Testing for the degree of commitment via set-identification

Laura Coroneo[*], Valentina Corradi and Paulo Santos Monteiro[†]

June 7, 2011

Abstract

We propose a new procedure for testing optimal monetary policy relying on set-identification. The identified set is characterized by means of moment inequality conditions nesting optimal policy under commitment and discretion. The approach is based on the derivation of bounds for inflation that are consistent with both forms of optimal policy and provide set identification of the economy's structural parameters. We derive testable implications that allow for specification tests and discrimination between the two alternative policy regimes. The methodology allows us to study the level of commitment exhibited by the United States monetary authority, and to set-identify the economy's structural parameters such that the target variables of the monetary authority (inflation and welfare relevant output gap) are consistent with optimal monetary policy.

JEL classification codes: E52; E58; C12.
Keywords: *Optimal Monetary Policy; Partial Identification.*

[*]University of Manchester, Economics - School of Social Sciences
[†]University of Warwick, Department of Economics.
We thank James Banks, Herman Bierens, Jean-Marie Dufour, Alastair Hall, Lynda Khalaf, Maral Kichian, Tatiana Kirsanova, Micheal McMahon, Sushama Murty, Carlos Noton, Denise Osborn, Adam Rosen. We also thank seminar participants at the University of Chile, the Central Bank of Chile, the University of Exeter, the University of Manchester, the University of Warwick, the 2009 Brussels Waseda Seminar at ECARES, the 2010 $(EC)^2$ conference at Toulouse School of Economics, the 2009 Econometric Society European Meeting at Universitat Autonoma de Barcelona and the 2009 North American Summer Meetings of the Econometric Society at Boston University, 2011 Time Series Conference at CIREQ-University of Montreal.

1 Introduction

This paper derives new results regarding the structural evaluation of monetary policy in the framework set by the New Keynesian model. Since the work of Kydland and Prescott (1977), the theory of optimal monetary policy is aware of the time inconsistency problem. An optimal state contingent plan announced ex-ante by the monetary authority fails to steer private sector expectations because, ex-post, past commitments are ignored. The theoretical literature has considered two alternative characterizations of optimal monetary policy: the commitment solution, whereby the optimal plan is history-dependent and the time-inconsistency problem is ignored; and the discretion solution, whereby the optimal policy is markov-perfect and the monetary authority re-optimizes each period. We describe a method for estimating and testing a model of optimal monetary policy without requiring an explicit choice of the relevant equilibrium concept. Our procedure considers a general specification of optimal monetary policy, nesting the commitment and the markov-perfect characterizations of optimal policy. The approach is based on the derivation of bounds for inflation that are consistent with both forms of optimal policy and yield set identification of the economy's structural parameters. We derive testable implications that allow for specification tests and discrimination between the monetary authority's modes of behavior.

In a discretionary regime the inflation rate on average exceeds the level that would be optimal if commitment to history dependent policy rules was feasible. This is the celebrated Barro and Gordon (1983) inflationary bias result, arising in the presence of distortions that imply a suboptimal natural output rate. In addition to this deterministic inflation-bias, under discretion there is a state-contingent inflation bias resulting from the fact that the monetary authority sets policy independently of the history of shocks. The upshot of this state-contingent bias is that when the output gap is negative, the inflation rate under discretion in the following period is higher than what it would be if the monetary authority was able to commit to history-dependent plans. This state-contingent inflationary bias allows for the derivation of an inflation lower-bound (obtained under commitment) and an upper-bound (obtained under discretion) based on the first order conditions that characterize optimal monetary policy under each policy regime. Our framework relies on these inflation bounds, which can be used to derive moment inequality conditions associated with optimal monetary policy, and identify the set of structural parameters for which the moment inequalities hold, i.e. the identified set. We estimate the identified set implied by optimal monetary policy and construct confidence regions that cover the identified set with a pre-specified probability using inference methods developed in Chernozhukov, Hong, and Tamer (2007) (CHT).

If the identified set is nonempty, we test whether the moment restrictions implied by a specific policy regime are satisfied. Assuming either discretion or commitment allows for point identification of the underlying structural parameters. Hence, parameters can be consistently estimated and it is possible to perform standard tests for overidentifying restrictions (Hansen, 1982). However, if our objective is to test for discretion or commitment under the maintained assumption of optimal monetary policy, the standard Hansen's J-test does not make use of all the available information. Instead, we propose a test for discretion and a test for commitment based on the null hypothesis that the structural parameters are contained in the identified set under discretion and commitment respectively. Thus, our approach explores the additional information obtained from the moment inequality conditions associated with the inflation bounds implied by optimal monetary policy.

Formally, the test is implemented using the criterion function approach of CHT and an extension of the Generalized Moment Selection method of Andrews and Soares (2010) that takes into account the contribution of parameter estimation error on the relevant covariance matrix.

We apply our testing procedure to investigate whether the time series of inflation and output gap in the United States are consistent with the New Keynesian model of optimal monetary policy that has been widely used in recent studies of monetary policy, following Rotemberg and Woodford (1997), Clarida, Galí and Gertler (1999), and Woodford (2003). The paper establishes the following three results: (i) the estimated identified set is nonempty and, therefore, the time series of inflation and output gap are consistent with optimal monetary policy; (ii) the estimated identified set under discretion is nonempty and contains the parameter vector estimated by GMM under discretion, providing evidence in favor of discretion; (iii) the estimated identified set under commitment is degenerate, allowing to formally reject the null hypothesis of commitment.

Moreover, our framework provides an estimated identified set for the parameters of the first order conditions describing the joint behavior of the monetary authority's target variables (inflation and welfare relevant output gap) under a very general characterization of optimal monetary policy. In particular, the estimated identified set is valid under a specification that nests a continuum of monetary policy rules characterized by differing degrees of commitment, and in which full commitment and discretion are the two extreme cases. This is the quasi-commitment model proposed by Schaumburg and Tambalotti (2007). Under quasi-commitment, the monetary authority deviates from commitment-based optimal plans with a fixed, exogenous probability, known to the public. The inflation bounds that we derive are compatible with periodic switches between commitment and discretion. Therefore, the moment inequality conditions used for partial identification of the economy's structural parameters are valid also under quasi-commitment.

The importance of being able to discriminate between different policy regimes on the basis of the observed time series of inflation and output is well recognized. In an early contribution, Baxter (1988) calls for the development of methods to analyze policy making in a maximizing framework, and suggests that *"what is required is the derivation of appropriate econometric specifications for the models, and the use of established statistical procedures for choosing between alternative, hypothesized models of policymaking"*.[1] This paper seeks to provide such an econometric specification. Our paper is also related to work by Ireland (1999), that tests and fails to reject the hypothesis that inflation and unemployment form a cointegrating relation, as implied by the Barro and Gordon model when the natural unemployment rate is non-stationary. Ruge-Murcia (2003) estimates a model that allows for asymmetric preferences, nesting the Barro and Gordon specification as a special case, and fails to reject the model of discretionary optimal monetary policy. Both these papers assume one equilibrium concept (discretion), and test whether some time series implications of discretionary policies, are rejected or not by the data. Our framework instead derives a general specification of optimal monetary policy, nesting the commitment and the discretion solutions as two special cases.

Using a full-information maximum-likelihood approach, Givens (2010) estimates a New Keynesian model for the US economy in which the monetary authority conducts optimal monetary policy. The model is estimated separately under the two alternatives of commitment and discretion, using quarterly data over the Volcker–Greenspan–Bernanke era; a comparison of the log-likelihood of the

[1] Baxter, 1988 (p.145).

two alternative models based on a Bayesian information criterion (to overcome the fact that the two models are non-nested) strongly favors discretion over commitment. A similar Bayesian approach has been used by Kirsanova and le Roux (2011), who also find evidence in favor of discretion for both monetary and fiscal policy in the UK. The partial identification framework we propose in this paper permits, instead, a general econometric specification that nests commitment and discretion as two special cases. Unlike full-information methods, our approach does not require a complete representation of the economy, nor strong assumptions about the nature of the forcing variables.

Simple monetary policy rules are often prescribed as useful guides for the conduct of monetary policy. For instance, a commitment to a Taylor rule (after Taylor, 1993)—according to which the short-term policy rate responds to fluctuations in inflation and some measure of the output gap—incorporates several features of an optimal monetary policy, from the standpoint of at least one simple class of optimizing models. Woodford (2001) shows that the response prescribed by these rules to fluctuations in inflation or the output gap tends to stabilize those variables, and stabilization of both variables is an appropriate goal, as long as the output gap is properly defined. Furthermore, the prescribed response to these variables guarantees determinate rational expectations equilibrium, and so prevents instability due to self-fulfilling expectations. Under certain simple conditions, a feedback rule that establishes a time-invariant relation between the path of inflation and of the output gap and the level of nominal interest rates can bring about an optimal pattern of equilibrium responses to real disturbances. Woodford and Gianonni (2010) show that it is possible to find simple target criteria that are fully optimal across a wide range of specifications of the economy stochastic disturbance processes. To the extent that the systematic behavior implied by simple rules takes into account private sector expectations, commitment-like behavior may be a good representation of monetary policy. Therefore, as Mcallumn (1999) forcefully argues, neither of the two modes of central bank behavior has as yet been established as empirically relevant. Our framework develops a new test procedure for null hypotheses concerning these two alternative policy regimes.

This paper also contributes to a growing literature proposing partial identification methods to overcome lack of information about the economic environment. For instance, Manski and Tamer (2002) examine inference on regressions with interval outcomes. In the industrial organization literature, Haile and Tamer (2003) use partial identification to construct bounds on valuation distributions in second price auctions. Blundell, Browning and Crawford (2008) derive bounds that allow to set-identify predicted demand responses in the study of consumer behavior. Ciliberto and Tamer (2009) study inference in entry games without requiring equilibrium selection assumptions. Galichon and Henry (2010) derive set-identifying restrictions for games with multiple equilibria in pure and mixed strategies.

The rest of the paper is organized as follows. Section 2 describes the theoretical model for the economy. Section 3 characterizes optimal monetary policy. Section 4 derives the bounds for inflation implied by the structural model of optimal monetary policy. Section 5 outlines the inference procedure. Section 6 describes how to test optimal monetary policy under discretion and under commitment. Finally, Section 7 reports the empirical findings, and Section 8 concludes. Appendix A cointains details about the theoretical model and Appendix B collects all proofs.

2 The Structural Model

The framework is that of the New Keynesian forward-looking model with monopolistic competition and Calvo price-setting exposed in Woodford (2010). The representative household seeks to maximize the following utility function

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ u\left(C_t; \xi_t\right) - \int_0^1 v\left(H_t\left(j\right); \xi_t\right) dj \right\},$$

where β is a discount factor, $H_t(j)$ is the quantity of labor of type j supplied, and ξ_t is a vector of exogenous disturbances that includes shocks to preferences; for each value of ξ_t, u is an increasing, concave function, and v is an increasing, convex function. The argument C_t is a Dixit and Stiglitz (1977) index of purchases of all the differentiated commodities

$$C_t = \left[\int_0^1 c_t\left(i\right)^{(\vartheta-1)/\vartheta} di \right]^{\vartheta/(\vartheta-1)}$$

where $c_t(i)$ is the quantity purchased of commodity i. The parameter $\vartheta > 1$ is the elasticity of substitution between goods. Each differentiated commodity i is supplied by a single monopolistically competitive producer. In each industry j, there are assumed to be many commodities. An industry j is a set of producers that use the same type of labor and always change their price contemporaneously. Thus, all producers from industry j produce the same quantity y_t^j. The representative household supplies all types of labor and consumes all varieties of goods. The optimal supply of labor of type j by the representative household is such that the following condition is satisfied

$$\frac{v_h\left(H_t\left(j\right); \xi_t\right)}{u_c\left(C_t; \xi_t\right)} \chi_t = w_t\left(j\right), \tag{1}$$

where $w_t(j)$ is the real wage of labor of type j in period t and $\chi_t \geq 1$ is a time-varying labor wedge, common to all labor markets and capturing the effect of taxes and labor market imperfections.

The aggregate resource constraint is given by

$$C_t + G_t \leq \left[\int_0^1 y_t\left(i\right)^{(\vartheta-1)/\vartheta} di \right]^{\vartheta/(\vartheta-1)} \equiv Y_t \tag{2}$$

where $y_t(i)$ is the quantity produced of commodity i and G_t (which is included in the vector of exogenous disturbances ξ_t) represents the government's expenditure on an exogenously given quantity of the same basket of commodities that is purchased by the households. In equilibrium condition (2) holds with equality.

Except for the fact that labor is immobile across industries, there is a common technology for the production of all goods: the commodity i is produced according to the production function

$$y_t\left(i\right) = f\left(h_t\left(i\right)\right),$$

where $h_t(i)$ is the quantity of labor employed by producer i and f is an increasing and concave function. Thus, for each producer i the relationship between the real marginal cost $s_t(i)$ and the quantity supplied is given by

$$s_t(i) = s\left(y_t(i), Y_t; \tilde{\xi}_t\right),$$

where the real marginal cost function is defined by

$$s\left(y, Y; \tilde{\xi}\right) = \chi \frac{v_h\left(f^{-1}(y); \xi\right)}{u_c(Y - G; \xi)} \left[f'\left(f^{-1}(y)\right)\right]^{-1}, \tag{3}$$

and $\tilde{\xi}_t$ augments the vector ξ_t with the the labor wedge χ_t. It follows that, if prices were fully flexible, each supplier would charge a relative price satisfying

$$\frac{p_t(i)}{P_t} = \mu s\left(y_t(i), Y_t; \tilde{\xi}_t\right),$$

where the aggregate price index P_t is defined by

$$P_t = \left[\int_0^1 p_t(i)^{1-\vartheta} di\right]^{1/(1-\vartheta)},$$

and $\mu = \vartheta(\vartheta - 1)^{-1} > 1$, is the producers' markup. Moreover, in the flexible price equilibrium all producers charge the same price and produce the same quantity Y_t^n, so that

$$\frac{1}{\mu} = s\left(Y_t^n, Y_t^n; \tilde{\xi}_t\right),$$

where Y_t^n is the natural rate of output, which corresponds to the equilibrium level of output under flexible prices. Nonetheless, it is assumed prices are sticky, so that the output of each commodity i differs. A log-linear approximation to the marginal cost function (3) around the deterministic equilibrium under flexible prices yields the condition

$$\widehat{s}_t(i) = \omega \widehat{y}_t(i) + \sigma^{-1}\widehat{Y}_t - \left(\omega + \sigma^{-1}\right)\widehat{Y}_t^n.$$

where $\widehat{X}$ denotes the log deviation from steady state of X; $\omega > 0$ is the elasticity of a firm's real marginal cost with respect to its own output level, and $\sigma > 0$ is the intertemporal elasticity of substitution of private expenditure. Averaging this condition over all goods i yields

$$\widehat{s}_t = \left(\omega + \sigma^{-1}\right)\left(\widehat{Y}_t - \widehat{Y}_t^n\right),$$

where $\widehat{s}_t$ is the logarithm of the average real marginal cost. Thus the welfare relevant measure of the output gap is related to variations in the economy-wide real marginal cost.

Definition 1 *The welfare relevant output gap $x_t \equiv \left(\widehat{Y}_t - \widehat{Y}_t^n\right)$ is the difference between the log of output and the log of natural output (the level of output that would arise if prices were fully flexible). The output gap x_t is proportional to the average real marginal cost in percentage deviation from steady state, and is given by $x_t = \left(\omega + \sigma^{-1}\right)^{-1}\widehat{s}_t$.*

To model the behavior of prices, we employ the probabilistic model due to Calvo (1983).[2] In any period each industry j has a fixed probability $1-\alpha$ that it may adjust its price during that period. Else, with probability α, the producers in that industry must keep their price unchanged. The resulting behavior of the price level is represented by the following two equations

$$\log P_t = \alpha \log P_{t-1} + (1-\alpha) \log p_t^\star \tag{4}$$

and

$$\sum_{s=t}^{\infty} (\alpha\beta)^{s-t} E_t \left[\log p_t^\star - \log P_s - \zeta x_s\right] = 0 \tag{5}$$

where $\zeta = \left(\omega + \sigma^{-1}\right) / (1+\omega\vartheta)$ and $p_t^\star$ is the price chosen by the reoptimizing firms in period t. Combining equations (4) and (5) yields the New Keynesian Phillips Curve (NKPC), given by

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t, \tag{6}$$

where $\kappa = (1-\alpha)(1-\alpha\beta)(\zeta/\alpha)$.

3 Optimal Monetary Policy

The efficient level of output satisfies the condition

$$s\left(Y_t^\star, Y_t^\star, \tilde{\xi}_t\right) = \chi_t, \tag{7}$$

and corresponds to the level of output under flexible prices and without distortions resulting from firm's market power and the labor wedge. Thus, from equations (2) and (7), we derive the following relationship

$$\log Y_t^n - \log Y_t^\star = \left(\omega + \sigma^{-1}\right)^{-1} \delta_t$$

where $\delta_t = -\log(\mu\chi_t) < 0$ is an exogenous stochastic shock resulting from time-varying distortions. If the distortions δ_t are small, the discounted sum of utility of the representative agent involving small fluctuations around the steady state can be approximated by a second-order Taylor expansion around the stable equilibrium associated with zero inflation, as follows

$$\mathcal{W} = E_0 \sum_{t=0}^{\infty} -\frac{\beta^t}{2} \left\{\pi_t^2 + \frac{\kappa}{\vartheta}\left[x_t + \left(\omega + \sigma^{-1}\right)^{-1} \delta_t\right]^2\right\}. \tag{8}$$

The derivation of (8) is given in Woodford (2003, chapter 6).

The model of optimal monetary policy under commitment is based on the assumption that the monetary authority maximizes (8) subject to the constraint imposed by the Phillips curve equation (6). If the monetary authority is able to commit to a state contingent path for inflation and the output gap, the conditions solving the monetary authority's problem at some given period s

[2] See Appendix A for a complete description of the model.

are

$$\begin{aligned}
\frac{\kappa}{\vartheta}\left[x_t + \left(\omega+\sigma^{-1}\right)^{-1}\delta_t\right] + \kappa\lambda_t &= 0, \qquad t = s, s+1, \ldots \\
\pi_t + \lambda_{t-1} - \lambda_t &= 0, \qquad t = s, s+1 \ldots \\
\beta E_t \pi_{t+1} &= \pi_t - \kappa x_t,
\end{aligned}$$

where λ_t is the Lagrangian multiplier associated with equation (6). The resulting joint path for inflation and output gap, assuming that the system has been initialized in period $s = 0$ and that $\lambda_{-1} = 0$, is given by

$$\pi_t = -\frac{1}{\vartheta}\left[x_t + \left(\omega+\sigma^{-1}\right)^{-1}\delta_t\right] + \frac{1}{\vartheta}\left[x_{t-1} + \left(\omega+\sigma^{-1}\right)^{-1}\delta_{t-1}\right]. \tag{9}$$

However, the commitment solution is time inconsistent in the Kydland and Prescott (1977) sense: each period t, the monetary authority is tempted to behave as if $\lambda_{t-1} = 0$, ignoring the impact of its current actions on the private sector expectations. When the monetary authority lacks a commitment technology, it must set policy sequentially. Under discretion, optimal monetary policy satisfies the markov property in the sense that the policy is chosen independently of past choices of the monetary authority. Thus, the policymaker acts as if $\lambda_{t-1} = 0$ and the resulting joint path for inflation and output gap is

$$\pi_t = -\frac{1}{\vartheta}\left[x_t + \left(\omega+\sigma^{-1}\right)^{-1}\delta_t\right] \tag{10}$$

There are several challenges to the empirical estimation of the parameters of the structural model of optimal monetary policy. First of all, the welfare relevant output gap x_t is not directly observable and there is no reason to believe that it can be proxied by the deviation of output from a smooth statistical trend. Sbordone (2002) and Gali and Gertler (1999) notice that the most direct way to measure time variation in the output gap is based on the variation in the production costs. This is indeed the case in our framework, as it can be seen from Definition 1. It follows that equations (9) and (10) can be expressed in terms of the economy-wide marginal cost, $\widehat{s}_t$. We define $\pi_t^{(c)}$, the inflation in period t under commitment, and $\pi_t^{(d)}$, the inflation in period t under discretion, as follows

$$\pi_t^{(c)} = -\phi^{-1}\left(\widehat{s}_t + \delta_t\right) + \phi^{-1}\left(\widehat{s}_{t-1} + \delta_{t-1}\right), \tag{11}$$

$$\pi_t^{(d)} = -\phi^{-1}\left(\widehat{s}_t + \delta_t\right), \tag{12}$$

where $\phi = \left(\omega+\sigma^{-1}\right)\vartheta$. The main empirical challenge is that modeling optimal monetary policy requires a decision about whether the monetary authority is endowed with a commitment technology or, instead, engages in discretionary monetary policy.

How does one decide whether the behavior of the monetary authority should be classified as discretionary or commitment-like? We propose a general characterization of optimal monetary policy that nests the commitment and the markov-perfect characterizations of optimal policy. This approach is based on the derivation of bounds for the inflation rate under the maintained assumption that the monetary authority implements optimal monetary policy.

4 Bounds for Inflation

Under a specific equilibrium concept, commitment or discretion, it is in principle possible to identify ϕ from observed data for inflation and output gap using equations (11) and (12). Thus, lack of knowledge about the equilibrium concept is what prevents exact identification. A general specification for optimal monetary policy, nesting the two alternative characterizations of optimality follows from the next simple result

Lemma 1 *Let $\delta_t \leq 0$ for all t almost surely. It follows that,* $\Pr\left(\pi_t^{(d)} \geq \pi_t^{(c)} \middle| \widehat{s}_{t-1} \leq 0\right) = 1$*, and optimal monetary policy implies that* $\Pr\left(\pi_t^{(c)} \leq \pi_t \leq \pi_t^{(d)} \middle| \widehat{s}_{t-1} \leq 0\right) = 1$.

The bounds for inflation in Lemma 1 are derived from equations (11) and (12). Since $\delta_t \leq 0$ for all t, it follows that

$$\Pr\left(\widehat{s}_{t-1} + \delta_{t-1} \leq 0 \middle| \widehat{s}_{t-1} \leq 0\right) = 1,$$

which implies that $\pi_t^{(d)} \geq \pi_t^{(c)}$ *almost surely*. The upshot is that we are able to find moment inequalities consistent with optimal monetary policy.

In the sequel we assume that the observed inflation rate differs from the actual inflation rate chosen by the monetary authority only through the presence of a zero mean measurement error.

Assumption 1 *Let π_t be the actual inflation rate in period t. The observed inflation rate, Π_t, is given by the sum of the actual rate of inflation and a zero mean exogenous measurement error v_t with finite variance.*

Notice that Assumption 1 does not require the measurement error to be independently distributed, on the contrary it can exhibit some time dependence. From Lemma 1 and Assumption 1, it follows that

$$\Pr\left(\pi_t^{(c)} + v_t \leq \Pi_t \leq \pi_t^{(d)} + v_t \middle| \widehat{s}_{t-1} \leq 0\right) = 1, \tag{13}$$

which establishes a lower bound and an upper bound for the observed inflation rate, Π_t. Thus we can establish the following result

Lemma 2 *Let δ_t a random disturbance with support in $\mathbb{R}^-$. Under Assumption 1, it follows that*

$$\Pr\left(-\phi^{-1}\left(\Delta\widehat{s}_t + \Delta\delta_t\right) + v_t \leq \Pi_t \leq -\phi^{-1}\left(\widehat{s}_t + \delta_t\right) + v_t \middle| \widehat{s}_{t-1} \leq 0\right) = 1.$$

Denoting by $1\left(\widehat{s}_{t-1} \leq 0\right)$ an indicator function taking value 1 if $\widehat{s}_{t-1} \leq 0$ and using Lemma 2 we derive moment inequalities that are implied by optimal monetary policy, as follows

Proposition 1 *Let $\phi \equiv \left(\omega + \sigma^{-1}\right)\vartheta$ be a striclty positive parameter. Under Assumption 1, the following moment inequalities*

$$-\operatorname{E}\left[\left(\phi\Pi_t + \widehat{s}_t + \delta_t - \phi v_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right)\right] \geq 0, \tag{14}$$

$$\operatorname{E}\left[\left(\phi\Pi_t + \Delta\widehat{s}_t + \Delta\delta_t - \phi v_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right)\right] \geq 0, \tag{15}$$

are consistent with optimal monetary policy under commitment or discretion.

Proposition 1 follows immediately from Lemma 2 noticing that the bounds on inflation are valid any time $\widehat{s}_{t-1} \leq 0$ and, therefore, they also hold when multiplied by $1(\widehat{s}_{t-1} \leq 0)$. The only variables which are not observed by the econometrician are the measurement error v_t and the exogenous disturbance δ_t. We define the following set of instruments

Assumption 2 *Let Z_t denote a p-dimensional vector of instruments, $p > 2$ such that*

1. *Z_t is strictly positive;*
2. $\mathrm{E}\left[v_t 1(\widehat{s}_{t-1} \leq 0)\ Z_t\right] = 0$*;*
3. $\mathrm{E}\left[\delta_{t-r} 1(\widehat{s}_{t-1} \leq 0) Z_t\right] = \overline{\delta}_r \mathrm{E}\left[Z_t\right]$*, for $r = 0, 1$, and where $\overline{\delta}_r \equiv \mathrm{E}\left[\delta_{t-r} 1(\widehat{s}_{t-1} \leq 0)\right] < 0$;*
4. $\mathrm{E}\left[\Pi_t Z_t\right] \neq 0$*,* $\mathrm{E}\left[\widehat{s}_t 1(\widehat{s}_{t-1} \leq 0) Z_t\right] \neq 0$ *and* $\mathrm{E}\left[\Delta\widehat{s}_t 1(\widehat{s}_{t-1} \leq 0)_t Z_t\right] \neq 0$*.*

The instrumental variables are assumed to be uncorrelated with the unobserved disturbance δ_{t-r} and with $\delta_{t-r} 1(\widehat{s}_{t-1} \leq 0)$ for $r = 0,\ 1$. In particular, Assumption 2.3 implies that

$$\mathrm{E}\left[(\Delta\delta_t)\ 1(\widehat{s}_t \leq 0) Z_t\right] = \mathrm{E}\left[(\Delta\delta_t)\ 1(\widehat{s}_t \leq 0)\right] \mathrm{E}\left[Z_t\right].$$

In what follows, we assume that $\overline{\delta}_1 \approx \overline{\delta}_2 \approx \overline{\delta}$, so that the term $E\left[(\Delta\delta_t)\ 1(\widehat{s}_t \leq 0)\right] \approx 0$. This approximation is accurate provided that either δ_t or $\widehat{s}_t$ are sufficiently persistent.[3] Finally, Assumption 2.4 requires that the instruments are not weak. Given Assumption 2, the moment inequalities in Proposition 1 can be written as

$$-\mathrm{E}\Big[\left((\phi\Pi_t + \widehat{s}_t)\ 1(\widehat{s}_{t-1} \leq 0) + \overline{\delta}\right) Z_t\Big] \quad \geq \quad 0, \tag{16}$$

$$\mathrm{E}\Big[(\phi\Pi_t + \Delta\widehat{s}_t)\ 1(\widehat{s}_{t-1} \leq 0) Z_t\Big] \quad \geq \quad 0. \tag{17}$$

Notice that $\overline{\delta}$ is a nuisance parameter to be estimated along with ϕ, the structural parameter of interest.

Definition 2 *Let $\theta = \left(\phi,\ \overline{\delta}\right) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^-$. The identified set under optimal monetary policy is given by*

$$\Theta^I = \{\theta \in \Theta : \ \textit{inequalities } (16) \textit{ and } (17) \textit{ hold}\}.$$

The moment conditions in (16)–(17) are linear in the parameters. Thus, as shown by Bontemps, Magnac and Maurin (2007), the identified set is closed, convex and, under mild requirements, bounded. Further, if the number of moment conditions is equal to the number of parameters, then the identified set is trivially nonempty. This is not our case, as we have $2p$ moment conditions, with $p > 2$. Note that the set of values ϕ for which the inequality condition (16) is satisfied increases

[3] If we do not make the approximation $E\left[(\Delta\delta_t)\ 1(\widehat{s}_t \leq 0)\right] \approx 0$, the problem is very similar but there is an additional nuisance parameter to be estimated, corresponding to $\overline{\delta}_2 \equiv \mathrm{E}\left[\delta_{t-1} 1(\widehat{s}_{t-1} \leq 0)\right]$.

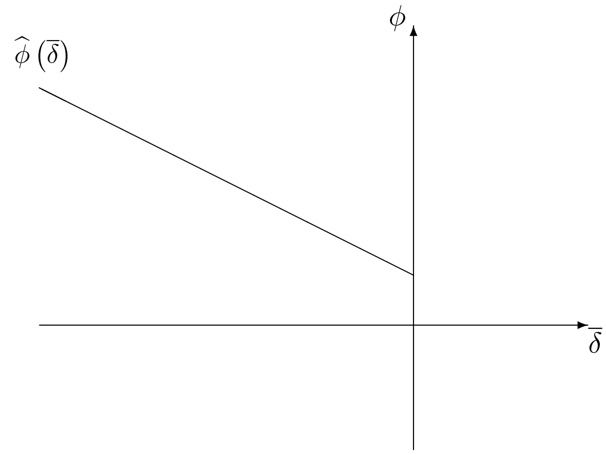


Figure 1: Discretion inequality condition in the $\overline{\delta}$ and ϕ space

linearly in $-\overline{\delta}$. In fact, since $\overline{\delta} \leq 0$, the smaller $\overline{\delta}$, the larger the set of values of ϕ satisfying the inequality constraint. Heuristically, the higher the level of distortions, the higher the level of inflation under discretion and, hence, the larger the range of inflation rates consistent with optimal monetary policy. This is illustrated graphically in Figure 1. The Figure represents the linear relation between ϕ and $\overline{\delta}$ under discretion and for $Z_t = 1$. The area below the line represents, for each value of $\overline{\delta}$, the set of ϕ compatible with optimal monetary policy.

Our first objective is to construct an estimator of Θ^I and provided the latter is not empty, to construct a confidence region, at a given level. If the identified set is nonempty, we have evidence that the monetary authority is implementing optimal monetary policy.

Before proceeding, notice that one may be tempted to reduce the two moment inequalities in (16)–(17) into a single moment equality condition, given by

$$\mathrm{E}\Big\{ \left[\phi\Pi_t + \left(\Delta\widehat{s}_t + \Delta\delta_t\right) + \psi_t\left(\widehat{s}_{t-1} + \delta_{t-1}\right)\right] 1\left(\widehat{s}_{t-1} \leq 0\right) Z_t \Big\} = 0,$$

where $\psi_t \in \{0, 1\}$ is a random variable taking value 0 in the case of commitment and 1 in the case of discretion. If ψ_t is degenerate, it may be treated as a fixed parameter ψ and the model can be estimated by GMM, provided appropriate instruments are available. This is an application of the conduct parameter method sometimes used in the industrial organization literature. However, if we allow for soft-commitment, so that periodic switches between commitment and discretion occur, ψ_t cannot be treated as a fixed parameter and this approach is no longer implementable (see Corts, 1999 and Rosen, 2006). Instead, the inflation bounds that we derive are compatible with a monetary authority that deviates from commitment-based optimal plans with some probability. Therefore, the moment inequality conditions used for partial identification of the economy's structural parameters are valid also under the quasi-commitment model proposed by Schaumburg and Tambalotti (2007).

5 Set Identification

In this section we describe how to estimate the identified set Θ^I using a partial identification approach. The basic idea underlying the estimation strategy is to use the bounds for the observed inflation rate derived from the theoretical model to generate a family of moment inequality conditions that are consistent with optimal policy. These moment inequality conditions are then used to construct a criterion function whose set of minimizers is the estimated identified set. Provided that the estimated identified set is nonempty, we use the criterion function to construct a confidence region for the identified set using the Generalized Moment Selection procedure proposed by Andrews and Soares (2010), and block-bootstrap. In what follows, we elaborate on each of the steps.

We define the following $2p$ vector of moment conditions associated with (16) and (17)

$$m_t(\theta) = \begin{bmatrix} \left(m_{d,t}^1(\theta), \ldots, m_{d,t}^p(\theta)\right)' \\ \\ \left(m_{c,t}^1(\theta), \ldots, m_{c,t}^p(\theta)\right)' \end{bmatrix},$$

where $m_{d,t}^i(\theta) = -\left[\left((\phi\Pi_t + \widehat{s}_t)\, 1(\widehat{s}_{t-1} \leq 0) + \overline{\delta}\right) Z_t^i\right]$, and $m_{c,t}^i(\theta) = \left[(\phi\Pi_t + \Delta\widehat{s}_t)\, 1(\widehat{s}_{t-1} \leq 0)\, Z_t^i\right]$. The sample analogs of the vector of moment conditions is

$$\begin{aligned} m_T(\theta) &= \left(m_T^1(\theta), \ldots, m_T^{2p}(\theta)\right)', \\ m_T^i(\theta) &= \frac{1}{T}\sum_{t=1}^{T} m_t^i(\theta) \quad \text{for} \quad i = 1, \ldots, 2p, \end{aligned}$$

where $m_t^i(\theta)$ is the i-th element of $m_t(\theta)$. Let $V(\theta)$ be the asymptotic variance of $\sqrt{T}m_T(\theta)$ and $\widehat{V}_T(\theta)$ the corresponding heteroscedasticity and autocorrelation consistent (HAC) estimator.[4] The criterion function we use for the inferential procedure is

$$Q_T(\theta) = \sum_{i=1}^{2p} \frac{\left[m_T^i(\theta)\right]_-^2}{\hat{v}_T^{i,i}(\theta)}, \tag{18}$$

where $[x]_- = x\,1(x \leq 0)$, and $\hat{v}_T^{i,i}(\theta)$ be the $i-$th element of the diagonal of $\widehat{V}_T(\theta)$. The probability limit of $Q_T(\theta)$ is given by $Q(\theta) = \mathrm{p}\lim_{T\to\infty} Q_T(\theta)$. The criterion function Q has the property that $Q(\theta) \geq 0$ for all $\theta \in \Theta$ and that $Q(\theta) = 0$ if and only if $\theta \in \Theta^I$, the identified set in Definition 2.

The estimator of the identified set $\widehat{\Theta}_T^I$ can be obtained as

$$\widehat{\Theta}_T^I = \left\{\theta \in \Theta \text{ s.t. } TQ_T(\theta) \leq d_T^2\right\}, \tag{19}$$

[4] $\widehat{V}_T(\theta)$ is constructed as follows

$$\widehat{V}_T(\theta) = \frac{1}{T}\sum_{k=-s_T}^{s_T}\sum_{t=s_T}^{T-s_T} \lambda_{k,T}\left(m_t(\theta) - m_T(\theta)\right)\left(m_{t+k}(\theta) - m_T(\theta)\right)',$$

where $\lambda_{k,T} = 1 - \frac{k}{s_T+1}$ and s_T is the lag truncation parameter, $s_T = o(T^{1/2})$.

where d_T satisfies the conditions in Proposition 2.

Assumption 3 *The following conditions are satisfied*

1. $W_t = (\Pi_t, \widehat{s}_t, Z_t)$ *is a strong mixing process with size* $-r/(r-2)$*, where* $r > 2$*;*
2. $\mathrm{E}\left(|W_{i,t}|^{2r+\iota}\right) < \infty$*,* $\iota > 0$ *and* $i = 1, 2, 3$*;*
3. $\mathrm{plim}_{T\to\infty} \widehat{V}_T(\theta) = V(\theta)$ *is positive definite for all* $\theta \in \Theta$*, where* Θ *is compact;*
4. $\sup_{\theta\in\Theta_I} |\nabla_\theta m_T(\theta) - D(\theta)| \overset{pr}{\to} 0$ *uniformly for all* θ *in* Θ^I*, where* $D(\theta)$ *is full rank.*

The following result establishes that, under Assumptions 1-3, the estimator $\widehat{\Theta}^I$ is a consistent estimator of the identified set.

Proposition 2 *Let Assumptions 1-3 hold. If as* $T \to \infty$*,* $\sqrt{\ln\ln T}/d_T \to 0$*, and* $d_T/\sqrt{T} \to 0$*, then*

$$P\left(\lim_{T\to\infty} \inf\left\{\Theta^I \subseteq \widehat{\Theta}^I_T\right\}\right) = 1,$$

and $\rho_H\left(\widehat{\Theta}^I_T, \Theta^I\right) = O_p\left(\frac{d_T}{\sqrt{T}}\right)$.[5]

It is easy to see that Proposition 2 holds e.g. for $d_T = \sqrt{\ln T}$.

To conduct inference in the moment inequality model, we construct a set $C_T^{1-\alpha}$ that asymptotically contains the identified set Θ^I with probability $1-\alpha$. This constitutes the confidence region.

Definition 3 *The* $(1-\alpha)$ *confidence region for the identified set* $C_T^{1-\alpha}$ *is given by*

$$\lim_{T\to\infty} P\left(\Theta^I \subseteq C_T^{1-\alpha}\right) = 1-\alpha,$$

where

$$C_T^{1-\alpha} = \left\{\theta \in \Theta : TQ_T(\theta) \le c_{\alpha,T}\right\},$$

and $c_{\alpha,T}$ *is the* $(1-\alpha)$*-percentile of the distribution of* $\sup_{\theta\in\Theta^I} TQ_T(\theta)$*.*

To compute the critical value $c_{\alpha,T}$ of the distribution of $\sup_{\theta\in\Theta^I} TQ_T(\theta)$, we replace the unknown set Θ^I by its consistent estimator $\widehat{\Theta}^I_T$, as shown in Proposition 2, and we use bootstrap critical values.[6] In order to reproduce the serial correlation of the moment conditions we rely on block bootstrap. In particular, let $T = bl$, where b denotes the number of blocks and l denotes the

[5] The Hausdorff distance between two sets A and B, is defined as $\rho_H(A,B) = \max\left[\sup_{a\in A} d(a,B), \sup_{b\in B} d(A,b)\right]$, with $d(a,B) = \inf_{b\in B} \|b-a\|$.

[6] Andrews and Soares (2010) and Bugni (2010) suggest the use of bootstrap percentiles over subsample based and asymptotic percentiles.

block length, and let $W_t^* = (\Pi_t^*, \widehat{s}_t^*, Z_t^*)$ denote the re-sampled observations. For each $\theta \in \widehat{\Theta}_T^I$, we construct

$$TQ_T^*(\theta) = \sum_{i=1}^{2p} \left(\sqrt{T} \left[\frac{m_T^{i*}(\theta) - m_T^i(\theta)}{\sqrt{\widehat{v}^{i,i*}(\theta)}} \right]_- 1 \left[m_T^i(\theta) \leq \sqrt{\widehat{v}^{i,i}(\theta)} \sqrt{2 \ln \ln T/T} \right] \right)^2, \qquad (20)$$

where $m_T^{i*}(\theta)$ is the bootstrap analog of the sample moment conditions $m_T^i(\theta)$, constructed using the bootstrapped data $(\Pi_t^*, \widehat{s}_t^*, Z_t^*)$, and $\widehat{v}^{i,i*}(\theta)$ is the $i-$th element in the diagonal of the bootstrap analog of the variance of the moment conditions V_θ^*. The indicator function in (20) implements the Generalized Moment Selection (GMS) procedure introduced by Andrews and Soares (2010), using information about the slackness of the sample moment conditions to infer which population moment conditions are binding, and thus enter into the limiting distribution. We perform B bootstrap replications and construct the $(1-\alpha)$-percentile $c_{\alpha,T,B}^*$. The following proposition can be established

Proposition 3 *Given Assumptions 1–3, and given $\widehat{\Theta}_T^I$ defined as in (19), as $T, l, b, B \to \infty$, $l^2/T \to 0$*

$$\lim_{T\to\infty} P\left(\Theta^I \subseteq \widehat{C}_T^{1-\alpha}\right) = 1 - \alpha,$$

where $\widehat{C}_T^{1-\alpha} = \left\{ \theta \in \Theta : TQ_T(\theta)) \leq c_{\alpha,T,B}^* \right\}$.

Estimation and inference based on the criterion function in (18) do not require any shape restriction on the identified set. For the case of convex and bounded identified set, sharper bounds can be found either via the estimation of the support function, as in Bontemps, Magnac and Maurin (2007) or via the set random variable approach of Beresteanu and Molinari (2008). The choice of the criterion function in (18) is mainly due to its computational simplicity.

6 Discretion vs Commitment

The logical next step of our analysis is to test the model of optimal monetary policy imposing either discretion or commitment. Heuristically, this implies testing whether there is a θ in the identified set, for which the moment inequality conditions associated with either discretion or commitment hold as equalities. If there is such a θ, then we have evidence in favor of discretion (commitment). Hence, we can discriminate between two alternative equilibrium concepts, maintaining the assumption of optimal monetary policy.

6.1 Testing for discretion

If the monetary authority implements optimal policy under discretion, the joint path of actual inflation and the economy-wide marginal cost is given by

$$\mathrm{E}\Big[\left(\left(\phi\Pi_t + \widehat{s}_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) + \overline{\delta}\right) Z_t\Big] \quad = \quad 0, \tag{21}$$

$$\mathrm{E}\Big[\left(\phi\Pi_t + \Delta\widehat{s}_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) Z_t\Big] \quad \geq \quad 0, \tag{22}$$

where the moment equality condition (21) follows from the assumption of discretion and the moment inequality condition (22) imposes a lower bound to the observed inflation rate as implied by optimal monetary policy. Notice that condition (21) point identifies both ϕ and the nuisance parameter $\overline{\delta}$. Therefore, under discretion, the model point identifies the parameter vector $\theta^d \equiv \left(\phi^d, \overline{\delta}^d\right)$.

Definition 4 *Let $\theta^d \equiv \left(\phi^d, \overline{\delta}^d\right) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^-$. The identified set under discretion and optimal monetary policy is a singleton and is given by*

$$\Theta_d^I = \left\{\theta^d \in \Theta : (21) \text{ and } (22) \text{ hold}\right\}.$$

Let $m_{d,t}^i\left(\theta^d\right) = \left[\left(\left(\phi^d\Pi_t + \widehat{s}_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) + \overline{\delta}^d\right) Z_t^i\right]$ be the moment condition associated with discretion and $m_{c,t}^i(\theta^d) = \left[\left(\phi^d\Pi_t + \Delta\widehat{s}_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) Z_t^i\right]$ the condition associated with the lower bound for observed inflation. We denote the corresponding sample means by $m_{d,T}^i\left(\theta^d\right)$ and $m_{c,T}^i\left(\theta^d\right)$, respectively. The asymptotic variance of $\sqrt{T}\left[m_{d,T}\left(\theta^d\right), m_{c,T}\left(\theta^d\right)\right]$ is $V^d\left(\theta^d\right)$ and $\widehat{V}_T^d\left(\theta^d\right)$ is the corresponding HAC estimator. To test for discretion, we base the inference procedure on the criterion function

$$TQ_T^d\left(\theta^d\right) = T\left(\sum_{i=1}^{p} \frac{m_{d,T}^i\left(\theta^d\right)^2}{\hat{v}_T^{i,i}\left(\theta^d\right)} + \sum_{i=p+1}^{2p} \frac{\left[m_{c,T}^i\left(\theta^d\right)\right]_-^2}{\hat{v}_T^{i,i}\left(\theta^d\right)}\right), \tag{23}$$

where $\hat{v}_T^{i,i}\left(\theta^d\right)$ is i-th element on the diagonal of $\widehat{V}_T^d\left(\theta^d\right)$. The estimated identified set under discretion and optimal monetary policy $\widehat{\Theta}_{d,T}^I$ and the estimated confidence region $\widehat{C}_{d,T}^{1-\alpha}$ are obtained using the criterion function (23) and the approach outlined in Section 5. Note that under discretion Θ_d^I is a singleton, thus $\widehat{\Theta}_{d,T}^I$ is a set shrinking towards Θ_d^I as $T \to \infty$. Since the first p moment conditions hold with equality, they all contribute to the asymptotic distribution of $TQ_T^d\left(\theta^d\right)$, thus we apply the GMS procedure only to the inequality moment conditions.

If the estimated identified set $\widehat{\Theta}_{d,T}^I$ is empty, we reject the model of optimal monetary policy under discretion. If instead the estimated identified set is nonempty, we construct a test statistic for the following null hypothesis

$$H_0^d : \exists\, \boldsymbol{\theta} \in \Theta_d^I$$

against the alternative

$$H_1^d: \nexists\, \boldsymbol{\theta} \in \Theta_d^I$$

where $\boldsymbol{\theta}$ is the true structural parameter vector which, under the null hypothesis of discretion, is point identified and can be consistently estimated using the Generalized Method of Moments (GMM) estimator applied to the moment equality condition (21).

To construct the test statistic, we first estimate $\widehat{\theta}_T^d$ by two-step GMM under discretion and then evaluate the criterion function (23) at the estimated parameter vector $\widehat{\theta}_T^d$

$$TQ_T^d\left(\widehat{\theta}_T^d\right) = T\left(\sum_{i=1}^{p} \frac{m_{d,T}^i\left(\widehat{\theta}_T^d\right)^2}{\hat{v}_T^{i,i}\left(\widehat{\theta}_T^d\right)} + \sum_{i=p+1}^{2p} \frac{\left[m_{c,T}^i\left(\widehat{\theta}_T^d\right)\right]_-^2}{\hat{v}_T^{i,i}\left(\widehat{\theta}_T^d\right)}\right), \tag{24}$$

where $\hat{v}_T^{i,i}\left(\widehat{\theta}_T^d\right)$ is the i-th diagonal element of $\widehat{V}_T\left(\widehat{\theta}_T^d\right)$, the HAC estimator of the asymptotic variance of $\sqrt{T}\left[m_{d,T}\left(\widehat{\theta}_T^d\right), m_{c,T}\left(\widehat{\theta}_T^d\right)\right]$ which takes into account the estimation error in $\widehat{\theta}_T^d$ (see Appendix).[7]

Since our inference is based on bootstrap critical values $c_{T,B,\alpha}^{*d}$, we use a bootstrap procedure that properly mimics the contribution of parameter estimation error.[8]

Proposition 4 *Let Assumptions 1–3 hold. Then,*

(i) under H_0^d, $\lim_{T,B\to\infty} \Pr\left(TQ_T^d\left(\widehat{\theta}_T^d\right) > c_{T,B,\alpha}^{*d}\right) = \alpha$,

(ii) under H_1^d, $\lim_{T,B\to\infty} \Pr\left(TQ_T^d\left(\widehat{\theta}_T^d\right) > c_{T,B,\alpha}^{*d}\right) = 1$,

where B *denotes the number of bootstrap replications.*

6.2 Testing for commitment

If the monetary authority implements optimal policy under commitment, the joint path of actual inflation and the economy-wide marginal cost is given by

$$-\,\mathrm{E}\Big[\left((\phi\Pi_t + \widehat{s}_t)\, 1\left(\widehat{s}_{t-1} \leq 0\right) + \overline{\delta}\right) Z_t\Big] \;\geq\; 0, \tag{25}$$

$$\mathrm{E}\Big[\left(\phi\Pi_t + \Delta\widehat{s}_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) Z_t\Big] \;=\; 0, \tag{26}$$

where the moment equality condition (26) follows from the assumption of commitment and the moment inequality condition (25) imposes an upper bound to the observed inflation rate, as implied by

[7] Andrews and Soares (2010) studies the limiting distribution of the statistic in (24) evaluated at a fixed θ. Hence, they do not take into account the variability of $\widehat{\theta}_T^d$.

[8] See Appendix for details.

optimal monetary policy. Notice that conditions (26)-(25) only set-identify $\overline{\delta} = E\left[\delta_t 1\left(\widehat{s}_{t-1} \leq 0\right)\right]$. Therefore, the model under commitment set-identifies the parameter vector $\theta^c \equiv \left(\phi^c, \overline{\delta}\right)$.

Definition 5 *Let $\theta^c \equiv \left(\phi^c, \overline{\delta}\right) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^-$. The identified set under commitment and optimal monetary policy is given by*

$$\Theta_c^I = \left\{\theta^c \in \Theta : (25) \text{ and } (26) \text{ hold}\right\}.$$

Let $m_{c,t}^i\left(\theta^c\right) = \left[\left(\phi^c \Pi_t + \Delta \widehat{s}_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) Z_t^i\right]$ be the moment condition associated with commitment and $m_{d,t}^i(\theta^c) = \left[\left(\left(\phi^c \Pi_t + \widehat{s}_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) + \overline{\delta}\right) Z_t^i\right]$ the moment condition associated with the upper bound for observed inflation. We denote the corresponding sample means by $m_{c,T}^i\left(\theta^c\right)$ and $m_{d,T}^i\left(\theta^c\right)$, respectively. The asymptotic variance of $\sqrt{T}\left[m_{d,T}\left(\theta^c\right), m_{c,T}\left(\theta^c\right)\right]$ is $V^c\left(\theta^c\right)$ and $\widehat{V}_T^c\left(\theta^c\right)$ is the corresponding HAC estimator. The estimated identified set under commitment and optimal monetary policy $\widehat{\Theta}_{c,T}^I$ and the estimated confidence region $\widehat{C}_{c,T}^{1-\alpha}$ are obtained using the same methodology as the one described in Section 6.1.

Testing for optimal monetary policy under commitment involves a different null hypothesis from the one used to test discretion. The reason is that, under the null of commitment, only the first element of the parameter vector θ is point identified, while the identified set is defined in the entire parameter space Θ. Intuitively, this happens because, under commitment, the average level of distortions is irrelevant for the conduct of optimal monetary police since there is no inflationary bias. Thus, we test for commitment as follows. If the estimated identified set $\widehat{\Theta}_{c,T}^I$ is empty, we reject the model of optimal monetary policy under commitment. If instead the estimated identified set is nonempty, we construct a test statistic based on the following null hypothesis

$$H_0^c: \exists\, \overline{\delta} \text{ such that } \boldsymbol{\phi} \in \Theta_c^I\left(\overline{\delta}\right)$$

against the alternative

$$H_1^c: \nexists\, \overline{\delta} \text{ such that } \boldsymbol{\phi} \in \Theta_c^I\left(\overline{\delta}\right)$$

where $\boldsymbol{\phi}$ is the true structural parameter which, under the null hypothesis of commitment, is point identified and can be consistently estimated using the GMM estimator applied to the moment equality condition (26). For a fixed $\overline{\delta}$, we construct the test statistic

$$TQ_T^c\left(\widehat{\phi}^c, \overline{\delta}\right) = T\left(\sum_{i=1}^{p} \frac{\left[m_{d,T}^i\left(\widehat{\phi}^c, \overline{\delta}\right)\right]_-^2}{\hat{v}_T^{i,i}\left(\widehat{\phi}^c, \overline{\delta}\right)} + \sum_{i=p+1}^{2p} \frac{m_{c,T}^i\left(\widehat{\phi}^c, \overline{\delta}\right)^2}{\hat{v}_T^{i,i}\left(\widehat{\phi}^c, \overline{\delta}\right)}\right), \tag{27}$$

and compute the corresponding critical value $c_{T,B,\alpha}^{*c}\left(\overline{\delta}\right)$, as discussed in Section 6.1.

Proposition 5 *Let Assumptions 1–3 hold. Then,*

(i) under H_0^c, $\lim_{T,B\to\infty} \Pr\left(TQ_T^c\left(\widehat{\phi}^c, \overline{\delta}\right) > c_{T,B,\alpha}^{*c}\left(\overline{\delta}\right)\right) = \alpha$,

(ii) under H_1^c, $\lim_{T,B\to\infty} \Pr\left(TQ_T^c\left(\widehat{\phi}^c, \overline{\delta}\right) > c_{T,B,\alpha}^{*c}\left(\overline{\delta}\right)\right) = 1$,

where B denotes the number of bootstrap replications.

Table 1: Unit Root Tests For Inflation

	ADF test t-Statistic	Phillips–Perron test t-Statistic
1960:2 – 2008:3	−2.180 (0.18)	−2.933 (0.04)
1970:1 – 2008:3	−2.428 (0.14)	−2.652 (0.08)
1979:3 – 2008:3	−2.902 (0.05)	−3.482 (0.01)
1983:1 – 2008:3	−5.313 (0.00)	−5.142 (0.00)

p-value in parentheses.

Note: The ADF statistic is computed using the Schwarz information criteria to select the lag length. The Phillips–Peron statistic, is computed using Andrews' (1991) method to select the value for the lag truncation parameter q required to form the HAC estimator. A constant is included in both test regressions.

7 Empirical Results

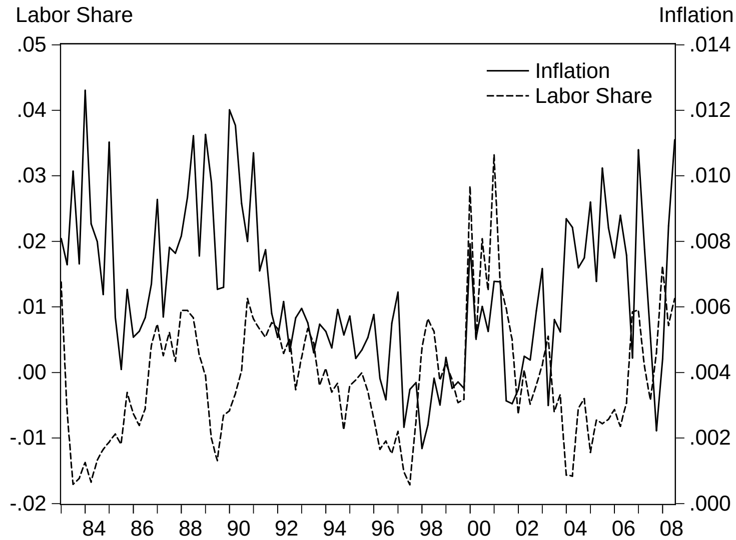
We apply our framework to test whether the behavior of the United States monetary authority is consistent with optimal monetary policy. Furthermore, we estimate an identified set for the economy's structural parameters. We start with a preliminary analysis of the data.

7.1 Data and preliminary analysis

In our empirical analysis, we use quarterly time-series for the US economy. As explained earlier, the theoretical output gap is not observable and there is no reason to believe that it can be proxied by the deviation of output from a smooth statistical trend. Thus, following Sbordone (2002) and Gali and Gertler (1999), we exploit the relationship between the theoretical output gap and the cyclical component of the economy-wide real marginal cost, $\widehat{s}_t$. In the theoretical economy, the real marginal cost is proportional to the unit labor cost (see Proposition 1). Hence, we use the cyclical component of the labor income share in the non-farm business sector (computed using the HP-Filter) to measure $\widehat{s}_t$.

Our measure of inflation is the annualized percentage change in the GDP deflator. The econometric framework developed in this paper relies on a stationarity assumption (see Assumption 3). Therefore, we begin by investigating whether the time series of US inflation is stationary. Our results show that there is a change in the behavior of inflation from $I(1)$ to $I(0)$ after 1982. Table 1 shows results from two alternative unit root tests, the augmented Dickey-Fuller (ADF) and the Phillips-Peron test. The two test statistics broadly concur in the conclusion that US inflation in non-stationary for samples beginning in 1960 and 1970. Turning to the sample period starting from 1979:3, the start of the tenure of Paul Volcker as Federal Reserve Chairman, the test statistics

Figure 2: Labor Share and Inflation in the US, 1983:1–2008:3.



are less conclusive. The ADF test has a p-value of 5% while the Phillips–Perron test leads to the rejection of the null hypothesis at conventional significance levels. Following the analysis in Clarida, Gali and Gertler (2000), we have decided to study the sample starting from 1983:1, that removes the first three years of the Volcker era.[9] For the post–1982 data, both test statistics clearly reject the null hypothesis of a unit root at the 1% confidence level. These findings are in line with the study of Halunga, Osborn and Sensier (2009) showing that there is a change in inflation persistence from $I(1)$ to $I(0)$ dated at June 1982. Thus, our empirical analysis focuses on the sample period 1983:1 to 2008:3, which spans the period starting after the disinflation and monetary policy shifts that occurred in the early 1980s and extends until the period when the interest rate zero lower bound becomes a binding constraint.[10] Figure 2 plots the time series of the US cyclical component of labor income share, $\widehat{s}_t$, and inflation for the sample period 1983:1 to 2008:3.

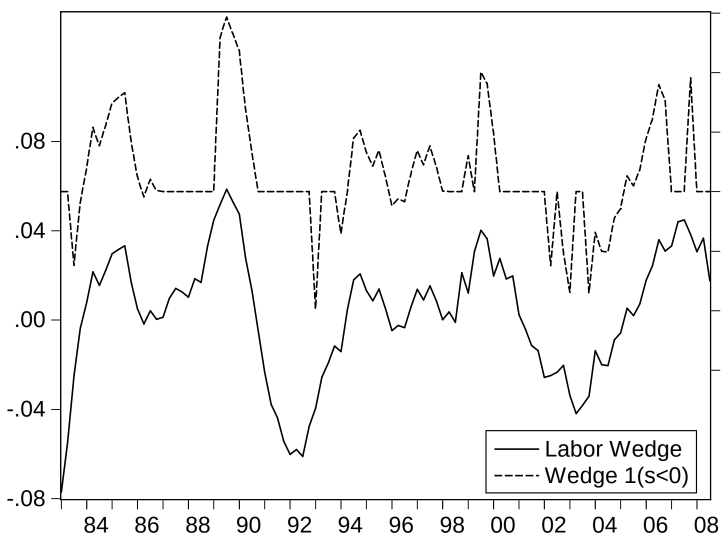
7.2 Instrumental variables

It is frequently assumed that movements in military purchases are exogenous; moreover, fluctuations in military spending account for much of the variation in government purchases (see Hall 2009). Thus, we use as instrument the variable *'Military Spending'*, given by the log of real government expenditure in national defense detrended using the HP-Filter. As a second instrument, we use the variable *'Oil Price Change'*, given by the log difference of the spot oil price. The instrumental variables are adjusted using the following transformation guaranteeing positiveness: $Z_+ = Z - min(Z)$.

[9] Clarida, Gali and Gertler (2000) offer two reasons for doing this. First, this period was characterized by a sudden and permanent disinflation episode bringing inflation down from about 10 percent to 4 percent. Second, over the period 1979:4-1982:4, the operating procedures of the Federal Reserve involved targeting non-borrowed reserves as opposed to the Federal Funds rate.

[10] After 2008:3, the federal funds rate rapidly fell toward the lower bound, signaling a period of unconventional monetary policy for which our econometric specification may be inadequate.

Figure 3: The Labor Wedge



The complete instrument list includes the variables *'Military Spending'*, *'Oil Price Change'*, and the constant, yielding $p = 3$ instruments and 6 moment conditions overall.

From Assumption 2, the instrumental variables are required to satisfy the exclusion restrictions

$$\mathrm{E}\left[v_t \, 1\left(\widehat{s}_{t-1} \leq 0\right) \, Z_t\right] = 0$$

$$\mathrm{E}\left[\delta_{t-r} \, 1\left(\widehat{s}_{t-1} \leq 0\right) Z_t\right] = \overline{\delta} \, \mathrm{E}\left[Z_t\right], \quad \text{for } r = 0, 1.$$

The variable $\delta_t = -\ln\left(\mu \chi_t\right)$ is an unobserved disturbance related to the time-varying labor wedge χ_t. However, it is possible to construct a theory-based series that proxies δ_t using data on aggregate consumption, hours worked and wages. The upshot is that we are able to gain valuable insights about the validity of the exclusion restrictions. In particular, following the procedure in Gali, Gertler and Lopez-Salido (2007) and Shimer (2009), we construct a measure of the labor wedge in terms of observables, conditional on some conventional assumptions about preferences and technology. By combining data on aggregate consumption, hours worked and wages, we construct a time series for the marginal rate of substitution between consumption and leisure and the real wage. As a result, we are able to obtain a time-series for the labor wedge χ_t, using equation (1) and specifying a parameterization for the household preferences. Thus, we assume the representative household has preferences given by

$$\frac{C_t^{1-1/\sigma}}{1-1/\sigma} - \int_0^1 \exp\left(\epsilon_t\right) \frac{H\left(j\right)^{1+\gamma}}{1+\gamma} dj.$$

where ϵ_t is an unobservable preferences shifter. Using equation (1), taking logs and averaging over

Table 2: Exclusion Restrictions

	Coefficient
Military Spending	-0.091
	(0.33)
Oil Price Change	-0.001
	(0.94)
R^2	0.01
F-stat	0.72
	(0.49)

p-value in parentheses.

Note: The dependent variable is $\widehat{\delta}_t 1\left(\widehat{s}_{t-1} < 0\right)$, and the regression includes a constant term (not reported).

j yields the condition

$$\delta_t = \sigma^{-1} c_t + \gamma h_t - w_t - \ln \mu + \epsilon_t, \tag{28}$$

where the average real wage w_t, the log of aggregate consumption c_t, and the log of aggregate hours h_t, are all observed in the data. If the volatility of ϵ_t is negligible at business cycle frequencies, we are able to identify the cyclical component of the labor wedge with the cyclical movements in $\sigma^{-1} c_t + \gamma h_t - w_t$, computed using the HP-Filter. Figure 3 shows the time series for the estimated cyclical component of the labor wedge $\widehat{\delta}_t$, and the state contingent variable $\widehat{\delta}_t 1\left(\widehat{s}_{t-1} < 0\right)$, under the assumption that σ and γ are both equal to unity.[11]

To obtain some insight about the validity of the exclusion restrictions, we regress the state contingent variable $\widehat{\delta}_t\, 1\left(\widehat{s}_{t-1} < 0\right)$ on the proposed list of instruments: *'Military Spending'* and *'Oil Price Change'* variable, including a constant term. The findings are shown in Table 2. None of the regressors is significant, the R^2 of the regression is only 1% and the F statistic for the joint hypothesis that all the regression coefficients are zero cannot be rejected at conventional significance levels. Overall, the evidence supports the exclusion restrictions.

7.3 Model specification tests

We begin by providing the estimation results for the identified set. For the implementation of the estimator, we specify a truncation parameter for the computation of the HAC variance estimator equal to four. The confidence sets are constructed using 1,000 block-bootstrap replications, also with block size equal to four.[12] In Figure 4, we show the estimated identified set implied by optimal monetary policy $\widehat{\Theta}_T^I$, and the corresponding confidence region at 95% confidence level, $\widehat{C}_T^{0.95}$. As discussed earlier, the upper-bound of the identified set grows linearly in the level of distortions

[11] This parametrization corresponds to the log-log specification for preferences, that is a popular choice in the business cycle literature as it implies balanced growth and a unit Frisch labor supply elasticity (see Cooley and Prescott 1995).

[12] We have also tried a few different values for the block length. Overall, our findings are quite robust to variation in the block length choice.

Figure 4: Identified Set and Confidence Set under Optimal Monetary Policy

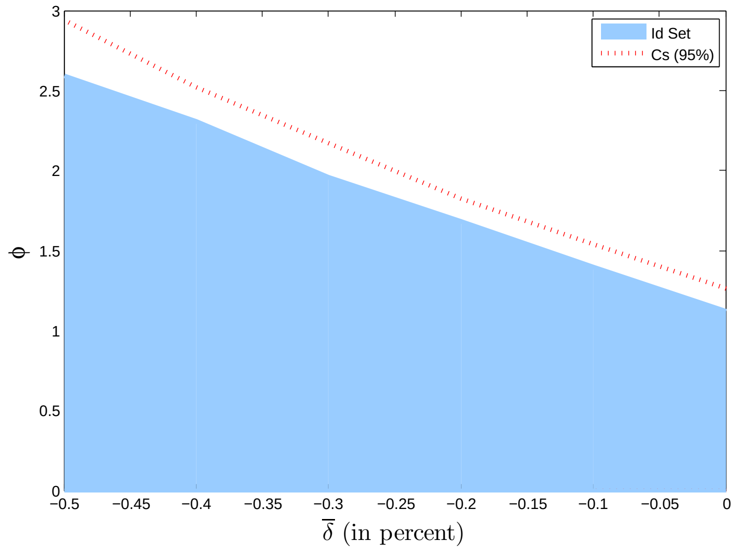
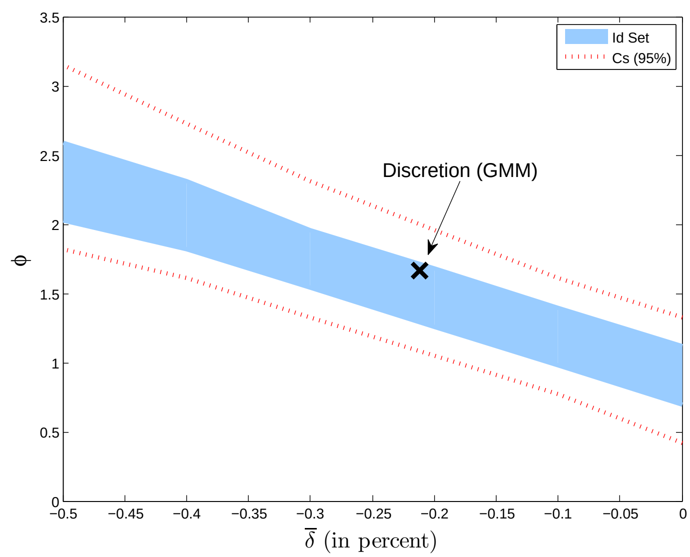


Figure 5: Identified Set and Confidence Sets under Discretion



Note: The shaded areas correspond to the estimated identified set. The dashed lines denote the confidence set at 95% confidence level.

Table 3: Model Specification Tests

	Discretion	Commitment
$J-$test	1.178	0.735
p-val	(0.278)	(0.693)
TQ_T	13.417	26.963
p-val*	(0.164)	(0.022)

Note: The upper panel shows results from the Hansen test of overidentifying restriction based on the GMM model under discretion or commitment. The statistic p-val is the p-value for the $J-$test computed from the χ_1^2 distribution for discretion and from the χ_2^2 distribution for commitment. The lower panel shows the results from the model specification test based on the test statistic TQ_T, defined by equations (24) and (27), for commitment and discretion. The statistic p-val*, is the p value computed via the bootstrapping procedure described in Section 6.1.

given by $-\overline{\delta}$. Since the identified set is not empty, we do not reject the hypothesis that the US monetary authority has implemented optimal monetary policy over the sample period 1983:1 to 2008:3.

In Figure 5, we report the estimated identified set and the corresponding confidence region imposing discretion and optimal monetary policy. As expected, the estimated identified set shrinks considerably but it is not empty. Moreover, the 95% confidence region contains the parameter vector estimated by GMM. This provides evidence in favor of discretion. Instead, the estimated identified set imposing commitment and optimal monetary policy is empty providing evidence against commitment. The following three results can, therefore, be established: (i) the identified set implied by optimal monetary policy is nonempty and, therefore, inflation and the theoretical output gap are consistent with optimal monetary policy; (ii) the identified set under discretion is nonempty and contains the parameter vector estimated by GMM under discretion, providing evidence in favor of discretion; (iii) the identified set under commitment is empty, providing evidence against the null hypothesis of commitment. We next examine the formal test statistics developed in sections 6.1 and 6.2 to test for discretion and commitment. Notice that since the estimated identified set under commitment is empty, the null of commitment and optimal monetary policy is already rejected. Nonetheless, we construct the formal test statistic for each null hypothesis.

The test is based on a two-step procedure. In particular, to test discretion we first estimate the parameter vector θ^d via GMM from condition (26); next, using the estimated vector we construct the test statistic for discretion $TQ_T^d\left(\widehat{\theta}_T^d\right)$ and compute bootstrap critical values. To test commitment, we proceed in the same way, except that the moment condition (26) only permits identification of the parameter ϕ^c. Therefore, after estimating this parameter by GMM, we construct the test statistic $TQ_T^c\left(\widehat{\phi}^c;\overline{\delta}\right)$ and compute bootstrap critical values. In the case of commitment, these

Table 4: Calibrated Parameters

Parameter	Definition	Value
α	Share of firms keeping prices fixed	0.6600
β	Discount factor	0.9914
ω	Output elasticity of real marginal cost	0.4400
σ	Intertemporal elasticity of substitution	6.2500

Note: Quarterly calibration. The parameters are calibrated following Rotemberg and Woodford (1997).

critical values depend on the nuisance parameter $\overline{\delta}$, that needs to be fixed. Therefore, to test commitment we consider the test statistic over a grid of values for $\overline{\delta}$.

Given that we are using a sufficient number of instrumental variables for overidentification, we start by reporting results from the standard Hansen test statistic for overidentifying restrictions. The results are reported on the upper panel of Table 3, for the null hypothesises of discretion (first column), and commitment (second column).[13] As can be seen, the standard J-test fails to reject either model.[14] Thus, by not making use of the full set of implications of optimal monetary policy, we are unable to discriminate between the two alternative policy regimes. However, using the additional information implied by the maintained assumption of optimal monetary policy, we can test the composite null hypothesis of optimal monetary policy and a specific policy regime—discretion or commitment.

The test statistic is based on equation (24) for the case of discretion, and equation (27) for the case of commitment. The results are shown in the lower panel of Table 3. For the case of discretion, the p-value associated with the test statistic is 16.4 percent and, therefore, we fail to reject the null hypothesis of discretion at all conventional levels. For the case of commitment, the parameter $\overline{\delta}$ is not identified by GMM and, therefore, needs to be fixed. We consider a dense grid of values for $\overline{\delta}$, and the resulting p-value associated with the test statistic (27) ranges between 1.5 percent and 2.2 percent. Hence, even choosing the most conservative case, a p-value of 2.2 percent allows for rejection of the null hypothesis of optimal policy under commitment at the 5% confidence level. This is in line with the fact that the estimated identified set under commitment and optimal monetary policy is empty.

7.4 Structural parameters estimated sets

The null hypothesis of commitment has been rejected and congruently, the identified set under commitment is empty. However, the identified set under discretion provides useful information about the economy's structural parameters and in particular ϕ. The upper panel of Table 5 provides estimated sets and 95% confidence regions for this parameter, for alternative values of $\overline{\delta}$. Notice

[13]The J–tests are constructed from the first order conditions in equations (12) and (11) respectively.

[14]Incidentally, the failure to reject either model using the standard test of overidentification provides evidence in favor of the instrumental variables used.

Table 5: Parameter Estimates and Confidence Regions

$\overline{\delta}$	Interval estimate for ϕ	Bootstrap 95% c.i.
0.00%	$\left[0.70,\ 1.12\right]$	$\left[0.42,\ 1.33\right]$
-0.21%[a]	$\left[1.26,\ 1.68\right]$	$\left[1.05,\ 1.96\right]$
-0.50%	$\left[2.03,\ 2.59\right]$	$\left[1.82,\ 3.08\right]$
$\overline{\delta}$	Implied interval estimate for κ	Bootstrap 95% c.i.
0.00%	$\left[0.0611,\ 0.0737\right]$	$\left[0.0563,\ 0.0854\right]$
-0.21%[a]	$\left[0.0498,\ 0.0578\right]$	$\left[0.0456,\ 0.0629\right]$
-0.50%	$\left[0.0383,\ 0.0446\right]$	$\left[0.0340,\ 0.0476\right]$

[a] GMM estimate for $\overline{\delta}$.

Note: The implied estimated for κ is obtained from the estimated set for ϕ and using equation (29), and the calibration described in the main text. The confidence intervals at the 95% are computed using the bootstrapping procedure described in Section 6.1.

that under the null hypothesis of discretion, the parameter $\overline{\delta}$ can actually be point-identified. The GMM estimate of $\overline{\delta}$ is -0.0021 percent, indicating that the monetary authority acts as if output is 0.21 percent beneath the efficient level. When $\overline{\delta} = -0.0021$, the implied 95% confidence region for ϕ is [1.05, 1.96].

There is a close connection between the parameter ϕ and the parameter κ which controls the response of inflation to changes in the output gap. In particular, the parameter κ is given by

$$\kappa = \left[\frac{(1-\alpha)(1-\alpha\beta)}{1+\omega\phi/\left(\omega+\sigma^{-1}\right)}\right]\frac{\left(\omega+\sigma^{-1}\right)}{\alpha} \tag{29}$$

Our model does not identify uniquely all the structural parameters. However, by setting values for the parameters $(\alpha, \beta, \omega, \sigma)$, the confidence region for ϕ implies a corresponding confidence region for the output gap elasticity of inflation κ, a parameter of considerable importance for monetary policy. A similar approach is followed by Rotemberg and Woodford (1997). Thus, we calibrate the parameter vector $(\alpha, \beta, \omega, \sigma)$ using the parameter values suggested by these authors, as described in Table 4, and consider the implied confidence region for κ.

The resulting estimated identified set and the 95% confidence region for κ are shown in the lower panel of Table 5 for alternative values for $\overline{\delta}$. Results are in line with previous findings in the literature (see Rotemberg and Woodford 1997), and suggest that the output gap elasticity of inflation is small. For a given duration of price stickiness a small value for κ indicates the existence of large strategic complementarities in price-setting, sometimes interpreted as a high degree of real rigidity (see Ball and Romer 1990).

8 Conclusion

This paper develops a method for estimating and testing a model of optimal monetary policy without requiring an explicit choice of equilibrium concept. The procedure considers a general specification of optimal monetary policy that nests discretion and commitment as two special cases. The general specification is obtained deriving bounds for inflation that are consistent with both forms of optimal policy and yield set identification of the economy's structural parameters. We propose a model specification test that makes use of the set of moment inequality and equality conditions implied by optimal monetary policy under a specific policy regime, and that allows testing the null hypothesises of discretion and commitment.

We apply our method to investigate whether the monetary policy in the United States is consistent with the New Keynesian model of optimal monetary policy over the sample period running from 1983:1 to 2008:3. The paper establishes the following three results: (i) the estimated identified set is nonempty and, therefore, the time series of inflation and output gap are consistent with optimal monetary policy; (ii) the estimated identified set under discretion is nonempty and contains the parameter vector estimated by GMM under discretion, providing evidence in favor of discretion; (iii) the estimated identified set under commitment is empty, allowing to formally reject the null hypothesis of commitment.

We also develop a formal test statistic that allows to test the null hypothesises of discretionary optimal monetary policy and of optimal monetary policy under commitment. The test fails to reject the null hypothesis of discretion but rejects the null hypothesis of commitment. In contrast, the standard J-test of overidentifying restrictions fails to reject either policy regime. Thus, by making use of the full set of implications of optimal monetary policy we are able to discriminate across policy regimes, rejecting commitment but not discretion.

A Calvo Pricing

A producer that changes its price in period t, chooses the new price to maximize the discounted flow of profits

$$E_t\left[\sum_{s=t}^{\infty}\alpha^{s-t}Q_{t,s}\Pi\left(p_t\left(i\right),p_s^j,P_s;Y_s,\tilde{\xi}_s\right)\right],$$

where $Q_{t,s}$ is the stochastic discount factor, given by

$$Q_{t,s}=\beta^{s-t}\frac{u_c\left(Y_s-G_s;\tilde{\xi}_t\right)}{u_c\left(Y_t-G_t;\tilde{\xi}_t\right)}\frac{P_t}{P_s}$$

and the profit function is given by

$$\Pi\left(p,p^j,P;Y,\tilde{\xi}\right)=pY\left(\frac{p}{P}\right)^{-\vartheta}-P\frac{v_h\left(f^{-1}\left(Y\left(p^j/P\right)^{-\vartheta}\right);\tilde{\xi}\right)}{u_c\left(Y-G;\tilde{\xi}\right)}\chi f^{-1}\left(Y\left(\frac{p}{P}\right)^{-\vartheta}\right),$$

The profit function Π is homogeneous of degree one in its first three arguments. Moreover, the price level evolves according to

$$P_t=\left[\alpha P_{t-1}^{1-\vartheta}+\left(1-\alpha\right)p_t^{\star 1-\vartheta}\right]^{1/(1-\vartheta)}. \tag{30}$$

The optimal price chosen in period t by the updating sellers, $p_t^\star$, satisfies the conditions

$$E_t\left[\sum_{s=t}^{\infty}\left(\alpha\beta\right)^{s-t}\Gamma\left(\frac{p_t^\star}{P_s},Y_s,\tilde{\xi}_s\right)\right]=0, \tag{31}$$

where the function Γ is given by

$$\begin{aligned}\Gamma\left(\frac{p_t^\star}{P_s},Y_s,\tilde{\xi}_s\right) &= u_c\left(Y_s-G_s;\tilde{\xi}_t\right)\frac{p_t^\star}{P_s}\Pi_1\left(\frac{p_t^\star}{P_s},\frac{p_t^\star}{P_s},1;Y_s,\tilde{\xi}_s\right)\\ &= u_c\left(Y_s-G_s;\tilde{\xi}_t\right)\left(1-\vartheta\right)Y_s\left(\frac{p_t^\star}{P_s}\right)^{-\vartheta}\left[\frac{p_t^\star}{P_s}-\mu\chi_t s\left(Y_s\left(\frac{p_t^\star}{P_s}\right)^{-\vartheta},Y_s;\tilde{\xi}_t\right)\right].\end{aligned}$$

By log-linearizing equations (30) and (31) around the deterministic steady state (under zero inflation), we obtain the following two conditions

$$\log P_t=\alpha\log P_{t-1}+\left(1-\alpha\right)\log p_t^\star \tag{32}$$

and

$$\sum_{s=t}^{\infty}\left(\alpha\beta\right)^{s-t}E_t\left[\log p_t^\star-\log P_s-\zeta\left(\widehat{Y}_s-\widehat{Y}_s^n\right)\right]=0 \tag{33}$$

where $\zeta = \left(\omega + \sigma^{-1}\right) / (1 + \omega\vartheta)$. Combining the equations (32) and (33) yields the New Keynesian Phillips curve, given by

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t,$$

where $\kappa = (1-\alpha)(1-\alpha\beta)(\zeta/\alpha)$ and the variable $x_t \equiv \left(\widehat{Y}_t - \widehat{Y}_t^n\right)$ is the relevant output gap, and is proportional to the average real marginal cost.

B Proofs

Proof of Lemma 1: Immediate from the definition of $\pi_t^{(c)}$ and $\pi_t^{(d)}$.

Proof of Lemma 2: Immediate by replacing (11) and (12) in the statement of Lemma 1 and using Assumption 1.

Proof of Proposition 1: From Lemma 2, provided $\phi > 0$,

$$\Pr\left(-\Delta\widehat{s}_t - \Delta\delta_t + \phi v_t \leq \phi\Pi_t \leq -\widehat{s}_t - \delta_t + \phi v_t \middle| \widehat{s}_{t-1} \leq 0\right) = 1,$$

which implies

$$\Pr\Big(\left(-\Delta\widehat{s}_t - \Delta\delta_t + \phi v_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) \leq \phi\Pi_t 1\left(\widehat{s}_{t-1} \leq 0\right) \leq \left(-\widehat{s}_t - \delta_t + \phi v_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right) \Big) = 1,$$

thus

$$-\mathrm{E}\left[\left(\phi\Pi_t + \widehat{s}_t + \delta_t - \phi v_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right)\right] \geq 0,$$

$$\mathrm{E}\left[\left(\phi\Pi_t + \Delta\widehat{s}_t + \Delta\delta_t - \phi v_t\right) 1\left(\widehat{s}_{t-1} \leq 0\right)\right] \geq 0,$$

as stated in the Proposition.

Proof of Proposition 2: The statement follows from Theorem 3.1 in Chernozhukov, Hong and Tamer (2007), with $\widehat{c} = d_T^2$, $a_T = T$, $\gamma = 2$, once we show that Assumptions 1–3 imply the satisfaction of their Conditions 1 and 2. Condition 1(a) holds as $\theta = \left(\phi, \overline{\delta}\right)$ lies in a compact set of $\mathbb{R}^+ \times \mathbb{R}^-$. Assumptions 1–2 allow to state the sample objective function as $Q_T(\theta)$ in (18). Given Assumption 3, as a straightforward consequence of the uniform law of large numbers for strong mixing processes, $Q_T(\theta)$ satisfies Condition 1(b)–(e) with $b_T = \sqrt{T}$ and $a_T = T$. Finally, it is immediate to see that $Q_T(\theta)$ in (18) satisfies Condition 2.

Proof of Proposition 3: The events $\left\{\Theta^I \subseteq C_T^{1-\alpha}\right\}$ and $\{\sup_{\theta\in\Theta^I} TQ_T(\theta) \leq c_{\alpha,T}\}$ are equivalent, and thus

$$\Pr\left(\Theta^I \subseteq C_T(1-\alpha)\right) = \Pr\left(\sup_{\theta\in\Theta^I} TQ_T(\theta) \leq c_{\alpha,T}\right),$$

where $c_{\alpha,T}$ is the $(1-\alpha)-$percentile of the limiting distribution of $\sup_{\theta\in\Theta^I} TQ_T(\theta)$. Given Assumptions 1–3, by Theorem 1 of Andrews and Guggenberger (2009), for any $\theta \in \Theta^I$,

$$TQ_T(\theta) \xrightarrow{d} \sum_{i=1}^{2p}\left(\left[\sum_{j=1}^{2p}\omega_{i,j}(\theta)\mathcal{Z}_i + h_i(\theta)\right]_-\right)^2$$

where $\mathcal{Z} = (\mathcal{Z}_1, \ldots, \mathcal{Z}_{2p}) \sim N(0, I_{2p})$ and $\omega_{i,j}$ is the generic element of the correlation matrix

$$\Omega(\theta) = D^{-1/2}(\theta)\, V(\theta)\, D^{-1/2}(\theta),$$

with $D(\theta) = diag(V(\theta))$ and $V(\theta) = p\lim_{T\to\infty}\widehat{V}_T(\theta)$, as defined in footnote 4. Finally, $h(\theta) = (h_1(\theta), ..., h_{2p}(\theta))'$ is a vector measuring the slackness of the moment conditions, given by

$$h_i(\theta) = \lim_{T\to\infty}\sqrt{T}\,\mathrm{E}\left(m_{i,T}(\theta)/\sqrt{v^{i,i}(\theta)}\right).$$

Given the stochastic equicontinuity on Θ^I of $TQ_T(\theta)$, because of Proposition 2, it also follows that

$$\sup_{\theta\in\widehat{\Theta}^I_T} TQ_T(\theta) \xrightarrow{d} \sup_{\theta\in\Theta^I}\sum_{i=1}^{2p}\left(\left[\sum_{j=1}^{2p}\omega_{i,j}(\theta)\mathcal{Z}_i + h_i(\theta)\right]_-\right)^2. \tag{34}$$

We need to show that the $(1-\alpha)$-percentile of the right-hand side of (34), $c_{\alpha,T}$, is accurately approximated by the $(1-\alpha)$-percentile of the bootstrap limiting distribution $\sup_{\theta\in\widehat{\Theta}^I_T} TQ^*_T(\theta)$, $c^*_{\alpha,T}$, conditional on the sample. By the law of the iterated logarithm as $T \to \infty$ and for $i = 1, ..., 2p$, we have that, almost surely,

$$\left(\frac{T}{2\ln\ln T}\right)^{1/2}\frac{m_{i,T}(\theta)}{\sqrt{v^{i,i}(\theta)}} \leq 1 \qquad \text{if } m_i(\theta) = 0$$

$$\left(\frac{T}{2\ln\ln T}\right)^{1/2}\frac{m_{i,T}(\theta)}{\sqrt{v^{i,i}(\theta)}} > 1 \qquad \text{if } m_i(\theta) > 0$$

As $\sup_{\theta\in\Theta^I}\left|\widehat{v}^{i,i}(\theta) - v^{i,i}(\theta)\right| = o_p(1)$, it follows that

$$\lim_{T\to\infty}\Pr\left(\left(\frac{T}{2\ln\ln T}\right)^{1/2}\frac{m_{i,T}(\theta)}{\sqrt{v^{i,i}(\theta)}} > 1\right) = 0 \qquad \text{if } m_i(\theta) = 0$$

$$\lim_{T\to\infty}\Pr\left(\left(\frac{T}{2\ln\ln T}\right)^{1/2}\frac{m_{i,T}(\theta)}{\sqrt{v^{i,i}(\theta)}} > 1\right) = 1 \qquad \text{if } m_i(\theta) > 0.$$

Hence, as T gets large, only the moment conditions holding with equality contribute to the bootstrap limiting distribution, and the probability of eliminating a non-slack moment condition approaches zero. Further, given the block resampling scheme, for all i, $\mathrm{E}^*\left(\sqrt{T}\left(m^*_{i,T}(\theta) - m_{i,T}(\theta)\right)\right) =$

$O_p(l/T)$ and $\text{var}^*\left(\sqrt{T}\left(m^*_{i,T}(\theta)\right)\right) = \widehat{v}^{i,i}(\theta) + O_p\left(l/\sqrt{T}\right)$, where E^* and var^* denote the mean and variance operator under the probability law governing the resampling scheme. Since $l = o\left(\sqrt{T}\right)$, as $T \to \infty$, conditional on the sample,

$$\left(\frac{\left(m^*_{1,T}(\theta) - m_{1,T}(\theta)\right)}{\sqrt{\widehat{v}_1^{1,1}(\theta)}}, ..., \frac{\left(m^*_{2p,T}(\theta) - m_{2p,T}(\theta)\right)}{\sqrt{\widehat{v}_1^{2p,2p}(\theta)}}\right) \simeq N\left(0, \widehat{\Omega}_T(\theta)\right).$$

Hence, conditionally on the sample, and for all samples but a set of probability measures approaching zero, $\sup_{\theta\in\widehat{\Theta}^I_T} TQ_T(\theta)$ and $\sup_{\theta\in\widehat{\Theta}^I_T} TQ^*_T(\theta)$ have the same limiting distribution, and so $c^*_{\alpha,T} - c_{\alpha,T} = o_p(1)$. The statement in the Proposition then follows.

Proof of Proposition 4: Letting $\theta = \left(\phi, \overline{\delta}\right)$, we construct the optimal GMM estimator

$$\widehat{\theta}^d_T = \arg\min_\theta m_{d,T}(\theta)' \widehat{\Omega}_{dd,T}\left(\widetilde{\theta}^d_T\right)^{-1} m_{d,T}(\theta),$$

where $\widetilde{\theta}^d_T = \arg\min_\theta m_{d,T}(\theta)' m_{d,T}(\theta)$, and $\widehat{\Omega}_{dd,T}\left(\widetilde{\theta}^d_T\right)$ is the HAC sample variance of $\sqrt{T}m_{d,T}\left(\widetilde{\theta}^{(d)}_T\right)$.

If we knew $\theta^d = \left(\phi^d, \delta^d\right)$, the statement would follow by a similar argument as in the proof of Proposition 3, simply comparing $TQ^d_T\left(\theta^d\right)$ with the $(1-\alpha)$-percentile of the empirical distribution of $TQ^{*d}_T\left(\theta^d\right)$. However, as we do not know θ^d, we replace it with the optimal GMM estimator, $\widehat{\theta}^d_T$. Thus, the parameter estimation error term, $\sqrt{T}\left(\widehat{\theta}^d_T - \theta^d\right)$, contributes to the limiting distribution of the statistics, as it contributes to its variance. Hence, we need a bootstrap procedure which is able to properly mimic that contribution. Now, via usual mean value expansion,

$$\sqrt{T}m_{d,T}\left(\widehat{\theta}^d_T\right) = \sqrt{T}m_{d,T}\left(\theta^d\right) + D_{d,T}\left(\overline{\theta}^d_T\right)\sqrt{T}\left(\widehat{\theta}^d_T - \theta^d\right) \tag{35}$$

$$\sqrt{T}m_{c,T}\left(\widehat{\theta}^d_T\right) = \sqrt{T}m_{c,T}\left(\theta^d\right) + D_{c,T}\left(\overline{\theta}^d_T\right)\sqrt{T}\left(\widehat{\theta}^d_T - \theta^d\right) \tag{36}$$

with $\overline{\theta}^d_T \in \left(\widehat{\theta}^d_T, \theta^d\right)$, $D_{d,T}(\theta) = \nabla_\theta m_{d,T}(\theta)$ and $D_{c,T}(\theta) = \nabla_\theta m_{c,T}(\theta)$. From (35) it follows that

$$\begin{aligned}\text{avar}\left(\sqrt{T}m_{d,T}\left(\widehat{\theta}^d_T\right)\right) &= \text{avar}\left(\sqrt{T}m_{d,T}\left(\theta^d\right)\right) + \text{avar}\left(D_{d,T}\left(\widehat{\theta}^d_T\right)\sqrt{T}\left(\widehat{\theta}^d_T - \theta^d\right)\right) + \\ &\quad +2\,\text{acov}\left(\sqrt{T}m_{d,T}\left(\theta^d\right), D_{d,T}\left(\widehat{\theta}^d_T\right)\sqrt{T}\left(\widehat{\theta}^d_T - \theta^d\right)\right).\end{aligned} \tag{37}$$

The asymptotic variance of the moment conditions $\sqrt{T}m_T\left(\theta^d\right)$ can be estimated by the HAC sample variance of $\left[\sqrt{T}m_{d,T}\left(\widehat{\theta}^d_T\right)\ \sqrt{T}m_{c,T}\left(\widehat{\theta}^d_T\right)\right]$

$$\widehat{\Omega}_T\left(\widehat{\theta}^d_T\right) = \begin{bmatrix} \widehat{\Omega}_{dd,T}\left(\widehat{\theta}^d_T\right) & \widehat{\Omega}_{dc,T}\left(\widehat{\theta}^d_T\right) \\ \widehat{\Omega}_{cd,T}\left(\widehat{\theta}^d_T\right) & \widehat{\Omega}_{cc,T}\left(\widehat{\theta}^d_T\right) \end{bmatrix}.$$

Via a mean value expansion of the GMM first order conditions around θ^d,

$$\sqrt{T}\left(\widehat{\theta}_T^d-\theta^d\right)=-\widehat{B}_{d,T}D_{d,T}\left(\widehat{\theta}_T^d\right)'\widehat{\Omega}_{dd,T}\left(\widehat{\theta}_T^d\right)^{-1}\sqrt{T}m_{d,T}\left(\theta^d\right). \tag{38}$$

with

$$\widehat{B}_{d,T}=\left(D'_{d,T}\left(\widehat{\theta}_T^d\right)\widehat{\Omega}_{dd,T}\left(\widehat{\theta}_T^d\right)^{-1}D_{d,T}\left(\widehat{\theta}_T^d\right)\right)^{-1},$$

hence, given Assumptions 1–3, $\widehat{B}_{d,T}^{-1/2}\sqrt{T}\left(\widehat{\theta}_T^d-\theta^d\right)\xrightarrow{d}N\left(0,I_2\right)$. We define the estimator of the asymptotic variance of the moment conditions evaluated at the optimal GMM estimator $\sqrt{T}m_T\left(\widehat{\theta}_T^d\right)$ as

$$\widehat{V}_T\left(\widehat{\theta}_T^d\right)=\begin{bmatrix}\widehat{V}_{dd,T}\left(\widehat{\theta}_T^d\right) & \widehat{V}_{dc,T}\left(\widehat{\theta}_T^d\right)\\ \widehat{V}_{cd,T}\left(\widehat{\theta}_T^d\right) & \widehat{V}_{cc,T}\left(\widehat{\theta}_T^d\right)\end{bmatrix},$$

where the first entry can be computed using (37) and and (38), i.e.

$$\widehat{V}_{dd,T}\left(\widehat{\theta}_T^d\right)=\widehat{\Omega}_{dd,T}\left(\widehat{\theta}_T^d\right)-D_{d,T}\left(\widehat{\theta}_T^d\right)\widehat{B}_{d,T}D'_{d,T}\left(\widehat{\theta}_T^d\right)$$

Also,

$$\begin{aligned}\widehat{V}_{cc,T}\left(\widehat{\theta}_T^d\right) &= \widehat{\Omega}_{cc,T}\left(\widehat{\theta}_T^d\right)+D_{c,T}\left(\widehat{\theta}_T^d\right)\widehat{B}_{d,T}D'_{c,T}\left(\widehat{\theta}_T^d\right)\\ &\quad -\widehat{\Omega}_{cd,T}\left(\widehat{\theta}_T^d\right)\widehat{\Omega}_{dd,T}\left(\widehat{\theta}_T^d\right)^{-1}D_{d,T}\left(\widehat{\theta}_T^d\right)\widehat{B}_{d,T}D'_{c,T}\left(\widehat{\theta}_T^d\right)\\ &\quad -D_{c,T}\left(\widehat{\theta}_T^d\right)\widehat{B}_{d,T}D'_{d,T}\left(\widehat{\theta}_T^d\right)\widehat{\Omega}_{dd,T}\left(\widehat{\theta}_T^d\right)^{-1}\widehat{\Omega}_{cd,T}\left(\widehat{\theta}_T^d\right).\end{aligned}$$

Note that, for the computation of statistic we need only an estimate of the diagonal element, hence we do not need a closed-form expression for $\widehat{V}_{dc,T}\left(\widehat{\theta}_T^d\right)$. Let

$$V_{dd}\left(\theta^d\right)=\text{plim}_{T\to\infty}\widehat{V}_{dd,T}\left(\widehat{\theta}_T^d\right),\ \ V_{cc}\left(\theta^d\right)=\text{plim}_{T\to\infty}\widehat{V}_{cc,T}\left(\widehat{\theta}_T^d\right).$$

It is easy to see that $V_{dd}\left(\theta^d\right)$ is of rank $p-2$, while $V_{cc}\left(\theta^d\right)$ is of full rank p, hence the asymptotic variance covariance matrix $V\left(\theta^d\right)$ is of rank $2p-2$. However, this is not a problem, as we are only concerned with the elements along the main diagonal.
We now outline how to construct bootstrap critical values. The bootstrap counterpart of $TQ_T^d\left(\widehat{\theta}_T^d\right)$

writes as:

$$
\begin{aligned}
TQ_T^{*d}\left(\widehat{\theta}_T^{*d}\right) &= T\sum_{i=1}^{p}\left(\frac{m_{d,T}^{i*}\left(\widehat{\theta}_T^{*d}\right)-m_{d,T}^{i}\left(\widehat{\theta}_T^{d}\right)}{\sqrt{\widehat{v}^{i,i\,*}\left(\widehat{\theta}_T^{*d}\right)}}\right)^2 \\
&+ T\sum_{i=p+1}^{2p}\left[\frac{m_{c,T}^{i*}\left(\widehat{\theta}_T^{*d}\right)-m_{c,T}\left(\widehat{\theta}_T^{d}\right)}{\sqrt{\widehat{v}^{i,i*}\left(\widehat{\theta}_T^{*d}\right)}}\right]_-^2 1\left[m_{c,T}^{i}\left(\widehat{\theta}_T^{d}\right)\leq\sqrt{\widehat{v}^{i,i}\left(\widehat{\theta}_T^{d}\right)}\sqrt{2\ln\ln T/T}\right],
\end{aligned}
$$

where $m_T^*(\theta)$ denote the moment conditions computed using the resampled observations. Moreover, $\widehat{\theta}_T^{*d}$ is the bootstrap analog of $\widehat{\theta}_T^{d}$, given by

$$
\widehat{\theta}_T^{*d} = \arg\min_\theta\left(m_{d,T}^*(\theta)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)'\widehat{\Omega}_{dd,T}^*\left(\widetilde{\theta}_T^{*d}\right)^{-1}\left(m_{d,T}^*(\theta)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right),
$$

with $\widetilde{\theta}_T^{*d} = \arg\min_\theta\left(m_{d,T}^*(\theta)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)'\left(m_{d,T}^*(\theta)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)$, and

$$
\widehat{\Omega}_{dd,T}^*\left(\widetilde{\theta}_T^{*d}\right) = \frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{d,I_k+i}\left(\widetilde{\theta}_T^{*d}\right)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)\left(m_{d,I_k+j}\left(\widetilde{\theta}_T^{*d}\right)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)', \quad (39)
$$

where I_i is an independent, identically distributed discrete uniform random variable on $[0, T-l-1]$. Finally, $\widehat{v}^{i,i*}\left(\widehat{\theta}_T^{*d}\right)$ is the $i-$th element on the diagonal of of $\widehat{V}_T^*\left(\widehat{\theta}_T^{*d}\right)$, the bootstrap counterpart of $\widehat{V}_T\left(\widehat{\theta}_T^{d}\right)$, which is given by

$$
\widehat{V}_T^*\left(\widehat{\theta}_T^{*(d)}\right) = \begin{pmatrix} \widehat{V}_{dd,T}^*\left(\widehat{\theta}_T^{*d}\right) & \widehat{V}_{dc,T}^*\left(\widehat{\theta}_T^{*d}\right) \\ \widehat{V}_{cd,T}^*\left(\widehat{\theta}_T^{*d}\right) & \widehat{V}_{cc,T}^*\left(\widehat{\theta}_T^{*d}\right) \end{pmatrix}.
$$

As for the computation of the bootstrap critical values, we need only the elements along the main diagonal, below we report only the expressions for $\widehat{V}_{dd,T}^*\left(\widehat{\theta}_T^{*d}\right)$ and $\widehat{V}_{cc,T}^*\left(\widehat{\theta}_T^{*d}\right)$, which are

$$
\widehat{V}_{dd,T}^*\left(\widehat{\theta}_T^{*d}\right) = \widehat{\Omega}_{dd,T}^*\left(\widehat{\theta}_T^{*d}\right) - \widehat{D}_{d,T}^*\left(\widehat{\theta}_T^{*d}\right)\widehat{B}_{d,T}^*\widehat{D}_{d,T}^{*\prime}\left(\widehat{\theta}_T^{*d}\right),
$$

where

$$
\widehat{B}_{d,T}^* = \left(\widehat{D}_{d,T}^{*\prime}\left(\widehat{\theta}_T^{*d}\right)\widehat{\Omega}_{dd,T}^*\left(\widehat{\theta}_T^{*d}\right)^{-1}\widehat{D}_{d,T}^*\left(\widehat{\theta}_T^{*d}\right)\right)^{-1},
$$

with $\widehat{D}_{d,T}^*\left(\widehat{\theta}_T^{*d}\right) = \nabla_\theta m_{d,T}^*\left(\widehat{\theta}_T^{*d}\right)$ and where $\widehat{\Omega}_{dd,T}^*\left(\widehat{\theta}_T^{*d}\right)$ is defined as in (39), but with $\widetilde{\theta}_T^{*d}$ replaced by $\widehat{\theta}_T^{*d}$, also

$$\begin{aligned}\widehat{V}_{cc,T}^{*}\left(\widehat{\theta}_T^{*d}\right) &= \widehat{\Omega}_{cc,T}^{*}\left(\widehat{\theta}_T^{*d}\right)+\widehat{D}_{c,T}^{*}\left(\widehat{\theta}_T^{*d}\right)\widehat{B}_{d,T}^{*}\widehat{D}_{c,T}^{*\prime}\left(\widehat{\theta}_T^{*d}\right)\\ &-\widehat{\Omega}_{cd,T}^{*}\left(\widehat{\theta}_T^{*d}\right)\widehat{\Omega}_{dd,T}^{*}\left(\widehat{\theta}_T^{*d}\right)^{-1}\widehat{D}_{d,T}^{*}\left(\widehat{\theta}_T^{*d}\right)\widehat{B}_{d,T}^{*}\widehat{D}_{c,T}^{*\prime}\left(\widehat{\theta}_T^{*d}\right)\\ &-\widehat{D}_{c,T}^{*}\left(\widehat{\theta}_T^{*d}\right)\widehat{B}_{d,T}^{*}\widehat{D}_{d,T}^{*\prime}\left(\widehat{\theta}_T^{*d}\right)\widehat{\Omega}_{dd,T}^{*}\left(\widehat{\theta}_T^{*d}\right)^{-1}\widehat{\Omega}_{cd,T}^{*}\left(\widehat{\theta}_T^{*d}\right),\end{aligned}$$

with $\widehat{D}_{d,T}^{*}\left(\widehat{\theta}_T^{*d}\right)=\nabla_\theta m_{c,T}^{*}\left(\widehat{\theta}_T^{*d}\right)$ and

$$\widehat{\Omega}_{cc,T}^{*}\left(\widehat{\theta}_T^{*d}\right)=\frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{c,I_k+i}\left(\widehat{\theta}_T^{*d}\right)-m_{c,T}\left(\widehat{\theta}_T^{d}\right)\right)\left(m_{c,I_k+j}\left(\widehat{\theta}_T^{*d}\right)-m_{c,T}\left(\widehat{\theta}_T^{d}\right)\right)',$$

$$\widehat{\Omega}_{cd,T}^{*}\left(\widehat{\theta}_T^{*d}\right)=\frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{c,I_k+i}\left(\widehat{\theta}_T^{*d}\right)-m_{c,T}\left(\widehat{\theta}_T^{d}\right)\right)\left(m_{d,I_k+j}\left(\widehat{\theta}_T^{*d}\right)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)'.$$

We compute B bootstrap replication of $TQ_T^{*d}\left(\widehat{\theta}_T^{*d}\right)$, say $TQ_{T,1}^{*d}\left(\widehat{\theta}_T^{*d}\right),...,TQ_{T,B}^{*d}\left(\widehat{\theta}_T^{*d}\right)$, and compute the $(1-\alpha)-$th percentile of its empirical distribution, $c_{T,B,\alpha}^{*d}\left(\widehat{\theta}_T^{*d}\right)$. We now need to establish the first order validity of the suggested bootstrap critical values. Broadly speaking, we need to show that to (do not) reject H_0^d whenever $TQ_T^d\left(\widehat{\theta}_T^{d}\right)$ is larger than (smaller than or equal to) $c_{T,B,\alpha}^{*d}\left(\widehat{\theta}_T^{*d}\right)$ provides a test with asymptotic size α and unit asymptotic power. To this end, we show that, under H_0^d, $TQ_T^{*d}\left(\widehat{\theta}_T^{*d}\right)$ has the same limiting distribution as $TQ_T^d\left(\widehat{\theta}_T^{d}\right)$, conditionally on the sample, and for all samples except a set of probability measure approaching zero. On the other hand, under H_1^d, $TQ_T^{*d}\left(\widehat{\theta}_T^{*d}\right)$ has same limiting distribution as under the null, while $TQ_T^d\left(\widehat{\theta}_T^{d}\right)$ diverges to infinity.

Now, a mean value expansion of the bootstrap GMM first order conditions around $\widehat{\theta}_T^{d}$, gives

$$\sqrt{T}\left(\widehat{\theta}_T^{*d}-\widehat{\theta}_T^{d}\right)=-\widehat{B}_{d,T}^{*}\widehat{D}_{d,T}^{*}\left(\widehat{\theta}_T^{*d}\right)\widehat{\Omega}_{dd,T}^{*}\left(\widehat{\theta}_T^{*d}\right)\sqrt{T}\left(m_{d,T}^{*}\left(\widehat{\theta}_T^{d}\right)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)$$

and

$$\sqrt{T}\left(m_{d,T}^{*}\left(\widehat{\theta}_T^{*d}\right)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)=\sqrt{T}\left(m_{d,T}^{*}\left(\widehat{\theta}_T^{d}\right)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)+\widehat{D}_{d,T}^{*}\sqrt{T}\left(\widehat{\theta}_T^{*d}-\widehat{\theta}_T^{d}\right).$$

Recalling that $l=o\left(T^{1/2}\right)$, straigthforward arithmetics gives that

$$\mathrm{E}^{*}\left(\sqrt{T}\left(m_{d,T}^{*}\left(\widehat{\theta}_T^{*d}\right)-m_{d,T}\left(\widehat{\theta}_T^{d}\right)\right)\right)=O_p\left(\frac{l}{\sqrt{T}}\right)=o_p(1),$$

$$\text{var}^* \left(\sqrt{T} \left(m^*_{d,T} \left(\widehat{\theta}^{*d}_T\right) - m_{d,T} \left(\widehat{\theta}^d_T\right)\right)\right) = \widehat{V}_T \left(\widehat{\theta}^d_T\right) + O_p \left(\frac{l}{\sqrt{T}}\right) = \widehat{V}_T \left(\widehat{\theta}^d_T\right) + o_p(1),$$

and

$$\widehat{V}^*_T \left(\widehat{\theta}^{*d}_T\right) - \widehat{V}_T \left(\widehat{\theta}^d_T\right) = o_{p^*}(1).$$

Hence, $\sqrt{T} \left(m^*_{d,T} \left(\widehat{\theta}^{*d}_T\right) - m_{d,T} \left(\widehat{\theta}^d_T\right)\right)$ has the same limiting distribution under both hypotheses, and such a limiting distribution coincides with that of $TQ^d_T \left(\widehat{\theta}^d_T\right)$ under the null.
As for the moment conditions under commitment, note that they contribute to the limiting distribution only when $m^i_{c,T} \left(\widehat{\theta}^d_T\right) \leq \sqrt{\widehat{v}^{i,i} \left(\widehat{\theta}^d_T\right)} \sqrt{2 \ln \ln T/T}$, and hence they properly mimic the limiting distribution of

$$\sum_{i=p+1}^{2p} \frac{\left[m^i_{c,T} \left(\widehat{\theta}^d_T\right)\right]^2_-}{\hat{v}^{i,i}_T \left(\widehat{\theta}^d_T\right)}.$$

The statement in the Proposition then follows.

Proof of Proposition 5: The moment equalities implied by commitment do not depend on $\overline{\delta}$, and

$$\widehat{\phi}^c_T = \arg\min_{\phi} m_{c,T}(\phi)' \, \widehat{\Omega}_{dd,T} \left(\widetilde{\phi}^c_T\right)^{-1} m_{c,T}(\phi),$$

where $\widetilde{\phi}^c_T = \arg\min_\phi m_{c,T}(\phi)' \, m_{c,T}(\phi)$ and $\widehat{\Omega}_{dd,T} \left(\widetilde{\phi}^c_T\right)$ is the HAC sample variance of $\sqrt{T} m_{c,T} \left(\widetilde{\phi}^c_T\right)$.
Via mean value expansion

$$\begin{aligned}
\sqrt{T} m_{d,T} \left(\widehat{\phi}^c_T, \overline{\delta}\right) &= \sqrt{T} m_{d,T} \left(\phi^c, \overline{\delta}\right) + D_{d,T}(\overline{\phi}, \overline{\delta}) \sqrt{T} \left(\widehat{\phi}^c_T - \phi^c\right) \\
\sqrt{T} m_{c,T} \left(\widehat{\phi}^c_T\right) &= \sqrt{T} m_{c,T} \left(\phi^c\right) + D_{c,T}(\overline{\phi}) \sqrt{T} \left(\widehat{\phi}^c_T - \phi^c\right)
\end{aligned}$$

where $\overline{\phi} \in (\widehat{\phi}^c_T, \phi)$, $D_{d,T}(\phi, \overline{\delta}) = \nabla_\phi m_{d,T} \left(\phi, \overline{\delta}\right)$ and $D_{c,T}(\phi) = \nabla_\phi m_{c,T}(\phi)$. Expanding the GMM fisrt order condition around ϕ^c

$$\sqrt{T} \left(\widehat{\phi}^c_T - \phi^c\right) = -\widehat{B}_{c,T} D_{c,T} \left(\widehat{\phi}^c_T\right)' \, \widehat{\Omega}_{cc,T} \left(\widehat{\phi}^c_T\right)^{-1} \sqrt{T} m_{c,T}(\phi^c)$$

where

$$\widehat{B}_{c,T} = \left(\widehat{D}'_{c,T} \left(\widehat{\phi}^c_T\right) \widehat{\Omega}_{cc,T} \left(\widehat{\phi}^c_T\right)^{-1} \widehat{D}_{c,T} \left(\widehat{\phi}^c_T\right)\right)^{-1}$$

The asymptotic variance of the moment conditions $\sqrt{T} m_T \left(\phi^c, \overline{\delta}\right)$ can be estimated by the HAC sample variance of $\left[\sqrt{T} m_{d,T} \left(\widehat{\phi}^c_T, \overline{\delta}\right) \ \sqrt{T} m_{c,T} \left(\widehat{\phi}^c_T\right)\right]$

$$\widehat{\Omega}_T \left(\widehat{\phi}^c_T, \overline{\delta}\right) = \begin{bmatrix} \widehat{\Omega}_{dd,T} \left(\widehat{\phi}^c_T, \overline{\delta}\right) & \widehat{\Omega}_{dc,T} \left(\widehat{\phi}^c_T, \overline{\delta}\right) \\ \widehat{\Omega}_{cd,T} \left(\widehat{\phi}^c_T, \overline{\delta}\right) & \widehat{\Omega}_{cc,T} \left(\widehat{\phi}^c_T\right) \end{bmatrix}.$$

Along the same lines as in the proof of Proposition 4, we define the estimator of the asymptotic variance of the moment conditions evaluated at the optimal GMM estimator as

$$\widehat{V}_T\left(\widehat{\phi}_T^c,\overline{\delta}\right)=\begin{pmatrix}\widehat{V}_{dd,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right) & \widehat{V}_{dc,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right)\\ \widehat{V}_{cd,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right) & \widehat{V}_{cc,T}\left(\widehat{\phi}_T^c\right)\end{pmatrix},$$

where

$$\widehat{V}_{cc,T}\left(\widehat{\phi}_T^c\right)=\widehat{\Omega}_{cc,T}\left(\widehat{\phi}_T^c\right)-\widehat{D}_{c,T}\left(\widehat{\phi}_T^c\right)\widehat{B}_{c,T}\widehat{D}'_{c,T}\left(\widehat{\phi}_T^c\right).$$

Also,

$$\begin{aligned}\widehat{V}_{dd,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right) &= \widehat{\Omega}_{dd,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right)+\widehat{D}_{d,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right)\widehat{B}_{c,T}\widehat{D}'_{d,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right)\\ &\quad -\widehat{\Omega}_{cd,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right)\widehat{\Omega}_{cc,T}\left(\widehat{\phi}_T^c\right)^{-1}\widehat{D}_{c,T}\left(\widehat{\phi}_T^c\right)\widehat{B}_{c,T}\widehat{D}'_{d,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right)\\ &\quad -\widehat{D}_{d,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right)\widehat{B}_{c,T}\widehat{D}'_{c,T}\left(\widehat{\phi}_T^c\right)\widehat{\Omega}_{cc,T}\left(\widehat{\phi}_T^c\right)^{-1}\widehat{\Omega}_{cd,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right).\end{aligned}$$

Let

$$V_{cc}\left(\phi^c\right) = \mathrm{plim}_{T\to\infty}\widehat{V}_{cc,T}\left(\widehat{\phi}_T^c\right),\quad V_{dd}\left(\phi^c,\overline{\delta}\right) = \mathrm{plim}_{T\to\infty}\widehat{V}_{dd,T}\left(\widehat{\phi}_T^c,\overline{\delta}\right).$$

Again, it is easy to see that $V_{cc}\left(\phi^c\right)$ is of rank $p-1$, while $V_{dd}\left(\phi^c,\overline{\delta}\right)$ is of full rank p, hence the asymptotic variance covariance matrix $\widehat{V}_T\left(\widehat{\phi}_T^c,\overline{\delta}\right)$ is of rank $p-1$. The bootstrap counterpart of $\widehat{V}_T\left(\widehat{\phi}_T^c,\overline{\delta}\right)$ is given by

$$\widehat{V}_T^*\left(\widehat{\phi}_T^{*c},\overline{\delta}\right)=\begin{pmatrix}\widehat{V}^*_{dd,T}\left(\widehat{\phi}_T^{*c},\overline{\delta}\right) & \widehat{V}^*_{dc,T}\left(\widehat{\phi}_T^{*c},\overline{\delta}\right)\\ \widehat{V}^*_{cd,T}\left(\widehat{\phi}_T^{*c},\overline{\delta}\right) & \widehat{V}^*_{cc,T}\left(\widehat{\phi}_T^{*c}\right)\end{pmatrix}.$$

As for the computation of the bootstrap critical values, we need only the element among the main diagonal, below we report only the expressions for $\widehat{V}^*_{cc,T}\left(\widehat{\phi}_T^{*c}\right)$ and $\widehat{V}^*_{dd,T}\left(\widehat{\phi}_T^{*c},\overline{\delta}\right)$, which are

$$\widehat{V}^*_{cc,T}\left(\widehat{\phi}_T^{*c}\right)=\widehat{\Omega}^*_{cc,T}\left(\widehat{\phi}_T^{*c}\right)-\widehat{D}^*_{c,T}\left(\widehat{\phi}_T^{*c}\right)\widehat{B}^*_{c,T}\widehat{D}^{*\prime}_{c,T}\left(\widehat{\phi}_T^{*c}\right),$$

where

$$\widehat{B}^*_{c,T}=\left(\widehat{D}^{*\prime}_{c,T}\left(\widehat{\phi}_T^{*c}\right)\widehat{\Omega}^*_{cc,T}\left(\widehat{\phi}_T^{*c}\right)^{-1}\widehat{D}^*_{c,T}\left(\widehat{\phi}_T^{*c}\right)\right)^{-1}$$

and where

$$\widehat{\Omega}^*_{cc,T}\left(\widehat{\theta}_T^{*c}\right) = \frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{c,I_k+i}\left(\widehat{\phi}_T^{*c}\right)-m_{c,T}\left(\widehat{\phi}_T^{c}\right)\right)\left(m_{c,I_k+j}\left(\widehat{\phi}_T^{*c}\right)-m_{c,T}\left(\widehat{\phi}_T^{c}\right)\right)',$$

and

$$\begin{aligned}
\widehat{V}^*_{dd,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right) &= \widehat{\Omega}^*_{dd,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)+\widehat{D}^*_{d,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)\widehat{B}^*_{c,T}\widehat{D}^{*\prime}_{d,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)- \\
&\quad \widehat{\Omega}^*_{cd,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)\widehat{\Omega}^*_{cc,T}\left(\widehat{\phi}^{*c}_T\right)^{-1}\widehat{D}^*_{c,T}\left(\widehat{\phi}^{*c}_T\right)\widehat{B}^*_{c,T}\widehat{D}^{*\prime}_{d,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right) \\
&\quad -\widehat{D}^*_{d,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)\widehat{B}^*_{c,T}\widehat{D}^{*\prime}_{c,T}\left(\widehat{\phi}^{*c}_T\right)\widehat{\Omega}^*_{cc,T}\left(\widehat{\phi}^{*c}_T\right)^{-1}\widehat{\Omega}^*_{cd,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right),
\end{aligned}$$

with

$$\begin{aligned}
\widehat{\Omega}^*_{dd,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right) &= \frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{d,I_k+i}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)-m_{d,T}\left(\widehat{\phi}^{c}_T,\overline{\delta}\right)\right)\left(m_{d,I_k+j}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)-m_{d,T}\left(\widehat{\phi}^{c}_T,\overline{\delta}\right)\right)', \\
\widehat{\Omega}^*_{cd,T}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right) &= \frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{c,I_k+i}\left(\widehat{\phi}^{*c}_T\right)-m_{c,T}\left(\widehat{\phi}^{c}_T\right)\right)\left(m_{d,I_k+j}\left(\widehat{\phi}^{*c}_T,\overline{\delta}\right)-m_{d,T}\left(\widehat{\phi}^{c}_T,\overline{\delta}\right)\right)'
\end{aligned}$$

The rest of the proof then follows by the same argument used in the proof of Proposition 4.

References

Andrews, D.W.K. and P. Guggenberger (2009). Validity of Subsampling and "Plug-In Asymptotic" Inference for Parameters Defined by Moment Inequalities. *Econometric Theory*, 25, 669-709.

Andrews, D.W.K. and G. Soares (2010). Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection. *Econometrica*, 78, 119-158.

Ball L. and D. Romer (1990). Real Rigidities and the Non-Neutrality of Money. *Review of Economic Studies*, 57, 183-203.

Barro, R. and D. Gordon (1983). A Positive Theory of Monetary Policy in a Natural Rate Model. *Journal of Political Economy,* 91, 589-610.

Baxter, M. (1988). Toward an Empirical Assessment of Game-Theoretic Models of Policy-Making. *Carnegie-Rochester Conference Series on Public Policy*, 28, 141-152.

Benigno P. and M. Woodford (2010). Inflation Stabilization and Welfare: The Case of Distorted Steady State. *Journal of the European Economic Association,* forthcoming.

Beresteanu, A. and F. Molinari (2008). Asymptotic Properties for a Class of Partially Identified Models. *Econometrica,* 76, 763-814.

Blundell R., M. Browning and I. Crawford (2008). Best Nonparametric Bounds on Demand Responses. *Econometrica*, 76, 1227-1262

Bontemps, C., T. Magnac and E. Maurin (2007). Set Identified Linear Models. Working Paper. University of Toulouse.

Bugni, F. (2010). Bootstrap Inference in Partially Identified Models. *Econometrica*, 78, 735-753.

Calvo, G. (1983). Staggered Prices in a Utility-Maximizing Framework, *Journal of Monetary Economics*, 12, 383-98.

Chernozhukov, V., H. Hong and E. Tamer (2007). Estimation and Confidence Regions for Parameters Sets in Econometric Models. *Econometrica*, 75, 1243-1284.

Ciliberto, F., and E. Tamer (2009). Market Structure and Multiple Equilibria in Airline Markets. *Econometrica*, 77, 1791-1828.

Clarida, R. J. Gali and M. Gertler (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*, 37, 1661-1707.

Clarida, R. J. Gali and M. Gertler (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and some Theory. *Quarterly Journal of Economics*115, 147-180.

Cooley, T. F. and E. C. Prescott (1995). Economic Growth and Business Cycles., In *Frontiers of Business Cycle Research*, edited by Thomas F. Cooley, 1-38. Princeton University Press.

Corts, K.S. (1999). Conduct parameters and the measurement of market power. *Journal of Econometrics*, 88, 227–250.

Dixit, A.K. and J. Stiglitz (1977). Monopolistic Competition and Optimum Price Diversity. *American Economic Review,* 67, 297-308.

Galí, J. and M. Gertler (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44, 195-222.

Galí, J., M. Gertler and J.D. Lopez-Salido (2007). Markups, Gaps and the Welfare Costs of Business Cycle Fluctuations. *Review of Economics and Statistics*, 89:1, 44-59.

Galichon, A. and M. Henry (2010). Set Identification in Models with Multipla Equilibria. *Review of Economic Studies,* forthcoming.

Kirsanova, T. and S. le Roux (2011). Commitment vs Discretion in the UK: An Empirical Investigation of the Monetary and Fiscal Policy Regime. *Unpublished Manuscript.*

Givens, G. (2010). Estimating Central Bank Preferences under Commitment and Discretion. *Unpublished Manuscript.*

Haile, P., and E. Tamer (2003). Inference With an Incomplete Model of English Auctions. *Journal of Political Economy*, 111, 1-51.

Hall B. (2009). By How Much Does GDP Rise If the Government Buys More Output? *Brookings Papers on Economic Activity*, Fall 2009.

Halunga A., D. Osborn and M. Sensier (2009). Changes in the Order of Integration of US and UK Inflation. *Economic Letters*, 102, 3032.

Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029-1054.

Hansen, L. P., and K. J. Singleton (1982). Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models. *Econometrica*, 50, 1269-1286.

Ireland, P. N. (1999). Does the Time-Consistency Problem Explain the Behavior of Inflation in the United States?, *Journal of Monetary Economics,* 44, 279-291.

Kydland, F.E. and E.C. Prescott (1977). Rules rather than Discretion: The Inconsistency of Optimal Monetary Plans. *Journal of Political Economy*, 85(3), 473-491.

Manski C. F. and E. Tamer (2002). Inference on Regressions With Interval Data on a Regressor or Outcome. *Econometrica*, 70, 519-547.

McCallum, B. (1999). Issues in the design of monetary policy rules. *Handbook of Macroeconomics*, in: J. B. Taylor and M. Woodford (ed.), Handbook of Macroeconomics, edition 1, volume 1, chapter 23, pages 1483-1530.

Rosen, A. (2007). Identification and estimation of firms marginal cost functions with incomplete knowledge of strategic behavior. CeMMAP working paper CWP03/07.

Rosen, A. (2008). Confidence Sets for Partially Identified Parameters that Satisfy a Finite Number of Moment Inequalities. *Journal of Econometrics,* 146, 107-117.

Rotemberg, J.J. and M. Woodford (1997). An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy. *NBER Macroeconmics Annual* 12, 297-346.

Ruge-Murcia, F. (2003). Does the Barro-Gordon Model Explain the Behaviour of US Inflation? A Reexamination of the Empirical Evidence. *Journal of Monetary Economics.* 50, 1375-1390.

Sbordone, A. (2002). Prices and Unit Labor Costs: A New Test of Price Stickiness. *Journal of Monetary Economics* 49(2), 265-292.

Schaumburg, E. and A. Tambalotti (2007). An Investigation of the Gains from Commitment in Optimal Monetary Policy. *Journal of Monetary Economics*, 54, 302-324.

Shimer, R. (2009). Convergence in Macroeconomics: The Labor Wedge. *American Economic Journal: Macroeconomics*, 1, 280-297.

Stock J.H. and M. Yogo (2002). Testing for Weak Instruments in Linear IV Regression. NBER Technical Working Paper 0284, National Bureau of Economic Research, Inc.

Taylor, J. (1993). Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.

Woodford, M. (2001). The Taylor Rule and Optimal Monetary Policy. *American Economic Review* 91, No. 2, 232-237.

Woodford, M. (2003). Interest and prices: Foundations of a theory of monetary policy. Princeton University Press.

Woodford, M. (2010). Optimal Monetary Stabilization Policy. B.M. Friedman and M. Woodford, eds., *forthcoming Handbook of Monetary Economics, vol. 3B*.

Woodford, M. and M. P. Giannoni (2010). Optimal Target Criteria for Stabilization Policy. Mimeo, Columbia University.