Macroeconomic Dynamics, **18**, 2014, 41–64. Printed in the United States of America.
doi:10.1017/S1365100513000278

ESTIMATED THRESHOLDS IN THE RESPONSE OF OUTPUT TO MONETARY POLICY: ARE LARGE POLICY CHANGES LESS EFFECTIVE?

LUIGGI DONAYRE
University of Minnesota–Duluth

This paper investigates the potential sources of the mixed evidence found in the empirical literature studying asymmetries in the response of output to monetary policy shocks of different magnitudes. Further, it argues that such mixed evidence is a consequence of the exogenous imposition of the threshold that classifies monetary shocks as small or large. To address this issue, I propose an unobserved-components model of output, augmented by a monetary policy variable, which allows the threshold to be endogenously estimated. The results show strong statistical evidence that the effect of monetary policy on output varies disproportionately with the size of the monetary shock once the threshold is estimated. Meanwhile, the estimates of the model are consistent with a key implication of menu-cost models: smaller monetary shocks trigger a larger response on output.

Keywords: Asymmetry, Monetary Policy, Regime Switching, Threshold Autoregressions, Unobserved Components Model

1. INTRODUCTION

An important area of research in the time series literature is the examination of potential asymmetric effects of monetary policy on output [DeLong and Summers (1988); Cover (1992); Teräsvirta and Anderson (1992); Beaudry and Koop (1993); Thoma (1994); Weise (1999); Rothman et al. (2001); Garcia and Schaller (2002); Lo and Piger (2005)]. Part of this interest has focused on asymmetries with respect to the size of the monetary policy shock, but the empirical evidence is limited and

I am very thankful to James Morley for valuable insights and suggestions. Helpful comments from two anonymous referees, Steven Fazzari, Kyu Ho Kang, Rody Manuelli, Werner Ploberger, Mark Vaughan, and the participants at the Applied Time-Series Research Group at Washington University, the 3rd Annual Graduate Student Conference at Washington University, the 45th Annual Meeting of the Missouri Valley Economic Association, the 9th Annual Missouri Economic Conference, the 2009 Far East and South Asian Meeting of the Econometric Society, the 2009 Latin American Meeting of the Econometric Society, and the 2009 meeting of the Latin American and Caribbean Economic Association are gratefully acknowledged. I also thank the Center for Research in Economics and Strategy (CRES), in the Olin Business School, Washington University in St. Louis, for financial support. All remaining errors are my own. Address correspondence to: Luiggi Donayre, Department of Economics, University of Minnesota–Duluth, 1318 Kirby Drive, Duluth, MN 55812, USA; e-mail: adonayre@d.umn.edu.

mixed. In consideration of this fact, and given the important policy implications of this type of asymmetry,[1] this paper examines whether the effects of monetary policy on output vary disproportionately with the size of the monetary shock and aims to shed some light on the reasons behind the mixed evidence found in the literature.

One of the first papers that addresses this type of asymmetry is Sensier (1996). She uses a Markov-switching approach and finds little support for asymmetries in the response of output to monetary policy shocks of different sizes in a study using data from the United Kingdom. Weise (1999), in turn, tests for asymmetries in a logistic smooth transition vector autoregression (LSTVAR) that includes U.S. output, prices, and money growth, using an estimation strategy that is consistent with a wide variety of structural macroeconomic models. Although his analysis includes other types of asymmetries, his results show that small and large monetary policy shocks have different effects on output, but which effects are larger depends on the time horizon under consideration. When viewed over a three-year horizon, small shocks have disporportionaly large effects on output, but over shorter horizons, the effects are disproportionately small.

Ravn and Sola (2004) use a Markov-switching model to test for asymmetries related to the size of monetary shocks that may arise in macroeconomic models with menu costs. Although their broader work considers small and large monetary policy shocks that can be expansive or contractive, they cannot reject that the effects of monetary shocks of different magnitude on output are symmetric when using M1 as the monetary instrument. They find strong evidence of this type of asymmetries only in the case when the Federal Funds rate (FFR) is used to measure monetary policy. Finally, Lo and Piger (2005) employ a time-varying Markov-switching model and find weak evidence of asymmetry in terms of the size of a policy shock when it is the only source of asymmetries in the model. However, that evidence disappears when taken into account together with asymmetries with respect to the state of the economy.

Although the findings of these studies are mixed, a common characteristic in all of them is that the threshold that classifies monetary shocks as small or large has been set exogenously, typically to one standard deviation. The appropriateness of this approach is unclear, because there is no economic reasoning supporting a particular threshold. This paper shows that the estimation of the threshold is important in helping to detect a significant asymmetric response of output to monetary shocks of different sizes. In particular, if the threshold is misspecified, tests for asymmetry have low power, leading to an inability to reject the null hypothesis of linearity. Thus, the main contribution of this paper is to establish an econometric strategy that allows estimation of the threshold, instead of imposing an ad hoc definition.

To do so, this paper proposes an approach that unifies various models that have dealt with this type of asymmetry. It is similar in spirit to the model proposed by Weise (1999), mentioned earlier, which tests for asymmetries using a multivariate LSTVAR. Although Weise only considers money-based indicators of monetary

Table 1. A summary of the empirical literature on asymmetries related to the size of monetary shocks

Paper/author(s)	Methodology	Threshold	Result
Sensier (1996)	Markov switching	One standard deviation (exogenous)	No evidence of asymmetry
Weise (1999)	Logistic smooth transition vector autoregression	Mean or zero (exogenous)	Weak evidence of asymmetry
Ravn and Sola (2004)	Markov switching	Variance of shock (exogenous)	Evidence for FFR case (not for M1)
Lo and Piger (2005)	Time-varying Markow switching	One standard deviation (exogenous)	Weak evidence when it is the only source of asymmetry; no evidence when considered together with state of economy

policy, an interest-rate measure of monetary policy is also used here. Further, his multivariate setting makes estimation of the threshold infeasible, and he imposes it arbitrarily at the mean of the threshold variable. Meanwhile, the model proposed here has the advantage over Weise's approach that it introduces a threshold autoregressive (TAR) process within an unobserved-components (UC) model, similar to what was considered in Lo and Piger (2005). Under this approach, the monetary policy variable has only transitory effects on output, consistent with the notion of money neutrality in the long run. This differs from the majority of the literature, which generally proceeds by regressing output growth on measures of policy actions. In contrast to Sensier (1996), Ravn and Sola (2004), and Lo and Piger (2005), the TAR process in this paper implies observable, but estimated thresholds, whereas they model the nonlinear relationship between money and output as a Markov-switching process, where the state variable is not observable. This, as explained previously, allows an estimate of the threshold that defines which shocks are considered "small" and which are considered "large." Table 1 summarizes and contrasts the methods used and the results found by the relevant papers in the literature.

By estimating the threshold, the results from the nonlinear UC model can shed light on the importance of the definition of size in determining the existence of an asymmetric response of output to monetary shocks of different magnitude. Furthermore, the results can be related to the implications of theoretical menu-cost models. These models provide a motivation for distinguishing monetary shocks in terms of their size (see Appendix A for a simple menu-cost model where the effect on output depends on the magnitude of the monetary policy shock). For example, Ball and Romer (1990) and Ball and Mankiw (1994) present standard menu-costs models in which only small monetary shocks have effects on output. In their models, when monetary shocks are large, the menu cost becomes a second-order cost and firms find it optimal to incur such cost and adjust their prices, leaving real output unchanged. Meanwhile, even small menu costs may be sufficient to generate aggregate nominal rigidity and large business cycle fluctuations [Mankiw (1985); Blanchard and Kiyotaki (1987); Caplin and Leahy (1991); Golosov and Lucas (2007)]. This paper, however, does not attempt to test the menu-cost theory formally.

The results found show strong statistical evidence of an asymmetric response of output to monetary shocks of different sizes when the threshold is estimated from the data. In contrast, when linearity is tested imposing the threshold at an ad hoc value—i.e., one standard deviation—instead of the estimated threshold, the null hypothesis cannot be rejected. This supports the claim that the threshold defining the size of shocks should be estimated from the data. Further, the results from the estimated model suggest that the response of output to small monetary policy shocks is larger than its response to large shocks. This is consistent with the key implications of menu-costs models, although the finding that large shocks still have effects on output is somewhat contrary to the simple version of the menu-cost story.

The remainder of this paper is organized as follows. The second section formally describes the nonlinear UC model, the empirical approach used to estimate it, and the bootstrap procedure for performing the linearity test. The third section reports the results for the empirical model and linearity tests. Some concluding remarks are provided in the fourth section.

2. EMPIRICAL APPROACH

This section describes the empirical model, the estimation approach, and the bootstrap procedure used to test whether the TAR-driven transitory component is statistically significant when compared with a linear specification.

2.1. An Unobserved-Components Model

This paper proposes a model that builds on the UC framework considered by Lo and Piger (2005). Previous studies testing for asymmetries have regressed output growth on measures of large and small policy actions using different models [Sensier (1996); Ravn and Sola (2004)]. Thus, the presence of asymmetries in those studies involves determining whether the coefficients associated with such policy measures are statistically different from each other. This latter approach is potentially inconsistent with the notion of long-run money neutrality. Because an UC model is employed in this paper, monetary shocks only affect the transitory component of output (i.e., they are neutral in the long run).

Further, the main difference from the approach taken by Lo and Piger (2005) is that, whereas they model the nonlinear relationship between money and output as a Markov-switching process, the regime switching here is driven by a TAR process. This process implies estimated, but observable thresholds, more in line with menu-cost theory. Thus, the model is given by

$$y_t = y_t^{\mathrm{T}} + y_t^{\mathrm{C}}, \tag{1}$$

$$y_t^{\mathrm{T}} = \mu + y_{t-1}^{\mathrm{T}} + \nu_t, \tag{2}$$

$$y_t^{\mathrm{C}} = \sum_{p=1}^{P} \phi_p y_{t-p}^{\mathrm{C}} + \sum_{j=1}^{J} \alpha_j^{\mathrm{S}} x_{t-j} I[q(t) \leq \gamma] + \sum_{j=1}^{J} \alpha_j^{\mathrm{L}} x_{t-j} I[q(t) > \gamma] + \epsilon_t, \tag{3}$$

where y_t is a measure of output, y_t^{T} is the permanent (or trend) component of output, y_t^{C} is the transitory (or cyclical) component of output, and x_t is a measure of monetary policy.

The system (1)–(3) is a modified version of the simple UC decomposition of real output into permanent and transitory components, as in Watson (1986). Following the original model, the permanent component of output, given in equation (2), is modeled as a random walk with a drift term, μ.

Equation (3) describes the dynamics of the transitory component of output, y_t^{C}. It is modeled as an autoregressive distributed-lag (ADL) process, where the independent variable is the monetary policy shock, x_t; $I[.]$ denotes the indicator function; $q(t)$ is the threshold variable; and γ is the threshold parameter. When $q(t) \leq \gamma$, the response coefficients are captured by the $J \times 1$ vector $\boldsymbol{\alpha}^{\mathrm{S}}$ and when $q(t) > \gamma$, they are given by the $J \times 1$ vector $\boldsymbol{\alpha}^{\mathrm{L}}$. To be consistent with the measures of monetary policy considered later, where the monetary variable does not affect output contemporaneously, only lags of x_t are allowed to enter equation (3). Note that the coefficients $\phi_p,\ p = 1, \ldots, P$, are not state-dependent. Because the question of interest is whether the response of output to monetary shocks varies with the size of the shock, the autoregressive dynamics is assumed to be the same in both regimes.

The innovations ϵ_t and ν_t have a joint normal distribution with mean zero and variance–covariance matrix Ω. To identify the model, the UC literature has traditionally assumed that the covariance between ϵ_t and ν_t is zero (i.e., that Ω is a diagonal matrix). Nonetheless, Morley et al. (2003) show that, under certain conditions, UC models can be estimated without imposing such a restriction.[2] Thus, to explore the possibility that the innovations from the transitory and permanent components of output are correlated, the covariance term is estimated. Further, the trend/cycle decomposition is interpreted here as an *estimate* of the unobserved components [see Morley (2011) for more details].

An additional feature of the model estimated here is that it allows for a one-time break in the variance–covariance matrix. The volatility of output, as well as that of many other macroeconomic aggregates, declined substantially since the mid-1980s. To account for this fact, an exogenous break date is set to the last quarter of 1983 to split the sample accordingly.[3]

2.2. Estimation Procedure

Following the approach to making inferences about TAR models discussed in Hansen (1997), the estimation of the coefficients of the system (1)–(3) involves an iterative procedure. The model is estimated sequentially for each possible value of the threshold parameter, yielding a γ-dependent loglikelihood function $\log L(\gamma)$ in each iteration. Thus, the maximum likelihood (ML) estimate of γ is the value of this parameter that maximizes $\log L(\gamma)$[4]. Formally, the ML estimate of γ is defined as

$$\hat{\gamma} = \underset{\gamma \in \Gamma}{\operatorname{argmax}}[\log L(\gamma)], \tag{4}$$

where $\Gamma = [\underline{\gamma}, \bar{\gamma}]$ is defined *a priori* to contain the middle 70% of all possible threshold values to ensure that the model is well identified.[5]

The ML estimates of $\boldsymbol{\alpha}^{\mathrm{S}}$, $\boldsymbol{\alpha}^{\mathrm{L}}$, σ_ϵ, σ_ν, and $\sigma_{\epsilon\nu}$ are thus given by the parameters associated with $\log L(\hat{\gamma})$. That is, the vector of ML estimates is given by $\hat{\theta} = \hat{\theta}(\hat{\gamma})$, where θ=($\boldsymbol{\alpha}^{\mathrm{S}'}$, $\boldsymbol{\alpha}^{\mathrm{L}'}$, $\sigma_\epsilon, \sigma_\nu, \sigma_{\epsilon\nu})'$.

2.3. Testing for a TAR-Driven Transitory Component

Computationally, estimation of the model is cumbersome because of the sequential iteration of the threshold parameter. Letting $|\Gamma|$ denote the cardinality of Γ, there are $|\Gamma|$ potential thresholds and, therefore, the same number of models to be estimated. This estimation routine becomes even more time-intensive when it is used to test whether the TAR-driven transitory component is statistically significant relative to a linear one. In particular, as will be explained later, the linearity test involves a bootstrap procedure in which data are generated under the null hypothesis B times. For each bootstrap sample, a grid search across possible threshold parameters is carried out. As a consequence, $B \times |\Gamma|$ potential models need to be estimated.[6]

The procedure developed by Hansen (1996, 1997) to test TAR processes against linear ones is modified to fit the UC framework (1)–(3). Considering these modifications, the relevant null hypothesis is given by $H_0 : \ \boldsymbol{\alpha}^{\mathrm{S}} = \boldsymbol{\alpha}^{\mathrm{L}}$. Because this problem is tainted by the existence of nuisance parameters (specifically, the threshold γ is not identified under the null hypothesis), a test with near-optimal power against a wide range of alternative hypotheses is given by the following likelihood ratio (LR) statistic:

$$\mathrm{LR} = \sup_{\gamma \in \Gamma} [\mathrm{LR}_n(\gamma)]\,, \tag{5}$$

where

$$\mathrm{LR}_n(\gamma) = 2\left[\hat{\log} L(\gamma) - \hat{\log} L_0\right] \tag{6}$$

is the LR statistic against the alternative $H_1 : \ \boldsymbol{\alpha}^{\mathrm{S}} \neq \boldsymbol{\alpha}^{\mathrm{L}}$ when γ is known. $\hat{\log} L_0$ and $\log \hat{L}(\gamma)$ correspond to the values of the log likelihood functions under the null hypothesis and under the alternative hypothesis for each γ, respectively.

Because γ is not identified, the distribution of the LR statistic (5) is nonstandard. However, its asymptotic distribution can be approximated through a bootstrap procedure [see Hansen (1996, 1997) for further details]. Following his work, the approach is modified to fit the system (1)–(3) and the asymptotic distribution of (5) is approximated by a bootstrap experiment in which $y_t^* = y_t^{\mathrm{NH}},\ t = 1, \ldots, n$ where y_t^{NH} is a new dependent variable generated under the null hypothesis. Using y_t^{NH}, a new LR statistic is calculated for this new dependent variable to form $\mathrm{LR}^* = \sup_{\gamma}\{2[\log\hat{L}_1^*(\gamma) - \log\hat{L}_0^*]\}$.

The procedure described here is similar in spirit to the one detailed in Hansen (1997). The asymptotical equivalence of the likelihood ratio statistic and the original F-statistic in his paper guarantees that his results are carried over to this framework. Moreover, because any UC model can be represented as an ARMA(p, q) process for the first differences, the regularity conditions in Hansen (1997) are satisfied and, therefore, the statistic (5) converges to the asymptotic distribution derived by him (see Appendix B for further details). This also guarantees that the bootstrap procedure is not invalidated because the permanent component of output is modeled as a random walk process.

Two Monte Carlo experiments were also conducted to evaluate the size and power of the bootstrap procedure in an UC setting. Regarding the size of the test, data are generated under the null hypothesis for 500 Monte Carlo repetitions and the size of the test is evaluated at the 1, 5, and 10% nominal levels. The results of the Monte Carlo experiment yield true sizes of 0.012, 0.062, and 0.106, respectively, suggesting that the test has relatively good size in all cases. Regarding the power of the test, data are generated according to the system (1)–(3) and based on the model estimates. Among the 500 Monte Carlo simulations, the test correctly rejected the null hypothesis 494 times, suggesting that the test has good power at the 5% significance level.

Concisely, the test to determine the significance of the UC model with a TAR-driven transitory component against the null hypothesis of a linear transitory component can be summarized in the following steps:

Step 1: Estimate the model under the null hypothesis $H_0: \ \boldsymbol{\alpha}^{\mathrm{S}} = \boldsymbol{\alpha}^{\mathrm{L}}$ and obtain the log likelihood function $\log\hat{L}_0$.

Step 2: Estimate the model under the alternative hypothesis $H_1: \ \boldsymbol{\alpha}^{\mathrm{S}} \neq \boldsymbol{\alpha}^{\mathrm{L}}$ and obtain the log likelihood function $\log\hat{L}_1(\gamma)$.

Step 3: Form the LR statistic $\mathrm{LR} = \sup_{\gamma\in\Gamma}\{2[\log\hat{L}_1(\gamma) - \log\hat{L}_0]\}$.

Step 4: Bootstrap distribution:

(a) Generate a new independent variable $y_t^* = y_t^{\mathrm{NH}}$ under the null hypothesis.

(b) Estimate the model under $H_0: \ \boldsymbol{\alpha}^{\mathrm{S}} = \boldsymbol{\alpha}^{\mathrm{L}}$ and obtain the log likelihood function $\log\hat{L}_0^*$.

(c) Estimate the model under $H_1: \ \boldsymbol{\alpha}^{\mathrm{S}} \neq \boldsymbol{\alpha}^{\mathrm{L}}$ and obtain the log likelihood function $\log\hat{L}_1^*(\gamma)$.

(d) Form the LR statistic $\mathrm{LR}^* = \sup_{\gamma}\{2[\log\hat{L}_1^*(\gamma) - \log\hat{L}_0^*]\}$.

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

Step 5: Obtain the bootstrapped p-value as the percentage of bootstrap samples for which $LR^* > LR$.

3. RESULTS

The iterative procedure for obtaining the threshold parameter described in (4) involves casting the model (1)–(3) in state-space form and applying the Kalman filter (see Appendix C for a general state-space representation of the model). For further details about the Kalman filter, refer to Hamilton (1994) and Kim and Nelson (1999).

According to the specification of the model, monetary shocks only enter the transitory component of output to be consistent with the notion of long-run money neutrality. Moreover, the monetary shock is assumed to affect the transitory component of output with a lag. This assumption is standard and well documented in the literature [Christiano et al. (1996, 2005); Leeper et al. (1996); Sims and Zha (1998); Thoma (2007)].

It is important to note that, because the interest of the paper is in the size of monetary shocks, the threshold variable $q(t)$ in equation (3) is the absolute value of the monetary policy shock. Considering the absolute value of the vector of monetary shocks implies a symmetric threshold around zero. To determine whether this assumption is too restrictive (in the sense that the asymmetry found could be driven by model misspecification), a model where the thresholds above and below zero are different in magnitude was also considered. The improvement in the likelihood, however, was not significant. In particular, the null hypothesis that the additional threshold is equal to the negative of the absolute value yielded a p-value of 0.862. Hence, only the symmetric threshold specification is reported.

The measure of monetary policy considered in this paper involves a shock where the Federal Funds Rate (FFR) is the monetary instrument.[7] All data are taken from the Federal Reserve Economic Data (FRED) database and were deseasonalized from the source. Output is measured as 100 times the natural logarithm of quarterly real GDP. The FFR series considered is the monthly effective Federal Funds Rate, converted to the quarterly frequency by computing a simple three-month arithmetic mean. Finally, the measure of prices is given by the logarithm of the deflator of quarterly real GDP. The sample period for the estimation of the model goes from 1954:Q3 through 2008:Q3, corresponding to 217 observations.

In the estimation of the model, the first 20 observations are used as a training sample to avoid the effects of the starting values associated with the initialization of the Kalman filter.[8] The first 10 and 30 observations were also considered as training samples, but the results were not very sensitive to these different specifications.

3.1. Results for the Federal Funds Rate Measure of Monetary Policy

An interest-rate-based monetary shock is constructed from the residuals of an identified VAR that contains three variables: the FFR, the logarithm of real GDP, and the logarithm of the GDP deflator. To identify the shock, the policy variable is ordered last in the VAR (i.e., monetary shocks do not affect output contemporaneously) and four lags of each variable are included.[9]

The identification of the policy shock is important to properly assess the effects of changes in monetary policy on the economy. The approach followed in this paper to identify such shocks is standard in the literature [Christiano et al. (1996, 1999, 2005)]. It is widely believed that a significant fraction of the variation in the Fed policy actions reflects the monetary authorities' systematic responses to variations in the state of the economy. Nonetheless, not all variations in monetary policy can be accounted for as reactions to the state of the economy. The fraction of this variation not accounted for is defined as a monetary policy shock.

Christiano et al. (2005) claim that this monetary shock can be correctly identified if one assumes that the policy shock is orthogonal to the Fed's information set. That is, time-t variables in the Fed's information set do not respond to time-t realizations of the monetary policy shock. Indeed, the authors argue that this assumption justifies the widely used two-step approach: estimating policy shocks as the fitted residuals of a VAR and using them to estimate the dynamic response of a variable to a monetary policy shock.

Although this identification scheme is standard, there are certainly advantages to abandoning the orthogonality assumption. However, there are also substantial costs. In particular, a broader set of economic relations must be identified. For example, Leeper et al. (1996) and Sims and Zha (1998, 2006) assume that the Fed does not look at the contemporaneous price level or output when setting its policy instrument. This assumption is clearly debatable.

Similarly, there are advantages to identifying the monetary policy shock within the UC model instead of using the two-step approach. Nonetheless, the two-step approach employed here allows a rich specification of the VAR to capture monetary policy shocks in the first step, in a way that attempts to control for the endogenous response of the FFR to output and inflation. Meanwhile, the relative simplicity of the second-step UC model allows specification of the asymmetries, together with estimation of the model and the bootstrap linearity test, which is implemented more easily than the high-dimensional environment that would be necessary to identify the monetary shock within the UC model. Furthermore, the joint estimation also raises potential simultaneity, correlation, and misspecification problems that could affect the estimates. For all of these reasons, and because the identification assumption made here generates qualitative effects of a monetary shock that are consistent with economic theory, the two-step approach is deemed appropriate.

Once the monetary shock is identified, the number of autoregressive coefficients for y_t^{T}, P, and the number of lags for the monetary shock, J, are determined. To do so, both the Akaike information criterion (AIC) and the Schwarz information criterion (SIC) are considered to select among models with different numbers of parameters. Given that quarterly data are used in the estimation, the maximum number of lags for each coefficient is set to 4. Both criteria select the model

TABLE 2. Parameter estimates: UC model with a TAR-driven transitory component ($\sigma_{\epsilon\nu}$ is estimated)

Parameter	Estimate	St. error	Parameter	Estimate	St. error
ϕ_1	0.690	0.097	σ_ν	1.746	0.219
ϕ_2	−0.119	0.033	$\sigma_{\epsilon\nu}$	−1.576	0.323
α_2^{S}	−0.222	0.090	λ	0.214	0.044
α_2^{L}	−0.031	0.052	μ	0.747	0.074
σ_ϵ	0.999	0.034	γ	0.157	
Log likelihood		−198.069	LR statistic		18.746

Note: This table reports estimated coefficients from the model given in (1)–(3) when the monetary instrument is the FFR and $\sigma_{\epsilon\nu}$ is estimated. The threshold variable, FFR-based monetary shock, was set to contain the 70% middle part of the observations to avoid overfitting. The sample period ranges from the third quarter of 1960 through the third quarter of 2008 after the first 20 observations are discarded to avoid distortions due to the initialization of the Kalman filter.

in which $P = J = 2$. After the model is specified, the estimation approach described in Section 2 is applied. It is important to mention that, when the model was estimated, the α_1^{S} and α_1^{L} coefficients were not statistically significant and, thus, they were removed from the final version of the model. Table 2 reports the estimated coefficients of the model (1)–(3) when the monetary instrument is the FFR.

In this table, regime S corresponds to the situation in which monetary shocks are small, as defined by the estimated threshold, whereas regime L corresponds to that in which monetary shocks are defined as large. The estimated coefficients linking monetary policy to output suggest that the estimated threshold divides policy shocks that have relatively small effects from those that have larger effects. For instance, suppose the monetary authority increases the FFR by 0.10 percentage points at time t. Given $\hat{\gamma}$, this corresponds to a situation in which regime S prevails. As a consequence of this increase in the FFR, output falls by 0.022 percentage points two quarters later. By contrast, suppose the monetary authority doubles the increase in FFR to 0.20 percentage points at time t. Then, this situation corresponds to regime L and implies a reduction in output by just 0.006 percentage points two quarters later. That is, the response coefficient in regime S, α_2^{S}, is proportionally larger than α_2^{L} (in absolute value). Note, also, that the response coefficient α_2^{L} is not statistically significant. As discussed in Appendix A, this result is consistent with the implications of the menu-costs models from Ball and Romer (1990) and Ball and Mankiw (1994).

The estimated threshold is $\hat{\gamma} = 0.157$, obtained as in equation (4). The standard deviation of the vector of shocks is 0.333, about two times larger than $\hat{\gamma}$. Thus, the estimates suggest that using one standard deviation as the threshold would classify some monetary shocks as being small when, in terms of their dynamic effects, they should have been considered large.

From Table 2, the estimated coefficients suggest that the variance of output is driven by changes in both the transitory and permanent components. Note,

TABLE 3. Parameter estimates: UC model with a TAR-driven transitory component ($\sigma_{\epsilon\nu}$ imposed to zero)

Parameter	Estimate	St. error	Parameter	Estimate	St. error
ϕ_1	1.379	0.083	σ_ν	0.902	0.090
ϕ_2	−0.476	0.057	$\sigma_{\epsilon\nu}$	0	
α_2^S	−0.174	0.102	λ	0.256	0.058
α_2^L	−0.069	0.046	μ	0.777	0.043
σ_ϵ	0.367	0.130	γ	0.157	
			Log likelihood	−207.442	

Note: This table reports estimated coefficients from the model given in (1)–(3) when the monetary instrument is the FFR and $\sigma_{\epsilon\nu}$ is imposed to be zero. The threshold variable, FFR-based monetary shock, was set to contain the 70% middle part of the observations to avoid overfitting. The sample period ranges from the third quarter of 1960 through the third quarter of 2008 after the first 20 observations are discarded to avoid distortions due to the initialization of the Kalman filter.

however, that σ_ν is larger than σ_ϵ, supporting the use of an UC framework instead of simple growth rates, as discussed earlier. To account for the break in variance, the variance–covariance matrix Ω was rescaled after the last quarter of 1983 by a factor λ [Sinclair (2009); Morley and Piger (2012)]. Hence, the fact that this factor is well below 1 and significant supports the notion that output growth volatility has fallen substantially since the mid-1980s.

The covariance of the innovation terms, $\sigma_{\epsilon\nu}$, is negative and statistically significant. Moreover, the correlation coefficient, $\rho = -0.903$, implies that the permanent and transitory components of output are strongly negatively correlated, consistent with the findings in the literature [Morley et al. (2003); Sinclair (2009)]. However, it is important to note that the estimates of the autoregressive coefficients ϕ_1 and ϕ_2 are different from the results in Morley et al. (2003) in that the period of the cycle implied by the AR parameters is even shorter. This difference is due to the inclusion of the monetary policy variable.

Given that the covariance term has typically been set to zero in other previous studies, it would be interesting to determine whether $\sigma_{\epsilon\nu}$ is statistically different from zero. To test this, the model is reestimated imposing $\sigma_{\epsilon\nu} = 0$. Using a simple likelihood ratio test, the p-value for the null hypothesis that $\sigma_{\epsilon\nu} = 0$ is zero to the fourth decimal. That is, the model where $\sigma_{\epsilon\nu} \neq 0$ is preferred.

Even though this restriction is rejected, the model that imposes $\sigma_{\epsilon\nu} = 0$ is also estimated, because it typically provides very different results. Table 3 reports the coefficients of the model (1)–(3) when the errors of the permanent and transitory components are not correlated. As can be observed from Table 3, the estimates associated with the response coefficients α_2^S and α_2^L do not change much. In particular, the magnitude of the response of output to a small monetary shock is larger than that of its response to a large shock, as before. Note, however, that the AR coefficients are now in line with typical UC models estimates, implying persistent, periodic, and ample cycles. In general, the results are robust to $\sigma_{\epsilon\nu}$ being estimated or imposed to be zero.

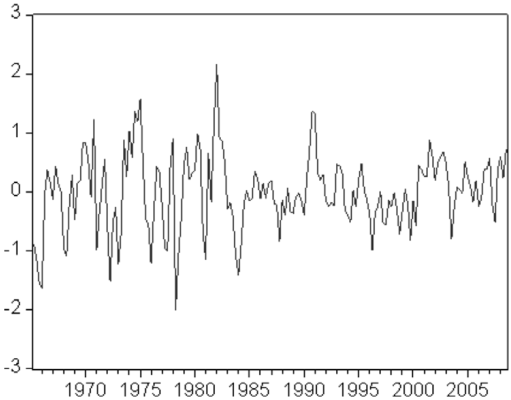

FIGURE 1. Estimated transitory component. Threshold variable: FFR-based monetary shock. Estimated transitory component when the FFR is the monetary instrument. The figure shows the estimated transitory component, y_t^C, from equation (3).

An additional concern that arises in models where the covariance term $\sigma_{\epsilon\nu}$ is estimated refers to the magnitude of the transitory component. When the restriction that $\sigma_{\epsilon\nu} = 0$ is relaxed, UC models cast in state-space form produce a trend-cycle decomposition that is identical to the Beveridge–Nelson (BN) approach [Morley et al. (2003); Sinclair (2009)]. The BN decomposition implies that a stochastic trend accounts for most of the variation in output. Figure 1 shows the estimated transitory component associated with the parameters from Table 2.

Consistent with the results found in Morley et al. (2003), the estimated cycle is small in amplitude compared to the traditional cyclical components from UC models where $\sigma_{\epsilon\nu} = 0$, and it is also much less persistent. However, the fact that the estimated cycle is relatively small and less persistent does not mean that monetary policy cannot play a role. Regardless of the size of the transitory component, the monetary authority does change the FFR to affect the state of the economy and, if the business cycle responds asymmetrically to those changes, it can have relevant policy implications. In fact, the results of the model where $\sigma_{\epsilon\nu}$ is estimated make a strong case for the role of the monetary authority. For example, a positive real shock that shifts the permanent component upward will leave actual output below trend, given the large negative correlation coefficient $\rho = -0.903$. This generates the need for a reduction in the FFR to close the gap.

A first test to determine whether the model with a threshold provides a significant improvement over a linear model involves a diagnostic test over the residuals of such models. This test produced evidence of remaining serial correlation in the case of the linear model, but not in the case of the model with a threshold.

More formally, to test whether the TAR-driven transitory component is significantly better than a linear one, the bootstrap procedure described in Section 2.3 is applied using 99 bootstrap samples.[10] The bootstrapped p-value that the procedure yields is 0.02. Thus, linearity is rejected at the 5% level and equation (3) is favored

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

over a simple linear specification. Moreover, when the same bootstrap procedure is applied to test linearity after the threshold of one standard deviation of the monetary shocks is imposed, linearity cannot be rejected. This result supports the idea that previous failures to reject linearity are a consequence of imposing the threshold in an ad hoc manner, instead of estimating it directly from the data.

If the Fed has to implement small changes in the FFR to have a large impact on output, a natural question that arises in this context is whether such a change must be carried out all at once or gradually. That is, does the size of the shock matter relative to the frequency of the data? This question is closely related to the interest-rate-smoothing literature and can have important policy implications. To address this issue, the model in (1)–(3) is reestimated using annual data from 1955 through 2007. If the size matters relative to the frequency of the data, the threshold estimated using annual data should be expected to be approximately four times the threshold estimated using quarterly data. When annual data are used, the estimated threshold is $\hat{\gamma}^{A} = 0.685$, slightly more than four times the threshold estimated using quarterly data, $\hat{\gamma}^{Q} = 0.157$.

This result supports the notion that the Fed and other central banks seek to smooth interest rates to obtain larger effects of policy by minimizing the volatility of the policy interest rate. For instance, suppose the Fed implemented a one-time increase in the FFR of 0.40 percentage points (40 basis points) in a given year. This change would be above $\hat{\gamma}^{Q} = 0.157$, and its effects on output would not be very big. A better approach, however, would be to "smooth" the change in the FFR and carry on four 0.1 percentage point increases in each quarter. In this way, each of those changes would be smaller than $\hat{\gamma}^{Q} = 0.157$, triggering a larger effect on output. Hence, by smoothing interest rates, the size of the change in the FFR required to reduce fluctuations in the economy can be smaller than it would be necessary otherwise.

Even though the response coefficients are larger when small monetary shocks hit the economy, it is important to evaluate these responses over time, given the nonlinear nature of the model. Simple impulse-response functions (IRF) are a convenient way to analyze the difference in the response of output to monetary shocks of different sizes over time. However, when the model is nonlinear, such as the one in equations (1)–(3), the IRFs are sensitive to the history of the system and the future shocks assumed to hit it.

To address these problems, generalized impulse-response functions (GIRFs) are constructed following Koop et al. (1996). The model is assumed to be known, so model uncertainty is not taken into account. Moreover, attention is restricted to the transitory component of output using the estimated parameters from Table 2. To compute the GIRFs, the following procedure is implemented (see Appendix D for a detailed description): First, monetary shocks and idiosyncratic shocks for periods 1 to q are drawn, with replacement, from the residuals of the identified VAR and the estimated transitory component, respectively, and, for a given history of the system, fed through equation (3) to produce a simulated data series.[11] This produces a forecast of the transitory component conditional on a particular history

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

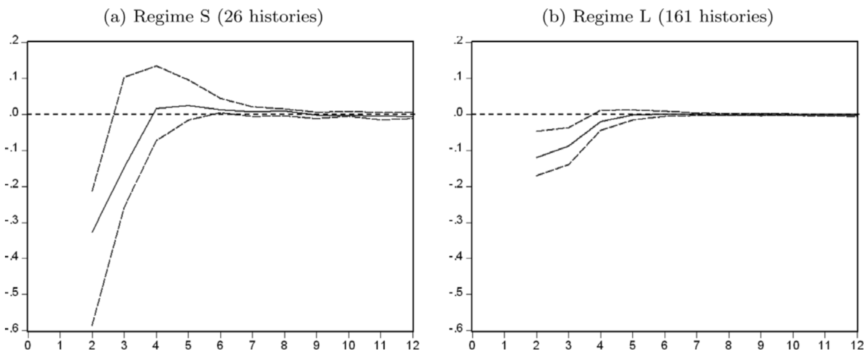

FIGURE 2. Generalized impulse-response functions. Threshold variable: FFR-based monetary shock. Generalized impulse-response functions of y_t^{C} to a positive shock to the monetary policy variable when the FFR is the monetary instrument, computed as described in Appendix D. The sizes of the shocks correspond to a standard deviation difference between the small and large shocks, with the estimated threshold as the middle point.

and sequence of shocks (both monetary and idiosyncratic) for q periods ahead. Second, the same procedure is carried out, given the same particular history and sequence of shocks, with the exception that the monetary shock to the transitory component of output in period 0 is fixed at a particular value. The shocks are fed through equation (3) and a forecast is produced as explained previously. Third, these steps are repeated 100 times and averaged out across individual Monte Carlo repetitions. Fourth, given the arbitrary shock and particular history, the difference between the averaged forecasts is taken to form a Monte Carlo estimate of the GIRF.

For regimes S and L, respectively, positive small and large monetary shocks to the transitory component of output are fixed in period zero so that they fall below or above the estimated threshold $\hat{\gamma} = 0.157$.[12] In particular, they are set to 0.08 and 0.23. Figure 2 presents the GIRFs for the transitory component of output when the FFR is the monetary instrument. Because each particular history generates a given forecast of y_t^{T}, the medians of these forecasts are reported, together with the 25th and 75th quantiles (dashed lines), which serve as bands for the GIRFs.

The left panel of Figure 2 plots the response of y_t^{T} for $q = 10$ periods ahead in regime S, that is, when the monetary shock hitting the system is small, corresponding to 26 possible histories. The right panel plots the response of y_t^{T} for the same number of periods ahead in regime L, corresponding to 161 possible histories.

It is important to mention that, because of outside lags of monetary policy, the shock at $t = 0$ will affect y_t^{C} at $t = 2$. As a consequence, the response of the transitory component of output at $t = 2$ is a combination of the response coefficient, α_2^{S} or α_2^{L}, depending on which regime prevails, and the effect of the particular history of the system, captured by the autoregressive coefficients, ϕ_1 and ϕ_2.

Given the estimated threshold, it is particularly interesting to note that most monetary policy shocks fall under the large-shock regime (i.e., there are few small-shock histories). Because the time distribution of those shocks does not concentrate on a specific time period in the sample, we infer from this result that the normal state of affairs in the U.S. economy is one of large monetary policy shocks. This makes a strong case for the importance of the smoothing of interest rates discussed previously. If most shocks are deemed relatively large by the estimated threshold, then the monetary authority will obtain a larger effect of policy by gradually carrying out smaller changes in the FFR.

From Figure 2, three observations can be made. First, the monetary shocks do not have an effect on output on impact, because only lags of x_t enter equation (3). When there is an effect, however, the response of the transitory component of output to "small" monetary shocks is much larger than its response to "large" monetary shocks. Indeed, the transitory component of output falls by 0.32 percentage points in regime S, when monetary shocks are "small," whereas it falls by only 0.12 percentage points in regime L, when monetary shocks are "large."

A second finding also supports the results found in Table 2. Over time, the response of the transitory component of output in regime S (when small monetary shocks hit the economy) is larger than that in regime L. Graphically, this is easily seen by comparing the magnitude of the median response of the transitory component of output in each regime. Such median responses reach their maxima in period 2 for both regimes (-0.32 in regime S and -0.12 in regime L). The difference in the magnitude of the response of the transitory component becomes even bigger if the 75th quantiles are considered. Their maxima are reached again in period 2 for both regimes (-0.58 in regime S and -0.17 in regime L). Furthermore, the accumulated median response of the transitory component of output in regime S is twice that in regime L (-0.48 and -0.24, respectively). Based on this evidence, the transitory response of output exhibits an overall larger response when small monetary shocks hit the economy, even controlling for future monetary and idiosyncratic shocks and for the history of the economic conditions. That is, they support the implications of menu-costs models.

Nonetheless, not all implications of menu-costs models are supported. As can be observed from Figure 2, a third finding shows that the response of output to large monetary shocks is different than zero at least in the first four periods, contrary to the predictions of menu costs. A potential explanation for this inconsistency between theory and data resides in the implicit assumption behind the implications of menu-costs models. According to these models, when large monetary shocks disturb the economy, the menu cost becomes secondary and relatively small and, thus, agents adjust their prices, leaving real balances and hence real output, unaffected. However, this result assumes that all firms adjust their prices and that they adjust them to the optimal price level. Neither of these assumptions seems to hold true, as shown in the data. Firms may be heterogeneous in the way they interpret monetary shocks. Moreover, they face imperfect information in the sense that, even when the menu cost is relatively low, their adjusted prices need not match the optimal price level.

4. CONCLUDING REMARKS

This paper investigates the potential sources of the mixed evidence in the empirical literature studying asymmetries in the response of output to monetary policy shocks of different sizes. It argues that such mixed evidence is a consequence of the exogenous imposition of the threshold that classifies monetary shocks as small or large. To overcome this situation, an unobserved-components model that allows estimation of the threshold parameter is proposed. Once this threshold is estimated from the data, there is strong evidence of asymmetry in the real effects of monetary policy shocks of different sizes. In contrast, when the bootstrap procedure for testing linearity is carried out imposing one standard deviation as an exogenous threshold, as is typical in the literature, linearity cannot be rejected. This supports the idea that the optimal threshold should be estimated endogenously, as the definition of size plays a crucial role in determining whether output varies disproportionately with the size of the monetary shock.

The estimated coefficients are consistent with the implications of menu-costs models. In particular, the response of the transitory component of output to small monetary policy shocks is larger than its response to large monetary policy shocks. Furthermore, the analysis of the GIRFs suggests that the difference in the response coefficients persists for many quarters. The policy implications of this finding are straightforward and significant. A reduction in the FFR always increases output, but this effect is both larger and more persistent when the decrease in the FFR is small, as defined by the endogenous threshold. As a consequence, during normal times, the Fed would impact output more effectively by carrying out gradual changes in the monetary instrument, a result that is consistent with the implications of the interest-rate-smoothing literature.

Given the estimated threshold, it is particularly interesting to note that most monetary policy shocks fall under the large-shock regime. Because the time distribution of those shocks is not concentrated on a specific time period in the sample, we infer from this result that the normal state of affairs in the U.S. economy, prior to the Great Recession, is one of large monetary policy shocks. This has, potentially, important policy implications. Given that the economy is typically in a large-shock regime, monetary authorities should exercise caution in determining the optimal size of policy changes to maximize their effects on the transitory component of output.

Meanwhile, the dynamic response of output to large monetary shocks reveals that such responses are significantly different from zero in the first few quarters following a shock. This result is at odds with standard menu-costs models that imply a neutral response of output to large monetary shocks. A potential explanation for this fact arises from the assumptions implicit in these models. In particular, they assume that when a shock is large, all firms adjust their prices. That is, it is assumed that all firms are homogeneous, which is debatable. Furthermore, menu-costs models assume that when firms do adjust their prices, they do so optimally (i.e., firms have perfect information). Along these lines, it would be interesting to evaluate the response of output to monetary shocks of different size when firms are allowed to be heterogeneous in their responses and face imperfect information. This issue is left for future research.

NOTES

1. Nonlinearities in the relationship between monetary policy and output, in general, have important implications for the way monetary policy is conducted. See, for example, Granger and Teräsvirta (1993); Anderson and Vahid (2001); Franses and Teräsvirta (2001); and Morley (2009).

2. In a univariate ARMA(p, q) process, the identification condition for estimating a nonzero covariance between the trend and transitory innovations is that $p \geq q + 2$. See Morley et al. (2003) for further details.

3. The focus of this paper is on thresholds, not break dates. Given that many authors have estimated the Great Moderation to have begun in the mid-1980s, the break date is set to the last quarter of 1983, broadly consistent with past findings.

4. Note that γ would be difficult to identify empirically if it were estimated simultaneously with the other parameters. Because small changes in the threshold would not lead to changes in the value of the likelihood function, the numerical optimization procedure would crash.

5. It is standard practice for TAR models to exclude 15% of the observations at each end of the vector of ordered thresholds to avoid distortions in inference. If possible thresholds that are too close to the beginning or the end of the ordered data were considered, there would not be enough observations to identify the subsample parameters strongly.

6. With quarterly data, $|\Gamma| = 138$ and for $B = 99$ bootstrap samples, 13,662 potential models need to be estimated using numerical optimization. Even if convergence for each model were achieved after only 30 seconds, the bootstrap procedure would require, approximately, 114 hours.

7. To investigate the robustness of the results to an alternative measure of monetary policy, the model is estimated considering M1 as the monetary instrument. The results, available from the author upon request, were not very different and thus are omitted from the final version of the document.

8. Because there is no unconditional expectation to initialize the Kalman filter for this model, a high variance is placed on the initial values. To avoid distortions arising from this initialization procedure, and to prevent sensitivity of the model to the initial values, the first 20 observations are disregarded when evaluating the likelihood. For further details about the initialization of the Kalman filter, refer to Kim and Nelson (1999).

9. Some authors would argue in favor of a VARMA specification to identify the monetary shock on the basis that, in their view, most macroeconomic time series follow an unknown ARMA process. To the extent that any invertible VARMA process can be approximated by a finite long order VAR, this is not a concern in this paper (see Athanasopoulos and Vahid (2008) for further details).

10. Given that a grid search over all possible values of the threshold parameters is necessary to estimate the model, bootstrapping the p-value is very time-consuming. As a consequence, only 99 bootstrap samples are used to test linearity. Even though a small number of bootstrap samples weakens power (but does not affect size), this is not a particular concern here because linearity is rejected.

11. The GIRFs are averaged over different histories taken from subsamples of the data. For instance, the GIRFs for the "small" monetary shock regime are computed averaging out over histories (or initial values for the first two lags of the transitory component) corresponding to all dates on which the given monetary shock was smaller than $\hat{\gamma}$.

12. This guarantees that the "small" ("large") shock is below (above) the estimated threshold, triggering a response of output captured by the $\alpha^{\mathrm{S}}(\alpha^{\mathrm{L}})$ coefficients.

13. That is, all agents set the price to $p_j^* = p_0$, for all $j = 1, ..., N$, so that their optimal price is consistent with the price level, leaving relative prices, and thus output, unchanged.

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

REFERENCES

Anderson, H.M. and F. Vahid (2001) Predicting the probability of a recession with nonlinear autoregressive leading-indicator models. *Macroeconomic Dynamics* 5, 482–505.

Athanasopoulos, G. and F. Vahid (2008) VARMA versus VAR for macroeconomic forecasting. *Journal of Business and Economic Statistics* 26, 237–252.

Ball, Lawrence and N. Gregory Mankiw (1994) Asymmetric price-adjustment and economic fluctuations. *Economic Journal* 423, 247–261.

Ball, Lawrence and David Romer (1990) Are prices too sticky? *Quarterly Journal of Economics* 104(3), 507–524.

Beaudry, Paul and Gary Koop (1993) Do recessions permanently change output? *Journal of Monetary Economics* 31, 149–163.

Blanchard, Olivier and N. Kiyotaki (1987) Monopolistic competition and the effects of aggregate demand. *American Economic Review* 77, 647–666.

Caplin, A.S. and J. Leahy (1991) State dependent pricing and the dynamics of money and output. *Quarterly Journal of Economics* 102, 703–725.

Christiano, L., M. Eichenbaum, and C. Evans (1996) The effects of monetary policy shocks: Evidence from the flow of funds. *Review of Economics and Statistics* 78(1), 16–34.

Christiano, L., M. Eichenbaum, and C. Evans (1999) Monetary policy shocks: What have we learned and to what end? In J. Taylor and Michael Woodford (eds.), *Handbook of Macroeconomics* vol. 1(2), pp. 65–148. Amsterdam: Elsevier.

Christiano, L., M. Eichenbaum, and C. Evans (2005) Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113(1), 1–45.

Cover, James P. (1992) Asymmetric effects of positive and negative money-supply shocks. *Quarterly Journal of Economics* 107, 1261–1282.

DeLong, J.B. and Lawrence Summers (1988) How does macroeconomics policy affect output. *Brookings Papers on Economic Activity* 2, 433–494.

Franses, P.H. and T. Teräsvirta (2001) Introduction to the special issue: Nonlinear modeling of multivariate macroeconomic relations. *Macroeconomic Dynamics* 5, 461–465.

Garcia, R. and H. Schaller (2002) Are the effects of monetary policy asymmetric? *Economic Inquiry* 40(1), 102–119.

Golosov, M. and Robert E. Lucas (2007) Menu Costs and Phillips Curves. NBER working paper 10187.

Granger, C.W.J. and T. Teräsvirta (1993) *Modelling Nonlinear Economic Relationships*. Oxford, UK: Oxford University Press.

Hamilton, James (1994) *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hansen, Bruce E. (1996) Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* 64, 413–430.

Hansen, Bruce E. (1997) Inference in TAR models. *Studies in Nonlinear Dynamics and Econometrics* 2(1), 1–14.

Kim, C.J. and Charles R. Nelson (1999) *State Space Models with Regime-Switching: Classical and Gibb-Sampling Approaches with Applications*. Cambridge, MA: MIT Press.

Koop, Gary, M.H. Pesaran, and Simon Potter (1996) Impulse-response analysis in nonlinear multivariate models. *Journal of Econometrics* 74, 119–147.

Leeper, E., C. Sims, and T. Zha (1996) What does monetary policy do? *Brookings Papers on Economic Activity* 2, 1–63.

Lo, M.C. and Jeremy Piger (2005) Is the response of output to monetary policy asymmmetric? Evidence from a regime-switching coefficients model. *Journal of Money, Credit and Banking* 37, 865–887.

Mankiw, N. Gregory (1985) Small menu costs and large business cycles: A macroeconomic model of monopoly. *Quarterly Journal of Economics* 100, 529–539.

Morley, J. (2009) Macroeconomics, nonlinear time series in. In R.A. Meyers (ed.), *Encyclopedia of Complexity and System Science*, pp. 5325–5348. Berlin: Springer.

Morley, J. (2011) The two interpretations of the Beveridge–Nelson decomposition. *Macroeconomic Dynamics* 15, 419–439.

Morley, J., C.R. Nelson, and E. Zivot (2003) Why are the Beveridge–Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics* 85(2), 235–243.

Morley, J. and J. Piger (2012) The asymmetric business cycle. *Review of Economics and Statistics* 94, 208–221.

Ravn, M. and M. Sola (2004) Asymmetric effects of monetary policy in the U.S.: Positive versus negative or big versus small? *Federal Reserve Bank of St. Louis Review* 86(5), 41–60.

Rothman, P., D. van Dijk, and P.H. Franses (2001) A multivariate STAR analysis of the relationship between money and output. *Macroeconomic Dynamics* 5, 506–532.

Sensier, M. (1996) The Asymmetric Effect of Monetary Policy in the U.K., Unpublished working paper, University of Oxford.

Sims, C. and T. Zha (1998) Comment on Glenn Rudebushch's "Do measures of monetary policy in a VAR make sense?" (with reply). *International Economic Review* 39, 933–948.

Sims, C. and T. Zha (2006) Does monetary policy generate recessions? *Macroeconomic Dynamics* 10(2), 231–272.

Sinclair, T. (2009) Asymmetry in the business cycle: Friedman's plucking model with correlated innovations. *Studies in Nonlinear Dynamics and Econometrics* 14(1), Article 3.

Teräsvirta, T. and H.M. Anderson (1992) Characterizing nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics* 7, S119–136.

Thoma, M. (1994) Subsample instability and asymmetries in money-income causality. *Journal of Econometrics* 64, 279–306.

Thoma, M. (2007) The Costs of Being Wrong about Lags in Monetary Policy. Working paper, University of Oregon.

Watson, M.W. (1986) Univariate detrending methods with stochastic trends. *Journal of Monetary Economics* 18, 49–75.

Weise, C. (1999) The asymmetric effects of monetary policy: A nonlinear vector autoregression approach. *Journal of Money, Credit and Banking* 31, 85–108.

APPENDIX A: A MENU COST MODEL TO DISTINGUISH BETWEEN SMALL AND LARGE SHOCKS

Why should we expect a different response of real economic activity when the size of monetary shocks differs? To motivate the distinction between small and large monetary shocks and to make the empirical analysis clear, this section presents a standard menu-cost model. The model implies that output responds differently, depending on whether monetary shocks are below or above a certain threshold.

The model presented here follows Ball and Romer (1990) and Ravn and Sola (2004). The economy consists of N producers–consumers, each of whom produces a differentiated good. The price of each good is set by each firm and equals $p_j,\ j = 1, \ldots, N$. The only friction in the model, which gives rise to nominal rigidities, is introduced through the assumption that firms must pay a menu cost if their price is adjusted. After observing the current price level p, each firm sets its price to p_j. However, this price will not necessarily prevail throughout the entire period. By paying a menu cost $c > 0$, firm j can adjust its

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

price at any point in time. Firm j's profit function π_j can be described according to

$$\pi_j = V(m/p, p_j/p) - cd_j, \qquad \textbf{(A.1)}$$

where d_j is a dummy variable that equals 1 if firm j adjusts its nominal price and 0 otherwise, $V(.)$ is increasing in both arguments, and m is the money stock in the economy.

In a symmetric equilibrium in which prices are not changed, the optimal price p_j^* for firm j is implicitly defined by the first-order condition

$$\frac{\partial \pi_j}{\partial p_j} \equiv \frac{\partial V(m/p, p_j^*/p)}{\partial(p_j/p)} = 0.$$

Because the equilibrium is symmetric, shocks to the stock of money m do not have real effects.[13] For simplicity, m is normalized to 1 and, thus, the equilibrium corresponds to the situation in which $m = p = p_j^* = 1$.

To show how the size of the shock to the stock of money can matter, firm j's profit function when its price is left unchanged is typically compared with its profit when it decides to adjust its price and incur the menu cost c, given that all other firms do not.

If the central bank unexpectedly changes the stock of money to $m \neq 1$, then firms need to evaluate whether it is optimal to leave their prices unchanged or adjust them. In the first case, profit for firm j when prices are not changed (sticky prices) is given by $\pi_{\text{sp}} = V(m, 1)$. However, if firm j decides to adjust its price, then it must pay the menu cost c and its profit is given by $\pi_{\text{ap}} = V(m, p_j^*/p) - c$. Hence, firms will not adjust their prices if $\pi_{\text{ap}} < \pi_{\text{sp}}$, which is equivalent to

$$V(m, p_j^*/p) - V(m, 1) < c.$$

Ball and Romer (1990) show that, when a second-order Taylor expansion is performed around $m = 1$, the range of money stock for which firms will not adjust their prices (that is, the monetary shock is so small that incurring the menu cost c is not optimal) is given by

$$m \in (1 - \bar{m}, 1 + \bar{m}), \qquad \textbf{(A.2)}$$

where $\bar{m} = (-2cV_{22}/V_{12}^2)^{1/2}$ and $V_{22}(V_{12})$ represents the second (mixed) derivative of $V(.)$ with respect to the second argument (first and second arguments). Similarly, the range of the money stock for which firms will adjust their prices (so that changes are neutral) is given by

$$m \in (-\infty, \bar{\bar{m}}) \cup (\bar{\bar{m}}, \infty), \qquad \textbf{(A.3)}$$

where $\bar{\bar{m}} = (-2c/V_{22})^{1/2}$. Hence, equations (A.2) and (A.3) show that real economic activity could behave differently if the size of the monetary shock differed. In particular, "small" monetary shocks can have real effects if $m \in (1 - \bar{m}, 1 + \bar{m})$ and "large" monetary shocks can be neutral if $m \in (-\infty, \bar{\bar{m}}) \cup (\bar{\bar{m}}, \infty)$.

APPENDIX B: REGULARITY CONDITIONS FOR THE UC MODEL

To guarantee that the statistic given in equation (5) converges to the asymptotic distribution derived by Hansen (1997), it can be shown that the UC model considered in (1)–(3) satisfies the regularity conditions he provides. For convenience, the model is reproduced here:

$$y_t = y_t^{\mathrm{T}} + y_t^{\mathrm{C}}, \tag{B.1}$$

$$y_t^{\mathrm{T}} = \mu + y_{t-1}^{\mathrm{T}} + \nu_t, \tag{B.2}$$

$$y_t^{\mathrm{C}} = \sum_{p=1}^{P} \phi_p y_{t-p}^{\mathrm{C}} + \sum_{j=1}^{J} \alpha_j^{\mathrm{S}} x_{t-j} I[q(t) \leq \gamma] + \sum_{j=1}^{J} \alpha_j^{\mathrm{L}} x_{t-j} I[q(t) > \gamma] + \epsilon_t. \tag{B.3}$$

It is known that this UC model implies an equivalent ARIMA representation for $\{y_t\}$. This can easily be seen by substituting equations (B.2) and (B.3) into (B.1) and then taking first differences:

$$(1-L)y_t = \mu + \nu_t + (1-L)\phi_p^{-1}(L) \left\{ \sum_{j=1}^{J} \alpha_j^{\mathrm{S}} I[q(t) \leq \gamma] + \sum_{j=1}^{J} \alpha_j^{\mathrm{L}} I[q(t) > \gamma] + \epsilon_t \right\}, \tag{B.4}$$

where L is the lag operator and $\phi_p(L)$ is a polynomial of order p in the lag operator. Then, from premultiplying the previous equation by $\phi_p(L)$ and rearranging terms, it follows that

$$\phi_p(L)\Delta y_t = \mu^* + \psi_t \tag{B.5}$$

and $\psi_t = \theta_p(L)\zeta_t$, with ζ_t being a linear combination of the original innovations ϵ_t and ν_t captured by the coefficients in the polynomial $\theta_p(L)$. For the specific case of the model considered in the paper, $p = 2$, and then (B.5)—and, as a consequence, also the UC model from (B.1)–(B.3)—becomes an ARMA(2,2) process:

$$\Delta y_t = \mu^* + \phi_1 \Delta y_{t-1} + \phi_2 \Delta y_{t-2} + \psi_t, \tag{B.6}$$

where $\psi_t = \zeta_t + \theta_1 \zeta_{t-1} + \theta_2 \zeta_{t-2}$. It is straightforward to check that (B.5) satisfies the regularity conditions provided in Hansen (1997). In particular,

1. $E(\psi_t) = 0$, $E(\psi_t^2) < \infty$, $E|\psi_t|^{2+\delta} < \infty$, and ψ_t has a density function $f(.)$ that is continuous and positive everywhere on $\mathbf{R}$, because ψ_t is a function of ϵ_t and ν_t.
2. Ergodicity is no longer required, because the regime switching occurs with respect to the x_t variable, which is exogenous, and not with respect to the autoregressive terms.
3. To ensure the identification of the regimes, it is imposed that $\alpha_j^{\mathrm{S}} \neq \alpha_j^{\mathrm{L}}$, for some j.

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

APPENDIX C: STATE-SPACE REPRESENTATION FOR THE UC MODEL WITH A TAR TRANSITORY COMPONENT

The state-space representation for the general $P = p$ and $J = j$ system given in equations (1)–(3) is provided here. The observation equation is given by

$$y_t = \begin{bmatrix} 1 & \mathbf{0}_{1\times(p-1)} & 1 \end{bmatrix} \begin{bmatrix} y_t^{\mathrm{C}} \\ y_{t-1}^{\mathrm{C}} \\ \vdots \\ y_{t-p+1}^{\mathrm{C}} \\ y_t^{\mathrm{T}} \end{bmatrix}.$$

The state equation is given by

$$\begin{bmatrix} y_t^{\mathrm{C}} \\ y_{t-1}^{\mathrm{C}} \\ \vdots \\ y_{t-p+1}^{\mathrm{C}} \\ y_t^{\mathrm{T}} \end{bmatrix} = F_{(p+1)\times(p+1)} \begin{bmatrix} y_{t-1}^{\mathrm{C}} \\ y_{t-2}^{\mathrm{C}} \\ \vdots \\ y_{t-p}^{\mathrm{C}} \\ y_{t-1}^{\mathrm{T}} \end{bmatrix}_{(p+1)\times 1} + \begin{bmatrix} \mathbf{0}_{p\times 1} \\ \mu \end{bmatrix} + G^i_{(p+1)\times j} \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-j} \end{bmatrix}_{j\times 1}$$

$$+ \begin{bmatrix} v_t \\ \mathbf{0}_{(p-1)\times 1} \\ \epsilon_t \end{bmatrix}_{(p+1)\times 1} \quad \text{for } i = 1, 2,$$

where

$$F = \begin{bmatrix} \Phi_{1\times p} & 0 \\ I_{p-1} & \mathbf{0}_{(p-1)\times 2} \\ \mathbf{0}_{1\times p} & 1 \end{bmatrix},$$

with $\Phi_{1\times p} = \begin{bmatrix} \phi_1 & \cdots & \phi_p \end{bmatrix}$, I_{p-1} being the identity matrix of order $(p-1)$ and $\mathbf{0}_{i\times j}$ being a matrix with i rows and j columns of zeros, and

$$G^1 = \begin{bmatrix} \alpha_1^{\mathrm{S}} & \cdots & \alpha_j^{\mathrm{S}} \\ 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{bmatrix}_{(p+1)\times j}$$

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

and

$$G^2 = \begin{bmatrix} \alpha_1^{\mathrm{L}} & \cdots & \alpha_j^{\mathrm{L}} \\ 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{bmatrix}_{(p+1)\times j}.$$

The variance–covariance matrix of the transitory component is given by

$$Q = E\left\{\begin{bmatrix} \nu_t \\ \mathbf{0}_{(p-1)\times 1} \\ \epsilon_t \end{bmatrix}\begin{bmatrix} \nu_t & \mathbf{0}_{1\times(p-1)} & \epsilon_t \end{bmatrix}\right\}$$

$$= \begin{bmatrix} \sigma_\nu^2 & 0 & \cdots & 0 & \sigma_{\epsilon\nu} \\ 0 & 0 & & 0 & 0 \\ & & \ddots & & \vdots \\ & & & & \\ \sigma_{\epsilon\nu} & 0 & \cdots & 0 & \sigma_\epsilon^2 \end{bmatrix}_{(p+1)\times(p+1)}.$$

APPENDIX D: COMPUTATION OF GENERALIZED IMPULSE-RESPONSE FUNCTIONS

The procedure to compute the generalized impulse-response functions (GIRFs) follows the one described in Koop et al. (1996). The reader is referred there for further details.

A GIRF can be defined as the effect of a one-time shock on the forecast of variables in a particular model, given a specific history. The response constructed must then be compared to a benchmark "no shock" scenario. In this way, the GIRF can be expressed as follows:

$$\mathrm{GI}_Y(q, \nu_t, \omega_t) = E\left[Y_{t+q}, \nu_t, \omega_{t-1}\right] - E\left[Y_{t+q}/\omega_{t-1}\right],$$

where GI_Y is the generalized impulse-response function of a variable Y for period q, given the specific history ω_{t-1} and initial shock ν_t, and $E[\,.\,]$ is the expectations operator.

To compute the GIRF, the conditional expectations in the preceding equation are simulated. The nonlinear model is assumed to be known (i.e., model uncertainty is ignored). The shock to Y, ν_0, occurs in period 0, and responses are computed q periods ahead. Thus, the GI_Y function is generated according to the following steps:

Step 1: Pick a history ω_{t-1}. The history is the actual value of the lagged endogenous variables at a particular date, or for a particular episode (e.g., those values of the endogenous variables that fall under regime S).

Step 2: Pick a sequence of two-dimensional shocks $\nu_{j,t+q}$, $q = 0, 1, \ldots, n$. This vector of shocks includes both monetary and idiosyncratic shocks. They are drawn with

https://doi.org/10.1017/S1365100513000278 Published online by Cambridge University Press

replacement from the vector of monetary shocks—the residuals from the identified VAR—and from the estimated residuals of the transitory component of the model.

Step 3: Using $\omega_{i,t-1}$ and $\nu_{j,t+q}$, simulate the path for y_{t+q} over n periods according to equation (3). This benchmark path is denoted as $Y_{t+q}(\omega_{i,t-1}, \nu_{j,t+q})$ for $q = 1, \ldots, n$.

Step 4: Using the same history, $\omega_{i,t-1}$, and shocks, $\nu_{j,t+q}$, as in the previous step, plus an additional initial shock ν_0 (the small or large monetary policy shock), simulate the path for y_{t+q} over $n+1$ periods according to the equation for the transitory component of output. This profile path is denoted $Y_{t+q}(\nu_0, \omega_{i,t-1}, \nu_{j,t+q})$ for $q = 0, 1, \ldots, n$.

Step 5: Repeat steps 2 to 4 B times.

Step 6: Repeat steps 1 to 5 R times and compute the quantiles of the difference between the profile and benchmark paths $Y_{t+q}(\nu_0, \omega_{i,t-1}, \nu_{j,t+q}) - Y_{t+q}(\omega_{i,t-1}, \nu_{j,t+q})$.