

Central Bank and asymmetric preferences: an application of sieve estimators to the U.S. and Brazil

Rodrigo de Sá*

Marcelo Savino Portugal†

Abstract

Whether central banks place the same weights on positive and negative deviations of inflation and of the output gap from their respective targets is an interesting question regarding monetary policy. To answer this question, the monetary authority's loss function was estimated non-parametrically using a sieve estimator to expand it through an orthogonal polynomial basis. The economy was modeled with forward-looking agents and with monetary authority's commitment. The method was applied to the United States from 1960 and to Brazil from 1999 onwards. No evidence of asymmetric preferences was found for the U.S. monetary authority. In Brazil, however, the central bank's preferences were asymmetric in terms of inflation, resulting in a larger loss of negative, rather than positive, deviations of inflation from its target.

Key words: Central Bank's preference. Asymmetric loss function. Sieve estimators. Orthogonal polynomials.

JEL: C14, E52

Resumo

Uma questão interessante na política monetária é se os Bancos Centrais dão pesos iguais para desvios positivos e negativos da inflação e do hiato do produto das suas respectivas metas. Para responder à esta questão, estimou-se as preferências da autoridade monetária não parametricamente através do método de sieve estimator. A economia foi modelada com agentes forward-looking e com comprometimento por parte da autoridade monetária. O método foi aplicado para a os Estados Unidos desde 1960 e para o Brasil a partir de 1999. Para a economia norte-americana não foram encontradas evidências de assimetria nas preferências da autoridade monetária. Já no Brasil o Banco Central mostrou ter preferências assimétricas quanto à inflação, auferindo uma maior perda de desvios negativos do que positivos em relação à meta.

Palavras chave: Preferências do Banco Central. Função perda assimétrica. Sieve estimators. Polinômios ortogonais.

*Economic researcher at Fundação de Economia e Estatística. E-mail: rodrigodesas@gmail.com.

†Professor at Graduate Program in Administration and Professor at Graduate Program in Economics, Universidade Federal do Rio Grande do Sul. E-mail: msp@ufrgs.br.

1 Introduction

The literature pays special attention to how central banks conduct the monetary policy. They act to control inflation and, owing to frictions¹ in the economy, they also affect the real side in the short run. As monetary authorities have the power to affect the output temporarily, they likely include output and employment stability among their goals, in addition to price stabilization.

Nonetheless, central banks do not always explicitly disclose their goals. Actually, their reaction to changes in economic variables – through the so-called reaction functions – is perceived more easily. In the seminal work by Taylor (1993), the monetary authority adjusts the nominal interest rate in response to deviations of inflation from its target and from the output gap.

The monetary authority's reaction function can be construed as the result of a model in which the central bank seeks to minimize a loss function, conditional on the restrictions imposed by the economy. Therefore, observing the reaction function, to what extent the interest rate responds to inflation and output deviations, for example, does not allow distinguishing between how much of this effect comes from the monetary authority's preferences and how much is imposed by the structure of the economy. Several studies try to make a distinction between these effects by estimating the central bank's preferences and not only the reaction function. The traditional approach to this problem consists in modeling the economy as linear equations, either backward-looking or forward-looking, and modeling the monetary authority as an agent that minimizes a loss that is a quadratic function of the difference between inflation and its target and of the output gap,

$$L(\pi_t, x_t) = (\pi_t - \pi^*)^2 + \lambda x_t^2,$$

where π is inflation, π^* is its target, x is the output gap and λ is its relative weight given by the monetary authority. A great advantage of this approach is its tractability. Even if one considers the model under commitment (i.e., the central bank's actions affect the future actions of private agents), one is dealing with a dynamic programming problem with quadratic objective function and linear transition functions, which can be solved analytically using Riccati equation². Another advantage of the quadratic loss function is that it can be seen as a second-order Taylor approximation to the utility function expected from a representative agent in a general equilibrium model with rational expectations and price frictions (Woodford, 2003, chapter 6).

It is implied that, by assuming that the central bank's loss can be represented by a quadratic function, the same weights are given to positive and negative deviations from its target variables. For instance, the loss the monetary authority would have with inflation one percentage point above the target would be the same as the one resulting from inflation one percentage point below the target. However, this is not necessarily true: to increase its credibility, the central bank might show greater aversion to inflation levels that exceed the target, as a precautionary measure. Second, the linear-quadratic approach produces linear reaction functions with a constant marginal effect

¹Frictions occur when changes in the general price level are not entirely transferred to the prices and wages of a given firm or sector, causing temporary changes in relative prices.

²See, for instance, Ljungqvist and Sargent (2004, chapter 5).

on the inflation interest rate, both for inflation rates close to or far away from their targets, and the same would occur with the output gap.

The first alternative to obtain reaction functions that are nonlinear is to assume the nonlinearity of the structural functions of the economy. Theoretically, Nobay and Peel (2000) demonstrate that a discretionary policy under a nonlinear Phillips curve produces an inflationary bias if the output corresponds to the potential output. Schaling (2004), based on a theoretical model in which the Phillips curve is convex (positive deviations of the variables are more inflationary than negative deviations are disinflationary), shows that the larger the economic uncertainty, the more aggressive the monetary policy conduct should be. Using U.S. data to calibrate their model, Chiarella et al. (2003) model the economy as a dynamic system with several equations, including a Phillips curve with downwardly rigid wages and prices, and find that the equilibrium of such system is unstable. Finally, Dolado et al. (2005), based on a nonlinear Phillips curve (quadratic in terms of output gap), obtain a reaction function in which one of the terms is the interaction (multiplication) between inflation expectation and the output gap. Using monthly data, the authors found evidence of asymmetry in the monetary policy conduct for Germany, France, Spain, and for the Eurozone, but they could not reject the hypothesis of linearity for the U.S. case.

The second alternative to obtain nonlinear reaction function is by modeling the monetary authority's loss function as a possibly asymmetric function, unlike the usual quadratic function described in the literature. In this vein, Orphanides and Wieland (2000) developed a model with nonlinearities both in the structure of the economy and in the monetary authority's preferences. They consider that the monetary authority's loss function contains an interval on which it is linear, representing the target zone for inflation. In addition, they also allow the Phillips curve to have a linear segment. As a result, they show that there is some incentive for the monetary authority to shift away from a linear policy. Cukierman et al. (2002) develop a model in which the specification of the loss function is generic and, from a first-order condition of the monetary policy, they carry out a comparative static exercise. Kim et al. (2005) estimated the central bank's loss function non-parametrically using the method proposed by Hamilton (2001)³, and found asymmetry in the function for the pre-Volcker period, but linearity for the whole sample and for the Volcker-Greenspan period.

In the same line, Surico (2007), based on Nobay and Peel (2003), estimated the asymmetry of the U.S. central bank's loss function, both for inflation and for the output gap, with linex specification⁴,

$$L(\pi_t, x_t) = \frac{e^{\alpha(\pi_t - \pi^*)} - \alpha(\pi_t - \pi^*) - 1}{\alpha^2} + \lambda \left(\frac{e^{\gamma x_t} - \gamma x_t - 1}{\gamma^2} \right),$$

in which parameters α and γ capture the asymmetry of the function. With $\alpha = \gamma = 0$, linex specification collapses in the quadratic loss function. By modeling the behavior of the monetary authority as discretionary, the author found evidence of asymmetry in the loss associated with the output gap for the pre-Volcker period, in which the central bank was far more concerned with

³Hamilton (2001) estimated a model of the form $y = \mu(x) + \epsilon$, where $\mu(\bullet)$ is an unknown function, considering μ , a random variable itself.

⁴In the original model, the function also depends on an quadratic parameter of smoothing of interest rate.

negative than with positive output deviations. The author, however, did not reject the hypothesis of symmetry of the loss function for inflation. The method proposed by Surico (2007) was applied to Brazil by Aragón and Portugal (2010), who found asymmetry in the Brazilian central bank's loss associated with inflation for the 2000-2007 period, with inflation rate below the target, causing a welfare loss higher than the above-target inflation. For the 2004-2007 subperiod, or for the output gap in the whole sample or in this subperiod, they did not find evidence of asymmetry.

The literature on the preferences regarding the monetary authority's reaction function shows that the hypotheses of the model have been relaxed, including the hypothesis that posits that the preferences are symmetric. The linex specification used by Nobay and Peel (2003), Surico (2007) and Aragón and Portugal (2010) generalizes the loss function, but it is still a parametric assumption about the central bank's behavior.

The natural path for the improvement of this literature is to further reduce restrictions to the functional form of the monetary authority's preferences. Thus, the aim of this paper is to estimate the central bank's loss function using a nonparametric estimation method. For that purpose, an optimization model with monetary authority's commitment is developed and, based on the resulting restrictions on the conditional moments of the variables, a sieve estimator is used to estimate the function that represents the central bank's preferences. The method is applied to U.S. and Brazilian data. This way, it will be possible to determine whether the preferences of these two authorities are asymmetric in relation to inflation and (or) to the output gap and which the loss functions are.

Estimating the form of a function implies minimizing the estimator's criterion function on an infinite-dimensional space, which poses several difficulties. The sieve method proposed by Grenander (1981) consists in minimizing the criterion function in a sequence of simpler spaces, generally of finite dimension, whose dimension increases with sample size.

Besides this introduction, this paper is organized as follows. In the second section the economy's structure is described. The method used for estimation is presented in the third section. In the fourth section are presented the results, for both the U.S. and for Brazil. A brief conclusion follows.

2 The economy

The economy is described by the method put forward by Svensson and Woodford (2005), characterized by two structural equations: an aggregate supply (AS) curve and an IS curve increased by expectations, both forward-looking. These two equations represent the log-linearization of the equilibrium of a microfounded model, as in Clarida et al. (1999), differing only as to the incorporation of an unpredictable random error unaffected by the monetary policy. The model follows the hypotheses below:

Hypothesis 1. *The economy is represented by equations*

$$\pi_{t+1} = \beta E_t \pi_{t+2} + \kappa E_t x_{t+1} + u_{t+1} \quad (1)$$

$$x_{t+1} = E_t x_{t+2} - \phi (E_t i_{t+1} - E_t \pi_{t+2} - r_{t+1}), \quad (2)$$

where π_{t+1} is the inflation between periods t and $t+1$, x_{t+1} is the output gap, i_t is the short-term nominal interest rate (deviation from its steady-state value) and u_{t+1} and r_{t+1} are error terms. In addition, $0 \leq \beta < 1$ is the intertemporal discount factor and $\kappa > 0$ and $\phi > 0$ are constants. E_t is the expectation operator conditional on information at t .

Note that, both in AS (equation (1)) and in IS (equation (2)), the expectation is obtained one period earlier than in the conventional literature, given that the prices are predefined in this economy. For the following analysis, it is also necessary to make an assumption about the evolution of the error terms.

Hypothesis 2. *The error terms u_{t+1} and r_{t+1} follow the transition equations*

$$\begin{aligned} u_{t+1} &= \rho u_t + \varepsilon_{t+1} \\ r_{t+1} &= \bar{r} + \omega (r_t - \bar{r}) + \eta_{t+1}, \end{aligned}$$

where ε_{t+1} and η_{t+1} are independent and identically distributed random variables, $0 \leq \rho < 1$ and $0 \leq \omega < 1$ are constants and \bar{r} is the mean natural real interest rate.

2.1 Optimal monetary policy

Given the equations that describe the economy, the monetary authority seeks to minimize its loss, which is assumed to be a function of the deviation of inflation from its target, of the output gap and of the deviation of the interest rate from its target. The first two goals are usual in the literature; the latter, which represents the preference over smaller interest rate variability, is justified by the monetary authority's concern with transactions frictions, as shown in Woodford (2003, 6.4.1, p. 420). Like utility in the consumer theory, this loss can be represented by a function with certain characteristics.

Hypothesis 3. *The monetary authority can be described by a separable loss function*

$$L(\pi_t - \pi_t^*, x_t, i_t) = \tilde{L}(\pi_t - \pi_t^*, x_t) + \frac{\gamma}{2} (i_t - i_t^*)^2, \quad (3)$$

where π_t^* and i_t^* are the inflation and interest rate targets⁵, respectively, for period t and $\gamma > 0$ is a parameter that measures the weight given by the monetary authority to interest rate smoothing.

Hypothesis 4. *The function $\tilde{L}(\bullet)$ that measures the loss associated with inflation and with the output gap,*

- (i) *has as minimum point $(0, 0)$;*
- (ii) *crosses $\tilde{L}(0, 0) = 0$;*
- (iii) *is limited;*
- (iv) *is supermodular, i.e.,*

⁵Note that the monetary authority's target for the output gap is assumed to be equal to zero, as usually described in the literature. As interests are expressed in deviations from their steady-state value, the interest target is also considered to be equal to zero.

$$\frac{\partial^2 \tilde{L}(\bullet)}{\partial \tilde{\pi}_t^2} \geq 0, \quad \frac{\partial^2 \tilde{L}(\bullet)}{\partial x_t^2} \geq 0, \quad \frac{\partial^2 \tilde{L}(\bullet)}{\partial \tilde{\pi}_t \partial x_t} \geq 0, \quad (4)$$

where $\tilde{\pi}_t = \pi_t - \pi_t^*$ is the deviation of inflation from its target;

(v) belongs to the square-integrable function space, $\tilde{L}(\bullet) \in \mathcal{L}^2(\chi)$, defined in a compact set χ .

Suppose that the minimum point of the monetary authority's loss function is on $(0, 0)$ is important because, otherwise, the function would imply a target (aimed to be achieved by the central bank) that differs from that specified earlier, which is currently used in the literature. The assumptions that the function is supermodular and that it is limited warrants the uniqueness of the reaction function in relation to optimal monetary policy.

Hypothesis (3) is the major contribution of this paper, as it does not assume any specific loss function. The fact that it is separable in the interest rate is a simplification of the present model, but it is recurrent in other studies in this area, also when a quadratic loss function is assumed. Moreover, the quadratic function also complies with the hypothesis of supermodularity.

Given the restrictions imposed by the structure of the economy, the monetary authority minimizes the current value of its loss function. Here it is assumed that the monetary authority does that with commitment, i.e., it takes into account the effect of its actions on the expectations of the private sector and the effects of these expectations on possible future actions. The central bank uses the short-term nominal interest rate as instrument, but this instrument also affects the expectations about inflation and about the output gap of the private sector initiative. Thus, its optimization problem is to minimize the current value of (3) subject to (1) and (2), represented by the following Lagrangian:

$$\begin{aligned} \min_{\{\pi_{t+1}, x_{t+1}, i_{t+1}\}} \mathfrak{L} &= E_{t_0} \sum_{t=t_0}^{\infty} E_t [\beta^{t+1-t_0} L(\tilde{\pi}_{t+1}, x_{t+1}, \tilde{i}_{t+1}) \\ &+ \lambda_{1,t+1} (\pi_{t+1} - \beta E_t \pi_{t+2} - \kappa E_t x_{t+1} - u_{t+1}) \\ &+ \lambda_{2,t+1} (x_{t+1} - E_t x_{t+2} + \phi (E_t i_{t+1} - E_t \pi_{t+2} - r_{t+1}))], \end{aligned} \quad (5)$$

where $\tilde{i}_t = i_t - i_t^*$ is the deviation of the interest rate from its target.

The first-order conditions of the problem are

$$\frac{\partial \mathfrak{L}}{\partial \pi_{t+1}} = E_t [\beta^{t+1-t_0} \tilde{L}_{\pi,t+1} + \lambda_{1,t+1} - \beta \lambda_{1,t} - \lambda_{2,t}] = 0 \quad (6)$$

$$\frac{\partial \mathfrak{L}}{\partial x_{t+1}} = E_t [\beta^{t+1-t_0} \tilde{L}_{x,t+1} - \kappa \lambda_{1,t+1} + \lambda_{2,t+1} - \lambda_{2,t}] = 0 \quad (7)$$

$$\frac{\partial \mathfrak{L}}{\partial i_{t+1}} = E_t [\beta^{t+1-t_0} \gamma \tilde{i}_{t+1} + \phi \lambda_{2,t+1}] = 0, \quad (8)$$

$\forall t = t_0 + 1, \dots$, where, for simplicity of notation, $\tilde{L}_{\pi,t+1} = \partial \tilde{L}(\tilde{\pi}_{t+1}, x_{t+1}) / \partial \pi_{t+1}$, analogous to $\tilde{L}_{x,t+1}$. By Isolating $\lambda_{2,t+1}$ in (8) and substituting it into (7), it is possible to solve this equation for $\lambda_{1,t+1}$, such that

$$\lambda_{1,t+1} = \kappa^{-1} \left(\beta^{t+1-t_0} \tilde{L}_{x,t+1} - \beta^{t+1-t_0} \frac{\gamma}{\phi} \tilde{i}_{t+1} + \beta^{t-t_0} \frac{\gamma}{\phi} \tilde{i}_t \right).$$

Substituting the multipliers in (6), one obtains Euler equation of the monetary authority under commitment for the economy described in this section,

$$E_t \left[L_{\pi,t+1} + \frac{(L_{x,t+1} - L_{x,t})}{\kappa} - \frac{\gamma}{\kappa\phi} \tilde{i}_{t+1} + \left(\frac{\gamma}{\kappa\beta\phi} + \frac{\gamma}{\kappa\phi} + \frac{\gamma}{\beta} \right) \tilde{i}_t - \frac{\gamma}{\kappa\beta\phi} \tilde{i}_{t-1} \right] = 0. \quad (9)$$

As equation (9) must be satisfied throughout the period $t = t_0 + 1, \dots$ so that the monetary authority optimizes its preferences, it is one candidate to optimal policy Woodford (2003, section 7.5.1).

2.2 Reaction function and comparative static

According to Woodford (2003), the central bank must choose its instrument, i_{t+1} , using a reaction function that guarantees the validity of the optimal policy rule. That is, i_{t+1} must guarantee that the projections of inflation and of the output gap, $E_t \pi_{t+1}$ and $E_t x_{t+1}$ are consistent with (9). A way to do that is by supposing that private agents form their expectations rationally, with the projections of these two variables given by the expectation in t defined in equations (1) and (2),

$$E_t x_{t+1} = E_t (x_{t+2} - \phi (i_{t+1} - \pi_{t+2} - r_{t+1})) \quad (10)$$

$$E_t \pi_{t+1} = E_t ((\beta + \kappa\phi) \pi_{t+2} + \kappa x_{t+2} - \kappa\phi (i_{t+1} - r_{t+1}) + u_{t+1}). \quad (11)$$

Therefore, the monetary authority must, at each time t , choose i_{t+1} that satisfies (9) with x_{t+1} and π_{t+1} substituted into (10) and (11), respectively. Given the nonparametric nature of the problem, there is no explicit solution for i_{t+1} and, besides, it is not possible to guarantee a priori that this choice represents an implicit reaction function in the optimal policy rule. The existence of the reaction function is shown in the following proposition.

Proposition 1. *There is a reaction function $i_{t+1} = f(\bullet)$ implied by optimal policy rule (9).*

Proof. Let $\varphi(i_{t+1}|\bullet)$ be the left-hand side of equation (9). The reaction function exists if and only if the i_{t+1} which solves $\varphi(i_{t+1}|\bullet) = 0$ exists and is unique. In order for i_{t+1} to exist and be unique it suffices that the function $\varphi(i_{t+1}|\bullet)$ has a sign change and that the sign of $\partial\varphi(i_{t+1}|\bullet)/\partial i_{t+1}$ is equal to the sign of $\lim_{i_{t+1} \rightarrow \infty} \varphi(\infty|\bullet)$. As $\lim_{i_{t+1} \rightarrow -\infty} \varphi(-\infty|\bullet) = \infty > 0$ and $\lim_{i_{t+1} \rightarrow \infty} \varphi(\infty|\bullet) = -\infty < 0$, the function has a solution. $\partial\varphi(i_{t+1}|\bullet)/\partial i_{t+1} = -\kappa\phi\tilde{L}_{\pi\pi} - 2\phi\tilde{L}_{\pi x} - \phi/\kappa\tilde{L}_{xx} - \gamma/\kappa\phi < 0$ since all parameters are positive (hypotheses (1) and (3)) and the loss function is supermodular (hypothesis (4)), guaranteeing the uniqueness of the solution (where $\tilde{L}_{\pi\pi} = \partial^2 \tilde{L}(\bullet)/\partial \pi^2$, being analogous to the other partial derivatives of the equation). \square

Once confirmed that the reaction function exists, it provides the optimal short-term nominal interest rate that the central bank must choose to satisfy its monetary policy rule. In the present model, the optimal interest rate at t is a function $i_{t+1} = f(\pi_t, x_t, i_t, i_{t-1}, E_t \pi_{t+2}, E_t x_{t+2}, E_t u_{t+1}, E_t r_{t+1})$

of the current economic status (inflation and output gap at t and interest rate at t and $t-1$) and of the expectation at t as to the future state of the economy (inflation, $E_t\pi_{t+2}$, output gap, E_tx_{t+2} , and economic shocks, E_tu_{t+1} and E_tr_{t+1}). By implicitly distinguishing (9), it is possible to obtain the marginal response of the function that defines the interest rate relative to its arguments,

$$\begin{aligned}
\frac{\partial i_{t+1}}{\partial E_t\pi_{t+2}} &= \frac{(\beta + \kappa\phi)\tilde{L}_{\pi\pi} + (2\phi + \beta/\kappa)\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx}}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial E_tx_{t+2}} &= \frac{\kappa\tilde{L}_{\pi\pi} + 2\tilde{L}_{\pi x} + 1/\kappa\tilde{L}_{xx}}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial E_tu_{t+1}} &= \frac{\tilde{L}_{\pi\pi} + 1/\kappa\tilde{L}_{\pi x}}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial E_tr_{t+1}} &= \frac{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx}}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} \in (0, 1) \\
\frac{\partial i_{t+1}}{\partial \pi_t} &= -\frac{1/\kappa\tilde{L}_{\pi x}}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} < 0 \\
\frac{\partial i_{t+1}}{\partial x_t} &= -\frac{1/\kappa\tilde{L}_{xx}}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} < 0 \\
\frac{\partial i_{t+1}}{\partial i_t} &= \frac{\gamma/\kappa\beta\phi + \gamma/\kappa\phi + \gamma/\beta}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial i_{t-1}} &= -\frac{\gamma/\kappa\beta\phi}{\kappa\phi\tilde{L}_{\pi\pi} + 2\phi\tilde{L}_{\pi x} + \phi/\kappa\tilde{L}_{xx} + \gamma/\kappa\phi} < 0.
\end{aligned}$$

Therefore, it should be noted that the response of interests in the model proposed here, which does not restrict the functional form of the monetary authority's preference, has the same signs as the traditional one, in which the central bank increases interests following an increase in inflation expectations, output gap and shocks. Furthermore, it has a negative response to the state variable represented by inflation and by the output gap of the previous period. An important fact is that the marginal effect of previous interest rates on the reaction function varies and decreases over time, $|\partial i_{t+1}/\partial i_t| > |\partial i_{t+1}/\partial i_{t-1}|$.

3 Sieve minimum distance estimate

As mentioned in the previous sections, the aim of this paper is to estimate the central bank's loss function. Estimating a function means estimating a parameter of an infinite-dimensional space, which could be difficult to implement, in addition to having undesirable asymptotic properties, such as inconsistency and a low convergence rate (Chen, 2007, p. 1).

Grenander (1981) proposed to estimate the parameters over a sequence of less complex spaces, usually of finite dimension, and let the dimension of these spaces grow according to the sample size, but slower so that the estimator is consistent.

To guarantee the estimator's desired properties, it is necessary, in addition to the growth rate of the space's complexity used in the approximation, that it well approximates the true space where

the population parameter is inserted. In order for it to be a good approximator, it is necessary that the approximate space be dense in the true space. The fact that a set \mathcal{M} is dense in another set, \mathcal{N} , implies that, for any point $n \in \mathcal{N}$, there is a point $m \in \mathcal{M}$ sufficiently close, or, equivalently, all the open balls of \mathcal{N} includes a point of \mathcal{M} . Thus, if the approximate space is asymptotically dense in the original space, one can find an estimator for the population parameter sufficiently close when the sample size increases.

As the loss function belongs to $\mathcal{L}_2(\chi)$, $\chi \subset \mathbb{R}^2$, compact, the space to be used as approximation should be dense in this set. A space generated by a basis containing orthogonal polynomials that go from χ to \mathbb{R} is dense in $\mathcal{L}_2(\chi)$.

The approximation of a function by a basis of its space has a global characteristic, in which one finds the function that is closest to the original one, with the concept of proximity given by a norm of the space that, in general, takes into account all points of the domain. This differs from what occurs with the Taylor approximation, which has a local characteristic, i.e., the level and slope of the original function are equaled in a given point of the domain, but this does not guarantee that this approximation is good in points outside this neighborhood.

3.1 The estimator

This section shows the generic version of the estimator, as described in Ai and Chen (2003). In rational expectations models, based on the maximizing behavior of the agents, one usually obtains restrictions on the conditional moments that must be satisfied by the variables.

Let (\mathbf{Y}, \mathbf{X}) be the variables of the economy, where \mathbf{Y} is the vector with the endogenous data and \mathbf{X} the vector with exogenous or predetermined data. Then the conditional moments can be written as

$$E[\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{X}] = \mathbf{0}, \quad (12)$$

where $\mathbf{Z} = (\mathbf{Y}', \mathbf{X}'_z)$, $\mathbf{X}_z \subseteq \mathbf{X}$, $\boldsymbol{\rho}(\bullet)$ is a vector of unknown functions that place restrictions on the moments (residuals, as the expectation is equal to zero). $E[\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{X}]$ is the expectation conditional on \mathbf{X} of $\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{X}$. $\boldsymbol{\theta}_0$ is a vector with parameters within a finite-dimensional space and $\mathbf{h}_0(\bullet) = (h_{01}(\bullet), \dots, h_{0q}(\bullet))$ is a vector of unknown functions. The intention is to estimate $\boldsymbol{\alpha}_0 = (\boldsymbol{\theta}_0, \mathbf{h}_0(\bullet))$, which makes the problem semiparametric, as one has parameters of interest both in the finite and infinite dimensions. This property perfectly suits the problem dealt with in the present paper, in which, besides the monetary authority's preference function, there are also structural economic parameters to be estimated.

As variables are time-dependent, it is necessary to establish conditions under which, one can asymptotically use the central limit theorem. In this case, the data have to be stationary β -mixing

Definition 1. Let $\{Y_t, t = 0, \pm 1, \pm 2, \dots\}$ be a strictly stationary series and define

$$\beta(n) = E \left[\sup_{B \in \mathcal{F}_n^\infty} |P(B) - P(B|Y_0, Y_{-1}, Y_{-2}, \dots)| \right],$$

where \mathcal{F}_n^∞ is the σ -algebra⁶ generated by $\{Y_t, n \leq t < \infty\}$. The process $\{Y_t\}$ is β -mixing if $\beta(n) \rightarrow 0$ when $n \rightarrow \infty$. (Fan and Yao, 2005, p. 68)

Intuitively, a process that is β -mixing is asymptotically independent. This because $\beta(n) \rightarrow 0$ when $n \rightarrow \infty$ implies that, for every possible event B (every possible path of Y_t for t between n and ∞), its unconditional and conditional probabilities (on previous Y_t) are the same, which characterizes the asymptotic independence between $\{Y_t, n \leq t < \infty\}$ and $\{Y_0, Y_{-1}, Y_{-2}, \dots\}$.

There is a sample $\{(\mathbf{Y}_t, \mathbf{X}_t), t = 1, \dots, N\}$ for the estimation of (12), where $(\mathbf{Y}_t, \mathbf{X}_t)$ is a stationary β -mixing process. The restrictions on conditional moments, $\boldsymbol{\rho} : \mathcal{Z} \times \mathcal{A} \rightarrow \mathbb{R}^{d_\rho}$, are known up to an unknown vector of parameters, $\boldsymbol{\alpha}_0 = (\boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) \in \mathcal{A} = \Theta \times \mathcal{H}$, where Θ is the space of finite-dimensional parameters and $\mathcal{H} = \mathcal{H}_1 \times \dots \times \mathcal{H}_q$ is the functional space where \mathbf{h}_0 is inserted. If the distribution function of \mathbf{Y} given \mathbf{X} , $F_{\mathbf{Y}|\mathbf{X}}$, were known, it would be possible to calculate each component of (12) as

$$E[\rho_i(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{x}] \equiv m_i(\mathbf{x}, \boldsymbol{\alpha}) = \int \rho_i(\mathbf{y}, \mathbf{x}_z, \boldsymbol{\theta}, \mathbf{h}(\bullet)) dF_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}(\mathbf{y}),$$

$i = 1, \dots, d_\rho$. Let $\mathbf{m}(\mathbf{x}, \boldsymbol{\alpha}) = (m_1(\mathbf{x}, \boldsymbol{\alpha}), \dots, m_{d_\rho}(\mathbf{x}, \boldsymbol{\alpha}))'$. Then the true parameter $\boldsymbol{\alpha}_0$ minimizes

$$\inf_{\boldsymbol{\alpha}=(\boldsymbol{\theta}, \mathbf{h}) \in \Theta \times \mathcal{H}} E[\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})' [\boldsymbol{\Sigma}(\mathbf{X})]^{-1} \mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})], \quad (13)$$

where $\boldsymbol{\Sigma}(\mathbf{X})$ is a positive definite matrix.

A natural estimator for $\boldsymbol{\alpha}_0$ would be the argument that minimizes the sample version of (13). However, to allow for this minimization, two issues have to be addressed. First, as the true functional form of $F_{\mathbf{Y}|\mathbf{X}}$, and consequently of $\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})$, is not known, this latter function must be estimated by any nonparametric method. Ai and Chen (2003) suggest also estimating it by the sieve estimator, more precisely by the linear sieve estimator, discussed in what follows. A second aspect is the fact that the minimization over the infinite-dimensional set \mathcal{H} can be complex and result in undesirable properties in the estimators. As discussed earlier, an alternative is the sieve method, replacing \mathcal{H} with a sequence of simpler sets, which are dense in the original space when the size of sample N increases. Thus, the optimization problem occurs over the set $\mathcal{H}_n = \mathcal{H}_n^1 \times \dots \times \mathcal{H}_n^q$, simpler than the original one, and usually of finite dimension. With the treatment of these two issues, we have the sieve minimum distance (SMD) estimator,

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_n &= (\hat{\boldsymbol{\theta}}_n, \hat{\mathbf{h}}_n) \\ &= \arg \min_{\boldsymbol{\alpha}=(\boldsymbol{\theta}, \mathbf{h}) \in \Theta \times \mathcal{H}_n} n^{-1} \sum_{t=1}^N \hat{\mathbf{m}}(\mathbf{X}_t, \boldsymbol{\alpha})' [\hat{\boldsymbol{\Sigma}}(\mathbf{X})]^{-1} \hat{\mathbf{m}}(\mathbf{X}_t, \boldsymbol{\alpha}), \end{aligned} \quad (14)$$

⁶A σ -algebra in a set Ω is a collection \mathcal{A} of subsets of Ω such that the empty set belongs to it and given $\mathcal{A}_n \in \mathcal{A}$, then $(\bigcup_{n=1}^\infty \mathcal{A}_n) \in \mathcal{A}$ e $\mathcal{A}_n^c \in \mathcal{A}$ (Isnard, 2007, p. 38). These characteristics of σ -algebra are fundamental in the probability theory as they guarantee that an event consisting of the combination of other events with definite probability will also have probability and, in addition, if an event has some probability associated with it, then its complementary event also has it.

where $\hat{\Sigma}(\mathbf{X})$ is a consistent estimator for $\Sigma(\mathbf{X})$, used to deal with some possible heteroskedasticity, and $\hat{\mathbf{m}}(\mathbf{X}, \boldsymbol{\alpha})$ is the linear sieve estimator for $\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})$.

Estimator $\hat{\mathbf{m}}(\mathbf{X}, \boldsymbol{\alpha})$ consists in expanding the original function into a basis of its space. Let $\{p_j(\mathbf{X}), j = 1, 2, \dots\}$ be a basis of \mathcal{L}^2 and $\mathbf{p}^{k_n}(\mathbf{X}) = (p_1(\mathbf{X}), \dots, p_{k_n}(\mathbf{X}))'$ the vector that contains all these bases of order smaller than or equal to k_n . This way, this basis can approximate the conditional moments

$$m_i(\mathbf{X}, \boldsymbol{\alpha}) = \sum_{j=0}^{k_n} c_{i,j} p_j(\mathbf{X}, \boldsymbol{\alpha}) = \mathbf{p}^{k_n}(\mathbf{X})' \mathbf{c}_i, \quad k_n \rightarrow \infty, \quad (15)$$

$i = 1, \dots, d_\rho$, where $\mathbf{c}_i = (c_{i,1}, \dots, c_{i,k_n})'$ is the vector with the coefficients of the i -th moment condition. The authors suggest as linear sieve estimator, for a given order of approximation k_n , estimating the coefficients $\{c_{i,j}, i = 1, \dots, d_\rho, j = 1, \dots, k_n\}$ by ordinary least squares⁷, regressing $\rho_i(\mathbf{Z}, \boldsymbol{\alpha})$ on $\mathbf{p}^{k_n}(\mathbf{X})$. Let $\boldsymbol{\rho}_i(\boldsymbol{\alpha}) = (\rho_i(\mathbf{Z}_1, \boldsymbol{\alpha}), \dots, \rho_i(\mathbf{Z}_N, \boldsymbol{\alpha}))'$ be the vector with the assessment of this moment condition in each point of the sample, analogous definition for $\mathbf{m}_i(\boldsymbol{\alpha})$, and \mathbf{P} the $N \times k_n$ matrix whose characteristic element is $P_{t,j} = p_j^{k_n}(\mathbf{X}_t)$, or, $\mathbf{P} = (\mathbf{p}^{k_n}(\mathbf{X}_1), \dots, \mathbf{p}^{k_n}(\mathbf{X}_N))'$. Then the regression for each moment restriction, in matrix notation, is given by

$$\hat{\mathbf{m}}_i(\boldsymbol{\alpha}) = \hat{\boldsymbol{\rho}}_i(\boldsymbol{\alpha}) = \mathbf{P} \hat{\mathbf{c}}_i, \quad i = 1, \dots, d_\rho,$$

with the estimator $\hat{\mathbf{c}}_i = (\mathbf{P}'\mathbf{P})^{-1} \mathbf{P}'\boldsymbol{\rho}_i(\boldsymbol{\alpha})$. Substituting it into (15) one obtains an explicit form for the expectation of moment restrictions⁸,

$$\hat{m}_i(\mathbf{X}, \boldsymbol{\alpha}) = \mathbf{p}^{k_n}(\mathbf{X}, \boldsymbol{\alpha})' (\mathbf{P}'\mathbf{P})^{-1} \mathbf{P}'\boldsymbol{\rho}_i(\boldsymbol{\alpha}), \quad i = 1, \dots, d_\rho. \quad (16)$$

In (16) we have an equation for each of the k_n moments. They can be combined into a single equation,

$$\hat{\mathbf{M}}(\mathbf{x}, \boldsymbol{\alpha}) = \Xi(\boldsymbol{\alpha}) \mathbf{P} (\mathbf{P}'\mathbf{P})^{-1} \mathbf{p}^{k_n}(\mathbf{x}), \quad (17)$$

where $\hat{\mathbf{M}}(\mathbf{x}, \boldsymbol{\alpha}) = (\hat{m}_1(\mathbf{x}, \boldsymbol{\alpha}), \dots, \hat{m}_{d_\rho}(\mathbf{x}, \boldsymbol{\alpha}))'$ and $\Xi(\boldsymbol{\alpha}) = (\boldsymbol{\rho}_1(\mathbf{X}, \boldsymbol{\alpha}), \dots, \boldsymbol{\rho}_{d_\rho}(\mathbf{X}, \boldsymbol{\alpha}))'$.

Note that the estimated expectation of conditional moments, evaluated in any point \mathbf{x} , is a function of unknown parameters of the model, both finite-dimensional and infinite-dimensional.

Substituting expectation (17) in (14) it has the criterion function that needs to be minimized in order to obtain the SMD,

$$n^{-1} \sum_{t=1}^N \hat{\mathbf{M}}(\mathbf{X}_t, \boldsymbol{\alpha})' \left[\hat{\Sigma}(\mathbf{X}) \right]^{-1} \hat{\mathbf{M}}(\mathbf{X}_t, \boldsymbol{\alpha}).$$

Using the identity matrix as weighting matrix, $\hat{\Sigma}(\mathbf{X}) = I$, the SMD estimator is identical with the generalized method of moments (GMM) estimator. This result is shown by Ai and Chen (2003, p.

⁷For a best description, see Newey (1997).

⁸An equivalent formulation for the equation is given by Ai and Chen (2003, p. 1799): $\hat{m}_i(\mathbf{X}, \boldsymbol{\alpha}) = \sum_{j=1}^N \rho_i(\mathbf{Z}_j, \boldsymbol{\alpha}) \mathbf{p}^{k_n}(\mathbf{X}_j)' (\mathbf{P}'\mathbf{P})^{-1} \mathbf{p}^{k_n}(\mathbf{X})$.

1799). With $\hat{\Sigma}(\mathbf{X}) = I$, the criterion function of the SMD can be written as

$$\left(\sum_{t=1}^N \boldsymbol{\rho}(\mathbf{Z}_t, \boldsymbol{\alpha}) \otimes \mathbf{p}^{k_n}(\mathbf{X}_t) \right) \left(\mathbf{I} \otimes (\mathbf{P}'\mathbf{P})^{-1} \right) \left(\sum_{t=1}^N \boldsymbol{\rho}(\mathbf{Z}_t, \boldsymbol{\alpha}) \otimes \mathbf{p}^{k_n}(\mathbf{X}_t) \right)', \quad (18)$$

where \mathbf{I} is the $d_\rho \times d_\rho$ identity matrix and \otimes denotes the Kronecker product. (18) is the criterion function of the GMM estimator of the system of unconditional moment restrictions

$$E[p_j(\mathbf{X}) \boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\alpha}_o)] = 0, \quad j = 1, \dots, k_n$$

with $\mathbf{I} \otimes (\mathbf{P}'\mathbf{P})$ weighting matrix. As highlighted by the authors, this alternative view of the SMD estimator provides a lower threshold to the approximation level k_n of linear sieve estimator,

$$d_\rho k_n \geq d_\theta + q_n,$$

where d_θ is the dimension of parameter space Θ and q_n is the dimension of the approximation of infinite-dimensional parameters, \mathcal{H}_n , only.

The next step is the choice of the sequence that approximates the space to functions \mathcal{H} . In this paper, we used the space generated by orthogonal polynomials, more specifically, by Chebyshev polynomials. For the problem analyzed here, the use of this basis, to the detriment of others that are commonly employed in the literature, such as splines and neural networks, is primarily due to the fact that polynomials have an explicit and simple form for their derivatives, which is how the monetary authority's loss function to be estimated enters the criterion function. A second aspect is that, according to the assumptions of the economic model to be estimated, the central bank's loss function is strictly concave in order to meet the bank's optimization requirements. Furthermore, the function is smooth, which eliminates the advantage of splines.

3.2 Loss function approximation

A basis for a space of functions defined in a two-dimensional set $\chi = \chi_1 \times \chi_2 \subset \mathbb{R}^2$ is a tensor of the bases of the univariate function spaces defined in χ_1 and χ_2 . Let B_1 be a basis of space $\mathcal{L}_2(\chi_1)$ and B_2 a basis of space $\mathcal{L}_2(\chi_2)$. Then a basis for $\chi = \chi_1 \times \chi_2$ is the tensor (set) $B_1 \otimes B_2 = \{f(\pi, x) = b_1(\pi) b_2(x) \mid b_1 \in B_1, b_2 \in B_2\}$. However, the amount of elements in a tensor increases exponentially with their dimension, leading Judd (1998) to argue that it is more efficient to use only the components of the tensor $f(\pi, x) = b_1(\pi) b_2(x)$ that have a lower or the same level as a k chosen in advance⁹. In this paper, this method was also used in an attempt to obtain a higher computational efficiency.

As mentioned in the previous subsection, the basis chosen to approximate the loss function was the sequence of Chebyshev polynomials. Because it is reasonable that the pair deviation of inflation from its target and output gap is contained in $[-1, 1]^2$, the domain of the polynomials did

⁹This is known as complete polynomials. A heuristic justification given by Judd (1998) is the multivariate Taylor approximation, in which the term of order k of the expansion is composed of polynomials $x_1^{i_1} \dots x_n^{i_n}$ such that $\sum_{l=1}^n i_l = k$.

not have to be altered. Let U_n be the Chebyshev polynomials of level n . The monetary authority's loss function is approximated by

$$L(\tilde{\pi}_t, x_t, \tilde{i}_t) = \tilde{L}(\tilde{\pi}_t, x_t) + \frac{\gamma}{2} (\tilde{i}_t)^2,$$

$$\tilde{L}(\tilde{\pi}_t, x_t) = \sum_{i=1}^{K_\pi} \psi_{i,0} U_i(\tilde{\pi}_t) + \sum_{i=1}^{K_x} \psi_{0,i} U_i(x_t) + \sum_{i=1}^{k_\pi} \sum_{j=1}^{k_x} \psi_{i,j} U_i(\tilde{\pi}_t) U_j(x_t) + \psi_{0,0}.$$

If the complete tensor had been used, one would have $K_\pi = K_x = k_\pi = k_x$. For the sake of efficiency, $K_\pi = K_x = 5$ and $k_\pi = k_x = 1$ were used here. This approximation level allows for sufficient asymmetry in the loss function, in addition to providing an interaction term between inflation and the output gap.

In order for the estimated loss function to comply with hypothesis (4), the optimization must occur with some restrictions. Guaranteeing that the minimum of the function occurs in point $(0, 0)$ allows reducing two parameters that need be estimated. Therefore, solving the equations that imply that $(0, 0)$ is the minimum, $\partial \tilde{L}(0,0)/\partial \tilde{\pi}_t = 0$ and $\partial \tilde{L}(0,0)/\partial x_t = 0$, it is possible to let $\psi_{1,0}$ and $\psi_{0,1}$ be a function of other parameters,

$$\psi_{1,0} = -(U_1(0))^{-1} \left(\sum_{i=2}^{K_\pi} \psi_{i,0} U_1'(0) + \sum_{i=1}^{k_\pi} \sum_{j=1}^{k_x} \psi_{i,j} U_i'(0) U_j(0) \right)$$

$$\psi_{0,1} = -(U_1(0))^{-1} \left(\sum_{i=2}^{K_x} \psi_{0,i} U_1'(0) + \sum_{i=1}^{k_\pi} \sum_{j=1}^{k_x} \psi_{i,j} U_i(0) U_j'(0) \right),$$

eliminating them.

To guarantee the supermodularity of the loss function, the estimated parameters must satisfy equation (4) $\forall \tilde{\pi}, x \in (-1, 1)$ ¹⁰.

As optimization problems are indifferent to affine transformations¹¹ in the objective function, for identifiability reasons, one must normalize the loss function to be estimated. In this paper, we normalized the function by allowing $\psi_{2,0} = 1$, which is equivalent to dividing it by $|\psi_{2,0}|$.¹²

3.3 Conditional moments and instruments

The major parameter of interest in this paper is the central bank's loss function. A natural choice for its estimation is Euler function (9). Nevertheless, this function does not contemplate all the information available in the monetary authority's optimization problem. To include this piece of information, it is necessary to jointly estimate the structural equations of the economy, (1) and

¹⁰In practice, as the values for inflation and output gap are around zero and far from -1 and 1 , supermodularity was demanded only in $(-0.9, 0.9)$. This prevents the rejection of candidate functions that do not satisfy this condition only in the limits of their support.

¹¹Affine transformation is a linear operator of the form $z \mapsto az + b$, $a > 0$.

¹²Based on several previous estimates, it is known that $\psi_{2,0} > 0$.

(2).

Euler equation is already in the conditional moment format required by the method. On the other hand, the other two equations have to be transformed. By applying the conditional expectation to the information set at time t in (1), one obtains $E_t[\pi_{t+1} - \beta\pi_{t+2} - \kappa x_{t+1}] = E_t[u_{t+1}]$. However, one still does not have a conditional moment whose expectation is equal to zero. It is necessary to subtract the amount $E_t[u_{t+1}] = E_t[\varrho u_t] = E_t[\varrho(\pi_t - \beta\pi_{t+1} - \kappa x_t)]$ from both sides of the equation, in which hypothesis (2) on error term structure was used. Thus, we have the conditional moment restriction

$$E_t[-\beta\pi_{t+2} + (1 + \varrho\beta)\pi_{t+1} - \varrho\pi_t - \kappa x_{t+1} + \varrho\kappa x_t] = 0. \quad (19)$$

Analogously, from (2), it is possible to obtain the third moment restriction to be estimated,

$$E_t[x_{t+2} + (1 + \omega x_{t+1}) - \omega x_t + \phi(i_{t+1} - \omega i_t) - \phi(\pi_{t+2} - \omega\pi_{t+1}) - \phi(1 + \omega)\bar{r}] = 0. \quad (20)$$

This way, the vector with restrictions on conditional moments that characterize the economy $E[\rho(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{X}] = \mathbf{0}$ is given by equations (9), (19) and (20).

In the present model, all the variables used, $\{\pi_t, x_t, i_t, t = t_0, \dots, N\}$, are endogenous, i.e., all of them are determined by the equations that describe it.

As outlined in the previous section, the first step is to estimate the expectation of conditional moments using the linear sieve estimator, as a function of the variables present in the information set. At time t , the information set is given by $I_t = \{\pi_\tau, x_\tau, i_\tau, \tau = t_0, \dots, t\}$. However, in practice, it is not necessary to use all variables present in the information set. In this paper, we used only the variables from time t , in addition to those from the previous period. In this way, one has a basis with 14 elements for the approximation of $\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})$,

$$\mathbf{p} = (1, \pi_t, x_t, i_t, \pi_t^2, x_t^2, i_t^2, \pi_t x_t, \pi_t i_t, x_t i_t, \pi_t x_t i_t, \pi_{t-1}, x_{t-1}, i_{t-1}).$$

Once the monetary authority's loss function has been estimated, one can verify whether or not it is quadratic. If so, its third derivative assumes value zero in all points of its domain and, therefore, the expectation of the third derivative is equal to zero.¹³ Then, it is possible to test whether the loss function is not quadratic by testing whether the expectation of its third derivative is different from zero. As the expectation is a smooth functional and the loss function is estimated by the sieve method, it is possible to use the result obtained by Chen and Shen (1998), who show that these functionals converge to a normal distribution at the rate $N^{1/2}$.¹⁴ So, the test statistics is $\mu_q = E(\partial^3 L / \partial q^3)$, $q = \{\pi, x\}$.

¹³Note that the opposite does not hold.

¹⁴The result is shown in theorem 2 proposed by Chen and Shen (1998).

4 Results

4.1 United States

Data from the first quarter of 1960 to the last quarter of 2006, totaling 188 observations, were used to model the preferences of the Federal Reserve System (Fed)¹⁵. The variables used followed Surico (2007). The logarithmic variation of the Personal Consumption Expenditure (PCE) deflator, available from the Federal Reserve Bank of St. Louis, was used as measure of inflation. The interest rate consisted of the Effective Federal Funds Rate and output was measured by the Industrial Production Index, both available at the Fed website. All the variables used were calculated as the value accumulated over the last four quarters. The inflation target was assumed to be constant and equal to 0.02. The output gap and the deviation of the interest rate from its steady-state value were calculated using the Hodrick-Prescott filter.

The histograms of these variables are shown in Figure (1a). Note that, considering the whole sampling period, most of the time, inflation was below its target. The output gap had both positive and negative periods. The historical series of these variables are displayed in Figure (2).

Over the nearly 5 decades included in the analysis, the economic policy conduct by the Fed has evidently gone through changes. Clarida et al. (2000) found evidence that the central bank's reaction function changed after Paul Volcker took over in 1979. According to the authors, during the Volcker-Greenspan period, the U.S. monetary authority began to respond more to inflation expectation, whereas in the pre-Volcker era the policy was less active. Based on this fact, as in Surico (2007), the Fed's loss function was estimated for the subsample corresponding to the Volcker-Greenspan period, more specifically to the period between the fourth quarter of 1982 and the last quarter of 2006. The author chose 1982 instead of 1979 because, according to him, the U.S. central bank's policy instrument between 1979 and 1982 was no longer the short-term interest rate, being replaced with compulsory deposits.

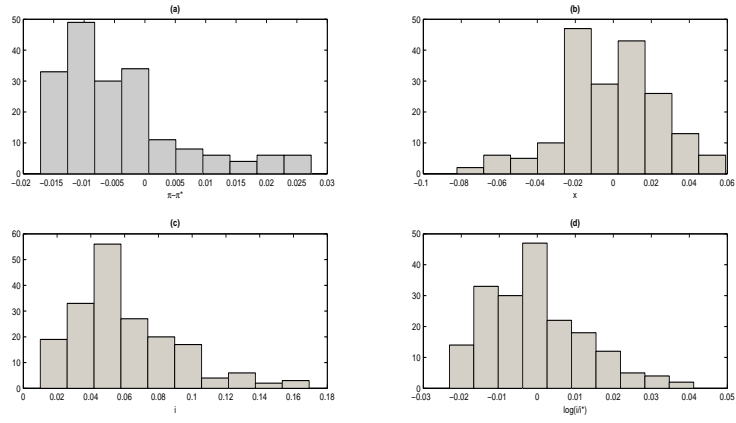
Table (1) shows the results for the estimated parameters in the whole sample. Among the parameters, it is important to underscore the low value for the parameter that confers sensitivity to the output gap to real interest rate ratio, $\phi = 0.009$.

The loss function estimated from the parameters shown in Table (1) is illustrated in Figure (3). As the parameter that shows the interaction between inflation and output gap was low, $\psi_{1,1} = 0.0023$, the analysis can be made by separately considering the losses from each of the

¹⁵More recent data were not used so as not to contaminate the sample with the international financial crisis of 2008.

Figure 1: Histogram of the variables used - USA

(a) 1960.I - 2006.IV



(b) 1982.IV - 2006.IV

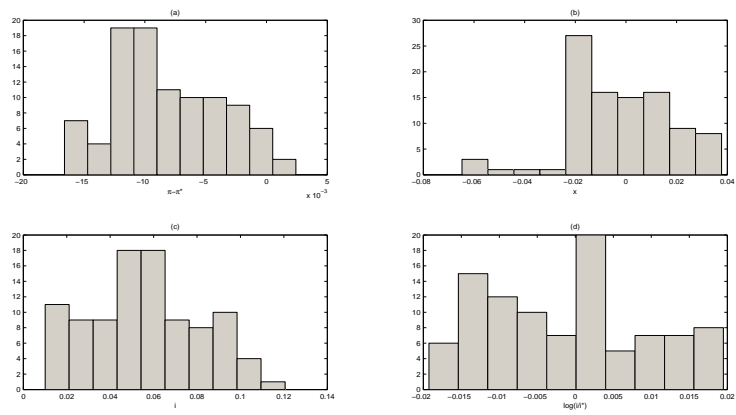


Figure 2: Historical series of the variables used - USA - 1960.I-2006.IV

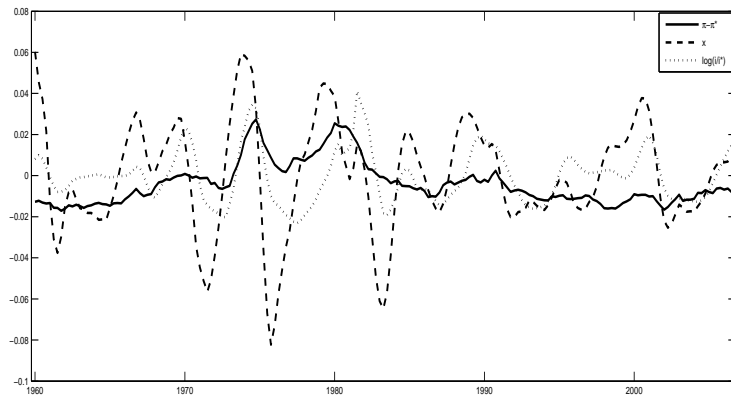
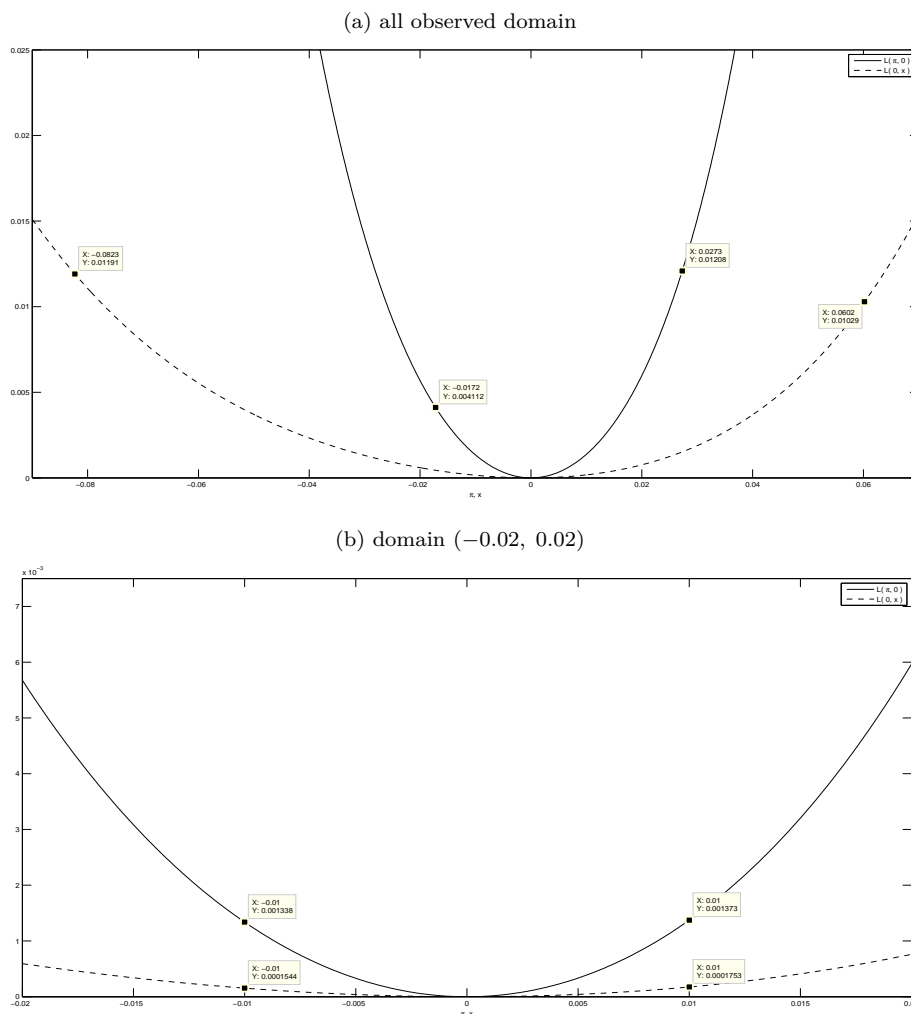


Table 1: Estimated parameters - USA

Parameter	1960.I-2006.IV		1982.IV-2006.IV	
	Value	SD	Value	SD
$\psi_{1,0}$	0.4278	R	-0.0172	R
$\psi_{2,0}$	1.0000	N	1.0000	N
$\psi_{3,0}$	0.4670	0.4167	-0.0324	0.4089
$\psi_{4,0}$	0.4210	0.0030	0.4176	0.0029
$\psi_{5,0}$	0.1403	0.1262	-0.0167	0.1240
$\psi_{0,1}$	0.0240	R	-0.0118	R
$\psi_{0,2}$	0.0467	0.0332	0.0311	0.0556
$\psi_{0,3}$	0.0249	0.0167	-0.0130	0.0245
$\psi_{0,4}$	0.0195	0.0136	0.0109	0.0233
$\psi_{0,5}$	0.0068	0.0052	-0.0040	0.0075
$\psi_{1,1}$	0.0023	0.0061	0.0086	0.0072
γ	0.0001	0.0052	0.0028	0.0016
β	0.9228	0.0105	0.9158	0.0327
κ	0.0792	0.06074	0.1925	0.0794
δ	0.0158	0.04962	0.1070	0.0214
ϕ	0.0090	R	0.0258	R
\bar{r}	0.0468	0.00421	0.0410	0.0073
ϱ	0.8776	0.04720	0.8218	0.1991
ω	0.8133	0.07076	0.7870	0.1689

Note: Value is the estimated value and SD is the estimated standard deviation. The standard deviation refers to bootstrapping with 500 replicates. Restricted parameters (R) are not estimated directly; i.e., a function is given by the model of the estimated parameters. N is the normalized parameter in the estimation.

Figure 3: Loss function - USA- 1960.I - 2006.IV



Note: $L(\pi, 0)$ shows the central bank's loss function value for several inflation values, with output gap equal to zero. Analogous to $L(0, x)$. The values displayed on the upper graph represent the maximum and minimum inflation and output gap. The lower graph shows only the interval $(-0.02, 0.02)$.

Table 2: Quadratic loss function test - USA

(a) 1960.I - 2006.IV			
Statistic	Estimated value	SD	<i>t-value</i>
$\mu_\pi = E(\partial^3 L / \partial \pi^3)$	-0.201	0.322	-0.625
$\mu_x = E(\partial^3 L / \partial x^3)$	0.066	0.040	1.634
(b) 1982.IV - 2006.IV			
Statistic	Estimated value	SD	<i>t-value</i>
$\mu_\pi = E(\partial^3 L / \partial \pi^3)$	-0.051	0.168	-0.305
$\mu_x = E(\partial^3 L / \partial x^3)$	0.000	0.052	-0.008

Note: The t-value corresponding to 5% is 1.65.
The standard deviation refers to bootstrapping with 500 replicates.

variables, without loss of generality. This result supports the current approaches available in the literature, in which the monetary authority's loss function is separable in inflation and in the output gap.

According to the results, the U.S central bank had a bigger concern with inflation than with the output gap between 1960 and 2006, as had already been pointed out in other studies. It should be remarked, however, that because the absolute values of the output gap are higher than inflation (Figure(2a)), its extreme values imply loss for the monetary authority similarly to the extreme values observed for inflation. Also note that the losses associated with inflation and output gap are both slightly asymmetric to the right, with a higher loss for the central bank when the variables are above their targets (the loss with deviation of inflation equal to 1% was 1.026 times greater than the loss with a -1% deviation; the loss with output gap equal to 1% was 1.135 times greater than with a -1% gap). However, it was not possible to reject, at a 5% significance level, the hypothesis that the true loss function is quadratic in each of the variables for this period, as outlined by the test in Table (2a).

As to the whole sample, the central bank in the Volcker-Greenspan period managed to keep inflation below its target for an even longer timeframe, with a maximum deviation of inflation of 0.24% (Figure (1b)). The results for this period were similar to those for the whole sample, with the estimated loss function shown in Figure (4). As with the whole sample, the loss function for inflation is above that obtained for the output gap. However, as in the previous case, because the absolute values of the output gap are higher than inflation, for a set of observations, the loss with the former is larger than with the latter. In the Volcker-Greenspan period, both the loss function related to inflation and the one related to output gap revealed slight asymmetry to the right. However, one should not reject that any of them is actually quadratic, as shown by the test in Table (2b). The higher amount of observations of positive rather than negative deviations of

inflation makes it harder to reject inflation, causing a significant variance in the estimation for the positive segment of the straight line.

Thus, both for the whole sample and the Volcker-Greenspan period, the loss function related to inflation and to output gap were asymmetric to the right, that is, the monetary authority had larger losses with positive deviations. Notwithstanding, it is not possible to reject that the true loss function is quadratic both for inflation and output gap in none of the tested periods. The lack of asymmetries in Fed's preferences for these periods was also found by Surico (2007), when modeling the asymmetry restricted to the linex equation.¹⁶

4.2 Brazil

Given that the Brazilian inflation targeting regime is more recent, having started in April 1999, it was used monthly data to estimate the central bank's loss function. The sample contains 143 observations, from April 1999 to February 2011. To remove the effect of U.S. dollar appreciation and of the rise in inflation between 2002 and 2003 from the estimation, we also estimated the model for the period between January 2004 and February 2011.

The broad consumer price index (IPCA), which indicates the central bank's actions, was used to measure inflation. Unlike the U.S. case, in Brazil, the inflation target is explicit and changes over time. As the target is for the annual inflation, it was interpolated as Aragón and Portugal (2010) did.¹⁷ For the output gap, it was used the industrial production index calculated by the Brazilian Institute of Geography and Statistics (IBGE), whereas for interest rates, it was used the over-Selic rate. As with the previous estimation, all variables were used with their accumulated value over the past 12 months and both the output gap and the deviation of interests from their targets were calculated using the Hodrick-Prescott filter.

Figure (5a) shows the histogram of Brazilian variables. In Brazil, differently from the U.S., most observations of inflation were higher than the target, and the amplitude of the deviations of inflation were also larger than the amplitude of the output gap. In addition, it is possible to observe the high interest rates found in the Brazilian economy, with most observations above 10% p.a. The paths of these variables are plotted in Figure (6).

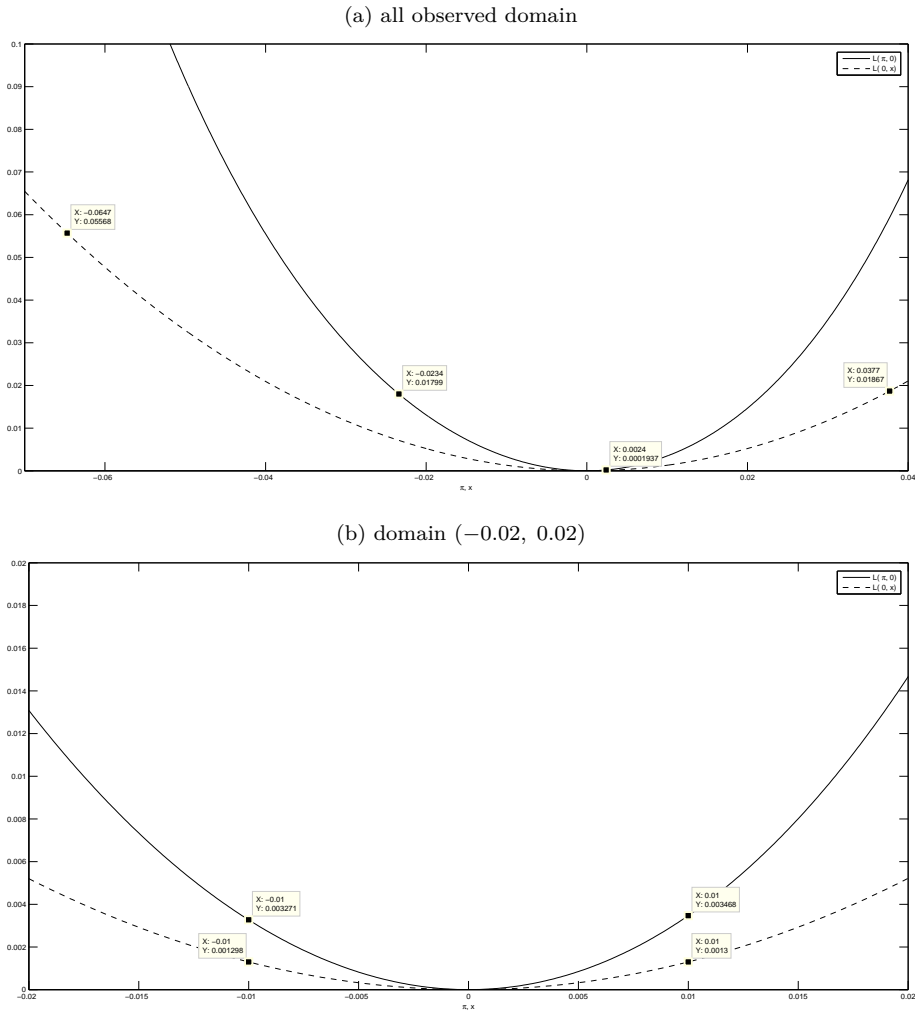
Table (3) presents the value of the parameters estimated for the Brazilian economy. As to the structure of the economy, the slopes of the Phillips and IS curves (0.1494 and 0.0335, respectively) are larger than those of the U.S. economy.

The estimation of the monetary authority's loss function during the targeting regime is shown in Figure (7). Again, the loss related to inflation is above that which was obtained for the output gap,

¹⁶The author found asymmetry only for output for the pre-Volcker period.

¹⁷Aragón and Portugal (2010) interpolate the inflation target according to equation $\pi_{m,t}^* = (12-m/12)\pi_{t-1}^* + (j/12)\pi_t^*$, where $\pi_{m,t}^*$ is the inflation target interpolated for month m of year t and π_t^* is the annual inflation target disclosed for year t . For 2003 and 2004, the adjusted targets of 8.5% and 5.5%, respectively, established by the 2003 Statement (*Carta Aberta*), were used.

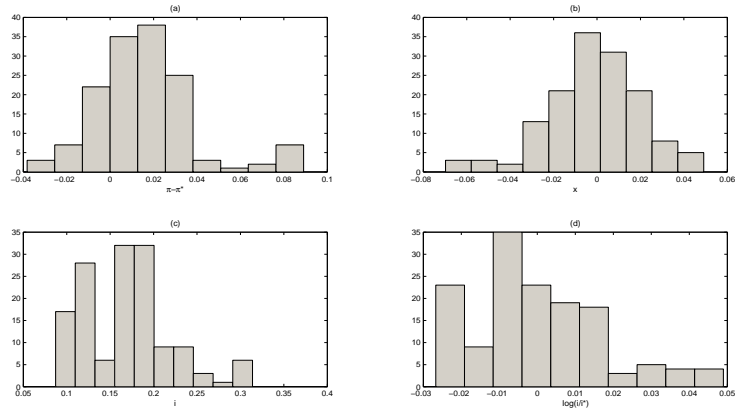
Figure 4: Loss function - USA- 1982.IV - 2006.IV



Note: $L(\pi, 0)$ shows the central bank's loss function value for several inflation values, with output gap equal to zero. Analogous to $L(0, x)$. The values displayed on the upper graph represent the maximum and minimum inflation and output gap. The lower graph shows only the interval $(-0.02, 0.02)$.

Figure 5: Histogram of the variables used - Brazil

(a) 1999.04 - 2011.02



(b) 2004.01 - 2011.02

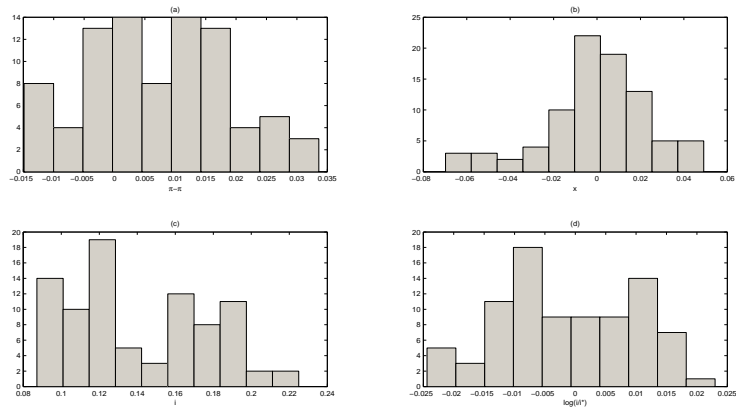


Figure 6: Historical series of the variables used - Brazil- 1999.04 - 2011.02

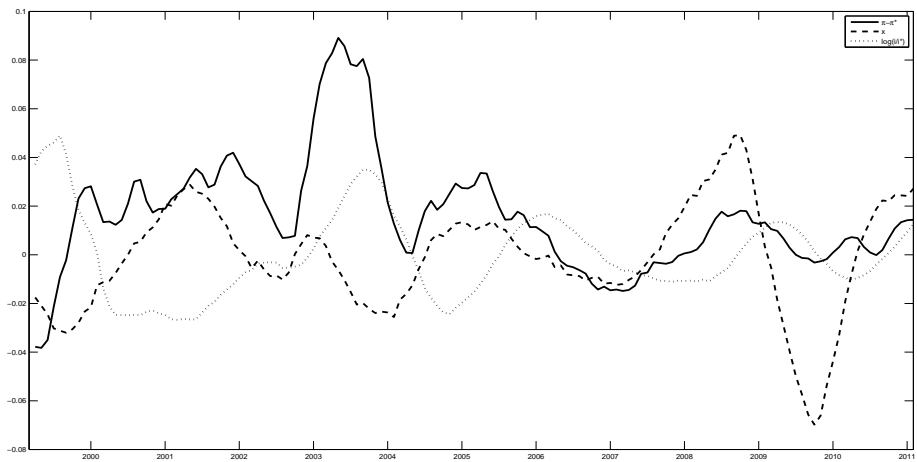
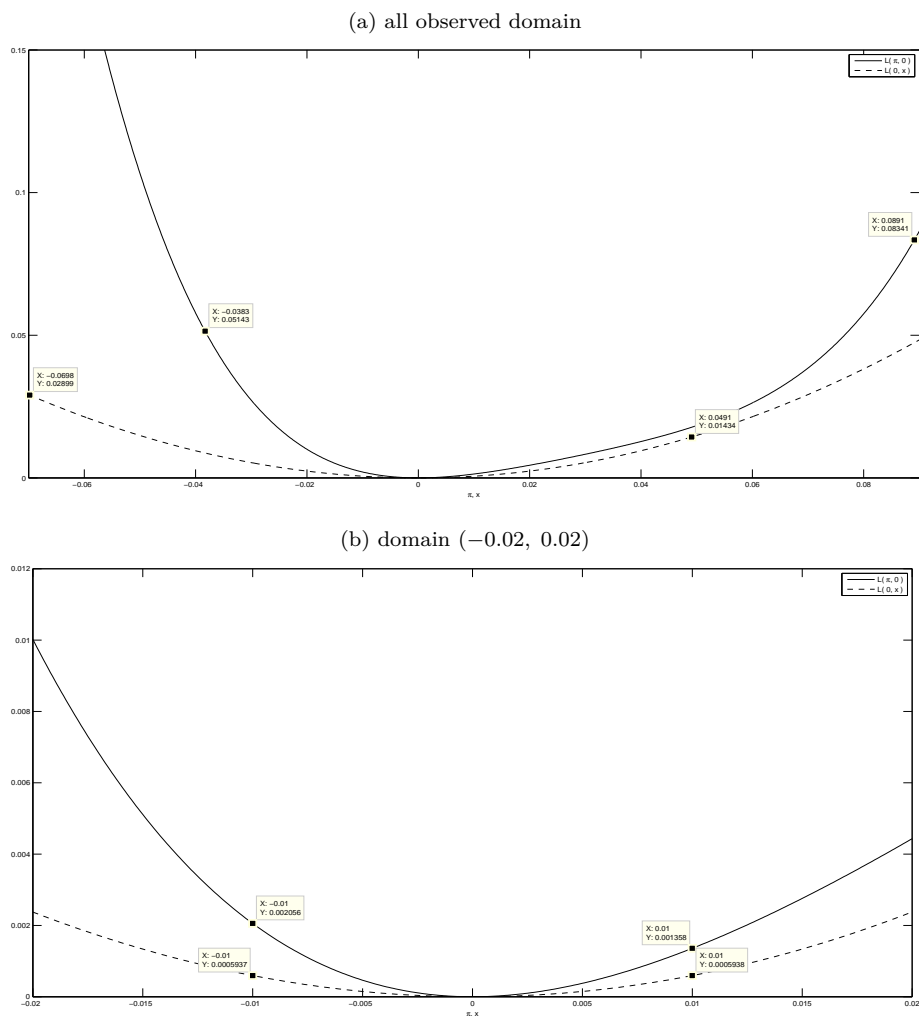


Table 3: Estimated parameters - Brazil

Parameter	1999.04-2011.02		2004.01-2011.02	
	Value	SD	Value	SD
$\psi_{1,0}$	-0.5475	R	-0.4064	R
$\psi_{2,0}$	1.0000	N	1.0000	N
$\psi_{3,0}$	-0.5529	0.2112	-0.4367	0.3154
$\psi_{4,0}$	0.4204	0.0019	0.4212	0.0008
$\psi_{5,0}$	-0.1434	0.0629	-0.1277	0.0958
$\psi_{0,1}$	0.0004	R	-0.0529	R
$\psi_{0,2}$	0.0035	0.8252	0.1280	0.1340
$\psi_{0,3}$	0.0004	0.4363	-0.0610	0.0713
$\psi_{0,4}$	0.0005	0.3490	0.0534	0.0566
$\psi_{0,5}$	0.0001	0.1297	-0.0200	0.0211
$\psi_{1,1}$	0.0044	0.0119	0.0032	0.0067
γ	0.0027	0.0045	0.0032	0.0009
β	0.9789	0.0509	0.9412	0.0230
κ	0.1494	0.1424	0.1282	0.1118
δ	0.0797	0.0365	0.1055	0.0130
ϕ	0.0335	R	0.0307	R
\bar{r}	0.0407	0.0091	0.0424	0.0122
ϱ	0.8622	0.1573	0.8756	0.1897
ω	0.9421	0.1300	0.9399	0.1794

Note: Value is the estimated value and SD is the estimated standard deviation. The standard deviation refers to bootstrapping with 500 replicates. Restricted parameters (R) are not estimated directly; i.e., a function is given by the model of the estimated parameters. N is the normalized parameter in the estimation.

Figure 7: Loss function - Brazil - 1999.04 - 2011.02



Note: $L(\pi, 0)$ shows the central bank's loss function value for several inflation values, with output gap equal to zero. Analogous to $L(0, x)$. The values displayed on the upper graph represent the maximum and minimum inflation and output gap. The lower graph shows only the interval $(-0.02, 0.02)$.

Table 4: Quadratic loss function test - Brazil

(a) 1999.04 - 2011.02			
Statistic	Estimated Value	SD	<i>t-value</i>
$\mu_{\pi} = E(\partial^3 L / \partial \pi^3)$	2.429	0.749	3.244*
$\mu_x = E(\partial^3 L / \partial x^3)$	0.000	0.949	0.000
(b) 2004.01 - 2011.02			
Statistic	Estimated Value	SD	<i>t-value</i>
$\mu_{\pi} = E(\partial^3 L / \partial \pi^3)$	3.469	0.190	18.299*
$\mu_x = E(\partial^3 L / \partial x^3)$	m.122	0.128	0.956

Note: * significant at 1%.

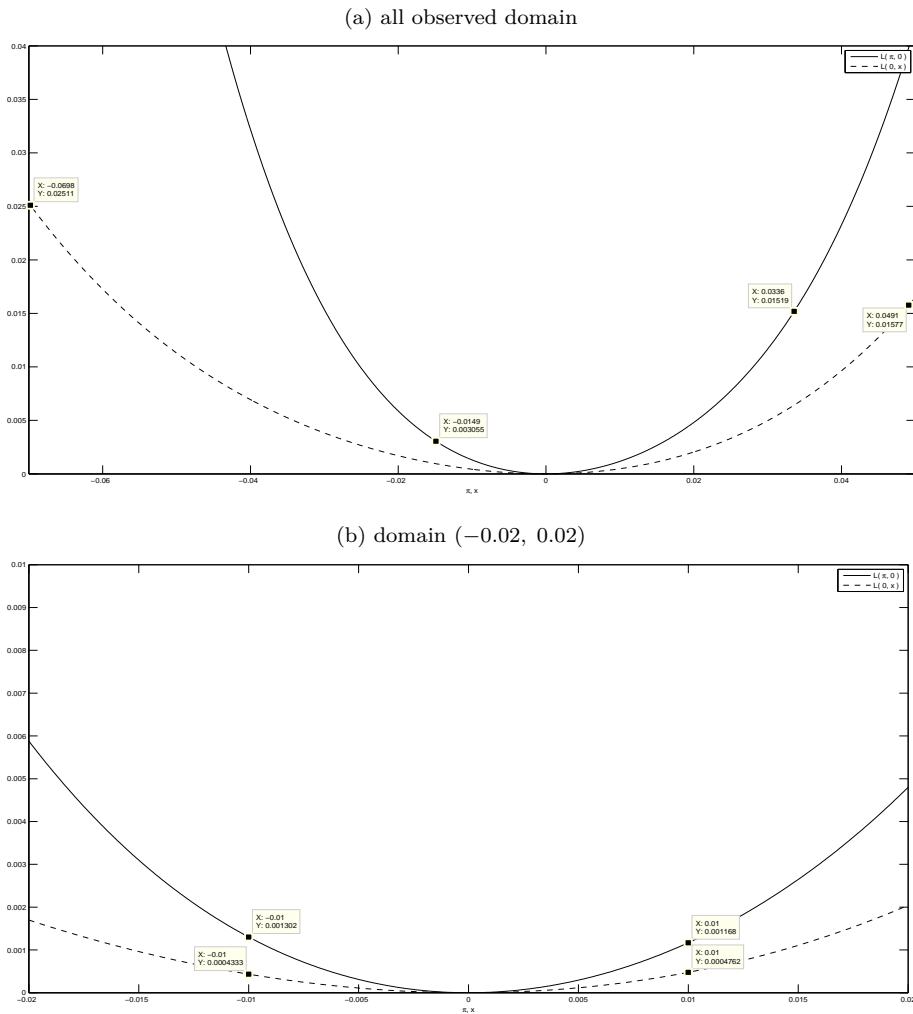
but within the interval (0, 0.06), they are relatively close. The loss associated with inflation proved to be quite asymmetric to the left, resulting in higher disutility for the monetary authority (the loss with a -1% deviation of inflation was 1.52 times greater than the loss with a 1% deviation). This lenience of the monetary authority towards positive deviations of inflation may hinder inflation control. The loss associated with the output gap was symmetric. The test for the true form of the loss function allows rejecting the quadratic form of the loss function related to inflation, confirming the asymmetry as previously highlighted, but does not allow rejecting it for the loss function associated with the output gap. The test result is shown in Table (4a).

The estimates for the structural economic parameters for the 2004-2011 subperiod were similar to those of the whole sample, as shown in Table (3). The estimated loss function, shown in Figure (8), is closer to the quadratic form, but with asymmetry to the left (a -1% deviation of inflation represented a loss 1.11 times greater for the monetary authority than the loss from a 1% deviation). Meanwhile, the loss function associated with the output gap was slightly symmetric to the right. Testing these results, again it is possible to reject that the Brazilian monetary authority's true loss function associated with inflation is quadratic, but the same does not hold for the output gap. Table (4b) summarizes these results.

In sum, it was rejected that the loss function associated with inflation is quadratic for the whole sample or for the 2004-2011 subsample in the Brazilian case. In both cases, the function was asymmetric to the left. This negative asymmetry for the whole sample had already been pointed out by Aragón and Portugal (2010), but the same did not occur with the final part of the sample.¹⁸ With respect to the loss associated with the output gap, one cannot reject that it is quadratic for the whole sample and also for the subsample, as Aragón and Portugal (2010) also observed.

¹⁸For the 2004-2007 subsample, Aragón and Portugal (2010) did not reject the hypothesis that the loss is symmetric. However, the subsample used in the present paper includes another four years.

Figure 8: Loss function - Brazil - 1999.04 - 2011.02



Note: $L(\pi, 0)$ shows the central bank's loss function value for several inflation values, with output gap equal to zero. Analogous to $L(0, x)$. The values displayed on the upper graph represent the maximum and minimum inflation and output gap. The lower graph shows only the interval $(-0.02, 0.02)$.

5 Conclusion

In this paper, we applied the nonparametric sieve estimator to determine the preferences of U.S. and Brazilian monetary policies. This allowed checking whether the central bank's loss function was asymmetric during the analyzed period. Therefore, it was possible to test whether these monetary authorities have a larger loss associated with positive or negative deviations of inflation and of the output gap.

The economy was modeled using forward-looking agents and the maximization of the monetary authority was conducted under commitment, minimizing a generic loss function. From the first-order conditions of the problem, an optimal policy rule was obtained, been a set of restrictions over the conditional moments of the economic variables. These restrictions include the finite-dimensional parameters of the structural equations of the economy and the central bank's loss function to be estimated. This way, it was possible to use the sieve method for their estimation, which informally consists in estimating the function within a less complex space than that in which it is inserted. The approximate space used here was the one generated by Chebyshev polynomials.

No significant evidence of asymmetry was found for Fed's preferences for the whole sample (1960-2004) and for the Volcker-Greenspan period. However, for both periods, the Fed was marginally more concerned with positive rather than with negative deviations associated with inflation and output gap.

On the other hand, the Brazilian central bank showed an asymmetric behavior towards inflation both for the whole sample (1999-2011) and for the 2004-2011 subperiod. And, in both samples, the loss was greater with the below-target inflation than with the above-target inflation. This is the opposite of what is expected for a more efficient inflation control, eventually leading to an inflationary bias in the Brazilian economy. These Brazilian central bank's preferences cause a sharper reduction in interest rates compared to upsurge periods.

This result encourages research into the reasons for this behavior of the Brazilian monetary authority. A hypothesis is that the independence of the Brazilian monetary authority is smaller than expected, causing the government's aversion for an economic downturn resulting from a below-target inflation to be partially transferred to the central bank as a rejection for negative deviations of inflation. Another hypothesis is that the method has problems with the estimation of the monetary authority's loss function due to the two crises (the pre-election crisis of 2002 and the international financial crisis in the late 2000s). In periods of crisis, the monetary authority uses other instruments than the interest rate to fight inflation, which is not captured directly by the structure of the economy adopted in this paper (where the central bank's instrument is nothing but the short-term interest rate), producing a downward bias towards the loss due to an above-target inflation.

A third hypothesis could be the short sampling period and, especially, the inflation rate above rather than below its target. Inflation was below its target in 23.1% of observations between 1999 and 2011 and in 31.4% between 2004 and 2011. This difference between positive and negative deviations of inflation might have resulted in poorer accuracy on the negative interval of the loss function, hindering comparisons with the positive interval. Nonetheless, one should recall that the

test presented here uses a quadratic loss function as null hypothesis, which can be rejected even if all observations happened to be positive.

This paper raises some questions that could be addressed by other studies. As to the modeling of the economy, new elements could be added, such as the exchange rate, in an attempt to explain the monetary authority's action by other means than the interest rate, as discussed for the Brazilian case. In addition, it is necessary to investigate the trajectories of the variables and the inflationary bias generated by this nonparametric approach. With regard to the sieve estimator, it would be interesting to include a Monte Carlo simulation to test its performance in finite samples. At a more advanced stage, one could try to develop a test for the functional form of the loss function whose null hypothesis includes the linex function, used in other empirical studies.

References

- AI, C., AND CHEN, X. Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71(6):1795–1843, 2003.
- ARAGÓN, E., AND PORTUGAL, M. Nonlinearities in central bank of brazil's reaction function: the case of asymmetric preferences. *Estudos Econômicos (São Paulo)*, 40(2):373–399, 2010.
- CHEN, X. Large sample sieve estimation of semi-nonparametric models. *Handbook of Econometrics*, 6:5549–5632, 2007.
- CHEN, X., AND LUDVIGSON, S. Land of addicts? an empirical investigation of habit-based asset pricing models. *Journal of Applied Econometrics*, 24(7):1057–1093, 2009.
- CHEN, X., AND SHEN, X. Sieve extremum estimates for weakly dependent data. *Econometrica*, 66(2):289–314, 1998.
- CHIARELLA, C., FLASCHEL, P., GONG, G., AND SEMMLER, W. Nonlinear phillips curves, complex dynamics and monetary policy in a keynesian macro model. *Chaos, Solitons & Fractals*, 18(3):613–634, 2003.
- CLARIDA, R., GALÍ, J., AND GERTLER, M. The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*, 37:1661–1707, 1999.
- CLARIDA, R., GALI, J., AND GERTLER, M. Monetary policy rules and macroeconomic stability: Evidence and some theory. *quarterly Journal of economics*, 115(1):147–180, 2000.
- CUKIERMAN, A., MUSCATELLI, V., AND OF GLASGOW. DEPT. OF ECONOMICS, U. *Do Central Banks Have Precautionary Demands for Expansions and for Price Stability?: Theory and Evidence*. CESifo, 2002.
- DOLADO, J., MARIA-DOLORES, R., AND NAVEIRA, M. Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the phillips curve. *European Economic Review*, 49(2): 485–503, 2005.

- FAN, J., AND YAO, Q. *Nonlinear time series: nonparametric and parametric methods*. Springer Verlag, 2005.
- GRENANDER, U. *Abstract inference*. 1981.
- HAMILTON, J. A parametric approach to flexible nonlinear inference. *Econometrica*, 69(3):537–573, 2001.
- ISNARD, C. *Introdução á Medida e Integração*. IMPA, 2007.
- JUDD, K. *Numerical methods in economics*. The MIT press, 1998.
- KIM, D. H., OSBORN, D. R., AND SENSIER, M. Nonlinearity in the fed’s monetary policy rule. *Journal of Applied Econometrics*, 20(5):621–639, 2005.
- LJUNGQVIST, L., AND SARGENT, T. *Recursive macroeconomic theory*. The MIT press, 2004.
- NEWKEY, W. Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics*, 79(1):147–168, 1997.
- NEWKEY, W., AND POWELL, J. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.
- NOBAY, A. R., AND PEEL, D. A. Optimal discretionary monetary policy in a model of asymmetric central bank preferences. *The Economic Journal*, 113(489):pp. 657–665, 2003.
- NOBAY, A., AND PEEL, D. Optimal monetary policy with a nonlinear phillips curve. *Economics Letters*, 67(2):159–164, 2000.
- ORPHANIDES, A., AND WIELAND, V. Inflation zone targeting. *European Economic Review*, 44(7):1351–1387, 2000.
- SCHALING, E. The nonlinear phillips curve and inflation forecast targeting: Symmetric versus asymmetric monetary policy rules. *Journal of Money, Credit, and Banking*, 36(3):361–386, 2004.
- SURICO, P. The fed’s monetary policy rule and u.s. inflation: The case of asymmetric preferences. *Journal of Economic Dynamics and Control*, 31(1):305 – 324, 2007.
- SVENSSON, L., AND WOODFORD, M. Implementing Optimal Policy through Inflation-Forecast Targeting. *The inflation-targeting debate*, page 19, 2005.
- TAYLOR, J. Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy*, volume 39, pages 195–214. Elsevier, 1993.
- WOODFORD, M. *Interest and prices: Foundations of a theory of monetary policy*. Princeton Univ Pr, 2003.

A Appendix

A.1 Consistency of the estimator

This section is based on the approach proposed by Chen and Ludvigson (2009) to establish the consistency of the estimator $\hat{\alpha}$ when the data are not independent. The authors do that by applying the A1 lemma of Newey and Powell (2003), which is described in what follows.

Lemma 1. *Suppose that (i) $Q(\alpha)$ has a single minimum in \mathcal{A} , located in α_0 ; (ii) $\hat{Q}(\alpha)$ and $Q(\alpha)$ are continuous, \mathcal{A} is compact and $\max_{\alpha \in \mathcal{A}} |\hat{Q}(\alpha) - Q(\alpha)| \xrightarrow{p} 0$; (iii) $\hat{\mathcal{A}}$ are compact subsets of \mathcal{A} such that for any $\alpha \in \mathcal{A}$ there is $\tilde{\alpha} \in \hat{\mathcal{A}}$ such that $\tilde{\alpha} \xrightarrow{p} \alpha$. Then $\hat{\alpha} = \arg \min_{\alpha \in \mathcal{A}} \hat{Q}(\alpha) \xrightarrow{p} \alpha_0$.*

Proof. See Newey and Powell (2003, p. 1576). □

Hypothesis (i) is concerned with the identifiability of the model. As to the continuity of the criterion functions of the sample and population required in (ii), it can be seen from equations (13) and (14) (recalling that the basis with which the loss function is approximated is continuous). The compactness of the finite-dimensional parameters is reasonable since it makes no economic sense if either it diverges to 1 or -1 and the compactness of the central bank's loss function is assumed as hypothesis (4). Hypothesis (iii) posits that the approximate space (sieve space) used is dense in the space where the true function to be estimated is inserted. The Stone-Weierstrass theorem shows that the space of all polynomials is dense in \mathcal{L}^2 . Following Chen and Ludvigson (2009), demonstrating the uniform convergence required in (iii) is the same as showing that equations

$$\sup_{\mathbf{X}_t, \alpha} |\hat{m}(\mathbf{X}_t, \alpha) - m(\mathbf{X}_t, \alpha)| = o_p(1) \quad (21)$$

and

$$\sup_{\mathcal{A} \times \Theta} \left| N^{-1} \sum_{t=1}^N m(\mathbf{X}_t, \alpha)' m(\mathbf{X}_t, \alpha) - E \left(m(\mathbf{X}_t, \alpha)' m(\mathbf{X}_t, \alpha) \right) \right| = o_p(1). \quad (22)$$

are valid.

O resultado (21) é obtido através do lema A1 de Ai and Chen (2003) modificado para permitir dados β -mixing estacionários. Já o resultado (22) pode ser obtido através de qualquer lei dos grandes números para dados β -mixing estacionários. Com as condições atendidas, tem-se $\|\alpha_N - \alpha_0\|_s = o_p(1)$. Chen and Ludvigson (2009, apêndice técnico, p. 7)

Result (21) is obtained by the A1 lemma of Ai and Chen (2003) which was modified to allow for stationary β -mixing data. Result (22) can be obtained by any law of large numbers for stationary β -mixing data. After the conditions have been met, one have $\|\alpha_N - \alpha_0\|_s = o_p(1)$. Chen and Ludvigson (2009, technical appendix, p. 7)