Do Monetary Policy Shocks Generate TAR or STAR Dynamics in Output?

Luiggi Donayre†

February 2014

Abstract

This paper studies whether the relationship between monetary policy shocks of different size and output is better described by threshold autoregressive (TAR) or smooth transition autoregressive (STAR) dynamics. Using a Bayesian framework, a TAR process and a STAR process are formally compared within an unobserved components model of output, augmented with a monetary policy variable. The Bayesian model comparison favors the notion that the dynamics are nonlinear and better described by a smooth transition between regimes, which suggests that aggregation plays a role in the dynamics between monetary policy and output. This evidence is further supported by the results of a model that uses output data at the sectoral level: when more disaggregated data are employed, the transition between regimes is more abrupt. Moreover, the results show that, when the transition between regimes is smooth, large monetary policy shocks identified as the residuals of a VAR are neutral, consistent with the implications of menu cost models.

JEL classification codes: C11, C22, E52

Keywords: Bayesian Analysis, Asymmetry, Monetary Policy, Smooth Transition Autoregressive Process, Threshold Autoregressive Process, Unobserved Components Model, MCMC Methods.

†I am thankful to James Morley for valuable suggestions and discussion. Helpful comments from two anonymous referees, Steven Fazzari, Irina Panovska and the participants at the Applied Time Series Research Group at Washington University in St. Louis, the Midwest Econometrics Group, the University of Minnesota-Duluth, the University of Wisconsin-Whitewater, the Society for Nonlinear Dynamic Economics Symposium, the NBER-NSF Seminar on Bayesian Inferences in Econometrics and Statistics and the Georgetown Center for Economic Research Conference are gratefully acknowledged. All remaining errors are my own. Author: Department of Economics, Labovitz School of Business and Economics, University of Minnesota - Duluth, 1318 Kirby Drive, Duluth, MN 55812. Phone: (218) 726-8463. Email: adonayre@d.umn.edu.
1 Introduction

Several studies have considered different nonlinear models to investigate whether the effects of monetary policy shocks on output vary disproportionately with the size of the shock (Sensier, 1996; Weise, 1999; Ravn and Sola, 2004; Lo and Piger, 2005). Despite earlier mixed results regarding the significance of this asymmetry, Donayre (2014) has recently found evidence that output responds asymmetrically to the magnitude of the monetary shock once the size threshold is endogenously estimated. Meanwhile, even though there are theoretical and empirical arguments that motivate this size asymmetry, little attention has been paid to the nature of the implied nonlinear relationship.

This paper adopts a Bayesian approach to analyze whether this nonlinear relationship is better described by threshold autoregressive (TAR) or smooth transition autoregressive (STAR) dynamics. In particular, a TAR process and a logistic STAR (LSTAR) process are formally compared within an unobserved components model of output, augmented by a monetary policy variable. The question of whether the transition between regimes is abrupt (as implied by TAR dynamics) or smooth (as implied by STAR dynamics) is important for two reasons. First, several authors have argued that a discrete switching may be appropriate when considering the effects of monetary policy on disaggregated prices, or when firms and goods are identical (Dumas, 1994; Teräsvirta, 1994; Sarno, 2003). This is less likely to occur at the aggregate level. Within the relationship between monetary policy and aggregate economic activity, a gradual adjustment, rather than an abrupt one, may be more appropriate. At the economy-wide level, the aggregate response of output is more likely to change smoothly, as firms adjust their prices non-simultaneously. Second, a smooth transition between regimes could intrinsically be related to the reasons central banks around the world have tended to change short-term interest rates in sequences of small steps, thus providing support for the interest-rate smoothing behavior of central banks.

From a methodological perspective, the Bayesian approach facilitates the estimation and comparison of TAR and LSTAR dynamics that would be imprecise or infeasible in a frequentist environment. First, all parameters can be easily jointly estimated without relying on grid-search procedures, which increases precision. Second, a Bayesian model comparison using marginal likelihoods is conceptually straightforward between TAR and LSTAR dynamics. Third, the Bayesian model comparison accounts for the degree of parameterization.1

Using data for the U.S., the results find strong evidence of nonlinear threshold dynamics in the relationship between monetary policy and output. That is, there is evidence that the effects of monetary

---

policy on output vary disproportionately with the size of the monetary shock. Further, the analysis of marginal likelihoods favors a smooth transition between regimes when the shock is identified from a VAR. When the UC-LSTAR model is re-estimated using disaggregated data on output, the estimated smoothing parameter indicates more discrete regime changes. This finding suggests that aggregation plays an important role determining the nature of the nonlinear relationship between monetary policy and output. Meanwhile, the results from the estimated model when the regime-switching is smooth suggest that the response of output to small monetary shocks is large and the response to large shocks is neutral, consistent with the key implications of menu-costs models. The results are robust to potentially asymmetric size thresholds and to monetary policy shocks identified using a Taylor rule. But caution should be exerted when interpreting the results in the case of policy shocks identified using the narrative approach from Romer and Romer (2004).

The remainder of this paper is organized as follows. In the second section, an UC model with linear and threshold-type nonlinear dynamics for the transitory component of output is described. The Bayesian approach and model comparison are also discussed. The third section describes the data and the results. Robustness is checked in the fourth section. The fifth section re-estimates the model using disaggregated data. Some concluding remarks are provided in the last section.

2 Empirical Models

Three different types of dynamics (linear, TAR and LSTAR) are considered within the framework of an UC model. Typically, economists are interested in the effects of monetary policy on the output gap (i.e., deviations of output from its potential level). In an UC framework, such effects are directly modeled by measuring the output gap as the transitory component of output. Further, decomposing output into a permanent and a transitory component allows the monetary policy shock to affect only the latter, consistent with the notion of long-run money neutrality. The general UC model can be described by:

\[ y_t = y_t^T + y_t^C \]  \hspace{1cm} (1)

\[ y_t^T = \mu + y_{t-1}^T + \nu_t \]  \hspace{1cm} (2)

\[ y_t^C = F(z) + \epsilon_t \]  \hspace{1cm} (3)
where \( y_t \) is a measure of output, \( y_t^P \) is the permanent (or trend) component of output, \( y_t^C \) is the transitory (or cyclical) component of output; \( z \) is a \( 1 \times k \) matrix that includes lags of \( y_t^C \) and lags of a measure of monetary policy; and \( F(\cdot) \) captures the functional form between output and monetary policy.

The system (1)-(3) is a modified version of the simple UC decomposition of real output into the 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, \( \mu \). However, the innovations \( \epsilon_t \) and \( \nu_t \) are allowed to be correlated here. In particular, they have a joint normal distribution with mean zero and non-diagonal variance-covariance matrix \( \Omega \), following Morley, Nelson, and Zivot (2003).

For the benchmark case, the UC model is characterized by a linear transitory component (UC-linear) and can be described by (1)-(2) and the following autoregressive distributed lag (ADL) process replacing (3):

\[
y_t^C = \sum_{p=1}^{P} \phi_p y_{t-p}^C + \sum_{j=1}^{J} \alpha_j x_{t-j} + \epsilon_t
\]

where \( x_t \), the independent variable, is a measure of monetary policy; all roots of the polynomial \( \phi(L) = 1 - \sum_{p=1}^{P} \phi_p L^p \) lie outside the unit circle; \( \phi_p, p = 1, \ldots, P \) are autoregressive coefficients; and \( \alpha_j, j = 1, \ldots, J \) captures the response of \( y_t^C \) to monetary policy.\(^2\)

In terms of the nonlinear dynamics, the UC model characterized by a TAR-driven transitory component (UC-TAR) can be described by (1)-(2) and the following ADL replacing (3):

\[
y_t^C = \sum_{p=1}^{P} \phi_p y_{t-p}^C + \sum_{j=1}^{J} \alpha_j^S x_{t-j} I(s_t \leq c) + \sum_{j=1}^{J} \alpha_j^L x_{t-j} I(s_t > c) + \epsilon_t
\]

where \( I(\cdot) \) denotes the indicator function; \( x_t \) and \( \phi_p, p = 1, \ldots, P \) have the same interpretation as in (4); \( s_t \) is the threshold variable; and \( c \) is the threshold parameter. When \( s_t \leq c \), the response-coefficients are captured by the \( J \times 1 \) vector \( \alpha^S \) and when \( s_t > c \), they are captured by the \( J \times 1 \) vector \( \alpha^L \). Note that the autoregressive coefficients \( \phi_p, p = 1, \ldots, P \) are not state-dependent.\(^3\)

\(^2\)In order to be consistent with the measures of monetary policy considered below, where the monetary variable does not affect output contemporaneously, only lags of \( x_t \) are allowed to enter equation (4).

\(^3\)The autoregressive dynamics are assumed to be the same in both regimes because the question of interest concerns the differences in the coefficients on the monetary shocks.
When the regime-switching is smooth, the UC model with a LSTAR-driven transitory component (UC-LSTAR) can be described by (1)-(2) and the following ADL replacing (3):

\[ y_t^C = \sum_{p=1}^{P} \phi_p y_{t-p}^C + \sum_{j=1}^{J} \alpha_j x_{t-j} + \sum_{j=1}^{J} \alpha_j^G x_{t-j} G(s_t) + \epsilon_t \] (6)

where \(x_t\) and \(\phi_p, \ p = 1, \ldots, P\) have the same interpretation as in (4); \(\alpha_j\) and \(\alpha_j^G, \ j = 1, \ldots, J\) are the response coefficients associated with the different regimes; and \(G(s_t)\) is a smooth function, bounded between 0 and 1, that can be described according to:

\[ G(s_t) = \left\{ 1 + \exp\left(-\gamma(s_t - c)\right) \right\}^{-1} \] (7)

where \(c\) is a location parameter and \(\gamma\) determines the smoothness of the change in the value of the logistic function (7). This transition function determines the weights put on each regime according to logistic specifications that depend on \(\gamma\). A desirable feature of this function is that it is the limiting case of either, a linear process (when \(\gamma \to 0\)), or a TAR process (\(\gamma \to \infty\)). To gain some insight regarding the importance of \(\gamma\) in determining the smoothness of the regime change, figure 1 displays the transition function for two values of the smoothness parameter, \(\gamma\). The left panel of figure 1 shows a logistic transition function that is relatively smooth, with associated \(\gamma = 2\). The right panel of figure 1 displays a more abrupt transition function, generated with \(\gamma = 15\). In both cases, \(c\) is set to 0.8 and the transition variable is \(s_t \in [0, 3]\).

An additional feature of the models estimated is that they allow for an exogenous, one-time break in the variance-covariance matrix to account for the Great Moderation.\(^5\) To reduce the dimensionality of the estimation, the parameter \(\lambda \in [0, 1]\) rescales the variance-covariance matrix \(\Omega\) to account for this reduction in volatility from a practical point of view. That is, the variance-covariance matrix after the break is given by \(\lambda\Omega\).\(^6\) This assumption is supported by the findings in Ahmed, Levin, and Wilson (2004), who cannot reject the hypothesis that the reduction in volatility in U.S. real GDP is

---

\(^4\)To the extent that the asymmetry studied in this paper is concerned with the size of the monetary policy shock, the threshold variable \(s_t\) is given by the absolute value of this shock in both nonlinear models. For this reason, the support for \(s_t\) in figure 1 is nonnegative.

\(^5\)The focus of this paper is not on break dates. Given that many authors have estimated the Great Moderation to begin in the mid 1980s, the break date is set to the first quarter of 1984, broadly consistent with previous findings.

\(^6\)This approach is also undertaken by Morley and Piger (2012) and Sinclair (2009) when accounting for the Great Moderation in UC models of U.S. real GDP.
proportional to the one prevailing during the Great Moderation.

2.1 Bayesian econometric approach

The Bayesian estimation of the models is conducted by means of a multiple-block Metropolis-Hastings (MH) algorithm with a random-walk chain proposal. To provide an accurate approximation of the target distribution, the proposal distribution is a multivariate Student t distribution, following much of the applied literature.

In the estimation of threshold-type nonlinear models, some issues arise in the implementation of the MH algorithm. The first one affects only the estimation time, as there is a need to grid-search across \( c \) and \( \gamma \) to find the posterior modes. However, this only applies to constructing the proposal distribution for Bayesian estimation for the first draw, given the random-walk chain proposal, and thus the grid-search procedure is innocuous. The second issue poses a problem to obtaining the scale of the proposal density, since the grid-search procedure makes it infeasible to use numerical derivatives to evaluate the curvature of the posterior. See appendix B for details on the approach followed to overcome these issues, the specification of the priors and the estimation procedure.

Once the models are estimated, they can be easily compared based on their marginal likelihoods. To formally test two models, the posterior odds ratio is calculated:
Pr(M_i/y) \Pr(M_j/y) = \frac{p(M_i) m(y/M_i)}{p(M_j) m(y/M_j)} \tag{8}

where the first factor in the right-hand side of (8) is the prior odds ratio (the ratio of the prior probability of model $i$ to the prior probability of model $j$). The second factor in the right-hand side of (8) is the Bayes factor, the ratio of the marginal likelihoods of the two models, which is given by $m(y/M_i) = \int_{-\infty}^{\infty} f(y/\theta, M_i)\pi(\theta/M_i)d\theta$ for any model $M_i$, and can be obtained following Chib and Jeliazkov (2001). Finally, for each iteration of the MH algorithm, 20,000 draws were considered after discarding the first 4,000 draws.

3 Results

3.1 Data and monetary shock identification

All data in this section are taken from the Federal Reserve Economic Data (FRED) database and are seasonally-adjusted from the source. Output is measured as one hundred times the natural logarithm of quarterly real Gross Domestic Product (GDP). The interest rate series considered is the monthly effective Federal Funds Rate (FFR), converted to quarterly frequency by means of a simple three-month arithmetic mean. Prices are measured using the natural logarithm of the GDP deflator. The sample goes from the third quarter of 1954 through the fourth quarter of 2007, corresponding to 214 observations.

To approximate the monetary policy variable, an interest rate-based monetary shock is constructed from the residuals of an identified VAR, which 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.

A correct identification of the policy shock is important to properly assess the effects of changes in monetary policy on the economy.\footnote{Not all variations in monetary policy can be accounted for as a reaction to the state of the economy. The fraction of this variation not accounted for is defined as a monetary policy shock.} The approach used to identify such shock is standard in the literature (Christiano, Eichenbaum, and Evans, 1996, 1999, 2005). 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.\footnote{That is, time $t$ variables in the Fed’s information set do not respond to time $t$ realizations of the monetary policy} The authors argue that this assumption justifies the broadly-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 (see Christiano et al. (2005) for further
discussion).

While the two-step approach makes the strong assumption that monetary policy shocks respond
linearly to economic conditions, a misspecification, in the end, would be reflected in a loss of asymmetry
in the response coefficients. To gain further insights, the robustness of the results to different measures
of $x_t$ is also evaluated in the next section.

The number of autoregressive coefficients for $y_t^C$, $P$, and the number of lags for the monetary shock,$J$, are obtained based on the Bayesian model comparison. Given quarterly data, the maximum number
of lags for each coefficient is set to 4.

3.2 Model comparison and posterior distribution

Table 1 below reports values of Bayes factors for all models estimated. Specifically, Bayes factors are
calculated as the ratio of the log marginal likelihood from the best model, in the numerator, and that
of a given model, in the denominator.

Table 1: Bayes factors for all models

<table>
<thead>
<tr>
<th></th>
<th>$p$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>UC-linear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.8851</td>
<td>0.9126</td>
<td>0.9043</td>
<td>0.8898</td>
</tr>
<tr>
<td>2</td>
<td>0.9070</td>
<td>0.9176</td>
<td>0.9113</td>
<td>0.8871</td>
</tr>
<tr>
<td>3</td>
<td>0.9022</td>
<td>0.9017</td>
<td>0.9109</td>
<td>0.8770</td>
</tr>
<tr>
<td>4</td>
<td>0.8924</td>
<td>0.8941</td>
<td>0.8973</td>
<td>0.8920</td>
</tr>
<tr>
<td>UC-TAR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.9142</td>
<td>0.9432</td>
<td>0.9625</td>
<td>0.8846</td>
</tr>
<tr>
<td>2</td>
<td>0.9515</td>
<td>0.9492</td>
<td>0.9521</td>
<td>0.9488</td>
</tr>
<tr>
<td>3</td>
<td>0.9077</td>
<td>0.8955</td>
<td>0.9551</td>
<td>0.8728</td>
</tr>
<tr>
<td>4</td>
<td>0.9028</td>
<td>0.8936</td>
<td>0.9500</td>
<td>0.9318</td>
</tr>
<tr>
<td>UC-LSTAR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.9810</td>
<td>0.9773</td>
<td>1.0000</td>
<td>0.9976</td>
</tr>
<tr>
<td>2</td>
<td>0.9924</td>
<td>0.9750</td>
<td>0.9928</td>
<td>0.8418</td>
</tr>
<tr>
<td>3</td>
<td>0.9930</td>
<td>0.9945</td>
<td>0.9722</td>
<td>0.8873</td>
</tr>
<tr>
<td>4</td>
<td>0.9877</td>
<td>0.9599</td>
<td>0.9602</td>
<td>0.8958</td>
</tr>
</tbody>
</table>

This table reports Bayes factors for the UC-linear, UC-TAR and
UC-LSTAR models for $p = 1, \ldots, 4$ and $j = 1, \ldots, 4$.

Based on the analysis of Bayes factors, the results in table 1 suggest that the model that best
describes the data is the UC-LSTAR model with $p = 3$ and $j = 1$. In general, there is strong evidence
shock.
that output varies disproportionately with the size of the monetary policy shock, as can be inferred from the fact that 22 out of the 23 best models are nonlinear in nature. Meanwhile, the evidence also suggests that the transition between regimes is smooth. In particular, directly comparing Bayes factors for UC-TAR and UC-LSTAR models, 11 out of the 12 best models support LSTAR-type dynamics. This result suggests that the smoothness of the regime-switching could be important in understanding whether the effects of monetary policy on output vary disproportionately with the size of the shock.

Table 2 summarizes the posterior distributions for the best nonlinear model. The policy response coefficient in regime 1 (when monetary policy shocks are small) is large and significant ($\alpha_1 = -0.831$). That is, a small monetary policy shock that increases the FFR in 25 basis points decreases output by 0.208 percentage points. Given a smooth transition between regimes, however, large monetary policy shocks are neutral: the policy response coefficient in regime 2 (when monetary policy shocks are large), $\alpha_2^G$, is not very different from zero, as suggested by the 90% Bayesian credibility intervals.

A possible explanation for these results arises within the context of menu cost models. To motivate the distinction between small and large monetary policy shocks, past studies have related their results to the implications of theoretical models with menu costs (Ball and Romer, 1990; Ball and Mankiw, 1994; Golosov and Lucas, 2007; Alvarez, Bihan, and Lippi, 2014). In such settings, only small monetary shocks have an effect on output since, in that case, keeping nominal prices fixed is associated with only a second-order cost. By contrast, monetary policy is neutral when the policy shock is large. In that situation, firms optimize by adjusting their prices, which leaves real output unchanged, since the menu cost becomes relatively small. Hence, the results suggest that, when the transition between regimes is smooth, the data are consistent with the implications of menu-cost models.

Table 2: Posterior distributions for the UC-LSTAR model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>90% interval</th>
<th>Parameter</th>
<th>Mean</th>
<th>90% interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>0.784</td>
<td>[0.489, 1.013]</td>
<td>$\sigma_{\epsilon n}$</td>
<td>-2.205</td>
<td>[-3.972, -0.980]</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>-0.096</td>
<td>[-0.405, 0.196]</td>
<td>$\lambda$</td>
<td>0.186</td>
<td>[0.128, 0.253]</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>-0.037</td>
<td>[-0.302, 0.210]</td>
<td>$\alpha_1$</td>
<td>-0.831</td>
<td>[-1.265, -0.467]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.771</td>
<td>[0.658, 0.891]</td>
<td>$\alpha_2^G$</td>
<td>0.151</td>
<td>[-0.004, 0.317]</td>
</tr>
<tr>
<td>$\sigma_\epsilon^2$</td>
<td>2.134</td>
<td>[0.615, 4.509]</td>
<td>$c$</td>
<td>0.154</td>
<td>[0.094, 0.342]</td>
</tr>
<tr>
<td>$\sigma_\nu^2$</td>
<td>3.061</td>
<td>[1.478, 4.893]</td>
<td>$\gamma$</td>
<td>5.782</td>
<td>[2.364, 11.245]</td>
</tr>
</tbody>
</table>

Posterior means and 90% credibility intervals for the UC-LSTAR model given by equations (1), (2) and (6), for $p = 3$ and $j = 1$. The threshold variable was set to contain the 70 percent middle part of the $s_t$ to avoid over-fitting. Sample period: 1954:Q3 through 2007:Q4.

The estimated threshold is $c = 0.154$, which implies that the lower regimes only become active when
the transition variable is low. The smoothness parameter, $\gamma = 5.782$, is relatively low, indicating a smooth transition between regimes. In this context, the question of whether the relationship between monetary policy and output can be better described by TAR or LSTAR dynamics becomes relevant, as it could shed light on the reasons behind the mixed empirical results regarding the response of output to monetary policy shocks of different size.

In the empirical literature, the theoretical implications of models with menu costs are only partially supported. Implicitly, they rely on an assumption of homogeneity across sectors. Nonetheless, this assumption may not necessarily be consistent with the features of the data. In particular, the mixed results could be explained by aggregation possibly influencing the nature of the regime switching, which can be related to recent developments in the exchange rate literature. Several studies have suggested that a smooth adjustment may be more appropriate in the presence of proportional transaction costs, sectoral aggregation and non-simultaneous price adjustments and, as a consequence, make use of STAR processes to study the degree of adjustment of real exchange rates to their purchase power parity (PPP) levels (Sarno, 2003; Sarno, Taylor, and Chowdhury, 2004; Bec, Ben-Salem, and Carrasco, 2010).

In light of this evidence, a gradual adjustment may be more appropriate within the relationship between monetary policy and aggregate economic activity. A discrete change in regimes may be appropriate when considering the effects of monetary policy on disaggregated goods prices, or when firms and traded goods are identical (Dumas, 1994; Teräsvirta, 1994; Sarno, 2003). At the aggregate level, however, where the economy is described by imperfect information or heterogeneous goods and sectors, the overall response of output to monetary policy is likely to reflect non-simultaneous adjustments and, consequently, to occur gradually. For example, a particular shock might be deemed small by the firms in some sectors and large by the firms in others. Because the aggregate response of output is an average of the individual responses of firms, it is, thus, likely to capture the heterogeneity of behaviors. Meanwhile, firms in some sectors may be able to react simultaneously to monetary shocks while those in other sectors might face different constraints that only allow them to adjust their prices with a lag. Hence, the overall response of output is more likely to change gradually at the aggregate level, as firms slowly adjust their prices.

By considering LSTAR-type dynamics in the relationship between monetary policy and output, the model can better capture the different responses that could occur in an environment in which goods are

---

9The results in Weise (1999), for example, show that small and large monetary shocks have different effects on output, but which effect is larger depends on the time horizon under consideration. Ravn and Sola (2004), in turn, find that small monetary shocks have disproportionately large effects on output when the instrument is the FFR, but not M1. More recently, Donayre (2014) finds that the response of output to small monetary shocks is large and significant. However, the response of output to large monetary shocks becomes neutral only after a few quarters.
not identical and sectors are, likely, heterogeneous. In this context, the evidence for a smooth transition between regimes suggests a potential role for aggregation in determining the nonlinear relationship between monetary policy and output. At the same time, it also offers a potential explanation for the mixed results regarding the implications of menu costs found in the empirical literature.

The rest of the model parameter estimates are consistent with previous findings in the literature. The rescaling factor $\lambda$ is well below one, supporting the reduction in output growth volatility since the mid 1980s. At the same time, $\sigma_\nu$ is larger than $\sigma_\epsilon$, reflecting a large permanent component of output. Meanwhile, $\sigma_{\epsilon \nu}$ is negative and statistically significant\(^{10}\) and the associated correlation coefficient, $\rho = -0.863$, implies that the permanent and transitory components of output are strongly negatively correlated, consistent with the findings in Morley et al. (2003) and Sinclair (2009).

Even though the response coefficient is larger when small monetary shocks hit the economy, it is important to evaluate these responses over time, given the nonlinear nature of the model. For this reason, generalized impulse-response functions (GIRFs) are constructed following Koop, Pesaran, and Potter (1996). The model is assumed to be known and attention is restricted to the transitory component of output, $y_t^C$. A detailed description of the computation of GIRFs is provided in appendix C.

A small, positive shock to $y_t^C$ is fixed in period 0 for regime 1 and a large monetary shock is fixed in period 0 for regime 2. These shocks are set so that they fall below or above the estimated threshold $c = 0.154$\(^{11}\). In particular, they are set to 0.08 and 0.22, so that the size of the shocks corresponds to a standard deviation difference between the ‘small’ and ‘large’ shocks, with the estimated threshold as the middle point. Figure 2 exhibits the accumulated GIRFs for the transitory component of output. Since each particular history generates a given forecast of $y_t^C$, the median of these forecasts are reported, together with the 25th and 75th quantiles (dashed lines), which serve as bands.

The left panel of figure 2 plots the accumulated response of $y_t^C$ for 20 periods ahead in regime 1 (i.e., when the monetary shock hitting the system is small). The right panel plots the accumulated response of $y_t^C$ in regime 2.

From figure 2, the accumulated response of $y_t^C$ to ‘small’ monetary shocks is much larger than that to ‘large’ shocks (−4.51 and 0.05, respectively). Moreover, the response of $y_t^C$ in the large-shock regime is neutral, given that the median response is not different from zero, as the estimated bands suggest.

\(^{10}\)Imposing $\sigma_{\epsilon \nu} = 0$ typically provides very different results in terms of the cycle-trend decomposition. Therefore, a natural question is whether the results are sensitive to $\sigma_{\epsilon \nu}$ being estimated or imposed to zero. In a frequentist environment, Donayre (2014) explores this possibility and finds that the model where $\sigma_{\epsilon \nu} \neq 0$ is preferred and that the results are robust to $\sigma_{\epsilon \nu}$ being estimated or imposed to zero. For this reason, the model where $\sigma_{\epsilon \nu} = 0$ is not considered here.

\(^{11}\)This guarantees that the ‘small’ (‘large’) shock is below (above) the estimated threshold, triggering a response of output captured by the $\alpha_1(\alpha_t^C)$ coefficient.
Figure 2: UC-LSTAR model: Generalized impulse-response functions

(a) Small-shock regime

(b) Large-shock regime

Accumulated response of $y_t^C$, within a UC-LSTAR model, to a positive shock to the monetary policy shock. The GIRFs are computed according to the details given in appendix C.

Based on this evidence, the transitory response of output exhibits an overall larger response when small monetary shocks hit the economy, even after controlling for future monetary and idiosyncratic shocks and for the history of the economic conditions. When the change of regimes occurs in a gradual fashion, the response of $y_t^C$ is much smaller and not very different from zero, in line with the implications of menu-costs models.

This finding is supported by the dynamic response of $y_t^C$ to ‘small’ and ‘large’ monetary policy shocks when the transition between regimes occurs more abruptly. Figure 3 reports the accumulated GIRFs for $y_t^C$ from a UC-TAR model, computed similarly to those reported in figure 2.

When the dynamics are described by an abrupt transition between regimes, the accumulated response of $y_t^C$ to ‘small’ and ‘large’ monetary shocks behaves asymmetrically, although the asymmetry is less clear than that in figure 2. The accumulated response of $y_t^C$ to a ‘small’ monetary policy shock is not much larger than that to a ‘large’ shock (-0.16 and -0.12, respectively). However, neither response is very different from zero, as suggested by the estimated bands. Further, the median of the response of $y_t^C$ to a ‘small’ monetary policy shock is considerably smaller than the median response of $y_t^C$ from the UC-LSTAR model in figure 2. At the same time, the small monetary shock seems to have a more permanent effect in figure 2, while the degree of uncertainty is higher for the UC-TAR case. These differences highlight how the dynamic response of $y_t^C$ behaves differently under different assumptions about the smoothness of the regime-switching. Moreover, these results support the notion that the mixed results in the literature could be driven by the fact that, at the aggregate level, the response of
output is an average of the individual response of firms.

4 Robustness

4.1 Alternative measures of monetary policy

The model from the previous section is re-estimated using two alternative measures of $x_t$. The first is obtained as the residuals from a Taylor rule which can, then, be thought of as a monetary policy shock in the sense that it represents the extent to which the FFR deviates from this Taylor rule (see Stock and Watson (2002b) for further details). The second measure of $x_t$ corresponds to the narrative approach of Romer and Romer (2004) [RR, henceforth] where the identification of the shocks does not impose linearity in the monetary policy response. Specifically, they develop a measure of policy relatively free of endogenous and anticipatory movements by first constructing a series of interest rate changes decided upon at meetings of the FOMC, and then isolating innovations to these policy rates that are orthogonal to the Fed’s internal forecasts.

In both cases, the Bayesian model comparison suggests that the model that best describes the data is a UC-LSTAR model with $p = 3$ and $j = 1$. Table 3 summarizes the posterior distributions for the model with each measure of $x_t$. The same sample period for the model with $x_t$ based on a Taylor rule is the same as in table 2. The sample period for $x_t$ based on RR goes from 1969:Q1 through 1996:Q4, given their data availability.
Table 3: Posterior distributions: alternative measures of monetary policy

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Taylor rule-based $x_t$</th>
<th>RR’s measure of $x_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>90% interval</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.891</td>
<td>[0.679, 1.110]</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>-0.141</td>
<td>[-0.441, 0.122]</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>0.019</td>
<td>[-0.198, 0.237]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.741</td>
<td>[0.551, 0.902]</td>
</tr>
<tr>
<td>$\sigma_\varepsilon$</td>
<td>3.491</td>
<td>[0.658, 6.340]</td>
</tr>
<tr>
<td>$\sigma_\nu$</td>
<td>4.220</td>
<td>[1.930, 6.487]</td>
</tr>
<tr>
<td>$\sigma_{\varepsilon\nu}$</td>
<td>-3.511</td>
<td>[-5.872, -1.003]</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.223</td>
<td>[0.154, 0.306]</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.489</td>
<td>[-0.937, -0.028]</td>
</tr>
<tr>
<td>$\alpha_{G1}$</td>
<td>0.092</td>
<td>[-0.149, 0.431]</td>
</tr>
<tr>
<td>$c$</td>
<td>0.135</td>
<td>[0.084, 0.298]</td>
</tr>
</tbody>
</table>

Posterior means and 90% credibility intervals for the UC-LSTAR model given by equations (1), (2) and (6), for $p = 3$ and $j = 1$, for $x_t$ based on a Taylor rule as in Stock and Watson (2002b), with sample period: 1954:Q3 through 2007:Q4, and for $x_t$ from Romer and Romer (2004), with sample period 1969:Q1 through 1996:Q4

The results for the model in which $x_t$ is measured based on a Taylor rule are, in general, very similar. From the left part of table 3, the policy response coefficient in the small-shock regime, $\alpha_1 = -0.489$, is significant and relatively large. The policy coefficient in the second regime, $\alpha_{G1} = 0.092$, is smaller in absolute value, and not very different from zero. Further, the estimated threshold, $c = 0.135$, and the smoothness parameter, $\gamma = 6.778$, are both very similar to the estimates when the monetary policy shock is obtained from the residuals of a VAR. The relatively low value of $\gamma$ suggests that the transition between regimes, in this case, also occurs smoothly. The rest of the parameter estimates are all in line with the results described before.

When $x_t$ is measured using the narrative approach from RR, the results are very similar with respect to autoregressive terms, the parameters of the variance-covariance matrix and those of the transition function. The results regarding the policy response coefficients are qualitatively similar, but slightly different in quantitative terms. Specifically, the policy response coefficient in the small-shock regime, $\alpha_1 = -1.402$, is larger than that in the second regime, $\alpha_{G1} = -0.959$, suggesting that smaller policy shocks still have a larger effect on output. Therefore, the evidence for nonlinearity using the RR shocks is still strong and significant. However, the response of output to large shocks is different from zero.

While this result is consistent with the findings in RR, who show that their shocks have effects that are substantially stronger than those identified using a VAR, it is worth noting that the differ-
ences between standard VAR and RR measures of $x_t$ could be driven by the fact that unanticipated policy contractions in a VAR are associated with smaller and less persistent increases in the FFR than equivalent innovations from the RR approach. Similarly, they could also be explained by the unusually dramatic movements in the FFR between 1979 and 1982 influencing the RR identification procedure (Coibion, 2011). Accounting for these factors, Coibion (2011) shows that the real effects of policy are more consistent across approaches and smaller than originally found by RR.\footnote{The results from RR imply not just greater effects of monetary policy shocks, but also a dramatically different historical interpretation of U.S. business cycles fluctuations. For example, the RR shocks imply that even the 1990-91 recession is attributed to contractionary monetary policy, a result at odds with the conventional consensus (Blanchard, 1993; Hall, 1993; Walsh, 1993).}

Overall, the results suggest that when the transition between regimes occurs smoothly, the data are consistent with the implication of menu-cost models in the sense that the response of output to small monetary policy shocks is larger. This result, together with the nonlinear dynamics, is robust across different measures of $x_t$. Nonetheless, caution should be exerted when interpreting the results in terms of the abruptness or smoothness of the regime-switching, as this may depend on the identification of the monetary shock.

### 4.2 Asymmetric thresholds

Considering $s_t = |x_t|$ implies a symmetric threshold around zero.\footnote{In the case where $s_t = |x_t|$, given the estimated threshold $c$, the response of $y_t^C$ to small monetary policy shocks that fall between $[-c, c]$ is different from its response to shocks outside that interval.} A natural question that arises is whether this assumption is too restrictive, in the sense that the asymmetry found before could be driven by model misspecification. To address this issue, the model summarized in table 2 is extended to allow for three regimes, using the observed monetary policy shock as the transition variable, $s_t = x_t$.

Specifically, equation (6) is replaced with the following three-regime LSTAR specification:

$$
y_t^C = \sum_{p=1}^{P} \phi_p y_{t-p} + \sum_{j=1}^{J} \alpha_{j} x_{t-j} + \sum_{j=1}^{J} \alpha_{j} G_1(s_t; c_1, \gamma) + \sum_{j=1}^{J} \alpha_{j} G_2(s_t; c_2, \gamma) + \epsilon_t \tag{9}
$$

where $y_t^C$, $x_t$ and $\epsilon_t$ are defined as in equation (6), $G_1(.)$ and $G_2(.)$ are logistic functions described according to (7), $\gamma$ is the smoothness parameter\footnote{To keep the model tractable, the transition functions $G_1(.)$ and $G_2(.)$ are restricted to be driven by the same smoothness parameter, $\gamma$. This does not affect the main results, however.}, $c_1$ is the location parameter associated with the first transition function, $G_1(.)$, while $c_2$ is the location parameter associated with the second transition...
function, $G_2(\cdot)$. It is assumed that $c_1 < c_2$, so that the policy coefficients from (9) change smoothly from $\alpha_1$ to $\alpha_3$, through $\alpha_2$, for increasing values of $s_t$, as $G_1(\cdot)$ changes from 0 to 1, followed by a similar change for $G_2(\cdot)$.

In this way, when $s_t = x_t$, equation (9) allows for three regimes. In the first regime, for $x_t < c_1$, relatively large, negative shocks affect the economy; in the second, when $c_1 \leq x_t \leq c_2$, relatively small shocks could generate dynamics that are different from large shocks; and the third regime, when $x_t > c_2$, is characterized by relatively large, positive monetary policy shocks. Because the two threshold parameters, $c_1$ and $c_2$, do not need to be symmetric around zero, a potential asymmetric behavior of $y_t^C$ with respect to the direction of the policy shock can also be studied. The posterior distributions for the model given by equations (1), (2) and (9) are summarized in table 4.

Table 4: Posterior distributions: three-regime LSTAR specification

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>90% interval</th>
<th>Parameter</th>
<th>Mean</th>
<th>90% interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>0.789</td>
<td>[0.530, 1.033]</td>
<td>$\lambda$</td>
<td>0.264</td>
<td>[0.184, 0.365]</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>-0.047</td>
<td>[-0.334, 0.189]</td>
<td>$\alpha_1^G_1$</td>
<td>-1.367</td>
<td>[-2.269, -0.392]</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>-0.051</td>
<td>[-0.298, 0.192]</td>
<td>$\alpha_1^G_2$</td>
<td>0.176</td>
<td>[-0.009, 0.377]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.767</td>
<td>[0.621, 0.915]</td>
<td>$\alpha_2$</td>
<td>0.195</td>
<td>[0.045, 0.407]</td>
</tr>
<tr>
<td>$\sigma_\epsilon^2$</td>
<td>2.323</td>
<td>[0.472, 5.319]</td>
<td>$c_1$</td>
<td>-0.101</td>
<td>[-0.284, -0.006]</td>
</tr>
<tr>
<td>$\sigma_\nu^2$</td>
<td>3.226</td>
<td>[1.633, 5.306]</td>
<td>$c_2$</td>
<td>0.185</td>
<td>[0.045, 0.407]</td>
</tr>
<tr>
<td>$\sigma_{\epsilon\nu}$</td>
<td>-2.421</td>
<td>[-4.677, -0.850]</td>
<td>$\gamma$</td>
<td>4.985</td>
<td>[1.749, 11.393]</td>
</tr>
</tbody>
</table>

Posterior means and 90% credibility intervals for the UC-LSTAR model given by equations (1), (2) and (9), for $p = 3$ and $j = 1$. The threshold variable was set to contain the 70 percent middle part of the $s_t$ to avoid overfitting. Sample period: 1954:Q3 through 2007:Q4.

In general, the results are not very different from those reported in table 2. The estimated thresholds, while different, retain a level of symmetry around zero. That is, $|c_1|$ and $|c_2|$ are not very different from $c = 0.154$ in table 2. This is consistent with the results in Donayre (2014), who finds, in a frequentist environment, that a UC-TAR model with asymmetric thresholds around zero is not statistically different from one with symmetric thresholds around zero.

At the same time, the estimated policy coefficients suggest an asymmetric response of $y_t^C$ to the direction of the policy shock. On the one hand, the response of cyclical output to negative, large monetary shocks, $\alpha_1$, is larger, in absolute value, than that to positive, large monetary shocks, $\alpha_1^G_2$. Nonetheless, neither of these two coefficients is very different from zero. On the other hand, $\alpha_1^G_1$, the response of $y_t^C$ to small monetary shocks, defined as $x_t \in [c_1, c_2]$, is large and different from zero, as evidenced by the Bayesian credibility set. Together, these estimates suggest that only small monetary policy shocks, when measured as the residuals from a VAR, have real effects on $y_t^C$, in line with the
menu-cost theory. While the results suggest an asymmetric behavior of cyclical output with respect to the direction of policy, evidence for such an asymmetry seems to be weaker, once size asymmetry is taken into consideration.\footnote{Size and sign asymmetries are possibly correlated. In that case, isolated asymmetric effects could be weakened when considered together with other types of asymmetries.} This finding is in line with the results obtained by Ravn and Sola (2004), who argue that only small, negative monetary shocks have real effects on economic activity.

The smoothness parameter, $\gamma = 4.985$, in this case, is smaller than the one reported in table 2, which suggests a smoother transition between regimes. The rest of the parameter estimates are in line with those reported in table 2. In this sense, the model is robust to potentially asymmetric thresholds.

5 A UC-LSTAR Model of Disaggregated Output

Given the evidence for a smooth regime-switching, a natural question refers to whether less aggregated data imply a more abrupt transition between regimes. If aggregation, in fact, suggests a smoother transition between regimes, then a natural implication of this finding is that the regime-switching should be less smooth for data that are more disaggregated. To address this issue, the UC-LSTAR model from table 2 is re-estimated using disaggregated data on output.

The data on disaggregated output are taken from the Federal Reserve Board database for the sample ranging between the first quarter of 1972 through the fourth quarter of 2007. In particular, data on manufacturing, mining and utilities are considered. Manufacturing comprises all industries included in the North American Industry Classification System (NAICS) definition of manufacturing (three-digit NAICS groups 311-316, 321-327, 331-337 and 339 and NAICS groups 1133 and 5111). Mining includes three-digit NAICS groups 211-213. Finally, utilities include electric utilities and natural gas distribution (NAICS groups 2211 and 2212).

While the data described still retain a level of aggregation, the purpose of re-estimating the model is to determine whether more disaggregated data suggests a change in the dynamics of the regime switching. If this is the case, the value of $\gamma$ should be higher, independently of the level of disaggregation.

Table 5 reports the estimated smoothing parameters $\gamma$ for the UC-LSTAR models using disaggregated data on manufacturing, mining and utilities, and for the UC-LSTAR model estimated used aggregate data on output. It is important to highlight that the prior distributions were not changed to re-estimate these models. That is, the view that the value of the smoothing parameter $\gamma$ is higher when using more disaggregated data is not imposed ex-ante. In all three cases, the estimated smooth-
Table 5: Smoothing parameters for UC-LSTAR models by sector

<table>
<thead>
<tr>
<th>Sector</th>
<th>Manufacturing</th>
<th>Mining</th>
<th>Utilities</th>
<th>Aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\gamma)</td>
<td>8.356</td>
<td>9.024</td>
<td>8.994</td>
<td>5.782</td>
</tr>
<tr>
<td>Percentile of aggregate (\gamma)</td>
<td>29</td>
<td>20</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>


The point estimates are larger in magnitude than the one estimated using aggregated data (\(\gamma = 5.782\))\(^{16}\). Therefore, in spite of the fact that the data retain some level of aggregation, the smoothness parameter increases in value for all three sectors, which suggests that the regime-switching is more abrupt when the data are more disaggregated, regardless of the level of disaggregation.

Meanwhile, table 5 also reports the percentile for 5.782 (the value of \(\gamma\) when aggregated data were used) for each posterior distribution of the disaggregated smoothness parameters. The fact that the percentiles for 5.782 for all three posterior distributions are well below the median provides further evidence that the smoothness parameters are consistently higher when more disaggregated data are used. These results support the findings from the model in table 2 and suggest that aggregation can play a key role in determining the nature of the nonlinear relationship between monetary policy and output.

Because point estimates are subject to variation, however, a more formal model comparison analysis is carried out for each sector, in the same spirit of the Bayesian model comparison at the aggregate level. Table 6 reports Bayes factors for models estimated for the mining, manufacturing and utilities sectors.

Based on the analysis of the Bayesian model comparison from table 6, UC-TAR model with \(p = 3\) and \(j = 2\) best fits the data for the mining sector. This result can be understood as providing further evidence that, with more disaggregated data, the change in regimes occurs more abruptly. It is important to mention, however, that the second and third best models are UC-LSTAR models, with \(p = j = 4\) and with \(p = j = 2\), respectively. In this sense, the results in this section are less compelling than those from the Bayesian model comparison at the aggregate level, where 11 out of the best 12 models supported LSTAR-type dynamics. However, this further suggests that sectoral aggregation plays a role in the smoothness of the change in regimes. At the aggregate level, there is strong evidence in favor of a

\(^{16}\text{Given that the data on disaggregated output were only available starting in 1972, the model with aggregated output was re-estimated for this subsample. The estimated smoothness parameter for this subsample was } \gamma = 5.972, \text{ similar to the one obtained using the whole sample.}\)
Table 6: Bayes for all models and sectors

<table>
<thead>
<tr>
<th>p</th>
<th>Mining</th>
<th>Manufacturing</th>
<th>Utilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>0.9656</td>
<td>0.9738</td>
<td>0.8148</td>
</tr>
<tr>
<td>2</td>
<td>0.9331</td>
<td>0.7939</td>
<td>0.9479</td>
</tr>
<tr>
<td>3</td>
<td>0.8188</td>
<td>0.8715</td>
<td>0.9412</td>
</tr>
<tr>
<td>4</td>
<td>0.9086</td>
<td>0.8782</td>
<td>0.9012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>j</th>
<th>UC-TAR</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7967</td>
<td>0.8033</td>
<td>0.7954</td>
</tr>
<tr>
<td>2</td>
<td>0.7758</td>
<td>0.7976</td>
<td>1.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.9387</td>
<td>0.8485</td>
<td>0.8852</td>
</tr>
<tr>
<td>4</td>
<td>0.9106</td>
<td>0.8112</td>
<td>0.9251</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>j</th>
<th>UC-LSTAR</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7999</td>
<td>0.8056</td>
<td>0.8794</td>
</tr>
<tr>
<td>2</td>
<td>0.9067</td>
<td>0.9036</td>
<td>0.7301</td>
</tr>
<tr>
<td>3</td>
<td>0.7530</td>
<td>0.8225</td>
<td>0.8048</td>
</tr>
<tr>
<td>4</td>
<td>0.8154</td>
<td>0.8048</td>
<td>0.8695</td>
</tr>
</tbody>
</table>

Bayes factors for the UC-linear, UC-TAR and UC-LSTAR models for $p = 1, \ldots, 4$ and $j = 1, \ldots, 4$ for output in the mining, manufacturing and utilities sectors.
smooth regime-switching. With less aggregated data, the evidence suggests an abrupt change, although less clearly.

For the manufacturing sector, the Bayesian model comparison suggests that a UC-TAR model with \( p = 3 \) and \( j = 4 \) best fits the data. Similarly to the case of the mining sector, the evidence favoring an abrupt regime-switching is less conclusive compared to the aggregated data case. However, the fact that 3 out of the 5 best models favor TAR-type dynamics provides further support to the notion that aggregation plays an important role in determining the dynamics of the change in regimes.

Finally, for the case of output in the utilities sector, a UC-LSTAR model with \( p = 4 \) and \( j = 3 \) is favored by the Bayes factors. Even though the results suggest that the best model implies a smooth transition between regimes, two UC-TAR models are still among the best five models. Part of the reason why the results are less conclusive could arise because the utilities sector is less likely to strongly respond to monetary policy than the other two sectors. Overall, however, the evidence from the Bayesian model comparison in this section suggests that, when the data are less aggregated, the regime-switching is more abrupt.

To better understand the nature of change in regimes, figure 4 plots the transition function (7) over the range of the threshold variable \( s_t \) for both aggregated and disaggregated data. When aggregated data on output are used to estimate the model, the change in the dynamics between regimes takes places in a smoother way. Specifically, figure 4 supports the LSTAR dynamics and the idea that aggregation plays a key role in understanding whether the effects of monetary policy on output vary with the size of the monetary shock. This can be inferred from the fact that the logistic function exhibits many points between 0 and 1. Meanwhile, the transition functions from the models with more disaggregated data exhibit a slightly more abrupt transition between regimes. Also, note that the transition function associated with the response of manufacturing output to monetary policy is above the other two transition functions. This is consistent with the fact that the manufacturing sector has typically responded more strongly to changes in monetary policy than the utilities and mining sectors.

One way to gain some additional insight about the smoothness of the transition functions is to determine the number of observations for which the weight implied by a given transition function was not far from 0.5. For example, the number of observations contained within \([0.3, 0.7]\) is 50 for the aggregate transition function, much higher than those for the mining, manufacturing and utilities sectors (28, 34 and 25, respectively).

Finally, it is important to mention that considering even more disaggregated databases typically
involves a tradeoff between the cross-sectional and time-series dimensions. That is, more disaggregated databases have less information over time. Because threshold-type nonlinear models are cumbersome to estimate, the efficiency of such estimates improves with longer data series. However, the results presented in this section suggest that the smoothness parameter, $\gamma$, tends to be higher for models with more disaggregated data.

6 Concluding Remarks

Using a Bayesian approach, this paper examines the role that smooth and abrupt transitions between regimes play in understanding whether the effects of monetary policy on output vary disproportionately with the size of the monetary shock.

Considering a gradual transition between regimes in the relationship between monetary policy and aggregate economic activity is important for two reasons. First, at a disaggregated level, firms within sectors are likely to be more homogeneous and, therefore, respond to shocks in a simultaneous fashion. At the aggregate level, however, firms from different sectors can be heterogeneous in their cost structure, their beliefs about the economy and the timing with which they respond to policy shocks. This heterogeneity implies that those firms will respond to shocks at different times and, thus, the transition between regimes is likely to occur more gradually. Second, allowing for the change in regimes
to be smooth could shed some light on why the empirical literature has found mixed results regarding implications of menu-cost models. This is because the response of output to monetary policy shocks at the aggregate level is the average of the individual responses of potentially heterogeneous sectors.

The results show that there is strong evidence of nonlinearities in the data and they are robust to the measure of policy shock. In particular, the response of output is asymmetric with respect to the magnitude of the monetary policy shock. At the same time, the Bayesian model comparison strongly suggests that the transition from one regime to the other is smooth, consistent with the notion that firms from different sectors are heterogeneous in their beliefs about the economy and the timing with which they respond to economic circumstances. This evidence is also supported by the fact that, when more disaggregated data are used, the transition between regimes becomes more abrupt. Meanwhile, the results from the estimated UC-LSTAR model suggest that the response of output in periods in which monetary shocks are small is large and significant. However, the response of output to large monetary shocks is neutral, when the monetary shock is identified as the residuals from a VAR, consistent with the implications of menu-cost models.

It is important to mention that one should be careful about the interpretation of the results. The evidence for neutral large monetary shocks is less conclusive when the RR narrative approach is used. Likewise, the results suggest that aggregation is likely to play a role in the dynamics between monetary policy and output. However, to assess the role that aggregation plays in such dynamics would require a more structural analysis. Further, this paper does not address the causes behind the asymmetric response of output to monetary policy shocks of different magnitude. Those issues are left for future research.
Appendix A

State-space representation for the UC model

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

$$y_t = \begin{bmatrix} 1 & 0_{1 \times (p-1)} & 1 \end{bmatrix}_{1 \times (p+1)} y_{tt}^{(p+1) \times 1}$$

The transition equation is given by:

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

where

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

with $\Phi_{1 \times p} = \begin{bmatrix} \phi_1 & \cdots & \phi_p \end{bmatrix}$, $I_{p-1}$ the identity matrix of order $(p - 1)$, $0_{i \times j}$ a matrix with $i$ rows and $j$ columns of zeros and $G^i$ a matrix of policy response coefficients with dynamics determined by $i$, so that

$$G^{linear} = \begin{bmatrix} \alpha_{1 \times j} \\ 0_{p \times j} \end{bmatrix}_{(p+1) \times j}$$
\[ G^{TAR} = \begin{bmatrix} \alpha_s^1 I(s_t \leq \gamma) + \alpha_l^1 I(s_t > \gamma) \\ 0_{p \times j} \end{bmatrix}_{(p+1) \times j} \]

and

\[ G^{STAR} = \begin{bmatrix} \alpha_1^1 + \alpha_l^1 G(s_t) \\ 0_{p \times j} \end{bmatrix}_{(p+1) \times j} \]

with \( G(s_t) = \{1 + \exp(-\gamma (s_t - c))\}^{-1} \).

The variance-covariance matrix of the transitory component is given by:

\[ Q = E \begin{bmatrix} \nu_t \\ 0_{(p-1) \times 1} \\ \epsilon_t \end{bmatrix} \begin{bmatrix} \nu_t & 0_{1 \times (p-1)} & \epsilon_t \end{bmatrix} = \begin{bmatrix} \sigma^2 \nu & 0 & \cdots & 0 & \sigma \epsilon \nu \\ 0 & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sigma \epsilon \nu & 0 & \cdots & 0 & \sigma^2 \epsilon \end{bmatrix}_{(p+1) \times (p+1)} \]
Appendix B

Markov-Chain Monte Carlo (MCMC) algorithm for the unobserved components models

To obtain the scale of the proposal density, this paper follows Lo and Morley (2013) and considers an alternative measure of the curvature of the posterior with respect to the location parameter $c$ and smoothness parameter $\gamma$. It involves inverting the likelihood ratio statistic for $c$, based on a $\chi^2(1)$ distribution assumption, to construct a 95% confidence interval and obtain a corresponding implied standard error. It is important to note that these issues, however, only involve the construction of the proposal distribution for the MH algorithm and do not pose a problem to the convergence of the chain. For further details, refer to Lo and Morley (2013).

An additional assumption made in the estimation of the threshold-type nonlinear models is that, for the UC-TAR (UC-LSTAR) case, $c$ (and $\gamma$ are) not correlated with the rest of the parameters in the model.\(^{17}\) Hence, for $\theta$ being the set of all parameters in the model, the proposal distribution is constructed as follows:

$$
\theta \sim mt(\mu_\theta, \kappa \Sigma_\theta, \eta)
$$

where $\mu_\theta$ is set to the posterior mode for the first drawing and to the previous draw from the random-walk chain for all other drawings; $\eta$ is the degrees of freedom parameter set exogenously to 15, following much of the applied Bayesian literature\(^{18}\); and $\kappa$ is a scaling factor for the proposal density, adjusted to attain an acceptance rate for the MH algorithm between 30% and 60%. Given the assumption described in the previous paragraph that $\hat{c}$ is ($\hat{c}$ and $\hat{\gamma}$ are) uncorrelated with the rest of parameters in the model, the proposal variance-covariance matrix $\Sigma_\theta$ for the UC-TAR model is given by:

$$
\Sigma_\theta = \begin{bmatrix}
\hat{\sigma}^2(\theta_{-c}) & 0 \\
0 & \hat{\sigma}^2_\gamma
\end{bmatrix}
$$

where $\hat{\sigma}^2(\theta_{-c})$ is the variance-covariance matrix of all parameters in the model, except for the threshold

---

\(^{17}\)Since it is infeasible to obtain numerical derivatives to evaluate the curvature of the posterior with respect to $c$ (and $\gamma$), the variance-covariance matrix of all parameters in the model cannot be jointly estimated based on the inverse Hessian matrix.

\(^{18}\)Using a multivariate Student $t$ distribution as the proposal allows for a wider set of possible draws, given fatter tails. To guarantee that the tails are, indeed fat, the degrees of freedom is set to 15.
parameter \( \hat{c} \), based on the estimated inverse Hessian matrix evaluated at the posterior mode. \( \hat{\sigma}^2_{\hat{c}} \) is the variance of the estimated location parameter \( \hat{c} \) based on the Lo and Morley (2013) approach discussed above. The proposal variance-covariance matrix \( \Sigma_\theta \) for the UC-LSTAR model is given by:

\[
\Sigma_\theta = \begin{bmatrix}
\hat{\sigma}^2(\hat{\theta}_{-\hat{c},-\hat{\gamma}}) & 0 \\
0 & \hat{\sigma}^2_{\hat{c},\hat{\gamma}}
\end{bmatrix}
\]

where \( \hat{\sigma}^2(\hat{\theta}_{-\hat{c},-\hat{\gamma}}) \) is the variance-covariance matrix of all parameters in the model, except for the location and smoothing parameters (\( \hat{c} \) and \( \hat{\gamma} \), respectively), and \( \hat{\sigma}^2_{\hat{c},\hat{\gamma}} = (\hat{\sigma}^2_{\hat{c}}, \hat{\sigma}^2_{\hat{\gamma}})' I \) is the variance-covariance matrix of (\( \hat{c}, \hat{\gamma} \)).

In terms of the priors, a number of considerations are taken into account. First, it is known that the degree of parameterization of the model influences the quality of inference in finite samples. Second, given that Bayes factors are functions of the prior normalizing constants, the prior settings can have a strong influence on the posterior model weights (Strachan and van Dijk, 2004). Generally, less informative priors tend to penalize more highly parameterized models. Third, the understanding of the behavior of economic variables is sometimes limited. Thus, there exists a conflict between the desire to specify uninformative priors and priors that are informative and would improve the efficiency of the estimation. Fourth, all models considered (i.e., UC-linear, UC-TAR and UC-LSTAR) are assumed to be equally likely a priori. Similarly, the use of larger, more parameterized models is neither avoided nor preferred. Taking into account these considerations, table 7 provides a summary for the prior distributions.

### Table 7: Prior distributions and parameters for all models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Support</th>
<th>Density</th>
<th>Mean</th>
<th>Variance</th>
<th>Parameter</th>
<th>Support</th>
<th>Density</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi_p )</td>
<td>( \mathbb{R} )</td>
<td>Normal*</td>
<td>0</td>
<td>0.65²</td>
<td>( \sigma_\epsilon )</td>
<td>( \mathbb{R}^+ )</td>
<td>Wishart</td>
<td>1.2</td>
<td>1.68</td>
</tr>
<tr>
<td>( \alpha_{ij} )</td>
<td>( \mathbb{R} )</td>
<td>Normal</td>
<td>-2</td>
<td>0.45²</td>
<td>( \sigma_{\nu} )</td>
<td>( \mathbb{R}^+ )</td>
<td>Wishart</td>
<td>3</td>
<td>4.20</td>
</tr>
<tr>
<td>( \alpha_{ij} )</td>
<td>( \mathbb{R} )</td>
<td>Normal</td>
<td>-2</td>
<td>0.45²</td>
<td>( \sigma_{\nu} )</td>
<td>( \mathbb{R}^+ )</td>
<td>Wishart</td>
<td>-1.2</td>
<td>1.68</td>
</tr>
<tr>
<td>( \alpha_{ij} )</td>
<td>( \mathbb{R} )</td>
<td>Normal</td>
<td>-2</td>
<td>0.45²</td>
<td>( \lambda )</td>
<td>( \mathbb{R}^+ )</td>
<td>Gamma</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>( \alpha_{ij} )</td>
<td>( \mathbb{R} )</td>
<td>Normal</td>
<td>-2</td>
<td>0.45²</td>
<td>( c )</td>
<td>( \mathbb{R}^{++} )</td>
<td>Gamma</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>( \mu )</td>
<td>( \mathbb{R} )</td>
<td>Normal</td>
<td>0</td>
<td>0.65³</td>
<td>( \gamma )</td>
<td>( \mathbb{R}^{++} )</td>
<td>Gamma*</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

*denotes a truncated distribution. The normal distribution of the autoregressive parameters \( \phi_p \) are truncated to guarantee stationarity. Similarly, the gamma distribution of the smoothness parameter \( \gamma \) excludes 0 from the support to guarantee that the model is identified, as explained below.

For all models considered, the priors for the autoregressive coefficients and for the drift in the permanent component follow a normal distribution. Each of the prior means is set to zero and each of the prior variances is set to 0.65³. These are relatively uninformative priors, considering the truncation.
of the normal distribution to ensure stationarity (i.e., the roots of the polynomial \( \phi(L) \) lie outside the unit circle).

Similarly, the policy coefficients are each assumed to follow a normal distribution with mean -2 and variance 0.45^2. The negative mean responds to economic theory: an increase in the FFR results in a reduction of output. It is important to note that all policy coefficients, independently of the regime prevailing, are assumed to have the same mean and variance. That is, the view that the effects of monetary shocks on output vary disproportionately with the size of the shock is not imposed a priori.

The variance-covariance matrix, \( \Omega \), is assumed to follow a Wishart distribution, since it is a nonnegative-definite matrix. For any \( n \times p \) matrix \( X \), with each row independently drawn from the \( p \)-variate normal distribution \( X_{(i)} = (x_1^i, \ldots, x_p^i) \sim N_p(0, R) \), \( X^T X \sim \text{Wishart}(R, \zeta) \), where \( R \) is a \( p \times p \) positive-definite matrix and \( \zeta \) is the number of degrees of freedom. Taking into account that the mean of \( X^T X \) satisfies \( \zeta R \) and the variances of its elements satisfy \( \zeta (r_{ij}^2 + r_{ii} r_{jj}) \) for \( i = 1, \ldots, p \) and \( j = 1, \ldots, p \), the distribution parameters are set to \( R = \begin{pmatrix} 0.4 & -0.4 \\ -0.4 & 1 \end{pmatrix} \) and \( \zeta = 3 \).

In the case of \( \lambda \), it is assumed to follow a gamma distribution with parameters \( a \) and \( b \). For any random variable \( X \sim \text{gamma}(a, b) \), the mean and variance satisfy the following equations:

\[
    a = \frac{[E(X)]^2}{Var(X)} \quad \text{and} \quad b = \frac{E(X)}{Var(X)} \quad (A-1)
\]

The prior mean for \( \lambda \) is set to 0.4 and the prior variance is set to 0.3. They, in turn, imply that the parameters for the gamma distribution are \( a = 0.533 \) and \( b = 1.333 \), according to equation (A-1). The supporting evidence for the Great Moderation (McConnel and Perez-Quiros, 2000; Blanchard and Simon, 2001; Stock and Watson, 2002a, 2005) is reflected on the mean for \( \lambda \), which is below 1.19

The threshold and smoothing parameters are also assumed to follow a gamma distribution. The prior mean for \( \gamma \) is set to 8, in line with the estimates in frequentist studies (Weise, 1999; van Dijk, Teräsvirta, and Franses, 2002; Rothman, van Dijk, and Franses, 2001; Öcal and Osborn, 2000; Sensier, Osborn, and Öcal, 2002). The prior variance for \( \gamma \) is set to 6. Given the equations in (A-1), this implies that the parameters for the gamma distribution for \( \gamma \) are \( a = 10.667 \) and \( b = 1.333 \).

For the prior of \( c \), a gamma distribution is assumed because the transition variable to be considered is the absolute value of the monetary policy shock. The prior mean is set to 0.5 and the prior variance is set to 0.2. For both, \( \gamma \) and \( c \), the priors are relatively uninformative.

To guarantee that the priors do not drive the results of the model, different prior specifications were

\footnote{It is important to note that the only restriction imposed on the parameter rescaling \( \Omega \) is that \( \lambda > 0 \), as implied by the assumed gamma distribution. \( \lambda \leq 1 \) is not imposed.}
considered. The results of the model, however, are robust across different parameterizations of the prior distributions.

Once priors are specified, the Bayesian estimation can be carried on in a standard way. For illustration purposes, the MCMC algorithm for the unobserved components model is explained here for the case of the UC-LSTAR model with parameters $\theta$, threshold $c$ and smoothness parameter $\gamma$. The cases of the UC-linear and UC-TAR models, and extension to a multiple block case, are straightforward.

The simulation from the likelihood function is carried out by means of the Metropolis-Hastings (MH) algorithm with a random-walk (RW) chain proposal. To generate the model parameters $\theta$ of the distribution from which the draws are made, the likelihood function is maximized conditional on $c$ and $\gamma$, at each iteration. This optimization procedure provides the mean and variance-covariance matrix for the random-walk chain proposal density, based on a multivariate Student-t distribution.

In this way, the MH algorithm consists of the following steps:

Step 1: Pick initial values for $\theta^{(0)}$, $c^{(0)}$ and $\gamma^{(0)}$.

Step 2: Iterate the MH algorithm for $j = 1, \ldots, M$. In the $(j + 1)^{th}$ iteration,

1. $\theta^{(j+1)}$ is generated according to:
   
   (a) Propose the draw:
   
   $$\theta' \sim q(\theta^{(j)}, \theta' | c^{(j)}, \gamma^{(j)}) = t_{c(j), \gamma(j)}(\hat{\theta}, \hat{\Sigma}, \nu)$$
   
   where $\hat{\theta} = \text{argmax} \ LogL(\theta | c^{(j)}, \gamma^{(j)})$ and $\hat{\Sigma} = -\left[ \frac{\partial^2 \ln L(\theta | c^{(j)}, \gamma^{(j)})}{\partial \theta \partial \theta'} \right]^{-1}$.

   (b) Calculate the transition probability:

   $$\alpha(\theta^{(j)}, \theta' | c^{(j)}, \gamma^{(j)}) = \min \left\{ \frac{\pi(\theta', c^{(j)}, \gamma^{(j)}) q(\theta', \theta^{(j)} | c^{(j)}, \gamma^{(j)})}{\pi(\theta^{(j)}, c^{(j)}, \gamma^{(j)}) q(\theta^{(j)}, \theta' | c^{(j)}, \gamma^{(j)})}, 1 \right\}$$

   (c) Drawing $u \sim U(0, 1)$, the acceptance or rejection rule for the new draw $\theta^{(j+1)}$ is given by:

   $$\theta^{(j+1)} = \begin{cases} 
   \theta', & \text{if } u \leq \alpha(\theta^{(j)}, \theta' | c^{(j)}, \gamma^{(j)}) \\
   \theta^{(j)}, & \text{otherwise}
   \end{cases}$$

2. $c^{(j+1)}$ is generated according to:

   $$c^{(j+1)} \sim \pi(c | \theta^{(j+1)}, \gamma^{(j)}, y)$$
where
\[
\pi(c|\theta^{(j+1)}, \gamma^{(j)}, y) = \frac{\pi(c) \int_y \pi(y|\theta^{(j+1)}, c, \gamma^{(j)}) f(y|\theta^{(j+1)}, c, \gamma^{(j)})}{\pi(y|\theta^{(j+1)}, \gamma^{(j)})} \\
= \frac{\pi(c) \int_y \pi(y|\theta^{(j+1)}, c, \gamma^{(j)})}{\int_y f(y|\theta^{(j+1)}, c, \gamma^{(j)})} \\
= \frac{\int_y \pi(y|\theta^{(j+1)}, c, \gamma^{(j)})}{f(y|\theta^{(j+1)}, c, \gamma^{(j)})}
\]

3. Similarly, \(\gamma^{(j+1)}\) is generated according to:

\[
\gamma^{(j+1)} \sim \pi(\gamma|\theta^{(j+1)}, c^{(j+1)}, y)
\]

where
\[
\pi(\gamma|\theta^{(j+1)}, c^{(j+1)}, y) = \frac{\pi(\gamma) \int_\gamma \pi(\gamma|\theta^{(j+1)}, c^{(j+1)}, \gamma) f(\gamma|\theta^{(j+1)}, c^{(j+1)}, \gamma)}{\pi(\gamma|\theta^{(j+1)}, c^{(j+1)})} \\
= \frac{\pi(\gamma) \int_\gamma \pi(\gamma|\theta^{(j+1)}, c^{(j+1)}, \gamma)}{\int_\gamma \pi(\gamma|\theta^{(j+1)}, c^{(j+1)}, \gamma)} \\
= \frac{\int_\gamma \pi(\gamma|\theta^{(j+1)}, c^{(j+1)}, \gamma)}{f(\gamma|\theta^{(j+1)}, c^{(j+1)}, \gamma)}
\]
Appendix C

Computation of generalized impulse-response functions

A generalized impulse-response function (GIRF) is defined as the effect of a one-time shock on the forecast of variables in a model, given a specific history. The response constructed is then compared to a benchmark “no shock” scenario. In this way, the GIRF can be expressed as follows:

\[ GI_Y(q, \nu_t, \omega_{t-1}) = E[Y_{t+q}, \nu_t, \omega_{t-1}] - E[Y_{t+q}/\omega_{t-1}] \]

where \( GI_Y \) is the generalized impulse-response function of a variable \( Y \) for period \( q \), given the specific history \( \omega_{t-1} \) and initial shock \( \nu_t \), and \( E[.] \) is the expectations operator. See Koop et al. (1996) for details.

To compute the GIRF, the conditional expectations are simulated for each draw from the posterior distribution. The nonlinear model is assumed to be known. The shock to \( Y \), \( \nu_0 \), occurs in period 0, and responses are computed for \( q \) periods ahead. Thus, the \( GI_Y \) function is generated according to the following steps:

Step 1: Pick a history \( \omega_{t-1} \). The history is the actual value of the lagged endogenous variables for a particular episode (e.g., the values associated with regime 1 or regime 2).

Step 2: Pick a sequence of two-dimensional monetary and idiosyncratic shocks \( \nu_{j,t+q} \), \( q = 0, 1, \ldots, n \). They are drawn with replacement from the vector of monetary shocks and from the estimated residuals of \( y_t^C \).

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 (6). 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 \( \nu_0 \) (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 \( Y_{t+q}(\nu_0, \omega_{i,t-1}, \nu_{j,t+q}) - Y_{t+q}(\omega_{i,t-1}, \nu_{j,t+q}) \).
References


