Abstract

In this paper, we construct confidence intervals for threshold parameters of TAR processes. While there have been important advances in the asymptotic theory for inference in TAR models, their small sample properties have received much less attention. In this context, following the work of Eo and Morley (2011), we propose a new method to construct confidence intervals for threshold parameters based on simulation of their marginal fiducial distributions (FD) from the likelihood function, by means of Markov-Chain Monte Carlo methods. This approach provides confidence intervals with accurate coverage and short expected length, even when the sample size is small. In particular, based on small-sample Monte Carlo experiments, we find that the FD confidence intervals for the threshold parameter have more accurate coverage rates and shorter expected length in comparison to bootstrap and asymptotic methods. We apply our method to a TAR model of the U.S. real GDP growth rate where the dynamics depend on whether credit conditions are tight or loose. Using our marginal fiducial approach, we find much tighter 95% confidence intervals for the threshold parameters that support the idea that recessions produced by credit crises exhibit largely transitory effects on the level of output.

JEL Classification Code: C15, C22
Keywords: Threshold Autoregressions, Fiducial Inference, Bootstrap Methods, Markov-Switching Monte Carlo Methods.
1 Introduction


While there have been important developments in the asymptotic theory for inference in TAR models (Chan, 1993; Hansen, 1997, 2000; Chan and Tsay, 1998), only a few studies have focused on their small sample properties. Cheng (2008), for example, develops confidence intervals for nonlinear regressions of the smooth transition type, based on standard and local asymptotics, that have good finite-sample coverage (although the nonlinear regressions she considers do not include autoregressive terms). Meanwhile, EFS, by means of Monte Carlo experiments, study the small sample coverage properties of confidence intervals for the threshold parameter constructed using asymptotic results and bootstrap methods. They find that none of the procedures performed satisfactorily across the full range of parameter values.

In this paper, following the approach developed in Eo and Morley (2011) for structural breaks, we propose a likelihood-based procedure to construct confidence intervals for the threshold parameter that provides high accuracy and short expected length, even in small samples. This approach involves the simulation of the marginal ‘fiducial’ distribution of the threshold parameter by integrating out other parameters from the likelihood function. To carry out this approach, we make use of Markov-chain Monte Carlo (MCMC) methods, widely applied in Bayesian analysis, even though we do so with a frequentist notion of inference in mind. The confidence intervals for the threshold parameter are inherently the same as Bayesian highest-posterior-density (HPD) intervals, given non-informative priors. However, we adopt a frequentist perspective in the sense that the inferences we make consider their coverage accuracy and expected length across repeated experiments.

The method proposed relies on fiducial inference, which was first suggested by Fisher (1930). It involves a frequentist viewpoint of probabilistic inferences about the parameters of a particular model, where the probabilities are proportional to the likelihood function (see Barnett (1999) for further discussion about fiducial inference). Despite the apparent contradiction between the probabilistic inferences about parameters and the standard frequentist approach, it is shown by Fisher (1930) that, in the case...
of a single location parameter for which the test statistic is pivotal\(^1\), fiducial and frequentist confidence intervals are the same. Based on this equivalence, we use fiducial distributions as intermediaries in our ultimate goal of constructing frequentist estimators of confidence intervals.\(^2\)

Our approach is related to some of the existing methods to construct confidence intervals for threshold parameters. It is analogous, in a large sample setting, to the asymptotic inverted likelihood ratio (LR) approach proposed by Hansen (1997). Both methods rely on the shape of the likelihood function and, thus, provide similar results when the sample is large enough. More generally, it has a similar motivation to bootstrap methods for constructing confidence intervals, as considered in EFS. Concretely, in small samples, bootstrap methods often perform better than those based on asymptotic distributions in settings when the bootstrapped distribution approximates finite-sample distributions that are close to being pivotal. In a broader way, it is also related to the literature in which nuisance parameters are integrated out of the likelihood, as in Andrews and Ploberger (1994), and to the extensive literature using MCMC methods for inference with TAR models (see, for example, Geweke and Terui (1993), Koop and Potter (1999, 2004), among others).

To determine the performance of the fiducial distribution (FD) approach, we compare it to a variety of asymptotic and bootstrap methods. With respect to asymptotic methods, Hansen (1997) derives the limiting distribution of the LR statistic in closed form and, thus, is able to provide the critical values necessary to construct confidence intervals by inverting the LR statistic. He applies his method to the U.S. unemployment rate, for a sample length of 463 months, and finds a tight 95% asymptotic confidence interval for the threshold. Nonetheless, this method will not necessarily provide confidence intervals with good coverage in settings where the sample is relatively small. Regarding bootstrap methods, we consider constructing confidence intervals by means of bootstrap percentile and bootstrap inverted likelihood ratio approaches, as in EFS. When they apply these methods to TAR models for the U.S. real GDP growth rate, they find that the confidence intervals are so wide that all candidate thresholds are included in these intervals. That is, all the methods they considered produce confidence intervals that are not particularly informative.

In order to make inferences about the threshold parameter and compare the various methods of constructing confidence intervals, we conduct a small-sample Monte Carlo experiment. This analysis suggests that the FD approach performs best overall in terms of producing relative short confidence intervals with accurate coverage rates compared to nominal confidence levels, even in small samples. In

\(^1\)A test statistic is said to be ‘pivotal’ if its distribution remains unchanged independently of the model parameters.

\(^2\)In a similar spirit, Hannig (2006) argues for the use of fiducial distributions as a tool for deriving frequentist inference procedures.
general terms, the asymptotic inverted likelihood ratio produces confidence intervals that are too narrow and undercover the true threshold parameter, while the bootstrapped inverted likelihood ratio produces confidence intervals that overcover the true threshold parameter are prohibitively wide. The bootstrap percentile approach generates relatively longer confidence intervals for the threshold parameter, and the coverage rate is never better than the one provided by our approach.

We apply our method to a TAR model for the U.S. real GDP growth rate when its dynamics depend on whether credit-market conditions, measured by a willingness-to-lend index, are tight or loose. We use the FD approach that we propose here, as well as the asymptotic and bootstrap methods, to construct 95% confidence intervals for the estimated threshold. The estimates suggest that the real GDP growth rate exhibits negative serial correlation when the willingness-to-lend index is below -5.6. Both asymptotic and bootstrap methods generate confidence intervals that are considerably wide around the estimated threshold. On the other hand, the FD confidence interval that we propose, [-5.69, 3.67], is the shortest and includes 0 (i.e., no change in willingness to lend) as a potential threshold. The estimates and confidence intervals support the idea that the U.S. real GDP growth rate exhibits negative serial correlation when banks are less willing to lend and positive serial correlation when banks are more willing to lend. In the context of the U.S. recession that began in December 2007, this result has important economic implications and argues against a literature (see, for instance, Reinhart and Rogoff (2008a, 2008b, 2008c) and the International Monetary Fund (IMF)'s World Economic Outlook for Spring 2009) that suggests that financial crises have large negative and permanent effects, as will be explained in the fourth section.

The remainder of the paper is organized as follows. Section 2 details the fiducial distribution approach that we propose to construct confidence intervals for the threshold parameter. Section 3 presents the Monte Carlo analysis the various methods of constructing confidence intervals for the threshold parameter, including our FD approach. In section 4, we estimate confidence intervals, using the FD method as well as asymptotic and bootstrap methods, for U.S. real GDP growth when the threshold is driven by the change in credit conditions. Finally, section 5 concludes.

2 The fiducial distribution approach

In this section, we describe the fiducial distribution (FD) approach we propose to constructing confidence intervals for threshold parameters. In providing the details for this method, we follow Eo and Morley (2011), who apply the FD approach to construct confidence intervals for the timing of structural breaks.
Meanwhile, developing the FD approach for the case of a TAR model departs from the structural break case in two ways: (i) a TAR process is inherently nonlinear, unlike the model addressed in Eo and Morley (2011); and (ii) the distribution of the threshold parameter, unlike a the case of a break date, is continuous.\(^3\)

In principle, our approach can be applied to any general TAR specification. For the sake of simplicity, we focus on a two-regime, TAR model, given by:

\[
y_t = (\phi_1 \gamma + \sum_{j=1}^{P} \phi_1^j y_{t-j})I(q_{t-1} \leq \gamma) + (\phi_0^\gamma + \sum_{j=1}^{P} \phi_2^j y_{t-j})I(q_{t-1} > \gamma) + \epsilon_t \tag{2.1}
\]

where \(I(.)\) is the indicator function defined according to:

\[
I = \begin{cases} 
1, & \text{if } q_{t-1} \leq \gamma \\
0, & \text{if } q_{t-1} > \gamma
\end{cases}
\]

The threshold variable \(q_{t-1} = q(y_{t-1}, \ldots, y_{t-p})\) is a known function of the data; \(\gamma\) is a threshold parameter determining the dynamics for each regime; \(\phi_1, \phi_2\) for \(i = 1, \ldots, p\) are the autoregressive coefficients prevailing in regimes 1 and 2, respectively; and \(\epsilon_t\) is assumed to be i.i.d. \((0, \sigma^2)\). Furthermore, it is assumed that \(\sigma^2 < \infty\) and that \(\phi_1, \phi_2\) satisfy regular stationarity conditions. When the model is implemented throughout the paper, the potential threshold values are restricted to the middle 70% of the threshold variable \(q_{t-1}\) to avoid an end of sample distortion.\(^4\)

Hansen (1997) shows that a grid search over potential thresholds yields a consistent estimator of \(\gamma\). In particular, the least squares estimate of \(\gamma\) is given by:

\[
\hat{\gamma} = \arg\min_{\gamma \in \Gamma} \sigma^2_n(\gamma)
\]

where \(\hat{\sigma}^2_n(\gamma)\) is the residual variance of the model conditional on a particular value of \(\gamma\), and \(\Gamma = [\underline{\gamma}, \bar{\gamma}]\) is defined to contain the middle 70% of the threshold variable, as explained above. For further details

\(^3\)A continuous variable requires a more precise sampler, a kernel estimator, and a bandwidth for it.

\(^4\)It is standard practice for TAR models to exclude the 15% of each end of the vector of ordered possible thresholds to avoid distortions in inference. If possible thresholds are too close to the beginning or the end of the ordered data, there will be insufficient observations to strongly identify the subsample parameters.
about the estimation of TAR models, refer to Hansen (1997).

In order to derive the marginal fiducial distribution for the threshold parameter, consider the probability density function (pdf) \( f(y/\theta, \psi) \) for model (2.1) with threshold and variance \( \theta = (\gamma, \psi) \in \Theta \) and coefficients \( \psi = (\phi_1, \phi_2) \in \Psi \), when evaluated at the observed data \( Y = y \). The likelihood function can then be defined as \( L(\theta, \psi/y) = f(y/\theta, \psi) \). Hence, the joint density for \( \theta \in \Theta \) and \( \psi \in \Psi \) can be constructed as the product of the likelihood function and the inverse of the integrated likelihood with respect to the model parameters:

\[
\pi(\theta, \psi/y) = L(\theta, \psi/y) \times \left[ \int_{\Theta} \int_{\Psi} L(\theta, \psi/y) d\psi d\theta \right]^{-1} \tag{2.2}
\]

Because fiducial inference involves the interpretation of these probabilities as being coherent from a frequentist viewpoint, it generates controversy.\(^5\) Despite this issue, we simply make use of fiducial distributions as an intermediate step towards our ultimate goal of making more traditional frequentist inferences. We do not seek to address the debate concerning the general coherence of fiducial inference in this paper.

Fisher (1930) showed that fiducial and frequentist confidence intervals can be the same in the case of a single location parameter for which the test statistic is pivotal. In the case of multidimensional parameter spaces, however, fiducial and frequentist confidence intervals can be at odds with each other. Thus, instead of considering the joint fiducial distribution from equation (2.2), we use the marginal fiducial distribution of the threshold parameter to construct confidence intervals:

\[
\pi(\gamma/y) = \int_{\Omega_{-1}} \pi(\theta, \psi/y) d\omega_{-1} \tag{2.3}
\]

where \( \omega_{-1} \in \Omega_{-1} \) includes all parameters from \( \theta \) and \( \psi \) with the exception of the threshold parameter, \( \gamma \). Thus, using the marginal fiducial distribution for \( \gamma \) from (2.3), the confidence intervals constructed using our proposed method will be similar to the frequentist ones, if we assume there exists a test statistic whose distribution is proportional to (2.3). The method we propose, however, does not require us to have such test statistic to construct the confidence intervals, unlike the frequentist approach.

\(^5\)Unlike the Bayesian approach, where parameters are thought of as random variables, assigning a distribution to them is at odds with the frequentist approach.
Theoretically, the marginal fiducial distribution in (2.3) is all we need to construct confidence intervals for $\gamma$. From a practical point of view, nonetheless, obtaining such a distribution is not generally easy. In particular, it is infeasible to use analytical methods to integrate the likelihood function to obtain the joint fiducial distribution in (2.2), and then integrate out model parameters to obtain (2.3). However, this marginal fiducial distribution (2.3) can be easily simulated using Markov-chain Monte Carlo (MCMC) methods.\(^6\)

As before, consider two blocks of parameters $\theta = (\gamma, \sigma)$ and $\psi = (\phi^1, \phi^2)$. The parameters in one block can be sampled conditional on the data and the parameters in the other block. Hence, given the data $y$ and parameters $\theta$, then:

$$
\psi \sim \pi(\psi, \theta/y)
$$

$$
\propto \pi(\psi)f(y/\psi, \theta)
$$

$$
\propto f(y/\psi, \theta)
$$

Similarly, for the threshold parameter, $\gamma \sim \int \pi(\theta/\psi, y)$, where:

$$
\pi(\theta/\psi, y) \sim \frac{\pi(\theta f(y/\psi, \theta))}{\pi(y/\psi)}
$$

$$
\propto \frac{\pi(\theta f(y/\psi, \theta))}{\int_{\Theta} \pi(\theta) f(y/\psi, \theta) d\theta}
$$

$$
\propto \frac{f(y/\psi, \theta)}{\int_{\Theta} f(y/\psi, \theta) d\theta}
$$

The MCMC approach simulates parameter values for $\theta$ and $\psi$ from their conditional distribution until the draws behave as if drawn from their joint and marginal distributions. For more details of the two-block, Metropolis-Hastings algorithm used in this paper, refer to the appendix.

Once the marginal fiducial distribution (2.3) for the threshold parameter is obtained, we can construct a confidence interval at the $1 - \alpha$ level in different ways. To the extent that the threshold variable is continuous, we would need to use a kernel density estimator if we wanted to apply the Bayesian highest-posterior-density (HPD) concept to the construction of the confidence interval, as in Eo and Morley (2011). The need to rely on kernel density estimators poses complications related to

---

\(^6\)For further details about the Metropolis-Hastings (MH) algorithm, the one we use in this paper, and MCMC methods, see Chib and Greenberg (1995).
the appropriate selection of the kernel function and the even more problematic choice of the bandwidth parameter. For these reasons, we opt for a different approach to the construction of our confidence intervals.

From a practical point of view, we are interested in constructing confidence intervals with the shortest expected length. Therefore, the confidence interval that we choose to construct can be defined as:

\[
[\gamma^*_L, \gamma^*_U] \tag{2.4}
\]

where \(\gamma^*_L, \gamma^*_U\) satisfy:

\[
\{\gamma^*_L, \gamma^*_U\} = \arg\min_{\gamma_L, \gamma_U \in \Xi} \{\gamma_U - \gamma_L\}
\]

and \(\Xi\) is the set of all thresholds \([\gamma_L, \gamma_U]\) such that \(\Pi(\gamma_U/y) - \Pi(\gamma_L/y) = 1 - \alpha\), for \(\Pi(\gamma/y)\) the cumulative posterior distribution function of \(\gamma\) for given \(y\).

It is important to note that if \(\pi(\gamma/y)\) is unimodal, then (2.4) is also a Bayesian highest-posterior-density (HPD) interval. Moreover, in that case, (2.4) will also exhibit the shortest possible expected length. If \(\pi(\gamma/y)\) is not unimodal, then (2.4) might exhibit longer intervals than those constructed under the HPD criterion, since it could include threshold parameters with relatively low simulated frequencies (see Eo and Morley (2011) for a detailed discussion about this issue). Despite this caveat, confidence intervals constructed using (2.4) have shorter expected length than other bootstrap and asymptotic methods, especially in small sample settings, as we will show in the Monte Carlo experiments.

Remarkably, the MCMC approach to construct confidence intervals is not limited to models of the type described in (2.1). As long as the likelihood function for the model can be specified, it is possible to construct FD confidence intervals. This feature is particularly useful in settings where models are complicated, such as multiple-threshold models, because deriving analytical distributions for such cases is very complicated, or infeasible. Likewise, even in the case of bootstrapping methods, the computational burden becomes very large, while it remains manageable for the FD approach.
3 Monte Carlo Analysis

In this section, we conduct Monte Carlo experiments to address two issues: the overall performance of our FD approach and how it relates to other methods to construct confidence intervals for threshold parameters. In particular, we compare our approach to the asymptotic inverted likelihood ratio (AIL), the bootstrapped inverted likelihood ratio (BSIL), and the bootstrap percentile (BP) approaches.\(^7\)

To compare all methods, we examine the coverage accuracy and the expected length of the confidence intervals. The coverage rate is computed as the proportion of instances in which the true threshold parameter falls into the constructed intervals across 1,000 Monte Carlo simulations. Its accuracy is determined by comparing it to specific nominal confidence levels. In our experiment, we construct 95% confidence intervals for each method. The expected length of the intervals is computed as the average length (the difference between the maximum and the minimum threshold values in the confidence interval) across Monte Carlo simulations. We calculate confidence intervals for different sample sizes of the Data Generator Process (DGP). In particular, we fix the sample size to \(T = 75, 100, 250, 500.\(^8\)

In the case of bootstrap methods, we use \(B = 199\) bootstrap samples to construct the confidence intervals. Note that to construct the bootstrap samples, we used the estimated coefficients and threshold parameters instead of the parameters from the DGP. That is, there are 199 bootstrap samples for each of the 1,000 generated \(y_t\) series. For the asymptotic inverted LR, we follow Hansen (1997) and use the asymptotic distribution that he derives, as well as the critical values that he provides. For the case of the FD approach, we employ a two-block MH algorithm with a Multivariate Student-t proposal density. We simulate the marginal fiducial distribution of the threshold parameter using 20,000 draws from the joint fiducial distribution, after discarding the first 4,000 burn-in draws.

In the literature, depending on which series is modeled as a TAR process, two main threshold variables are considered: the change in the lagged dependent variable, \(y_{t-1} - y_{t-d}\), or a lag level of it, \(y_{t-d}\), where \(d\) is the delay lag. Therefore, we carry on two different Monte Carlo experiments for two different threshold variables and associated DGP. The results of the Monte Carlo experiments are presented in the next subsections.

\(^7\)The asymptotic and bootstrap approaches we use in this experiment are standard. Refer to Hansen (1997), for further details regarding the asymptotic inverted LR method, and to EFS, for details regarding the bootstrap approaches.

\(^8\)Each series was generated for \(T + 200\) observations and the first 200 observations were used as a burn-in sample to avoid any initial effects distortion.
3.1 Monte Carlo results for the model with $y_{t-1} - y_{t-d}$ as the threshold variable

When the threshold is the change in the lagged dependent variable, we set $d = 2$ and generate data according to the following two-regime TAR(1) process:

$$y_t = (0.5 + 0.6y_{t-1})I(y_{t-1} - y_{t-2} \leq 0) + (0.9 + 0.8y_{t-1})I(y_{t-1} - y_{t-2} > 0) + \epsilon_t \quad (3.1)$$

where $\epsilon_t \sim N(0, 1)$ and $I(.)$ is the indicator function. Hence, if $y_{t-1} - y_{t-2} \leq 0$, the intercept and autoregressive coefficient are 0.5 and 0.6, respectively. If $y_{t-1} - y_{t-2} > 0$, they are 0.9 and 0.8, respectively. When the data is generated according to the DGP in (3.1), the results of the Monte Carlo experiment are summarized in figure (1).

In figure 1, the coverage rates for different sample sizes correspond to the lines at the top (measured in the right axis) while the expected lengths correspond to the bars at the bottom (measured in the left axis). The results of the Monte Carlo experiment show that the confidence intervals constructed
using the BSIL approach perform poorly: they massively overcover the true threshold parameter. In particular, the coverage rate is 100% (or very close to it) for all cases. This can be explained by the excessively long confidence intervals that this method provides. The expected length of the BSIL confidence intervals is much greater than the expected length of any of the other approaches for all sample sizes.

All other three approaches (BP, AIL and FD) exhibit coverage rates whose accuracy increases with the sample size. However, when considering both criteria together, the FD we propose approach performs the best. While the confidence intervals constructed using the AIL approach exhibit the shortest expected length, their accuracy in terms of coverage rates is not satisfactory. Specifically, they undercover the true threshold parameter in all cases, except in relatively large samples (when $T = 500$). Likewise, the confidence intervals constructed using the BP approach undercover the true threshold parameter in all cases, except for when $T = 500$. The FD approach, on the other hand, provides accurate coverage rates even for small sample sizes. Only when the sample size is very small (when $T = 75$) does it slightly undercover the true threshold parameter. But even then, its coverage rate is much closer to the 95% confidence level than any of the other approaches. Furthermore, their expected length is shorter than the expected length from the BP confidence intervals and only slightly longer than the expected length from the asymptotic approach.

3.2 Monte Carlo results for the model with $y_{t-d}$ as the threshold variable

When the threshold variable is the lag of the dependent variable, we set $d = 1$ and generate data according to the following two-regime TAR(1) process\footnote{The DGP process we use in this case is the same one considered in Enders, Falk and Siklos (2007).}:

$$y_t = (0.7 - 0.5y_{t-1})I(y_{t-1} \leq 0.5) + (-1.8 + 0.7y_{t-1})I(y_{t-1} > 0.5) + \epsilon_t$$

(3.2)

where $\epsilon_t \sim N(0, 1)$ and $I(\cdot)$ is the indicator function. Hence, if $y_{t-1} \leq 0.5$, the intercept and autoregressive coefficient are 0.7 and -0.5, respectively. If $y_{t-1} > 0.5$, they are -1.8 and 0.7, respectively. When the data is generated according to the DGP in (3.2), the results of the experiment are summarized in figure 2.

As in the previous case, the results of the Monte Carlo experiment show that the confidence intervals constructed using the BSIL approach perform poorly. They massively overcover the true threshold parameter.
parameter, with coverage rates of 100% or very close to it for all cases. Likewise, the expected length of the BSIL confidence intervals is much greater than the expected length of any of the other approaches for all sample sizes.

The BP, AIL and FD approaches exhibit confidence intervals with much shorter expected lengths. While the asymptotic approach provides the shortest confidence intervals of all, their coverage rate is not accurate, even for large samples in this case. The BP and FD approaches have very good coverage rates for all sample sizes. However, the expected length of the FD approach is shorter than that of the BP approach in all cases. Therefore, as in the previous subsection, the FD approach performs the best when both, the expected length and coverage rate are taken into consideration.

4 Application: U.S. real GDP growth rate and credit conditions

In this section, we use the FD approach to construct confidence intervals for the threshold parameter of a TAR model of U.S. real GDP growth and to examine the behavior of its relationship with credit
The idea that output might behave differently depending on the conditions of the credit market is an old one. Models of this type have been proposed by several authors, notably Blinder (1987), McCallum (1991), Bernanke and Gertler (1995), Galbraith (1996) and Kiyotaki and Moore (1997). More recently, the intensification of the 2008-2009 recession has been explained by the propagation of negative real shocks through dysfunctions in the credit market. Mendoza and Terrones (2008) show that output, consumption, and investment rise significantly above trend during the expansionary phase of credit booms, and fall below trend during the contractionary phase. Reinhart and Rogoff (2008a) show that standard economic indicators for the U.S., such as asset price inflation, rising leverage, large sustained current account deficits, and a lessening economic growth rate, exhibited all the signs of an imminent financial crisis. Similarly, Dunkelberg and Wade (2009) claim that credit worthiness deteriorated over 2008, leading to higher loan rejection rates, and therefore, lower output.

Confronted with the idea that the current recession is driven by credit-market conditions, we estimated a TAR model of the U.S. real GDP growth rate where the dynamics depend on whether credit conditions are tight or loose. Once the model is estimated, we construct confidence intervals for the threshold parameter for all methods that we considered in the Monte Carlo experiment. The next two subsections detail the procedures and results.

4.1 Model specification and linearity test

We measure the growth rate of output as the logarithmic change in U.S. real GDP and proxy credit conditions using the index of willingness to make consumer installment loans, from the Senior Loan Officer Opinion Survey carried out by the Federal Reserve Board. This is a diffusion index that reflects underwriting standards for approving new credit card applications. In particular, a positive value of the index reflects a positive change in willingness to lend, whereas a negative value of the index reflects a negative change in willingness to lend. The data go from 1966:Q2 through 2009:Q1, corresponding to 171 observations. Figure 3 plots both series over the sample period considered.

The general representation of the model can be described by:

\[ \begin{align*}
\Delta Y_t & = \begin{cases} 
\beta_1 Y_{t-1} & \text{if } C_t \leq c \\
\beta_2 Y_{t-1} & \text{if } C_t > c
\end{cases} 
\end{align*} \]

\[ \text{where } Y_t \text{ is the output growth rate, } C_t \text{ is the credit condition index, and } c \text{ is the threshold.}\]

Many studies (Potter, 1995; Pesaran and Potter, 1997; Enders, Falk and Siklos, 2007) consider so-called self-exciting threshold autoregressive (SETAR) models in which the dynamics of GDP growth depend only on its lags. There are two reasons why our application considers alternative threshold variables. First, while there is evidence of non-linearities using Markov-Switching processes, we are unable to reject linearity when considering a SETAR model for GDP growth. Second, we are also interested in the economic question of what other variables drive the nonlinearity in real GDP dynamics.

The Federal Reserve Board calculates this index by dividing the net number of respondents more willing to make loans by the total number of respondents. Using Haver codes, the formula is: \(((FT121TN1+FT121TN2)-(FT121TN4+FT121TN5))/FT121TN6)*100.
Figure 3
Postwar real U.S. GDP and Credit Conditions
(1966:Q3 through 2009:Q1)

U.S. real GDP growth rate

Willingness to lend

The left panel shows the growth rate of GDP, measured as the logarithmic change in real U.S. GDP. The right panel plots willingness to make consumer installment loans, from the Senior Loan Officer Opinion Survey carried out by the Federal Reserve Board. Recessions, as defined by the NBER, are shown in shaded areas.

Figure 3:

\[
y_t = \left( \alpha_0^1 + \sum_{p=1}^{P} \alpha_p^1 y_{t-p} \right) I(q_{t-d} \leq \gamma) + \left( \alpha_0^2 + \sum_{p=1}^{P} \alpha_p^2 y_{t-p-1} \right) I(q_{t-d} > \gamma) + \epsilon_t \tag{4.1}
\]

where \( y_t \) is the U.S. real GDP growth rate; \( q_t \) is a measure of credit conditions; \( d \) is a delay lag with which credit conditions affect the economy; and \( I(.) \) is the indicator function. In regime 1 (regime 2), when credit conditions are tight (loose), the dynamics of the real GDP growth rate are captured by the vector of coefficients \( \alpha^1 \) (\( \alpha^2 \)).

To specify the model, we first consider the number of autoregressive lags to be included. We choose among lag lengths for \( p = 1, \ldots, 4 \) and the Akaike Information Criterion (AIC) determines that the optimal number of lags is \( P = 4 \). We next test the model for linearity, following Hansen’s (1997) approach, and determine the optimal delay lag, \( d \). The linearity test involves a bootstrap procedure to approximate the asymptotic distribution of relevant sup F-statistic. (For further details about the linearity test, refer to Hansen (1996, 1997)).

Table 1 shows the result of the linearity test. It reports the model sum of squared errors (SSE) from the various models, and the bootstrap-calculated p-values for the test of the null of linearity against the particular threshold model. The least-squares principle suggests that the optimal delay lag should be obtained through the minimization of the SSE. It is then clear from table 1 that the optimal threshold
Table 1: Linearity test results for different TAR models of the U.S. growth rate (threshold variable: \( q_{t-d} \))

<table>
<thead>
<tr>
<th>d</th>
<th>SSE</th>
<th>F-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0100</td>
<td>8.026</td>
<td>0.734</td>
</tr>
<tr>
<td>1</td>
<td>0.0099</td>
<td>11.847</td>
<td>0.358</td>
</tr>
<tr>
<td>2</td>
<td>0.0088</td>
<td>34.099</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>0.0088</td>
<td>36.656</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>0.0085</td>
<td>29.281</td>
<td>0.003</td>
</tr>
</tbody>
</table>

This table reports the sum of squared residuals (SSE), the sup F-test statistic and bootstrap-calculated p-values for the null of linearity against for different model specifications. The bootstrap procedure to calculate p-values used 1,000 replications.

variable corresponds to \( \hat{d} = 4 \).

Note that the model with \( q_{t-4} \) as the threshold variable is highly statistically significant. The associated p-value suggests than linearity can be rejected even at the 1 percent significance level. Furthermore, the implicit assumption that credit conditions affect the growth rate of output with a delay is supported by the results in table 1. In particular, the test cannot reject linearity when credit conditions affect the growth rate of output contemporaneously or with a lag of 1 quarter.

4.2 Estimation results

Once we set the optimal delay lag to \( \hat{d} \), we proceed to estimate the model by OLS, as explained in section 2. The results of the estimated model are presented in table 2.

Table 2: Parameter estimates: TAR model of the U.S. real GDP growth rate. Threshold variable: willingness to lend (lagged 4 periods)

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>( y_{t-1} )</th>
<th>( y_{t-2} )</th>
<th>( y_{t-3} )</th>
<th>( y_{t-4} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient in regime 1</td>
<td>0.0045</td>
<td>0.1308</td>
<td>-0.3892</td>
<td>-0.4071</td>
<td>-0.1876</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
<td>(0.2238)</td>
<td>(0.2761)</td>
<td>(0.2068)</td>
<td>(0.2198)</td>
</tr>
<tr>
<td>Coefficient in regime 2</td>
<td>0.0031</td>
<td>0.2625</td>
<td>0.2135</td>
<td>-0.0037</td>
<td>0.0820</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
<td>(0.2238)</td>
<td>(0.2761)</td>
<td>(0.2068)</td>
<td>(0.2198)</td>
</tr>
</tbody>
</table>

This table reports estimated coefficients from the model given in (4.1). The threshold variable was set to contain the 70 percent middle part of the observations to avoid overfitting. The sample period ranges from the third quarter of 1966 through the first quarter of 2009. Standard errors are reported in parentheses.

The least squares estimate of the threshold is \( \hat{\gamma} = -5.60 \). That is, the TAR model splits the observations into 2 regimes, depending on whether credit conditions are tight or loose. (i.e, when \( q_{t-4} \leq -5.60 \) or \( q_{t-4} > -5.60 \), respectively). A priori, we can think of the two regimes as episodes.
of expansion and contraction. In light of this, we would expect the real GDP growth rate to behave differently depending on whether financial institutions are willing to lend or not. Because the index we use to measure the willingness of financial institution is positive (negative) when they are willing (not willing) to lend, the natural candidate for the threshold parameter is zero.\footnote{Intuitively, as the economy begins to show signs of a contraction, financial institutions are less willing to lend to consumers, reducing the index, consumption and, therefore, output. Conversely, when the economy starts to expand, financial institutions are more likely to lend.}

More importantly, the dynamics of the model are different in each regime. For instance, when credit conditions are tight, the sum of the autoregressive coefficients for the U.S. real GDP growth rate is -0.853. On the other hand, when credit conditions are loose, the sum of the autoregressive coefficients is 0.554. The implication of this is that the U.S. real GDP growth rate exhibits negative serial correlation in times in which financial institutions are less inclined to lend loans to consumers and positive serial correlation in times in which their willingness to lend is higher.

The results are in line with those in Galbraith (1996), who finds substantial evidence of threshold-type effects in U.S. data. Furthermore, given that the sum of autoregressive coefficients in regime 1 (when credit is tight) is large and negative (-0.853), our empirical findings imply that the effects of credit crises on output are only temporary. This supports the results of Kim and Nelson (1999) and Kim, Morley and Piger (2005), who find that the permanent effects of recessions are small. It is worth stressing, however, that such implications are at odds with a literature based on cross-country analysis that suggests that financial crises have large negative and permanent effects. Reinhart and Rogoff (2008a, 2008b, 2008c) and the IMF World Economic Outlook, for instance, analyze several economic indicators (e.g., housing and asset prices, unemployment, output) in a cross-section of countries and find that the effects of financial crises are deep and prolonged.

As the recession that began in December 2007 comes to an end, the path that the economy follows during its recovery phase becomes relevant, and our results and those mentioned in the preceding paragraph have different implications in this matter. If it was true that the effects of financial crises are deep and permanent, then output would be expected to grow along a permanently lower path than the one prevailing before the recession. Meanwhile, our results imply that, when the economy is in recession, the effects will be temporary and output will return back to its path prior to recession.

To illustrate these different implications, we calculate generalized impulse-response functions (GIRF) to evaluate the dynamic response of output to idiosyncratic shocks hitting the economy in each regime. We fix the idiosyncratic shock to be one standard deviation of the vector of residuals from the model. Then, we generate a forecast for $y_t$ given a particular history of the model -whether credit has been
Figure 4
Generalized Impulse-Response Functions
Threshold: Willingness-to-lend index

Regime 1 (26 possible histories)  Regime 2 (141 possible histories)

The graphs show the generalized impulse-response functions of output to a negative idiosyncratic shock for a horizon of 20 quarters. The size of the shock is one standard deviation of the vector of residuals from the estimated model. The GIRF were calculated using 1,000 Monte Carlo repetitions.

The graphs show the generalized impulse-response functions of output to a negative idiosyncratic shock for a horizon of 20 quarters. The size of the shock is one standard deviation of the vector of residuals from the estimated model. The GIRF were calculated using 1,000 Monte Carlo repetitions.

Based on this evidence, our empirical findings support the notion that the effects of credit crises on output growth are temporary.
4.3 Confidence intervals for the estimated threshold

Once the model is estimated, we construct 95% confidence intervals for the estimated threshold parameter based on the methods analyzed in the Monte Carlo experiment. The results are summarized in figure 5. Across the different methods, the threshold parameters included in the confidence intervals range from -79.00 to 63.00.

The bootstrapped inverted LR (BSIL) approach generates a prohibitively wide confidence interval, spanning the whole support of the threshold variable: \([-79.00, 63.00]\). This result is consistent with our findings in the Monte Carlo experiments. The asymptotic inverted LR (AIL) approach, on the other hand, generates a relatively shorter confidence interval, \([-5.70, 27.00]\). The confidence interval generated by the bootstrap percentile approach is \([-36.20, -4.20]\). Notably, all of them are much longer than the confidence interval generated by the FD approach, \([-5.70, 3.67]\), consistent with the results presented in the Monte Carlo experiments. Note also that there is a bias in the sense that the interval is not symmetric around the estimated threshold, \(\hat{\gamma} = -5.60\). With the exception of the BP confidence interval, most of the values included in the interval are above the estimated threshold. Furthermore, the BP interval does not contain the obvious candidate threshold value, zero. All other approaches generate confidence intervals where zero is included as a potential threshold value at the 95% confidence level.
Given the relatively small sample size, and based on the results of the Monte Carlo experiments presented in the previous section, the confidence interval constructed using the FD approach gives us a better idea of the dynamics of the real GDP growth rate when credit conditions change. In particular, the results from this section support the idea that the growth rate is negatively serially correlated when credit conditions over the previous 4 quarters have been tight (corresponding to a negative willingness-to-lend index) and positively correlated when they have been improving (i.e., when the willingness-to-lend index is positive).

5 Concluding Remarks

In this paper, we propose a new approach to constructing confidence intervals for threshold parameters that provides high accuracy and short expected length, even in small samples. In practice, our approach simulates the marginal 'fiducial' distribution of the threshold by integrating out other parameters from the likelihood function. To carry out this methodology, we make use of Markov-chain Monte Carlo (MCMC) methods often employed in Bayesian methods, even though we do so with a frequentist notion of inference in mind. The confidence intervals for the threshold parameter are inherently the same as Bayesian highest-posterior-density (HPD) ones, given non-informative priors (if the marginal fiducial distribution is unimodal and symmetric). However, we adopt a frequentist perspective in the sense that our inferences consider their coverage accuracy and expected length across repeated samples.

To examine the overall performance of our method and to compare its performance to that of bootstrap and asymptotic methods to construct confidence intervals, we conduct a Monte Carlo experiment. The results show that, when considering the expected length and coverage rates of the confidence intervals, the FD approach performs better than the other approaches considered.

We apply our method to a TAR model of U.S. real GDP growth where the dynamics depend on the way credit conditions change. The estimated model supports the existence of threshold-type nonlinearities. Moreover, using our marginal fiducial approach, we find much tighter 95% confidence intervals for the threshold parameter that support the economic theory behind them. In particular, the confidence intervals we construct support the notion that credit crises correspond to recessions with largely transitory effects on the level of output.
Appendix A

Markov-Chain Monte Carlo (MCMC) Algorithm for the Fiducial Distribution

For illustration purposes, we explain here the MCMC algorithm for a threshold autoregressive process with model parameters $\theta$ and threshold parameter $\gamma$. The extension to other blocking schemes is straightforward.

The simulation from the likelihood function is carried out by means of the Metropolis-Hastings (MH) algorithm with an independent chain proposal. To generate the model parameters $\theta$ of the distribution from which the draws are made, we minimize the sum of squared residuals (SSR) conditional on each possible threshold parameter, $\gamma$, at each iteration. This optimization procedure provides the mean and variance-covariance matrix for the independent chain proposal density, based on a multivariate Student-t distribution.

Taking into account these considerations, the Metropolis-Hastings algorithm consists of the following steps:

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

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

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

   (b) Calculate the transition probability:
   
   $$ \alpha(\theta^{(j)}, \theta' | \gamma^{(j)}) = \min \left\{ \frac{\pi(\theta', \gamma^{(j)}) q(\theta', \theta^{(j)} | \gamma^{(j)})}{\pi(\theta^{(j)}, \gamma^{(j)}) q(\theta^{(j)}, \theta' | \gamma^{(j)})}, 1 \right\} $$

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

   $$ \theta^{(j+1)} = \begin{cases} \theta', & \text{if } u \leq \alpha(\theta^{(j)}, \theta' | \gamma^{(j)}) \\ \theta^{(j)}, & \text{otherwise} \end{cases} $$
2. \( \gamma^{(j+1)} \) is generated according to:

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

where

\[
\pi(\gamma^{(j+1)}, y) = \frac{\pi(\gamma) f(y^{(j+1)}, \gamma)}{\pi(y^{(j+1)})} = \frac{\pi(\gamma) f(y^{(j+1)}, \gamma)}{\int_\gamma \pi(\gamma) f(y^{(j+1)}, \gamma)}
\]

\[
= \frac{f(y^{(j+1)}, \gamma)}{\int_\gamma f(y^{(j+1)}, \gamma)}
\]