all intermediate-goods-producing firms, as well as households, make identical choices when solving their optimization problem. Therefore, we can drop the index k and j. In slight abuse of notation, let $\Delta X_t = X_t/X_{t-1}$ and $\pi_t = \Delta P_t$. We use $w_t = W_t/P_t$ to denote the real wage. Since the non-stationary technology process A_t induces a stochastic trend in output, consumption, and real wages, it is convenient to express the model in terms of detrended variables $y_t = Y_t/A_t$, $c_t = C_t/A_t$ and $\tilde{w}_t = w_t/A_t$.

Intermediate Goods Producers. Using the above notation, multiplying (A.13) by P_t, and replacing Y_t by $A_t y_t$. we can simplify the first-order condition for $P_t(j)$ as follows:

$$0 = \big(1 - \Phi_p(\pi_t)\big)A_t y_t - \pi_t \Phi_p'(\pi_t) A_t y_t - \frac{\mu_t}{\lambda_{p,t}} A_t y_t + \beta \mathbb{E}_t \big[Q_{t+1|t} \pi_{t+1} \Phi_p'(\pi_{t+1}) A_{t+1} y_{t+1}\big].$$

Dividing by $A_t y_t$ and replacing A_{t+1}/A_t by $\gamma \exp(z_{t+1})$ we obtain

$$0 = \big(1 - \Phi_p(\pi_t)\big) - \pi_t \Phi_p'(\pi_t) - \frac{\mu_t}{\lambda_{p,t}} + \beta \mathbb{E}_t \big[Q_{t+1|t} \pi_{t+1} \Phi_p'(\pi_{t+1}) \Delta y_{t+1} \gamma \exp(z_{t+1})\big].$$

We proceed by rewriting (A.14) as

$$\tilde{w}_t = \big(1 - \Phi_p(\pi_t)\big) - \mu_t. \tag{A.20}$$

Households. In terms of detrended consumption we can express $Q_{t+1|t}$ as

$$Q_{t+1|t} = \left(\frac{c_{t+1}}{c_t}\right)^{-\tau} \frac{1}{\gamma} \exp(-z_{t+1}). \tag{A.21}$$

The consumption Euler equation remains unchanged:

$$1 = \beta \mathbb{E}_t \left[Q_{t+1|t} \frac{R_t}{\pi_{t+1}}\right]. \tag{A.22}$$

We now divide (A.19) by λ_t and replace λ_t by $c_t^{-\tau}/(A_t P_t)$:

$$\begin{aligned}
0 = & \ \frac{\chi_H}{\lambda_w} \frac{1}{\tilde{w}_t} c_t^\tau H_t^{1+1/\nu} + H_t\big(1 - \Phi_w(\pi_t \Delta w_t)\big) - \frac{1}{\lambda_w} H_t\big(1 - \Phi_w(\pi_t \Delta w_t)\big) \\
& - \pi_t \Delta w_t H_t \Phi_w'(\pi_t \Delta w_t) + \beta \mathbb{E}_t \big[Q_{t+1|t} \pi_{t+1} \Delta w_{t+1}^2 H_{t+1} \Phi_w'(\pi_{t+1} \Delta w_{t+1})\big].
\end{aligned}$$

Aggregate Resource Constraint. The aggregate production function (in terms of detrended output) is

$$y_t = H_t. \tag{A.23}$$

The intermediate goods producers' dividend payments to the households are given by

$$D_t = \big(1 - \Phi_p(\pi_t)\big)Y_t - w_t H_t. \tag{A.24}$$

Combining the household budget constraint and the government budget constraint and detrending all variables leads to the aggregate resource constraint

$$c_t + \zeta y_t = \big(1 - \Phi_p(\pi_t)\big)y_t - \tilde{w}_t y_t \Phi_w(\pi_t \Delta w_t),$$

where $\Delta w_t = \Delta \tilde{w}_t \gamma \exp(z_t)$.

The model economy has a unique steady state in terms of the detrended variables that is attained if the innovations $\epsilon_{R,t}$, $\epsilon_{g,t}$, and $\epsilon_{z,t}$ are zero at all times. The steady-state inflation π equals the target rate π^* and

$$R = \frac{\gamma}{\beta}\pi^*, \ \mu = \lambda_p, \ c = \left(\frac{(1 - \lambda_p)(1 - \lambda_w)g^{-\frac{1}{\nu}}}{\chi_H}\right)^{\frac{1}{\tau + 1/\nu}}, \ y = g\tilde{c}, \ H = y, \ \tilde{w} = (1 - \lambda_p).$$

B.3 Posterior Simulator

We first estimate a log-linearized version of the DSGE model using the random walk Metropolis (RWM) algorithm described in An and Schorfheide (2007). Using the same covariance matrix for the proposal distribution as for the linearized DSGE model, we then run the RWM algorithm based on the likelihood function associated with the second-order approximation of the DSGE model. The covariance matrix of the proposal distribution is scaled such that the RWM algorithm has an acceptance rate of approximately 50%. We use 80,000 particles to approximate the likelihood function of the nonlinear DSGE model, while the variance of measurement errors is set to 10% of the sample variance of the observables. We generate 120,000 draws from the posterior distribution of the nonlinear DSGE model. The summary statistics reported in Table 3 in the main paper are based on the last 100,000 draws of this sequence.

Table A-8: Posterior Estimates for DSGE Model Parameters: Linear Model

Parameter	1960:Q1 to 2007:Q4		1984:Q1 to 2007:Q4	
	Mean	90% Interval	Mean	90% Interval
$400\left(\frac{1}{\beta}-1\right)$	0.48	[0.06, 1.01]	1.31	[0.60, 2.17]
π^A	3.46	[2.94, 3.97]	2.80	[2.33, 3.29]
γ^A	1.86	[1.39, 2.34]	1.88	[1.53, 2.24]
τ	6.54	[4.37, 9.24]	4.78	[2.57, 8.70]
ν	0.09	[0.06, 0.13]	0.08	[0.03, 0.15]
$\kappa(\varphi_p)$	0.01	[0.01, 0.02]	0.18	[0.09, 0.30]
φ_w	62.33	[44.48, 83.14]	14.89	[6.15, 25.88]
ψ_w		N/A		
ψ_p		N/A		
ψ_1	1.45	[1.24, 1.68]	2.67	[2.10, 3.30]
ψ_2	0.80	[0.54, 1.09]	0.76	[0.41, 1.11]
ρ_r	0.77	[0.73, 0.82]	0.71	[0.61, 0.79]
ρ_g	0.97	[0.96, 0.98]	0.96	[0.93, 0.98]
ρ_z	0.26	[0.10, 0.41]	0.07	[0.01, 0.19]
ρ_p	0.99	[0.98, 0.99]	0.93	[0.87, 0.98]
$100\sigma_r$	0.18	[0.14, 0.22]	0.18	[0.13, 0.25]
$100\sigma_g$	0.65	[0.44, 0.95]	0.76	[0.39, 1.34]
$100\sigma_z$	0.75	[0.64, 0.85]	0.47	[0.37, 0.56]
$100\sigma_p$	15.28	[12.66, 18.18]	7.63	[5.96, 9.48]

Notes: Estimation sample is 1984:Q1 to 2010:Q4. For 90% credible interval we are reporting the 5th and 95th percentile of the posterior distribution.