Testing for optimal monetary policy via moment inequalities*

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

[†]University of York, Department of Economics and Related Studies

[‡]University of Surrey, School of Economics

September 18, 2017

Abstract

The specification of an optimizing model of the monetary transmission mechanism requires selecting a policy regime, commonly commitment or discretion. In this paper we propose a new procedure for testing optimal monetary policy, relying on moment inequalities that nest commitment and discretion as two special cases. The approach is based on the derivation of bounds for inflation that are consistent with optimal policy under either policy regime. We derive testable implications that allow for specification tests and discrimination between the two alternative regimes. The proposed procedure is implemented to examine the conduct of monetary policy in the United States economy.

JEL classification codes: C12; C52; E52; E58.

Keywords: *Bootstrap; GMM; Moment Inequalities; Optimal Monetary Policy.*

*We thank the Editor, Barbara Rossi, and three anonymous referees for their very constructive and insightful comments that helped improve the paper substantially. We are also grateful to Herman Bierens, Davide Debortoli, Jean-Marie Dufour, Alastair Hall, George Kapetanios, Lynda Khalaf, Maral Kichian, Tatiana Kirsanova, Paul Levine, Christian Mattes, Michael McMahon, Karel Mertens, Milan Nedeljkovic, Carlos Noton, Denise Osborn and Adam Rosen for useful comments. We also thank seminar participants at the University of Chile, the University of Exeter, the University of Lancaster, KU Leuven, the University of Manchester, Universitat Pompeu Fabra, the University of Warwick, the Bank of Serbia, the Federal Reserve Board, the Central Bank of Chile, the Bank of England, the 2011 Time Series Conference at CIREQ, the 2011 International Conference on Computing in Economics and Finance, the 2012 North America Winter Meeting of the Econometric Society, the 2013 Royal Economic Society Meeting, the MMF Annual Conference 2013 and the 2016 Conference on New Trends and Developments in Econometrics in the Central Bank of Portugal. Laura Coroneo gratefully acknowledges the support of the ESRC grant ES/K001345/1.

1 Introduction

This paper proposes new methods for the evaluation of monetary policy within 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 may fail to steer private sector expectations because, ex-post, past commitments are ignored. The theoretical literature has considered two alternative characterizations of optimal monetary policy: commitment, whereby the optimal plan is history dependent and the time-inconsistency problem is ignored; and discretion, whereby the monetary authority re-optimizes each period. We propose a method for estimating and testing a structural model of optimal monetary policy, without requiring an explicit choice of the relevant equilibrium concept. Our procedure considers a general specification, that nests optimal policy under commitment and discretion. The approach is based on the derivation of bounds for the inflation rate that are consistent with both forms of optimal policy and yield set identification of the economy structural parameters. We derive testable implications that allow for specification tests and discrimination between the monetary authority's modes of behavior.

Under discretion there exists a state-contingent inflation bias resulting from the fact that the monetary authority must set policy independently of the history of shocks (Svensson, 1997). 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.

More generally, our framework applies to the optimal linear regulator problem, and relies on state-contingent bounds for a target variable that are used to derive moment inequality conditions associated with optimal policy, and to identify the set of structural parameters for which the moment inequalities hold, i.e. the identified set. We characterize the identified set implied by optimal monetary policy using inference methods developed in Chernozhukov, Hong, and Tamer (2007). We then test whether the moment restrictions implied by a specific policy regime are satisfied.

Assuming a specific policy regime enables point identification of the underlying structural parameters. Thus, parameters can be consistently estimated and standard tests of overidentifying restrictions (Hansen, 1982) can be performed. 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 which explore 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 Chernozhukov et al. (2007) 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.

In addition, the moment inequality conditions implied by optimal monetary policy under discretion and commitment, respectively, can be used to perform a model selection test to discriminate between the two alternative policy regimes, maintaining the assumption of optimal monetary policy. Following Shi (2015), we compare the two models and select the one that is closer to the truth in terms of a pseudo-distance measure based on the Kullback-Leibler divergence measure.

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 the work of Rotemberg and Woodford (1997), Clarida, Galí, and Gertler (1999), and Woodford (2003). Using the sample period running from 1983Q1 until 2008Q3, we find evidence in favor of discretionary optimal monetary policy, and against commitment. In contrast, the standard J–test of overidentifying restrictions lacks power and fails to reject either policy regime.[1] Thus, by making use of the full set of implications of optimal monetary policy, we construct a more powerful model specification test, allowing the rejection of commitment but not discretion. This finding is further supported by the model selection test based on Shi (2015).

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 pioneering work, Baxter (1988)

[1]The lack of power of the J–test in the context of forward-looking models estimated using GMM is discussed in Mavroeidis (2005). Our test of the monetary policy regime explores a larger set of moment inequality restrictions implied by optimal monetary policy and, therefore, contributes to increasing the power of the specification test.

calls for the development of methods to analyze policy making in a maximizing framework, and says 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"*.[2] 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 policy, nesting the commitment and the discretion solutions as two special cases.

Using a full-information maximum-likelihood approach, Givens (2012) 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 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 (2013), who also find evidence in favor of discretion for monetary and fiscal policy in the UK. Debortoli and Lakdawala (2015) estimate a medium-scale DSGE model allowing for deviations from commitment plans that follow a regime switching process. They reject both the full commitment and the discretion model, which are nested special cases of their model. The partial identification framework that we propose in this paper also constitutes a general econometric specification that nests commitment and discretion as two special cases. Unlike full-information methods, our approach does not require strong assumptions about the nature of the forcing variables (shock processes).

Simple monetary policy rules are often prescribed as guides for the conduct of monetary policy. For instance, a commitment to a Taylor rule (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

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

of optimizing models. Woodford (2001) shows the response prescribed by these rules tends to stabilize inflation and the output gap, 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. Giannoni and Woodford (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 McCallum (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 testing procedure for hypotheses concerning these two alternative policy regimes.

This paper also contributes to a growing literature that proposes 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. 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 set-identification of predicted demand responses in the study of consumer behavior. Ciliberto and Tamer (2009) propose new methods for inference in entry games without requiring assumption about the equilibrium selection. Galichon and Henry (2011) 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 class of optimal linear regulator problems to which out framework applies. Section 3 derives the bounds for inflation implied by optimal monetary policy and outlines the inference procedure. Section 4 describes the proposed test for optimal monetary policy. Section 5 describes the model selection test. Section 6 presents Montecarlo evidence on the small sample performance of the tests. Finally, Section 7 reports the empirical findings and Section 8 concludes.

2 Optimal monetary policy

Our methodology applies to the optimal linear regulator problem obtained when the policymaker's objective function is quadratic and the structural equations describing the economy's equilibrium dynamics are linear. This framework is widely used to study optimal monetary policy in the New Keynesian model with staggered prices and monopolistic competition.[3] The objective function of the monetary authority, which in the canonical case is derived as a second order approximation to the utility of a stand-in agent around the stable equilibrium with zero inflation (Woodford, 2003), takes the form

$$\begin{aligned} \mathbf{U} &= \mathbb{E}_0 \left[-\frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left(y_t' \mathbf{W} y_t \right) \right], \\ &= \mathbb{E}_0 \left[-\frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left(\pi_t^2 + s_t' \mathbf{Q} s_t + x_t' \mathbf{R} x_t \right) \right], \end{aligned} \tag{1}$$

where $\mathbb{E}_t$ denotes agents's expectations at date t, $y_t = [\pi_t, s_t, x_t]'$ is a $n \times 1$ vector of endogenous variables with $n \geq 2$; π_t is a scalar random variable (the inflation rate in our benchmark example), s_t is an $m \times 1$ vector, with $m \geq 1$, and x_t is of dimension $(n - m - 1) \times 1$; $\beta \in (0, 1)$ is a parameter scalar representing the discount factor. The matrix $\mathbf{W}$ is a $n \times n$ symmetric positive semidefinite matrix containing the weights on the individual target variables, with the following block diagonal structure

$$\mathbf{W} = \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix}, \tag{2}$$

with $\mathbf{Q}$ and $\mathbf{R}$ conformable square matrices.

The constraints on possible equilibrium outcomes (the structural equations) are represented by the following m-dimensional system

$$\begin{aligned} u_t &= \mathbf{A} y_{t-1} + \mathbf{B} y_t + \beta \mathbf{C} \mathbb{E}_t \left(y_{t+1} \right), \\ &= \begin{bmatrix} \mathbf{a} & \mathbf{A}_s & \mathbf{A}_x \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ s_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{b} & \mathbf{B}_s & \mathbf{B}_x \end{bmatrix} \begin{bmatrix} \pi_t \\ s_t \\ x_t \end{bmatrix} + \beta \begin{bmatrix} \mathbf{c} & \mathbf{C}_s & \mathbf{C}_x \end{bmatrix} \mathbb{E}_t \begin{bmatrix} \pi_{t+1} \\ s_{t+1} \\ x_{t+1} \end{bmatrix}, \end{aligned} \tag{3}$$

[3] See Woodford (2003) for a detailed description of this class of structural models.

for all t, where u_t is a vector of exogenous disturbances, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are $m \times n$ matrices, $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are $m \times 1$ vectors, and $\mathbf{A}_s$, $\mathbf{A}_x$, $\mathbf{B}_s$, $\mathbf{B}_x$, $\mathbf{C}_s$ and $\mathbf{C}_x$ are conformable matrices.[4] In particular, $\mathbf{B}_s$ is an $m \times m$ square matrix.

In the sequel, we restrict attention to models admitting a representation such that $\mathbf{a} = \mathbf{0}$ and $\mathbf{A}_s = \mathbf{0}$, so that the vector of target variables $[\pi_t, s_t]'$ does not include predetermined variables, with all the endogenous predetermined variables included in x_t. Moreover, we require that $\mathbf{C}_s = 0$ and the matrix $\mathbf{B}_s$ to be nonsingular. These restrictions allow the Lagrange multipliers associated with each of the m constraints to be mapped into the contemporaneous values of s_t. Many influential models used to study optimal monetary policy satisfy these restriction. For example, Clarida et al. (1999), Svensson and Woodford (2004) and Giannoni and Woodford (2004) model of inflation inertia, all admit such representation.

The problem of the monetary authority under commitment is to choose bounded state contingent sequences $\{y_t\}_{t \geq 0}$ to maximize (1) subject to (3). The Lagrangian formulation of this problem is given by

$$\mathbb{E}_0 \left\{ -\frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left[\pi_t^2 + s_t' \mathbf{Q} s_t + x_t' \mathbf{R} x_t - \lambda_t' \left(\mathbf{A} y_{t-1} + \mathbf{B} y_t + \beta \mathbf{C} y_{t+1} \right) \right] \right\}, \tag{4}$$

where λ_t is a h-dimensional vector of Lagrange multipliers, with initial condition $\lambda_{-1} = \mathbf{0}$. The first order conditions solving the monetary authority's problem under commitment are

$$\pi_t - \mathbf{c}' \lambda_{t-1} - \mathbf{b}' \lambda_t = 0, \tag{5}$$

$$\mathbf{Q} s_t - \mathbf{B}_s' \lambda_t = \mathbf{0}, \tag{6}$$

$$\mathbf{R} x_t - \mathbf{C}_x' \lambda_{t-1} - \mathbf{B}_x' \lambda_t - \beta \mathbf{A}_x' \mathbb{E}_t \left(\lambda_{t+1} \right) = \mathbf{0}, \tag{7}$$

for all $t \geq 0$, together with the constraint (3) and the initial condition $\lambda_{-1} = \mathbf{0}$. From equation (5) the necessary conditions for optimal policy under commitment require that the target variable π_t satisfies the condition

$$\pi_t = \mathbf{c}' \lambda_{t-1} + \mathbf{b}' \lambda_t. \tag{8}$$

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 $\lambda_{t-1} = \mathbf{0}$, ignoring the impact of

[4] This formulation follows Dennis (2007) and Debortoli and Lakdawala (2015). The matrix $\mathbf{C}$ in the system of equations (3) is premultiplied by the discount factor β for convenience, without any loss of generality.

its current actions on the private sector expectations. Under discretion, the policymaker acts as if $\lambda_{t-1} = \mathbf{0}$, and the resulting path for target variable π_t satisfies the condition

$$\pi_t = \mathbf{b}'\lambda_t. \tag{9}$$

Finally, both under discretion and commitment, from equation (6) it is possible to obtain the Lagrange multipliers as follows

$$\begin{aligned} \lambda_t &= \mathbf{B}_s'^{-1}\mathbf{Q}s_t, \\ &= \mathbf{D}s_t, \end{aligned} \tag{10}$$

In what follows, we define the sublist of structural parameters $\phi = \{\mathbf{b}, \mathbf{c}, \mathbf{D}\}$, and let $\phi_0 = \{\mathbf{b}_0, \mathbf{c}_0, \mathbf{D}_0\}$ denote the "true" value of ϕ. In addition, we define $\pi_t^c(\phi_0)$ as the inflation in period t consistent with the first order conditions for optimal policy under commitment, given knowledge of s_t and the structural parameters in ϕ_0. In the same way, $\pi_t^d(\phi_0)$ is the inflation in period t consistent with the first order conditions under discretion. Making use of (8), (9) and (10), $\pi_t^c(\phi_0)$ and $\pi_t^d(\phi_0)$ are, respectively, given by

$$\pi_t^c(\phi_0) = \mathbf{c}_0'\mathbf{D}_0 s_{t-1} + \mathbf{b}_0'\mathbf{D}_0 s_t, \tag{11}$$

$$\pi_t^d(\phi_0) = \mathbf{b}_0'\mathbf{D}_0 s_t. \tag{12}$$

To model optimal monetary policy requires a decision about whether the first order conditions of the policy maker are represented by (11) or, instead, by (12). But how does one decide whether the behavior of the monetary authority should be classified as discretion or commitment-like? We propose a general characterization of optimal monetary policy nesting both modes of behavior. The approach is based on the derivation of bounds for the inflation rate under the maintained assumption that the monetary authority implements optimal monetary policy, in the sense that at any point in time either (11) or (12) is satisfied.

3 Bounds for inflation

Under a specific equilibrium concept, commitment or discretion, it is in principle possible to identify ϕ_0 from observed data for inflation and the output gap using, respectively, equation (11) or (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. *Consider an economy whose structural equations can be represented by the system* (3), *with* $\mathbf{a} = \mathbf{0}$, $\mathbf{A}_s = \mathbf{0}$, $\mathbf{C}_s = \mathbf{0}$ *and* $\mathbf{B}_s$ *a nonsingular matrix. Optimal policy implies that*

$$\Pr\left(\pi_t^c(\phi_0) \le \pi_t(\phi_0) \le \pi_t^d(\phi_0) \middle| \mathbf{c}_0'\mathbf{D}_0 s_{t-1} \le 0\right) = 1,$$
$$\Pr\left(\pi_t^d(\phi_0) \le \pi_t(\phi_0) \le \pi_t^c(\phi_0) \middle| \mathbf{c}_0'\mathbf{D}_0 s_{t-1} > 0\right) = 1,$$

where $\pi_t(\phi_0)$ *is the actual inflation rate in period* t.

The bounds for inflation in Lemma 1 follow immediately from equations (11) and (12).

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 measurement error with mean $\bar{\Pi}_0$, possibly different from zero, thus allowing for the presence of a trend in measured inflation.[5]

Assumption 1. *Let* $\pi_t(\phi_0)$ *be the actual inflation rate in period* t. *The observed inflation rate is* $\Pi_t = \pi_t(\phi_0) + v_t$, *where* v_t *has mean* $\bar{\Pi}_0$ *and variance* σ_v^2.

3.1 Moment inequalities

The upshot of Lemma 1 is that we are able to derive moment inequality conditions implied by optimal monetary policy and nesting commitment and discretion as two special cases. From Lemma 1 it is immediate to see that

$$\begin{aligned}\Pr\left(\pi_t^c(\phi_0) + v_t \le \Pi_t \le \pi_t^d(\phi_0) + v_t \middle| \mathbf{c}_0'\mathbf{D}_0 s_{t-1} \le 0\right) = 1,\\ \Pr\left(\pi_t^d(\phi_0) + v_t \le \Pi_t \le \pi_t^c(\phi_0) + v_t \middle| \mathbf{c}_0'\mathbf{D}_0 s_{t-1} > 0\right) = 1,\end{aligned} \tag{13}$$

which establishes a lower and upper bound for the observed inflation rate, Π_t.

We assume that enough is known about the structural parameters of the economy so that the sign of each element of ϕ_0 is known with certainty, and denote $\mathbf{S} = \mathbf{sign}(\mathbf{c}_0'\mathbf{D}_0)$ the $1 \times p$ vector

[5] For instance, studies for the United States estimate the overstatement of true inflation to be in the range of 0.5 to 2.0 percentage points per year (Bernanke and Mishkin, 1997).

which is obtained after applying the sign function to each element of $\mathbf{c}_0'\mathbf{D}_0$. Then, we define the p-dimensional vector $\mathbf{S}_t = \mathbf{S}' \circ s_t$, where $\circ$ denotes the Schur product (element by element vector multiplication). Next, we obtain $1(\mathbf{S}_{t-1} \leq 0)$ and $1(\mathbf{S}_{t-1} > 0)$, the indicator functions taking value one when, respectively, each element of $\mathbf{S}_{t-1}$ is non-positive and each element of $\mathbf{S}_{t-1}$ is positive, and zero otherwise. We are thus able to derive the following moment inequalities that are implied by optimal monetary policy

Proposition 1. *Under Assumption 1, the following moment inequalities*

$$\mathrm{E}\begin{bmatrix} -\left(\Pi_t - \mathbf{b}_0'\mathbf{D}_0 s_t - v_t\right) 1\left(\mathbf{S}_{t-1} \leq 0\right) \\ \left(\Pi_t - \mathbf{b}_0'\mathbf{D}_0 s_t - v_t\right) 1\left(\mathbf{S}_{t-1} > 0\right) \\ \left(\Pi_t - \mathbf{c}_0'\mathbf{D}_0 s_{t-1} - \mathbf{b}_0'\mathbf{D}_0 s_t - v_t\right) 1\left(\mathbf{S}_{t-1} \leq 0\right) \\ -\left(\Pi_t - \mathbf{c}_0'\mathbf{D}_0 s_{t-1} - \mathbf{b}_0'\mathbf{D}_0 s_t - v_t\right) 1\left(\mathbf{S}_{t-1} > 0\right) \end{bmatrix} \geq 0, \tag{14}$$

are implied by optimal monetary policy under either commitment or discretion, where $\{\mathbf{b}_0, \mathbf{c}_0, \mathbf{D}_0\}$ *denote the "true" structural parameter and* E *is the unconditional expectation operator.*

Proposition 1 follows immediately from (13) and the fact that $1(\mathbf{S}_{t-1} \leq 0) = 1$ is a sufficient condition for $\mathbf{c}_0'\mathbf{D}_0 s_{t-1} \leq 0$ and, similarly, that $1(\mathbf{S}_{t-1} > 0) = 1$ is a sufficient condition for $\mathbf{c}_0'\mathbf{D}_0 s_{t-1} > 0$.[6]

Next, we define the following set of instruments

Assumption 2. *Let Z_t denote a p-dimensional vector of instruments such that*

1. *Z_t has bounded support;*
2. $\mathrm{E}\left[v_t 1\left(\mathbf{S}_{t-1} \leq 0\right) Z_t\right] = \bar{\Pi}\mathrm{E}\left[1\left(\mathbf{S}_{t-1} \leq 0\right) Z_t\right]$, *and* $\mathrm{E}\left[v_t 1\left(\mathbf{S}_{t-1} > 0\right) Z_t\right] = \bar{\Pi}\mathrm{E}\left[1\left(\mathbf{S}_{t-1} > 0\right) Z_t\right]$;
3. $\mathrm{E}\left[\left(\Pi_t - \bar{\Pi}\right) 1\left(\mathbf{S}_{t-1} \leq 0\right) Z_t\right] \neq 0$, $\mathrm{E}\left[s_t 1\left(\mathbf{S}_{t-1} \leq 0\right) Z_t\right] \neq 0$, $\mathrm{E}\left[s_{t-1} 1\left(\mathbf{S}_{t-1} \leq 0\right) Z_t\right] \neq 0$.

Assumption 2.1. guarantees that, without loss of generality, the vector of instruments can be restricted to have positive support. Assumption 2.2 requires the instrumental variables to be uncorrelated with the measurement error v_t. Finally, Assumption 2.3 requires that the instruments are relevant.

[6] We construct the moment functions in 14 using $1(\mathbf{S}_{t-1} \leq 0)$ instead of $1(\mathbf{c}_0'\mathbf{D}_0 s_{t-1} \leq 0)$ to obtain moment functions which are differentiable in the parameters ϕ (and in fact linear given an appropriate reparameterization), thus avoiding complications to do with non-smooth moment functions. In particular, we explore the fact that $1(\mathbf{S}_{t-1} \leq 0) = 1$ is a sufficient condition for $1(\mathbf{c}_0'\mathbf{D}_0 s_{t-1} \leq 0)$=1.

3.2 The identified set

Given Assumption 2, the moment inequalities in Proposition 1 can be written as

$$\mathrm{E}\left[m_{d,t}\left(\phi_0,\bar{\Pi}_0\right)\right] \equiv \mathrm{E}\begin{bmatrix} -\left(\Pi_t-\bar{\Pi}_0-\mathbf{b}_0'\mathbf{D}_0 s_t\right)1\left(\mathbf{S}_{t-1}\leq 0\right)Z_t \\ \left(\Pi_t-\bar{\Pi}_0-\mathbf{b}_0'\mathbf{D}_0 s_t\right)1\left(\mathbf{S}_{t-1}>0\right)Z_t \end{bmatrix} \geq 0, \tag{15}$$

$$\mathrm{E}\left[m_{c,t}\left(\phi_0,\bar{\Pi}_0\right)\right] \equiv \mathrm{E}\begin{bmatrix} \left(\Pi_t-\bar{\Pi}_0-\mathbf{c}_0'\mathbf{D}_0 s_{t-1}-\mathbf{b}_0'\mathbf{D}_0 s_t\right)1\left(\mathbf{S}_{t-1}\leq 0\right)Z_t \\ -\left(\Pi_t-\bar{\Pi}_0-\mathbf{c}_0'\mathbf{D}_0 s_{t-1}-\mathbf{b}_0'\mathbf{D}_0 s_t\right)1\left(\mathbf{S}_{t-1}>0\right)Z_t \end{bmatrix} \geq 0. \tag{16}$$

We use $\theta = \left(\phi,\bar{\Pi}\right) \in \Theta$ to denote a representative value of the parameter space. The "true" underlying vector value of θ in the model is denoted θ_0 which, in general, is not point identified by the conditions (15) and (16). Thus, we define the identified set consistent with optimal monetary policy as follows

Definition 1. *Let $\theta = \left(\phi,\ \bar{\Pi}\right) \in \Theta$. The identified set is defined as*

$$\Theta^I \equiv \left\{\theta \in \Theta : \textit{such that } \mathrm{E}\left[m_t\left(\theta\right)\right] \geq 0\right\}$$

with $m_t\left(\theta\right) \equiv \left[m_{d,t}\left(\phi,\bar{\Pi}\right),\ \ m_{c,t}\left(\phi,\bar{\Pi}\right)\right]'$.

Under optimal monetary policy, Θ^I is never empty. From the linearity of the moment functions

$$E\left[m_t\left(\theta\right)\right] = E\left[m\left(\theta_0\right)\right] + \nabla_\theta m'\left(\theta-\theta_0\right), \tag{17}$$

where $\nabla_\theta m$ denotes the gradient of the moment functions. The first terms on the RHS of (17) is non-negative because of (15) and (16). Hence, by construction $\theta_0 \in \Theta^I$. On the other hand, Θ^I may be non-empty even if (13) does not hold. In fact, violation of (13) does not necessary imply a violation of (15) and/or (16). Thus, θ_0 may belong to the identified set even in the case of no optimal monetary policy. In this sense, a non-empty identified set, while necessary for optimal monetary policy, is not sufficient.

Although our moment inequalities are linear in the transformed parameter space $\tilde{\theta} = \left\{\phi^\dagger,\phi^\ddagger,\bar{\Pi}\right\}$, with $\phi^\dagger = \mathbf{c}'\mathbf{D}$ and $\phi^\ddagger = \mathbf{b}'\mathbf{D}$, our set-up is rather different from Bontemps, Magnac, and Maurin (2012). In their case, lack of point identification arises because one can observe only lower and upper bounds for the dependent variable. In our case, we observe Π_t, s_t and s_{t-1}, and lack of identification arises because we do not know which model generated the observed series. In particular, their characterization of the identified set relies on the boundedness of the intervals defined by the upper and lower bound of the observed variables, and thus does not necessarily apply to our set-up.[7] Beresteanu and Molinari (2008) random set approach also applies to models which are incomplete because the dependent variable and/or the regressors are interval-valued. For this reason, in the sequel we use the criterion function of Chernozhukov et al. (2007).

Before proceeding, notice that one may be tempted to reduce the moment inequalities (15) and (16) into a single moment equality condition, given by

$$\mathrm{E}\Big[\left(\Pi_t - \bar{\Pi}_0 - \varphi_t \mathbf{c}_0' \mathbf{D}_0 s_{t-1} - \mathbf{b}_0' \mathbf{D}_0 s_t\right) Z_t\Big] = 0,$$

where $\varphi_t \in \{0, 1\}$ is a random variable taking value 1 in the case of commitment and 0 in the case of discretion. If φ_t is degenerate, it may be treated as a fixed parameter φ and the model can be estimated by GMM, provided appropriate instruments are available. This is an application of the conduct parameter method (CPM) sometimes used in the industrial organization literature. But, this approach is problematic since optimal monetary policy is characterized by either commitment or discretion, and the standard regularity condition for consistency are violated.[8]

3.3 Preliminaries on inference

Before describing the model specification test in Section 4, we describe some preliminary notions related to inference on the identified set Θ^I. The basic idea underlying the specification tests is to use the bounds for the observed inflation rate derived above to generate a family of moment inequality conditions that are consistent with optimal policy. These moment inequality conditions may be used to obtain a criterion function whose set of minimizers is the estimated identified set. If the estimated identified set is non-empty, we construct the corresponding confidence region.

[7] In the Appendix E.3, included as supplementary material, we show how to adapt our framework to provide an interpretation of the model specification test based on the set-up developed by Bontemps et al. (2012).

[8] In the IO literature, the CPM methods is usually applied to obtain an average estimate of market power across segmented markets with different structures (Corts, 1999, for a discussion of existing applications). Only in this context it is possible to interpret an estimator for φ with continuous support.

We define the following $4p$ moment functions associated with (15) and (16)

$$\begin{aligned}
m_{i,d,t}^{-}\left(\phi,\bar{\Pi}\right) &= -\left(\Pi_t-\bar{\Pi}-\mathbf{b}'\mathbf{D}s_t\right)1\left(\mathbf{S}_{t-1}\leq 0\right)Z_t^i,\\
m_{i,d,t}^{+}\left(\phi,\bar{\Pi}\right) &= \quad\left(\Pi_t-\bar{\Pi}-\mathbf{b}'\mathbf{D}s_t\right)1\left(\mathbf{S}_{t-1}> 0\right)Z_t^i,\\
m_{i,c,t}^{-}\left(\phi,\bar{\Pi}\right) &= \quad\left(\Pi_t-\bar{\Pi}-\mathbf{c}'\mathbf{D}s_{t-1}-\mathbf{b}'\mathbf{D}s_t\right)1\left(\mathbf{S}_{t-1}\leq 0\right)Z_t^i,\\
m_{i,c,t}^{+}\left(\phi,\bar{\Pi}\right) &= -\left(\Pi_t-\bar{\Pi}-\mathbf{c}'\mathbf{D}s_{t-1}-\mathbf{b}'\mathbf{D}s_t\right)1\left(\mathbf{S}_{t-1}> 0\right)Z_t^i,
\end{aligned}$$

with Z_t^i the i^{th} element of Z_t. The corresponding sample moment functions are

$$\begin{aligned}
m_{i,d,T}^{-}\left(\phi,\bar{\Pi}\right) &= T^{-1}\sum_{t=1}^{T} m_{i,d,t}^{-}\left(\phi,\bar{\Pi}\right), & m_{i,d,T}^{+}\left(\phi,\bar{\Pi}\right) &= T^{-1}\sum_{t=1}^{T} m_{i,d,t}^{+}\left(\phi,\bar{\Pi}\right),\\
m_{i,c,T}^{-}\left(\phi,\bar{\Pi}\right) &= T^{-1}\sum_{t=1}^{T} m_{i,d,t}^{-}\left(\phi,\bar{\Pi}\right), & m_{i,c,T}^{+}\left(\phi,\bar{\Pi}\right) &= T^{-1}\sum_{t=1}^{T} m_{i,c,t}^{+}\left(\phi,\bar{\Pi}\right),
\end{aligned}$$

and are collected in the $4p$-dimensional vector of sample moment functions

$$m_T\left(\theta\right)=\begin{bmatrix}\left(m_{1,d,T}^{-}\left(\phi,\bar{\Pi}\right),\ldots,m_{p,d,T}^{-}\left(\phi,\bar{\Pi}\right)\right)'\\ \left(m_{1,d,T}^{+}\left(\phi,\bar{\Pi}\right),\ldots,m_{p,d,T}^{+}\left(\phi,\bar{\Pi}\right)\right)'\\ \left(m_{1,c,T}^{-}\left(\phi,\bar{\Pi}\right),\ldots,m_{p,c,T}^{-}\left(\phi,\bar{\Pi}\right)\right)'\\ \left(m_{1,c,T}^{+}\left(\phi,\bar{\Pi}\right),\ldots,m_{p,c,T}^{+}\left(\phi,\bar{\Pi}\right)\right)'\end{bmatrix}. \tag{18}$$

We let $m_{i,T}\left(\theta\right)$ denote the i-th element of $m_T\left(\theta\right)$, and define $V\left(\theta\right)$, the asymptotic variance of $\sqrt{T}m_T\left(\theta\right)$, and $\widehat{V}_T\left(\theta\right)$ the corresponding heteroscedasticity and autocorrelation consistent (HAC) estimator.[9] Finally, we impose the following assumption

Assumption 3. *The following conditions are satisfied*

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

[9] This is obtained as $\widehat{V}_T\left(\theta\right)=\frac{1}{T}\sum_{k=-s_T}^{s_T}\sum_{t=s_T}^{T-s_T}\lambda_{k,T}\left(m_t\left(\theta\right)-m_T\left(\theta\right)\right)\left(m_{t+k}\left(\theta\right)-m_T\left(\theta\right)\right)'$, s_T is a lag truncation parameter such that $s_T=o(T^{1/2})$ and $\lambda_{k,T}=1-k/\left(s_T+1\right)$.

The criterion function we use for the inferential procedure is

$$Q_T(\theta) = \sum_{i=1}^{4p} \frac{[m_{i,T}(\theta)]_-^2}{\widehat{v}^{i,i}(\theta)}, \tag{19}$$

where $[x]_- = x\,1(x \leq 0)$, and $\widehat{v}^{i,i}(\theta)$ is the $i-$th element on the diagonal of $\widehat{V}_T(\theta)$. The probability limit of $Q_T(\theta)$ is given by $Q(\theta) = \mathrm{p\,lim}_{\mathrm{T}\to\infty} Q_T(\theta)$. 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 $\theta \in \Theta^I$, where Θ^I is as in Definition 1. Under Assumptions 1–3 a consistent estimator of the identified set $\widehat{\Theta}_T^I$ can be obtained as

$$\widehat{\Theta}_T^I = \left\{\theta \in \Theta \text{ s.t. } TQ_T(\theta) \leq d_T^2\right\}, \tag{20}$$

where d_T satisfies the conditions $\sqrt{\ln\ln T}/d_T \to 0$ and $d_T/\sqrt{T} \to 0$. In Appendix B we show how to obtain an estimator for the identified set and construct a confidence region $C_T^{1-\alpha}$ that asymptotically contains the identified set Θ^I with probability $1-\alpha$.

4 Specification tests

The next step in our analysis is to test for the null hypothesis of discretion (commitment), taking into account the lower (upper) bound imposed by optimal monetary policy. 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 θ, then we have evidence in favor of discretion (commitment). The test consists of a two-step procedure: in the first step the structural parameters are estimated under either discretion or commitment; in the second step we test if the estimated parameters are in the identified set implied by optimal monetary policy under either discretion or commitment.

In the sequel we consider our benchmark application, the New Keynesian model with staggered prices and monopolistic competition that has become widely used to study optimal monetary policy.[10] As is well known, the optimizing model of staggered price-setting proposed by Calvo (1983) results in the following equation relating the inflation rate to the economy-wide real marginal

[10]See Appendix A for a more detailed description of the structural model.

cost and expected inflation

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \psi s_t + u_t, \tag{21}$$

where ψ and β are positive parameters related to technology and preferences, π_t is the inflation rate, s_t the real marginal cost in deviation from the flexible-price steady state, and u_t is an exogenous stochastic shock resulting from time-varying markups and other distortions.

The objective function of the monetary authority is derived as a second order approximation to the utility of a stand-in agent around the stable equilibrium associated with zero inflation, and takes the form

$$\mathbf{U} = \mathbb{E}_0 \left[-\frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left(\pi_t^2 + \zeta s_t^2 \right) \right], \tag{22}$$

with α a positive parameter that relates to technology and preferences. Thus, in the benchmark model we obtain $\mathbf{b} = 1$, $\mathbf{B}_s = -\psi$, $\mathbf{c} = -1$, $\mathbf{Q} = \alpha$, and $\mathbf{D} = -(\zeta/\psi)$, and the moment inequality conditions that characterize optimal monetary policy corresponding to (15) and (16) specialize as follows

$$\mathrm{E}\left[m_{d,t}\left(\mathbf{D}, \bar{\Pi}\right)\right] \equiv \mathrm{E}\begin{bmatrix} -\left(\Pi_t - \bar{\Pi} - \mathbf{D} s_t\right) 1\left(s_{t-1} \leq 0\right) Z_t \\ \left(\Pi_t - \bar{\Pi} - \mathbf{D} s_t\right) 1\left(s_{t-1} > 0\right) Z_t \end{bmatrix} \geq 0, \tag{23}$$

$$\mathrm{E}\left[m_{c,t}\left(\mathbf{D}, \bar{\Pi}\right)\right] \equiv \mathrm{E}\begin{bmatrix} \left(\Pi_t - \bar{\Pi} - \mathbf{D} \Delta s_t\right) 1\left(s_{t-1} \leq 0\right) Z_t \\ -\left(\Pi_t - \bar{\Pi} - \mathbf{D} \Delta s_t\right) 1\left(s_{t-1} > 0\right) Z_t \end{bmatrix} \geq 0, \tag{24}$$

with $\theta = \left\{\mathbf{D}, \bar{\Pi}\right\}$, the parameter space. In more general applications, the parameter vectors in $\mathbf{b}$ and $\mathbf{c}$ may be unknown, and $\theta = \left\{\mathbf{b}, \mathbf{c}, \mathbf{D}, \bar{\Pi}\right\}$. In such cases, $\mathbf{b}$ and $\mathbf{c}$ may be pre-estimated from the system (3) as they are invariant across policy regimes, and the covariance estimator $\widehat{V}_T(\theta)$ needs to capture the estimation error due to the estimators $\widehat{\mathbf{b}}$ and $\widehat{\mathbf{c}}$.

4.1 Testing for discretion

If the monetary authority implements optimal policy under discretion the joint path of actual inflation and the economy-wide real marginal cost satisfies the moment conditions

$$\mathrm{E}\left[m_{d,t}^0\left(\theta_0\right)\right] = \mathrm{E}\left[\left(\Pi_t - \bar{\Pi}_0 - \mathbf{D}_0 s_t\right) Z_t\right] = 0, \tag{25}$$

$$\mathrm{E}\left[m_{c,t}\left(\theta_0\right)\right]=\mathrm{E}\left[\begin{array}{c}\left(\Pi_t-\bar{\Pi}_0-\mathbf{D}_0\Delta s_t\right)1\left(\mathbf{S}_{t-1}\leq 0\right)Z_t\\ -\left(\Pi_t-\bar{\Pi}_0-\mathbf{D}_0\Delta s_t\right)1\left(\mathbf{S}_{t-1}>0\right)Z_t\end{array}\right]\geq 0, \tag{26}$$

with m^0 denoting the moment functions that do not include the indicator on s_{t-1}. The moment equality conditions in (25) follow from the assumption of discretion and the moment inequality conditions (26) impose a lower bound to the observed inflation rate as implied by optimal monetary policy. As already mentioned, conditions (25) point identify θ_0, provided we can find at least one instrument, in addition to the intercept, satisfying Assumption 2. We define the following test for optimal monetary policy under discretion.

Definition 2. *Let* $\theta_0 \equiv \left(\mathbf{D}_0,\ \bar{\Pi}_0\right) \in \Theta$. *We define the null hypothesis of discretion and optimal monetary policy as,*

$$H_0^d:\ \theta_0 \textit{ satisfies conditions } (25)\text{–}(26),$$

against the alternative $H_1^d:\ \theta_0$ *does not satisfy conditions* (25)–(26).

To test the null hypothesis of discretion we follow a two-step procedure. Under the null hypothesis, the structural parameter vector θ_0 is point-identified and it can be consistently estimated via the optimal GMM estimator using the moment conditions (25).[11] Thus, to test the null hypothesis of discretion we first obtain an estimate for the structural parameter vector using the optimal GMM estimator, denoted $\widehat{\theta}_d$. In the second step, we construct the following test statistic

$$TQ_T^d\left(\widehat{\theta}_d\right)=T\left[\sum_{i=1}^{p}\frac{m_{i,d,T}^0\left(\widehat{\theta}_d\right)^2}{\widehat{v}^{i,i}\left(\widehat{\theta}_d\right)}+\sum_{i=1}^{2p}\frac{\left[m_{i,c,T}\left(\widehat{\theta}_d\right)\right]_-^2}{\widehat{v}^{i,i}\left(\widehat{\theta}_d\right)}\right], \tag{27}$$

where $\widehat{v}^{i,i}\left(\widehat{\theta}_d\right)$ is the i-th diagonal element of $\widehat{V}_T\left(\widehat{\theta}_d\right)$, the HAC estimator of the asymptotic variance of $\sqrt{T}\left[m_{d,T}^0\left(\widehat{\theta}_d\right), m_{c,T}\left(\widehat{\theta}_d\right)\right]$, which takes into account the estimation error in $\widehat{\theta}_d$.[12]

Notice that since the first p moment conditions hold with equality, they all contribute to the asymptotic distribution of $TQ_T^d\left(\widehat{\theta}_d\right)$. Thus, we apply the Generalized Moment Selection (GMS)

[11] If we assume that θ_0 satisfies (25)–(26), then it is possible to obtain an estimator using the approach of Moon and Schorfheide (2009), who consider the case in which the set of moment equalities point identify the parameters of interest, and use the additional information provided by the set of moment inequalities to improve efficiency. However, our objective is to test whether there exists θ_0 satisfying (25)–(26).

[12] See Appendix C for the definition of $\widehat{V}_T\left(\widehat{\theta}_d\right)$.

procedure introduced by Andrews and Soares (2010) only to the inequality conditions.[13] Andrews and Soares (2010) study the limiting distribution of the statistic in (27) evaluated at a fixed θ. In our case, due to the two-step testing procedure, we need to take into account the contribution of the estimation error to the asymptotic variance of the moment conditions, and compute bootstrap critical values that properly mimic the contribution of parameter estimation error. The first order validity of the bootstrap percentiles is established in the following Proposition.

Proposition 2. *Let Assumptions 1, 2 and 3 hold. Let $c^d_{B,\alpha}$ be the $(1-\alpha)$ percentile of the empirical distribution of $TQ^{*d}_T\left(\widehat{\theta}^*_d\right)$, the bootstrap counterpart of $TQ^d_T\left(\widehat{\theta}_d\right)$, which is defined in the proof of the Proposition. Then, as $T\to\infty$, $B\to\infty$, $l\to\infty$, and $l^2/T\to 0$, we have that:*

(i) under H^d_0, $\limsup_{T,B\to\infty}\Pr\left(TQ^d_T\left(\widehat{\theta}_d\right)>c^d_{B,\alpha}\right)=\alpha$,

(ii) under H^d_1, $\lim_{T,B\to\infty}\Pr\left(TQ^d_T\left(\widehat{\theta}_d\right)>c^d_{B,\alpha}\right)=1$,

where B denotes the number of bootstrap replications.

4.2 Testing for commitment

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

$$\mathrm{E}\left[m_{d,t}\left(\theta_0\right)\right]=\mathrm{E}\begin{bmatrix}-\left(\Pi_t-\bar{\Pi}_0-\mathbf{D}_0 s_t\right)1\left(\mathbf{S}_{t-1}\leq 0\right)Z_t\\ \left(\Pi_t-\bar{\Pi}_0-\mathbf{D}_0 s_t\right)1\left(\mathbf{S}_{t-1}>0\right)Z_t\end{bmatrix}\geq 0, \tag{28}$$

$$\mathrm{E}\left[m^0_{c,t}\left(\theta_0\right)\right]=\mathrm{E}\left[\left(\Pi_t-\bar{\Pi}_0-\mathbf{D}_0\Delta s_t\right)Z_t\right]=0, \tag{29}$$

where the moment equality condition (29) follows from the assumption of commitment and the moment inequality condition (28) imposes an upper bound to the observed inflation rate, as implied by optimal monetary policy. We define the following test for optimal policy under commitment.

Definition 3. *Let $\theta_0\equiv\left(\mathbf{D}_0,\ \bar{\Pi}_0\right)\in\Theta$. We define the null hypothesis of commitment and optimal monetary policy as,*

$$H^c_0:\ \theta_0 \textit{ satisfies conditions } (28)\text{–}(29).$$

[13] See Appendix B for details on the GMS method.

against the alternative H_1^c *:* θ_0 *does not satisfy conditions* (28)–(29).

The test of optimal monetary policy under commitment has the same structure as the test under discretion, with an analogous test statistic, given by

$$TQ_T^c\left(\widehat{\theta}_c\right) = T\left[\sum_{i=1}^{2p}\frac{\left[m_{i,d,T}\left(\widehat{\theta}_c\right)\right]_-^2}{\widehat{v}^{i,i}\left(\widehat{\theta}_c\right)} + \sum_{i=1}^{p}\frac{m_{i,c,T}^0\left(\widehat{\theta}_c\right)^2}{\widehat{v}^{i,i}\left(\widehat{\theta}_c\right)}\right], \tag{30}$$

with $\widehat{\theta}_c$ the optimal GMM estimator under commitment. We establish the following Proposition.

Proposition 3. *Let Assumptions 1, 2 and 3 hold. Let* $c_{\alpha,B}^c$ *be the* $(1-\alpha)$ *percentile of the empirical distribution of* $TQ_T^{*c}\left(\widehat{\theta}_c^*\right)$*, the bootstrap counterpart of* $TQ_T^c\left(\widehat{\theta}_c\right)$*. Then, as* $T \to \infty$*,* $B \to \infty$*,* $l \to \infty$*, and* $l^2/T \to 0$*, we have that:*

(i) under H_0^c*,* $\lim \sup_{T,B\to\infty} \Pr\left(TQ_T^c\left(\widehat{\theta}_c\right) > c_{\alpha,B}^c\right) = \alpha$*,*

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

where B *denotes the number of bootstrap replications.*

5 Model selection

The moment conditions (25), (26), (28) and (29) can also be used to perform model selection tests to discriminate between discretion and commitment, maintaining the assumption of optimal monetary policy. Following Shi (2015), we construct a quasi likelihood ratio test for the null hypothesis that both models (discretion and commitment) are equally close to the true data. If the null hypothesis is rejected, we select the one closer to the true model in terms of a pseudo-distance measure.

We test the following null hypothesis

$$H_0 : d(\mathcal{D}, \mu) = d(\mathcal{C}, \mu), \tag{31}$$

against the alternative $H_1 : d(\mathcal{D}, \mu) < d(\mathcal{C}, \mu)$, where $\mathcal{D}$ is the model for discretion and optimal policy in (25) and (26), $\mathcal{C}$ is the model for commitment and optimal policy in (28) and (29), and

μ is the true model. To test the null hypothesis in (31), we construct the test statistic

$$QLR_T = \max_{\theta \in \Theta} \frac{1}{T} \sum_{t=1}^{T} \mathcal{M}_t^d \left(\theta, \widehat{\gamma}_d \left(\theta\right)\right) - \max_{\theta \in \Theta} \frac{1}{T} \sum_{t=1}^{T} \mathcal{M}_t^c \left(\theta, \widehat{\gamma}_c \left(\theta\right)\right), \tag{32}$$

where $\mathcal{M}_t^d \left(\theta, \gamma \left(\theta\right)\right) = \exp \left(\gamma \left(\theta\right)' \begin{bmatrix} m_{d,t}^0 \left(\theta\right) \\ m_{c,t} \left(\theta\right) \end{bmatrix} \right)$, $\mathcal{M}_t^c \left(\theta, \gamma \left(\theta\right)\right) = \exp \left(\gamma \left(\theta\right)' \begin{bmatrix} m_{c,t}^0 \left(\theta\right) \\ m_{d,t} \left(\theta\right) \end{bmatrix} \right)$, and

$$\widehat{\gamma}_i \left(\theta_i\right) = \arg \min_{\gamma \in \mathbb{R}^p \times \mathbb{R}_+^{2p}} T^{-1} \sum_{t=1}^{T} \mathcal{M}_t^i \left(\theta_i, \gamma\right), \tag{33}$$

with $i \in \{d, c\}$. In turn, the pseudo true set of parameters can be estimated as

$$\widehat{\Theta}_T^i = \arg \max_{\theta \in \Theta} T^{-1} \sum_{t=1}^{T} \mathcal{M}_t^i \left(\theta, \widehat{\gamma}_i \left(\theta\right)\right). \tag{34}$$

Under discretion and optimal policy the conditions (25) and (26) point identify $\widehat{\theta}_T^d$ and, similarly, under commitment and optimal policy, the conditions (28) and (29) point-identify $\widehat{\theta}_T^c$. Having the estimated parameters, we define

$$T \widehat{\omega}_T^2 \left(\widehat{\theta}_d, \widehat{\theta}_c\right) = \sum_{k=-s_T}^{s_T} \sum_{t=s_T}^{T-s_T} \lambda_{k,T} \left(\Delta_t - \bar{\Delta}\right) \left(\Delta_{t+k} - \bar{\Delta}\right)', \tag{35}$$

with $\Delta_t = \mathcal{M}_t^d \left(\widehat{\theta}_d, \widehat{\gamma}_d \left(\widehat{\theta}_d\right)\right) - \mathcal{M}_t^c \left(\widehat{\theta}_c, \widehat{\gamma}_c \left(\widehat{\theta}_c\right)\right)$, $\bar{\Delta} = T^{-1} \sum_{t=1}^{T} \Delta_t$, and where s_T and $\lambda_{k,T}$ are as defined in footnote 9.

Shi (2015) shows that under H_0 we have that

$$\sqrt{T} \frac{QLR_T}{\widehat{\omega}_T \left(\widehat{\theta}_d, \widehat{\theta}_c\right)} \rightarrow_d N(0, 1), \tag{36}$$

and, therefore, we reject the null hypothesis (31) in favor of the alternative at the $1 - \alpha$ level if $\sqrt{T} QLR_T / \widehat{\omega}_T \left(\widehat{\theta}_d, \widehat{\theta}_c\right) > z_\alpha$, where z_α is the $1 - \alpha$ quantile of the standard normal distribution.

6 Montecarlo experiments

In this section, we perform Montecarlo simulations to analyze the small sample properties of the model specification test presented in Section 4. The data generating process (DGP) used in the Montecarlo experiment is described in Appendix D. We simulate 1,000 vectors of time-series, each with 1,100 observations and we discard the first 1,000 observations to eliminate the influence of the initial values. The resulting time-series length is 100, which is similar to our empirical application and a typical sample size in empirical studies of monetary policy using quarterly observations. We consider both discretion and commitment, and we seek to analyze the size and power properties of the tests described in Propositions 2 and 3. We also examine how the performance of the proposed tests varies with the strength of the instruments, by varying the length of the lags used as instruments. In particular, the instrumental variables used in the Monte Carlo are lagged values of inflation and the labor income share, and we look at the performance of the test when the instrument list includes the lags: $(t-2, t-3, t-4)$; $(t-3, t-4, t-5)$; $(t-4, t-5, t-6)$; and $(t-5, t-6, t-7)$.

For each sample, we obtain the critical values $c^d_{\alpha,B}$ and $c^c_{\alpha,B}$ following the bootstrap procedure described in Propositions 2 and 3. Table 1 reports the percentage of times the null hypothesis is rejected, obtained from the critical values based on the nominal level $\alpha = 0.10$. The results show that our test is undersized, rejecting the true DGP about 1% of the times instead of 10%. The power properties of the test are good. When the DGP is discretion, the false model is rejected very frequently, with rejection rates ranging between 99%, when instruments are strong, and 33%, for the weakest set of instruments. In turn, when the DGP is commitment, the rejection rates for the false model range between 76% and 10%. As expected, the power of the test declines as the instruments become weaker. However, the test is found to still perform well when very long lags are used as instrumental variables.

7 Empirical application

In this section, we apply the model specification and model selection tests proposed above to study the monetary policy in the United States since the start of the 1980s, and using the benchmark model of optimal monetary policy described in Section 6. The sample spans a period in which

Table 1: Monte Carlo experiments: rejection rates (nominal level $\alpha = 0.10$)

	DGP: Discretion ($T = 250$)		
Instrument lags:	$t-1 \dots t-3$	$t-2 \dots t-4$	$t-3 \dots t-5$
H_0: discretion	0.118	0.112	0.126
H_0: commitment	1.000	1.000	0.992
	DGP: Commitment ($T = 250$)		
Instrument lags:	$t-1 \dots t-3$	$t-2 \dots t-4$	$t-3 \dots t-5$
H_0: discretion	0.980	0.954	0.926
H_0: commitment	0.148	0.192	0.174
	DGP: Discretion ($T = 500$)		
Instrument lags:	$t-1 \dots t-3$	$t-2 \dots t-4$	$t-3 \dots t-5$
H_0: discretion	0.086	0.092	0.122
H_0: commitment	1.000	1.000	1.000
	DGP: Commitment ($T = 500$)		
Instrument lags:	$t-1 \dots t-3$	$t-2 \dots t-4$	$t-3 \dots t-5$
H_0: discretion	1.000	0.994	1.000
H_0: commitment	0.126	0.160	0.152

The table reports the rejection rates of the test statistics TQ_T, in (27) and (30), with 10% nominal level. Each Monte Carlo simulation has T observations and "burn-in" sample of size 1,000. The critical values $c^d_{\alpha,B}$ and $c^c_{\alpha,B}$ are based on 500 block-bootstrap replications of block size 4.

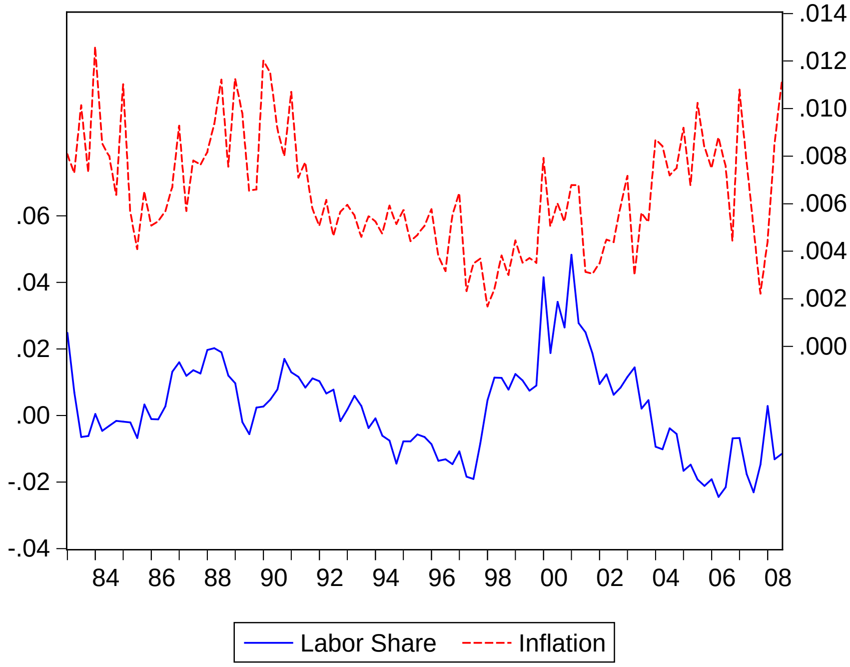
monetary policy has been perceived as good (Clarida, Gali, and Gertler, 2000).[14]

7.1 Data and sample

We use quarterly time-series for the US economy over the sample period 1983Q1 to 2008Q3. Following Galí and Gertler (1999) and Sbordone (2002), we exploit the proportional relationship between the output gap and the labor income share (equivalently, real unit labor costs). Hence, we use the labor income share in the non-farm business sector, detrended using a quadratic polynomial, to measure s_t. The measure of inflation is the percentage change in the GDP deflator.

[14]The term "good" is used loosely to describe a period in which monetary policy is consistent with achieving stable and low inflation. Clarida et al. (2000) argue that this is due to a stronger systematic reaction of monetary policy to changes in expected inflation.

Figure 1: labor share and inflation in the US, 1983Q1–2008Q3.



The econometric framework developed in this paper is for stationarity data (see Assumption 3). Halunga, Osborn, and Sensier (2009) show that there is a change in inflation persistence from $I(1)$ to $I(0)$ dated at June 1982. This result is related to the study of Lubik and Schorfheide (2004) who estimate a structural model of monetary policy for the US using full-information methods, and find that only after 1982 the estimated interest-rate feedback rule that characterizes monetary policy is consistent with equilibrium determinacy. Moreover, following the analysis in Clarida et al. (2000), we study the sample starting from 1983Q1, that removes the first three years of the Volcker era. Clarida et al. (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 1979Q4 – 1982Q4, the operating procedures of the Federal Reserve involved targeting non-borrowed reserves as opposed to the Federal Funds rate. Thus, our empirical analysis focuses on the sample period 1983Q1 to 2008Q3, 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.[15] Figure 1 plots the time-series of the US labor income share and inflation for the sample period 1983Q1 to 2008Q3.

[15]After 2008Q3, 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.

Following standard practice in the literature (see, for example, Galí and Gertler, 1999), we include in the instrument set lagged values of the labor income share and inflation, assumed to be orthogonal to the measurement error in inflation (Assumption 2.2). The instrument set used comprises the first, second and third-order lags of the labor income share and inflation. These instrumental variables are adjusted using the transformation $Z_+ = Z - \min(Z)$, guaranteeing positiveness. Notice that Assumption 2.1 guarantees this transformation always exists.

The complete instrument set also includes the unit vector, yielding $p = 7$ instruments and 28 moment conditions overall. Of course, weak instruments are a potential problem given that the first step in our test requires a consistent estimator of the structural parameter vector θ_0. For example, when we are testing discretion we require that

$$\mathrm{E}\left[m_{d,t}^{0}(\theta_0)\right] = \mathrm{E}\left[\left(\Pi_t - \bar{\Pi}_0 - \mathbf{D}_0 s_t\right) Z_t\right] = 0, \tag{37}$$

holds at the "true" value $\theta_0 = \left(\mathbf{D}_0, \bar{\Pi}_0\right)'$ and no other value of θ. If the instrument is irrelevant, in the sense that the correlation between Π_t and Z_t is zero (or weakly different from zero), then $\theta_0 = \left(\mathbf{D}_0, \bar{\Pi}_0\right)'$ is not identified since, given $\bar{\Pi}_0$, any value of $\mathbf{D}$ satisfies the moment condition. Thus, instrument relevance requires the correlation between Π_t and Z_t to be strong, as indicated by Assumption 2.3.

Reassuringly, the instruments used (which include lags of inflation and of the real marginal cost) pass the standard tests of weak instruments. In particular, the Kleibergen and Paap (2006) Wald statistic (the robust counterpart of the Cragg-Donald Wald statistic) is 14.079, which suggests that weak identification should not be considered a problem. Finally, the null hypothesis of underidentification is confortably rejected based on the Kleibergen and Paap (2006) rank test, that yields a p-value of 0.001.

7.2 Baseline empirical results

We first examine the formal test statistics developed in sections 4.1 and 4.2 to test for discretion and commitment, under the maintained assumption of optimal monetary policy. The tests are based on a two-step procedure. In particular, to test discretion we first estimate the parameter vector θ_d via optimal GMM from condition (25). Next, using the estimated vector of parameters

Table 2: model specification and model selection tests

	Panel A: model specification tests	
	H_0: discretion	H_0: commitment
J–test	11.730	11.637
p–val	(0.14)	(0.20)
TQ_T	16.900	21.830
p–val	(0.40)	(0.04)
	Panel B: model selection test	
	$H_0 : d(\mathcal{D}, \mu) = d(\mathcal{C}, \mu)$	
QLR_T	4.916	
p–val	(0.00)	

The p–values for the J test and for TQ_T are obtained from 1,000 block-bootstrap replications with blocks of size 4. The J test is based on the moment condition (25) for discretion, and (29) for commitment. The test statistics TQ_T correspond to (27) and (30). The test statistic QLR_T is given by (32). The instrument list includes $\Pi_{t-1}, \Pi_{t-2}, \Pi_{t-3}$, and $s_{t-1}, s_{t-2}, s_{t-3}$.

$\widehat{\theta}_d$ we construct the test statistic for discretion $TQ_T^d\left(\widehat{\theta}_d\right)$ and compute the bootstrap critical value. To test commitment, we proceed in an analogous way, making use of condition (29) to obtain $\widehat{\theta}_c$.

The results are reported in the Panel A of Table 2. Since we use enough instrumental variables for overidentification, we start by obtaining results from the standard Hansen J–test statistic for overidentifying restrictions. The table reports the J–tests and the corresponding p–values, for the null hypotheses of discretion (first column) and commitment (second column) based, respectively, on the moment conditions in (25) and (29). The p–value of the J–test for discretion is 14% and that for commitment is 21%. Thus, the standard test for overidentifying restrictions fails to reject either model.

By not making use of the full set of implications of optimal monetary policy, we are unable to reject either 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, by constructing the test statistic TQ_T. The test statistic is based on equation (27) for the case of discretion and equation (30) for the case of commitment. For the case of discretion, the p–value associated with the test statistic is 40%. Instead, for the case of commitment, the p–value is 4%, allowing for rejection of the null hypothesis of optimal policy under commitment at the 5% level. Thus, we reject commitment but fail to reject discretion at all conventional levels.

Finally, Panel B of Table 2 reports the results from the Shi (2015) model selection test presented in Section 5. We consider the null hypothesis that the distance between the commitment and discretion model is zero, against the two-sided alternative, and construct the test statistic so that a positive realization of QLR_T constitutes a rejection of the null hypothesis in favor of discretion. The test statistic is indeed positive, with $QLR_T = 4.916$ and the p-value is less than 1%. Thus, the null hypothesis is rejected decisively in favor of the alternative of discretion.

7.3 Inflation indexation

Our baseline model corresponds to a simple version of the NK model. However, it is possible to apply the specification test proposed in this paper to more general versions of the model that include sources of endogenous persistence that have been found to be empirically relevant. To illustrate this point, we now consider a version of the NK model that includes inflation inertia.[16]

We incorporate inflation inertia by considering a variant of the Calvo model in which firms index to lagged inflation if they cannot re-optimize their price, as in Giannoni and Woodford (2004) and Christiano, Eichenbaum, and Evans (2005). The optimizing model of staggered price-setting with partial indexation results in the following equation relating the rate of inflation to the economy-wide real marginal cost, lagged inflation and expected inflation

$$\pi_t - \gamma\pi_{t-1} = \beta\mathbb{E}_t\left(\pi_{t+1} - \gamma\pi_t\right) + \psi s_t + u_t, \tag{38}$$

where $\gamma \in [0, 1]$ indicates the degree of indexation. This hybrid version of the NK Phillips Curve is widely used in empirical work. Giannoni and Woodford (2004) show that the welfare-theoretic stabilization objective function that corresponds to the formulation of the Phillips Curve with

[16] In Appendix E.1 (available as supplementary material), we show how our methodology can be applied to a special (but empirically salient) case of the Erceg, Henderson, and Levin (2000) model with sticky wages as well as prices.

inflation indexation includes as a target variable the "quasi-differenced" inflation, given by

$$\widetilde{\pi}_t = \pi_t - \gamma \pi_{t-1}. \tag{39}$$

Thus, the policymaker's problem is to maximize

$$\mathbf{U} = -\mathbb{E}_0 \left[\frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left(\widetilde{\pi}_t^2 + \zeta s_t^2 \right) \right]. \tag{40}$$

subject to the Phillips Curve (38), which can be written in terms of $\widetilde{\pi}_t$ as follows

$$\widetilde{\pi}_t = \beta \mathbb{E}_t \widetilde{\pi}_{t+1} + \psi s_t + u_t, \tag{41}$$

This model is analogous to our baseline model, except that π_t is everywhere replaced by the quasi-differenced inflation rate $\widetilde{\pi}_t$. Therefore, the moment inequality conditions that characterize optimal monetary policy are given by the moment conditions analogous to (15) and (16), but with $\widetilde{\Pi}_t = \Pi_t - \gamma_0 \Pi_{t-1}$ in place of Π_t and $(1-\gamma_0)\bar{\Pi}_0$ in place of $\bar{\Pi}_0$, as follows

$$\mathrm{E}\left[m_{d,t}\left(\mathbf{D}_0, \gamma_0, \bar{\Pi}_0\right)\right] \equiv \mathrm{E}\begin{bmatrix} -\left(\widetilde{\Pi}_t - (1-\gamma_0)\bar{\Pi}_0 - \mathbf{D}_0 s_t\right) 1\left(s_{t-1} \leq 0\right) Z_t \\ \left(\widetilde{\Pi}_t - (1-\gamma_0)\bar{\Pi}_0 - \mathbf{D}_0 s_t\right) 1\left(s_{t-1} > 0\right) Z_t \end{bmatrix} \geq 0, \tag{42}$$

$$\mathrm{E}\left[m_{c,t}\left(\mathbf{D}_0, \gamma_0, \bar{\Pi}_0\right)\right] \equiv \mathrm{E}\begin{bmatrix} \left(\widetilde{\Pi}_t - (1-\gamma_0)\bar{\Pi}_0 - \mathbf{D}_0 \Delta s_t\right) 1\left(s_{t-1} \leq 0\right) Z_t \\ -\left(\widetilde{\Pi}_t - (1-\gamma_0)\bar{\Pi}_0 - \mathbf{D}_0 \Delta s_t\right) 1\left(s_{t-1} > 0\right) Z_t \end{bmatrix} \geq 0. \tag{43}$$

Given Asumption 1, the measurement error in quasi-differenced inflation is given by the first-order moving average $\widetilde{v}_t = v_t - \gamma v_{t-1}$, and has mean $(1-\gamma)\Pi_0$ and variance $(1+\gamma^2)\sigma_v^2$. Therefore, the instrumental variables in Z_t should not include the first lag of measured inflation, since they need to be independent from v_{t-1}.

In Table 3 we show results for the standard J-test, our model specification test and the model selection test of Shi (2015) constructed based on the moment inequality conditions (42) and (43). The results are shown for different levels of inflation indexation, including $\gamma = 0.2, 0.4, 0.6$ and 0.8, for discretion (Panel A) and commitment (Panel B). We notice first that the J-test fails to reject at the 10% level any of the 8 models considered. Instead, our test fails to reject discretion for each level of indexation considered but rejects commitment for $\gamma = 0.2$ and $\gamma = 0.4$.

Table 3: specification test (model with inflation indexation)

	Panel A: H_0 is Discretion			
indexation:	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$
J-test (p–val)	0.223	0.195	0.141	0.100
TQ_T (p–val)	0.279	0.248	0.276	0.403
	Panel B: H_0 is Commitment			
indexation:	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$
J-test (p–val)	0.433	0.392	0.405	0.314
TQ_T (p–val)	0.068	0.073	0.130	0.160
	Panel C: model selection test			
indexation:	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$
QLR_T (p–val)	0.008	0.019	0.026	0.010

The p–values for the J test and for TQ_T are obtained from 1,000 block-bootstrap replications with blocks of size 4. The instrument list includes $\Pi_{t-2}, \Pi_{t-3}, \Pi_{t-4}$, and $s_{t-2}, s_{t-3}, s_{t-4}$.

Another important finding is that it becomes more difficult to reject optimal policy (either discretion or commitment) as the degree of inflation indexation is increased. This finding has a natural interpretation. When there is no inflation indexation, optimal policy yields a process for inflation which has a low degree of persistence. This counterfactual feature leads to the empirical rejection of the model. Instead, with higher degrees of inflation indexation, optimal policy is consistent with some persistency in the level of inflation. Thus, it is harder to reject both discretion and commitment if we allow for a high degree of inflation indexation.

In these circumstances, the model selection test of Shi (2015) presented in Section 5 is specially relevant. Results are reported in Panel C of Table 3. We consider the null hypothesis that the distance between the commitment and discretion model is zero, against the two-sided alternative, and construct the test statistic so that a positive realization of QLR_T constitutes a rejection of the null hypothesis in favor of discretion. For all four levels of inflation indexation considered, the p-value is less than 5%. Thus, the null hypothesis is rejected decisively in favor of the alternative of discretion, no matter the level of inflation indexation.

8 Conclusion

This paper develops a methodology for estimating and testing a model of optimal monetary policy without requiring an explicit choice of the relevant 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 by deriving bounds for inflation that are consistent with both forms of optimal policy. This allows for the construction of a test statistic based on the combination of moment equality and inequality conditions that incorporate a wider set of implications of optimal monetary policy and, therefore, provides a more powerful specification test. Yet, unlike full-information methods, our approach does not require strong assumptions about the nature of the forcing variables.

We apply our method to investigate if the behavior of the United States monetary authority is consistent with the New Keynesian model of optimal monetary policy. Our 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.

Our two-step testing procedure can be used more generally to test the validity of models that combine moment equality and inequality conditions, when the parameters of the model can be consistently estimated under the null hypothesis.

Appendix

A Benchmark structural model

The framework is the New Keynesian forward-looking model with monopolistic competition and Calvo price-setting described in Clarida et al. (1999) and Woodford (2010), with an efficient steady state. In the log-linear form model, the inflation rate π_t, the economy's log average real marginal cost in deviation from its flexible price steady state s_t, and the output gap x_t, are determined by an aggregate supply and an aggregate demand relation, as follows

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \psi s_t + u_t, \tag{A.1}$$

$$z_t = \mathbb{E}_t z_{t+1} - \sigma \left(i_t - \mathbb{E}_t \pi_{t+1} \right) + \nu_t, \tag{A.2}$$

where $z_t = \ln \left(Y_t / Y_t^n \right)$ is the output gap , $i_t \geq -i^\star$ denotes the nominal interest rate in deviation from its steady state $i^\star$, u_t is an exogenous stochastic shock resulting from time-varying desired markups and other distortions, ν_t captures shocks to the natural real interest rate, β is the agent's discount factor, $\sigma > 0$ is the elasticity of intertemporal substitution and ψ is a nonlinear function of the relevant structural parameters, given by

$$\psi = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha(1+\vartheta\varsigma)}, \tag{A.3}$$

with $\alpha \in (0,1)$ the fraction of prices that are not reset optimally each period, $\vartheta > 1$ the elasticity of substitution across differentiated goods, and $\varsigma > 0$ the elasticity of each firm's real marginal cost with respect to its own output level.

In turn, the relation between the output gap and the economy's log average real marginal cost is given by

$$s_t = \left(\varsigma + \sigma^{-1} \right) z_t, \tag{A.4}$$

and the aggregate supply relation can be written as

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa z_t + u_t, \tag{A.5}$$

with

$$\kappa = \psi\left(\varsigma + \sigma^{-1}\right) = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha}\frac{\varsigma + \sigma^{-1}}{1+\vartheta\varsigma}. \tag{A.6}$$

Finally, the second order approximation to the utility of a stand-in agent around the steady state equilibrium associated with zero inflation takes the form

$$\mathbf{U} = \mathbb{E}_0\left[-\frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left(\pi_t^2 + (\kappa/\vartheta)\, z_t^2\right)\right], \tag{A.7}$$

and, using (A.4) to substitute for the output gap, yields

$$\mathbf{U} = \mathbb{E}_0\left[-\frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left(\pi_t^2 + \zeta s_t^2\right)\right], \tag{A.8}$$

with

$$\zeta = \frac{\kappa}{\left(\varsigma + \sigma^{-1}\right)^2\vartheta} = \frac{\psi}{\left(\varsigma + \sigma^{-1}\right)\vartheta}, \tag{A.9}$$

the relative target weight on the log average real marginal cost.

B An estimator of the identified set

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

Proposition B.1. *Let Assumptions 1–3 hold. If* $\sqrt{\ln\ln T}/d_T \to 0$ *and* $d_T/\sqrt{T} \to 0$*, we have:*

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

and $\rho_H\left(\widehat{\Theta}_T^I, \Theta^I\right) = O_p\left(\frac{d_T}{\sqrt{T}}\right)$.[17]

Proof of Proposition B.1

It is easy to see that Proposition B.1 holds for example with $d_T = \sqrt{\ln T}$. This follows from

[17]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 $d(a,B) = \inf_{b\in B}\|b-a\|$.

Theorem 3.1 in Chernozhukov et al. (2007, CHT), with $\widehat{c} = d_T^2$, $a_T = T$, $\gamma = 2$, once we show that Assumptions 1, 2 and 3 imply the satisfaction of Conditions 1 and 2 in CHT. Condition 1(a) holds as $\theta = \left(\phi, \bar{\Pi}\right)$ lies in a compact subset of Θ. Assumptions 1 and 2 allow to state the sample objective function as $Q_T(\theta)$ in (19). 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 (19) satisfies Condition 2.[18]

To conduct inference, 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 4. *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 $C_T^{1-\alpha} = \{\theta \in \Theta : TQ_T(\theta) \leq c_\alpha\}$*, and* c_α *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 $\widehat{\Theta}_T^I$ and we use bootstrap critical values.[19] 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 the block length, let (Π_t^*, s_t^*, Z_t^*) denote the

[18] The moment conditions (23) and (24) are linear in the parameters. Thus the identified set is convex. When the identified set is convex, one can rely on the result of Kaido and Santos (2014) in order to obtain a more efficient estimation of the identified set. The key point, is that in the convex case there is a one to one correspondence between the support function and the identified set Θ^I, with the support function defined as

$$\delta^\star\left(q|\Theta^I\right) = \sup_{\theta\in\Theta^I}\left(q'\theta\right)$$

for all q, such that $||\,q\,|| = 1$. For Θ^I convex,

$$\theta \in \Theta^I \text{ if and only if for all } q,\ q'\theta \leq \delta^\star\left(q|\Theta^I\right),$$

so identification of Θ^I is equivalent to the identification of its support function. The support function can be estimated at a parametric rate. Now, in the general convex case estimation of the support function is not straightforward, as it requires a rather complex constrained optimization. But, for linear moment conditions and exact identification, the estimation of the support function can be implemented as detailed in Bontemps et al. (2012).

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

resampled observations. For each $\theta \in \widehat{\Theta}_T^I$, we construct

$$TQ_T^*(\theta) = \sum_{i=1}^{4p} \left(\sqrt{T} \left[\frac{m_{i,T}^*(\theta) - m_{i,T}(\theta)}{\sqrt{\widehat{v}_*^{i,i}(\theta)}} \right]_- 1 \left[\frac{m_{i,T}(\theta)}{\sqrt{\widehat{v}^{i,i}(\theta)}} \leq \sqrt{2 \ln \ln T / T} \right] \right)^2, \tag{B.1}$$

where $m_{i,T}^*(\theta)$ is the bootstrap analog of the sample moment conditions $m_{i,T}(\theta)$, constructed using the bootstrapped data (Π_t^*, s_t^*, Z_t^*), and $\widehat{v}_*^{i,i}(\theta)$ is the $i-$th element on the diagonal of the bootstrap analog of the variance of the moment conditions $\widehat{V}_T^*(\theta)$. The indicator function in (B.1) implements the Generalized Moment Selection (GMS) procedure introduced by Andrews and Soares (2010), that uses 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 of $\sup_{\theta \in \widehat{\Theta}^I} TQ_T^*(\theta)$, and obtain the $(1-\alpha)$ percentile, $c_{B,\alpha}$.

The following proposition can be established:

Proposition B.2. *Let Assumptions 1, 2 and 3 hold, and let $\widehat{\Theta}_T^I$ be defined as in (20). Then as $T \to \infty$, $B \to \infty$, $l \to \infty$, and $l^2/T \to 0$, we have that*

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

where $C_{T,B}^{1-\alpha} = \{\theta \in \Theta : TQ_T(\theta) \leq c_{B,\alpha}\}$.

Proof of Proposition B.2

The events $\{\Theta^I \subseteq C_T^{1-\alpha}\}$ 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 $\theta \in \Theta^I$,

$$TQ_T(\theta) \overset{d}{\to} \sum_{i=1}^{4p} \left(\left[\sum_{j=1}^{4p} \omega_{i,j}(\theta) \mathcal{Z}_i + h_i(\theta) \right]_- \right)^2$$

where $\mathcal{Z} = (\mathcal{Z}_1, \ldots, \mathcal{Z}_{4p}) \sim N(0, I_{2p})$ 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 9. Finally, $h(\theta) = (h_1(\theta), ..., h_{4p}(\theta))'$ 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

$$\sup_{\theta\in\widehat{\Theta}_T^I} TQ_T(\theta) \xrightarrow{d} \sup_{\theta\in\Theta^I} \sum_{i=1}^{4p} \left(\left[\sum_{j=1}^{4p} \omega_{i,j}(\theta) \mathcal{Z}_i + h_i(\theta)\right]_-\right)^2. \tag{B.2}$$

We need to show that the $(1-\alpha)$-percentile of the right-hand side of (B.2), $c_{\alpha,T}$, is accurately approximated by the $(1-\alpha)$-percentile of the bootstrap limiting distribution $\sup_{\theta\in\widehat{\Theta}_T^I} TQ_T^*(\theta)$, $c^*_{\alpha,T}$, conditional on the sample. By the law of the iterated logarithm as $T\to\infty$ and for $i = 1, ..., 4p$, we have that, almost surely,

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

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

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

Hence, as T gets large, only the moment conditions that hold 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, $\mathrm{E}^*\left(\sqrt{T}\left(m^*_{i,T}(\theta) - m_{i,T}(\theta)\right)\right) = O_p\left(l/\sqrt{T}\right)$

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,

$$\left(\frac{\left(m^*_{1,T}(\theta) - m_{1,T}(\theta)\right)}{\sqrt{\widehat{v}^{1,1}(\theta)}}, \ldots, \frac{\left(m^*_{4p,T}(\theta) - m_{4p,T}(\theta)\right)}{\sqrt{\widehat{v}^{4p,4p}(\theta)}}\right) \simeq N\left(0, \widehat{\Omega}_T(\theta)\right).$$

Hence, conditionally on the sample, for all samples except a set of probability measure 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.

C Additional proofs

Proof of Lemma 1

Immediate from the definition of $\pi^c_t(\phi_0)$ and $\pi^d_t(\phi_0)$, Equations (11) and (12).

Proof of Proposition 1

Given Assumption 1, it follows immediately from Equation (13) and the fact that $1(\mathbf{S}_{t-1} \leq 0) = 1$ is a sufficient condition for $\mathbf{c}_0'\mathbf{D}_0 s_{t-1} \leq 0$ and, similarly, that $1(\mathbf{S}_{t-1} > 0) = 1$ is a sufficient condition for $\mathbf{c}_0'\mathbf{D}_0 s_{t-1} > 0$.

Proof of Proposition 2

Letting $\theta = \left(\phi, \bar{\Pi}\right)$, we construct the optimal GMM estimator

$$\widehat{\theta}_d = \arg\min_{\theta \in \Theta} m^0_{d,T}(\theta)' \,\widehat{\Omega}_{dd,T}\left(\widetilde{\theta}_d\right)^{-1} m^0_{d,T}(\theta),$$

where $\widetilde{\theta}_d = \arg\min_\theta m^0_{d,T}(\theta)' m^0_{d,T}(\theta)$, and $\widehat{\Omega}_{dd,T}\left(\widetilde{\theta}_d\right)$ is the HAC estimator of the variance of $\sqrt{T} m^0_{d,T}(\theta_0)$. If we knew $\theta_0 = \left(\phi_0, \bar{\Pi}_0\right)$, the statement would follow by a similar argument as in the proof of Proposition B.2, simply comparing $TQ^d_T(\theta_0)$ with the $(1-\alpha)$-percentile of the empirical distribution of $TQ^{*d}_T(\theta_0)$. However, as we do not know θ_0 we replace it with the optimal GMM estimator, $\widehat{\theta}_d$. Thus, the parameter estimation error term, $\sqrt{T}\left(\widehat{\theta}_d - \theta_0\right)$, 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,

$$\sqrt{T}m_{d,T}^{0}\left(\widehat{\theta}_d\right)=\sqrt{T}m_{d,T}^{0}\left(\theta_0\right)+D_{d,T}\left(\overline{\theta}_d\right)\sqrt{T}\left(\widehat{\theta}_d-\theta_0\right) \tag{C.1}$$

$$\sqrt{T}m_{c,T}\left(\widehat{\theta}_d\right)=\sqrt{T}m_{c,T}\left(\theta_0\right)+D_{c,T}\left(\overline{\theta}_d\right)\sqrt{T}\left(\widehat{\theta}_d-\theta_0\right) \tag{C.2}$$

with $\overline{\theta}_d\in\left(\widehat{\theta}_d,\theta_0\right)$, $D_{d,T}\left(\theta\right)=\nabla_{\theta}m_{d,T}^{0}\left(\theta\right)$ and $D_{c,T}\left(\theta\right)=\nabla_{\theta}m_{c,T}\left(\theta\right)$. From (C.1) it follows that

$$\begin{aligned}\text{avar}\left(\sqrt{T}m_{d,T}^{0}\left(\widehat{\theta}_d\right)\right) &= \text{avar}\left(\sqrt{T}m_{d,T}^{0}\left(\theta_0\right)\right)+\text{avar}\left(D_{d,T}\left(\widehat{\theta}_d\right)\sqrt{T}\left(\widehat{\theta}_d-\theta_0\right)\right)+\\ &\quad +2\ \text{acov}\left(\sqrt{T}m_{d,T}^{0}\left(\theta_0\right),D_{d,T}\left(\widehat{\theta}_d\right)\sqrt{T}\left(\widehat{\theta}_d-\theta_0\right)\right).\end{aligned} \tag{C.3}$$

The asymptotic variance of the moment conditions $\sqrt{T}m_T\left(\theta_0\right)$ can be estimated by

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

Via a mean value expansion of the GMM first order conditions around θ_0,

$$\sqrt{T}\left(\widehat{\theta}_d-\theta_0\right)=-\widehat{B}_{d,T}D_{d,T}\left(\widehat{\theta}_d\right)'\widehat{\Omega}_{dd,T}\left(\widehat{\theta}_d\right)^{-1}\sqrt{T}m_{d,T}^{0}\left(\theta_0\right), \tag{C.4}$$

with

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

hence, given Assumptions 1–3, $\widehat{B}_{d,T}^{-1/2}\sqrt{T}\left(\widehat{\theta}_d-\theta_0\right)\xrightarrow{d}N\left(0,I_2\right)$. We define the estimator of the asymptotic variance of the moment conditions evaluated at the optimal GMM, $\sqrt{T}m_T\left(\widehat{\theta}_d\right)$, as

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

where the first entry can be computed using (C.3) and (C.4), as follows

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

Also,

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

Note that, for the computation of the test statistic we need only an estimate of the diagonal elements of the asymptotic variance of the moment conditions, hence we do not need a closed-form expression for $\widehat{V}_{dc,T}\left(\widehat{\theta}_d\right)$. Let

$$V_{dd}\left(\theta_0\right) = \text{plim}_{T\to\infty}\widehat{V}_{dd,T}\left(\widehat{\theta}_d\right), \; V_{cc}\left(\theta_0\right) = \text{plim}_{T\to\infty}\widehat{V}_{cc,T}\left(\widehat{\theta}_d\right).$$

Since $V_{dd}\left(\theta_0\right)$ is of rank $p-2$, while $V_{cc}\left(\theta_0\right)$ is of full rank $2p$, the asymptotic variance covariance matrix $V\left(\theta_0\right)$ is of rank $3p-2$. 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 $TQ_T^d\left(\widehat{\theta}_d\right)$ writes as

$$\begin{aligned}TQ_T^{*d}\left(\widehat{\theta}_d^*\right) &= T\sum_{i=1}^{p}\left(\frac{m_{i,d,T}^{0*}\left(\widehat{\theta}_d^*\right)-m_{i,d,T}^{0}\left(\widehat{\theta}_d\right)}{\sqrt{\widehat{v}_*^{i,i}\left(\widehat{\theta}_d^*\right)}}\right)^2\\ &\quad +T\sum_{i=1}^{2p}\left[\frac{m_{i,c,T}^{*}\left(\widehat{\theta}_d^*\right)-m_{i,c,T}\left(\widehat{\theta}_d\right)}{\sqrt{\widehat{v}_*^{i,i}\left(\widehat{\theta}_d^*\right)}}\right]_-^2 1\left[m_{i,c,T}\left(\widehat{\theta}_d\right)\leq\sqrt{\widehat{v}^{i,i}\left(\widehat{\theta}_d\right)}\sqrt{2\ln\ln T/T}\right],\end{aligned}$$

where $m_T^*\left(\theta\right)$ denote the moment conditions computed using the resampled observations, and $\widehat{\theta}_d^*$ is the bootstrap analog of $\widehat{\theta}_d$, given by

$$\widehat{\theta}_d^* = \arg\min_{\theta\in\Theta}\left(m_{d,T}^{0*}\left(\theta\right)-m_{d,T}^{0}\left(\widehat{\theta}_d\right)\right)'\widehat{\Omega}_{dd,T}^{*}\left(\widetilde{\theta}_d^*\right)^{-1}\left(m_{d,T}^{0*}\left(\theta\right)-m_{d,T}^{0}\left(\widehat{\theta}_d\right)\right),$$

with $\widetilde{\theta}_d^* = \arg\min_{\theta\in\Theta} \left(m_{d,T}^{0*}(\theta) - m_{d,T}^0\left(\widehat{\theta}_d\right)\right)' \left(m_{d,T}^{0*}(\theta) - m_{d,T}^0\left(\widehat{\theta}_d\right)\right)$, and

$$\widehat{\Omega}_{dd,T}^*\left(\widetilde{\theta}_d^*\right) = \frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{d,I_k+i}^0\left(\widetilde{\theta}_d^*\right) - m_{d,T}^0\left(\widehat{\theta}_d\right)\right)\left(m_{d,I_k+j}^0\left(\widetilde{\theta}_d^*\right) - m_{d,T}^0\left(\widehat{\theta}_d\right)\right)', \tag{C.5}$$

where I_i is an independent, identically distributed discrete uniform random variable on $[0, T-l-1]$.

Finally, $\widehat{v}^{i,i*}\left(\widehat{\theta}_d^*\right)$ is the $i-$th element on the diagonal of of $\widehat{V}_T^*\left(\widehat{\theta}_d^*\right)$, the bootstrap counterpart of $\widehat{V}_T\left(\widehat{\theta}_d\right)$, given by

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

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 $\widehat{V}_{dd,T}^*\left(\widehat{\theta}_d^*\right)$ and $\widehat{V}_{cc,T}^*\left(\widehat{\theta}_d^*\right)$, which are

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

where

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

with $\widehat{D}_{d,T}^*\left(\widehat{\theta}_d^*\right) = \nabla_\theta m_{d,T}^{0*}\left(\widehat{\theta}_d^*\right)$ and where $\widehat{\Omega}_{dd,T}^*\left(\widehat{\theta}_d^*\right)$ is defined as in (C.5), but with $\widetilde{\theta}_d^*$ replaced by $\widehat{\theta}_d^*$, also

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

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

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

$$\widehat{\Omega}_{cd,T}^*\left(\widehat{\theta}_d^*\right) = \frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{c,I_k+i}\left(\widehat{\theta}_d^*\right) - m_{c,T}\left(\widehat{\theta}_d\right)\right)\left(m_{d,I_k+j}^0\left(\widehat{\theta}_d^*\right) - m_{d,T}^0\left(\widehat{\theta}_d\right)\right)'.$$

We compute B bootstrap replication of $TQ_T^{*d}\left(\widehat{\theta}_d^*\right)$, say $TQ_{T,1}^{*d}\left(\widehat{\theta}_d^*\right),...,TQ_{T,B}^{*d}\left(\widehat{\theta}_d^*\right)$, and compute the $(1-\alpha)-$th percentile of its empirical distribution, $c_{T,B,\alpha}^{*d}\left(\widehat{\theta}_d^*\right)$. 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 $TQ_T^d\left(\widehat{\theta}_d\right)$ is larger than (smaller than or equal to) $c_{T,B,\alpha}^{*d}\left(\widehat{\theta}_d^*\right)$ provides an asymptotically non conservative test and unit asymptotic power. To this end, we show that, under H_0^d, $TQ_T^{*d}\left(\widehat{\theta}_d^*\right)$ has the same limiting distribution as $TQ_T^d\left(\widehat{\theta}_d\right)$, conditionally on the sample, and for all samples except a set of probability measure approaching zero. On the other hand, under H_1^d, $TQ_T^{*d}\left(\widehat{\theta}_d^*\right)$ has a well defined limiting distribution, while $TQ_T^d\left(\widehat{\theta}_d\right)$ diverges to infinity.

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

$$\sqrt{T}\left(\widehat{\theta}_d^*-\widehat{\theta}_d\right)=-\widehat{B}_{d,T}^*\widehat{D}_{d,T}^*\left(\widehat{\theta}_d^*\right)\widehat{\Omega}_{dd,T}^*\left(\widehat{\theta}_d^*\right)\sqrt{T}\left(m_{d,T}^{0*}\left(\widehat{\theta}_d\right)-m_{d,T}^0\left(\widehat{\theta}_d\right)\right)$$

$$\sqrt{T}\left(m_{d,T}^{0*}\left(\widehat{\theta}_d^*\right)-m_{d,T}^0\left(\widehat{\theta}_d\right)\right)=\sqrt{T}\left(m_{d,T}^{0*}\left(\widehat{\theta}_d\right)-m_{d,T}^0\left(\widehat{\theta}_d\right)\right)+\widehat{D}_{d,T}^*\sqrt{T}\left(\widehat{\theta}_d^*-\widehat{\theta}_d\right).$$

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

$$\mathrm{E}^*\left(\sqrt{T}\left(m_{d,T}^{0*}\left(\widehat{\theta}_d^*\right)-m_{d,T}^0\left(\widehat{\theta}_d\right)\right)\right)=O_p\left(\frac{l}{\sqrt{T}}\right)=o_p(1),$$

$$\mathrm{var}^*\left(\sqrt{T}\left(m_{d,T}^{0*}\left(\widehat{\theta}_d^*\right)-m_{d,T}^0\left(\widehat{\theta}_d\right)\right)\right)=\widehat{V}_T\left(\widehat{\theta}_d\right)+O_p\left(\frac{l}{\sqrt{T}}\right)=\widehat{V}_T\left(\widehat{\theta}_d\right)+o_p(1),$$

and

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

Thus, under both hypotheses, $\sqrt{T}\left(m_{d,T}^{0*}\left(\widehat{\theta}_d^*\right)-m_{d,T}^0\left(\widehat{\theta}_d\right)\right)$ has a limiting distribution which is well defined, and coincides with that of $TQ_T^d\left(\widehat{\theta}_d\right)$ under the null.

As for the moment conditions under commitment, they contribute to the limiting distribution only when $m_{i,c,T}\left(\widehat{\theta}_d\right)\leq\sqrt{\widehat{v}^{i,i}\left(\widehat{\theta}_d\right)}\sqrt{2\ln\ln T/T}$, and hence they mimic the limiting distribution of

$$\sum_{i=1}^{2p}\frac{\left[m_{i,c,T}\left(\widehat{\theta}_d\right)\right]_-^2}{\widehat{v}^{i,i}\left(\widehat{\theta}_d\right)}.$$

The statement in the Proposition then follows.

Proof of Proposition 3

The proof is analogous to that of Proposition 2.

D DGP used in the Montecarlo

The data generating process (DGP) for the inflation rate π_t and log average real marginal cost in deviation from the flexible-price steady state s_t used for the Montecarlo simulations is the NK model described in Appendix A. In particular, the DGP under commitment is given by

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \psi s_t + u_t, \tag{D.1}$$

$$\phi \pi_t = -\Delta s_t, \tag{D.2}$$

with u_t an $\mathtt{AR}(1)$ process; $\phi = (\psi/\zeta)$ where ζ is the target weight on s_t relative to inflation in the loss function of the central bank. The DGP under discretion is the same, except that equation (D.2) is replaced with

$$\phi \pi_t = -s_t. \tag{D.3}$$

The choice of parameter values is guided by evidence reported in the NK business cycle literature. In particular, we set $\psi = 0.0230$ and $\beta = 0.9420$, based on Galí and Gertler (1999). As shown in Appendix A, from the theoretical foundations of the loss function (22), $\zeta = \left(\varsigma + \sigma^{-1}\right)^{-1} (\psi/\vartheta)$ and, hence, $\phi = (\psi/\zeta) = \left(\varsigma + \sigma^{-1}\right) \vartheta$. We set $\vartheta = 7$, a value for the elasticity of substitution across goods which is commonly found in the business cycle literature (Golosov and Lucas, 2007), and we set $\sigma = 1$ and $\varsigma = 1.25$, following the baseline calibration in Chari, Kehoe, and McGrattan (2000). Thus, the implied value for ϕ is 15.75.

Finally, both DGP include measurement error in inflation, so that measured inflation is given by

$$\Pi_t = \pi_t + v_t, \quad \text{with } v_t \sim \mathtt{N}\left(\Pi_0, \sigma_v^2\right). \tag{D.4}$$

The mean of the measurement error is set to $\Pi_0 = 0.005$, which corresponds to annual inflation being overstated on average by 2% (as reported in Bernanke and Mishkin, 1997), and the variance is set to $\sigma_v^2 = 0.001$.

In turn, the stochastic process for the cost-push shock is given by

$$u_t = \rho u_{t-1} + \epsilon_t. \tag{D.5}$$

We set $\rho = 0.8$ and choose the variance of ϵ_t, σ_ϵ^2 to match exactly the volatility of the real marginal cost, s_t. We use the volatility of s_t as a target, because the real marginal cost is assumed to be measured without error.

E Supplementary material

E.1 An example with sticky wages and prices

We consider the model of sticky wages and prices of Erceg et al. (2000). When both wages and prices are sticky, the objective function of the policymaker is

$$\mathbf{U} = -\mathbb{E}_0 \left[\frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left(\zeta_p \pi_{p,t}^2 + \zeta_w \pi_{w,t}^2 + \zeta_x x_t^2 \right) \right], \tag{E.1}$$

and the relevant structural equations of the economy are

$$\pi_{p,t} = \beta \mathbb{E}_t \pi_{p,t+1} + \kappa_p x_t + \xi_p \left(\omega_t - \omega_t^n \right), \tag{E.2}$$

$$\pi_{w,t} = \beta \mathbb{E}_t \pi_{w,t+1} + \kappa_w x_t + \xi_w \left(\omega_t - \omega_t^n \right), \tag{E.3}$$

$$\omega_t = \omega_{t-1} + \pi_{w,t} - \pi_{p,t}, \tag{E.4}$$

where x_t is the output gap, $\pi_{p,t}$ and $\pi_{w,t}$ are the price and wage inflation, respectively, ω_t is the log of the real wage, and ω_t^n is the natural real wage (the real wage in the absence of nominal rigidities), and is an exogenous process that depends on technology, preferences and other such disturbances. Equations (E.2) and (E.3) are the price and wage Phillips Curves, and (E.4) captures the real wage dynamics.

Of course, this model has an endogenous predetermined variable (real wage ω_t) and, hence, Lemma 1 cannot be applied immediately. However, as shown in Woodford (2003, chapter 3) and Giannoni and Woodford (2004), for the special case in which $\kappa_p = \kappa_w$, the evolution of the real wage is independent of monetary policy. To see this, let $\kappa_p = \kappa_w = \kappa > 0$, and subtract (E.2) from (E.3), to obtain

$$\pi_{w,t} - \pi_{p,t} = \beta \mathbb{E}_t \left(\pi_{w,t+1} - \pi_{p,t+1} \right) + \left(\xi_w - \xi_p \right) \left(\omega_t - \omega_t^n \right), \tag{E.5}$$

where the output gap, x_t, no longer appears. Using (E.4) to substitute in (E.5), yields

$$\Delta \omega_t = \beta \mathbb{E}_t \Delta \omega_{t+1} + \left(\xi_p - \xi_w \right) \left(\omega_t - \omega_t^n \right), \tag{E.6}$$

implying a unique bounded solution for the path of the real wage, independent of monetary policy, given a bounded process for the exogenous disturbance ω_t^n.

Thus, for the special case $\kappa_p = \kappa_w = \kappa$, the structural equations that represent the economy's equilibrium conditions may be written in a form which does not include endogenous predetermined variables, as follows

$$\begin{bmatrix} 1 & 0 & -\kappa \\ 0 & 1 & -\kappa \end{bmatrix} y_t + \begin{bmatrix} -\beta & 0 & 0 \\ 0 & -\beta & 0 \end{bmatrix} \mathbb{E}_t y_{t+1} = \begin{bmatrix} \epsilon_{p,t} \\ \epsilon_{w,t} \end{bmatrix}, \tag{E.7}$$

where $y_t = \begin{bmatrix} \pi_{p,t}, & \pi_{w,t}, & x_t \end{bmatrix}'$ is the vector of target variables and

$$\begin{bmatrix} \epsilon_{p,t} \\ \epsilon_{w,t} \end{bmatrix} = \begin{bmatrix} \xi_p \left(\omega_t - \omega_t^n\right) \\ \xi_w \left(\omega_t - \omega_t^n\right) \end{bmatrix}, \tag{E.8}$$

are exogenous disturbances to the real wage, independent of monetary policy, given by (E.6). In turn, the objective function of the policy maker is given by (1), with the matrix W of policy weights which is diagonal, with diagonal elements $\begin{bmatrix} \zeta_p & \zeta_w & \zeta_x \end{bmatrix}' \geq 0$.

The first-order conditions solving for optimal policy under commitment are

$$\begin{bmatrix} \zeta_p & 0 & 0 \\ 0 & \zeta_w & 0 \\ 0 & 0 & \zeta_x \end{bmatrix} \begin{bmatrix} \pi_{p,t} \\ \pi_{w,t} \\ x_t \end{bmatrix} + \beta^{-1} \begin{bmatrix} -\beta & 0 \\ 0 & -\beta \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{t-1}^1 \\ \lambda_{t-1}^2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\kappa & -\kappa \end{bmatrix} \begin{bmatrix} \lambda_t^1 \\ \lambda_t^2 \end{bmatrix} = 0, \tag{E.9}$$

with $\lambda_t = \begin{bmatrix} \lambda_t^1 & \lambda_t^2 \end{bmatrix}'$, the Lagrange multipliers. Instead, optimal policy under discretion requires

$$\begin{bmatrix} \zeta_p & 0 & 0 \\ 0 & \zeta_w & 0 \\ 0 & 0 & \zeta_x \end{bmatrix} \begin{bmatrix} \pi_{p,t} \\ \pi_{w,t} \\ x_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\kappa & -\kappa \end{bmatrix} \begin{bmatrix} \lambda_t^1 \\ \lambda_t^2 \end{bmatrix} = 0. \tag{E.10}$$

Finally, consider the vector $T = \begin{bmatrix} \gamma, & \gamma, & 0 \end{bmatrix}'$; then we have that

$$n_t = T'Wy_t = \begin{bmatrix} \gamma, & \gamma, & 0 \end{bmatrix} \begin{bmatrix} \zeta_p & 0 & 0 \\ 0 & \zeta_w & 0 \\ 0 & 0 & \zeta_x \end{bmatrix} \begin{bmatrix} \pi_{p,t} \\ \pi_{w,t} \\ x_t \end{bmatrix}, \\ = \gamma \left(\zeta_p \pi_{p,t} + \zeta_w \pi_{w,t} \right) = \psi_t, \tag{E.11}$$

with $\gamma = \left(\zeta_p + \zeta_w \right)^{-1}$, so that ψ_t is a weighted average of price and wage inflation. By applying Lemma 1 we obtain the following bounds for this target variable

$$\Pr\left(\psi_t^c \leq \psi_t \leq \psi_t^d \middle| \gamma\lambda_{t-1}^1 + \gamma\lambda_{t-1}^2 \leq 0 \right) = 1, \tag{E.12}$$

$$\Pr\left(\psi_t^d \leq \psi_t \leq \psi_t^c \middle| \gamma\lambda_{t-1}^1 + \gamma\lambda_{t-1}^2 > 0 \right) = 1, \tag{E.13}$$

with

$$\psi_t^c = \gamma \left(\lambda_t^1 + \lambda_t^2 \right) - \gamma \left(\lambda_{t-1}^1 + \lambda_{t-1}^2 \right), \tag{E.14}$$

$$\psi_t^d = \gamma \left(\lambda_t^1 + \lambda_t^2 \right), \tag{E.15}$$

and where, from (E.9), we obtain $\gamma\lambda_t^1 + \gamma\lambda_t^2 = \gamma \left(\zeta_x / \kappa \right) x_t$.

The target variable ψ is measured with error, with $\Psi = \psi + v_t$ the measured target variable with mean $\bar{\Psi}_0$, as in Assumption 1. Given a vector of instrumental variables, $Z_t > 0$, we obtain the following moment inequality conditions

$$\mathrm{E} \begin{bmatrix} -\left(\Psi_t - \bar{\Psi}_0 - \psi_t^d \right) 1 \left(x_{t-1} \leq 0 \right) Z_t \\ \left(\Psi_t - \bar{\Psi}_0 - \psi_t^d \right) 1 \left(x_{t-1} > 0 \right) Z_t \\ \left(\Psi_t - \bar{\Psi}_0 - \psi_t^c \right) 1 \left(x_{t-1} \leq 0 \right) Z_t \\ -\left(\Psi_t - \bar{\Psi}_0 - \psi_t^c \right) 1 \left(x_{t-1} > 0 \right) Z_t \end{bmatrix} \geq 0, \tag{E.16}$$

where we made use of the fact that $\gamma \left(\lambda_t^1 + \lambda_t^2 \right) > 0$ if and only if $x_t > 0$. Thus, we can apply the same tests of optimal policy as in the baseline model, but where instead of measured inflation Π_t, the moment inequalities are based on the measured target variable $\Psi_t = \gamma \left(\zeta_p \Pi_{p,t} + \zeta_w \Pi_{w,t} \right)$, a weighted average of measured price and wage inflation.

We are only able to apply our method to the version of the Erceg et al. (2000) model with $\kappa_p = \kappa_w$, implying that wage and price inflation are equally responsive to changes in the output gap. However, this parameter restriction is empirically salient as shown in Amato and Laubach (2003), and is consistent with the finding in many empirical studies that the estimated response of the real wage to monetary policy shocks is close to zero (e.g. Christiano et al., 2005).

E.2 An alternative formulation of the specification test

We can follow Andrews and Soares (2010) and test whether there is a parameter value for which the set of moment equalities and inequalities are satisfied. For example, when testing discretion we could proceed as follows.

Let $\theta = \left(\mathbf{D}, \overline{\Pi}\right)$, and $H_0^d : \exists\ \theta = \theta_0$ satisfying

$$\begin{aligned}
\mathrm{E}\left[m_{d,t}^0\left(\theta\right)\right] &= \mathrm{E}\left[\left(\Pi_t - \bar{\Pi} - \mathbf{D}s_t\right) Z_t\right] = 0, \\
\mathrm{E}\left[m_{c,t}\left(\theta\right)\right] &= \mathrm{E}\left[\begin{array}{c} \left(\Pi_t - \bar{\Pi} - \mathbf{D}\Delta s_t\right) 1\left(\mathbf{S}_{t-1} \leq 0\right) Z_t \\ -\left(\Pi_t - \bar{\Pi} - \mathbf{D}\Delta s_t\right) 1\left(\mathbf{S}_{t-1} > 0\right) Z_t \end{array}\right] \geq 0,
\end{aligned}$$

and to test H_0, we can construct the following statistics

$$TQ_T^d(\theta) = T\left[\sum_{i=1}^{p} \frac{m_{i,d,T}^0\left(\theta\right)^2}{\widehat{v}^{i,i}\left(\theta\right)} + \sum_{i=1}^{p} \frac{\left[m_{i,c,T}^c\left(\theta\right)\right]_-^2}{\widehat{v}^{i,i}\left(\theta\right)}\right]$$

and the following bootstrap statistics

$$\begin{aligned}
TQ_T^{*d}\left(\theta\right) &= T\sum_{i=1}^{p}\left(\frac{m_{i,d,T}^{0*}\left(\theta\right) - m_{i,d,T}^0\left(\theta\right)}{\sqrt{\widehat{v}_*^{i,i}\left(\theta\right)}}\right)^2 \\
&\quad + T\sum_{i=1}^{2p}\left[\frac{m_{i,c,T}^*\left(\theta\right) - m_{i,c,T}\left(\theta\right)}{\sqrt{\widehat{v}_*^{i,i}\left(\theta\right)}}\right]_-^2 1\left[m_{i,c,T}\left(\theta\right) \leq \sqrt{\widehat{v}^{i,i}\left(\theta\right)}\sqrt{2\ln\ln T/T}\right],
\end{aligned}$$

and let $c_{T,1-\alpha}^*\left(\theta\right)$ be the $(1-\alpha)$ bootstrap critical value. The $(1-\alpha)$ confidence set for is

$$CS_n = \left\{\theta : TQ_T^d(\theta) \leq c_{T,1-\alpha}^*\left(\theta\right)\right\}$$

We do not reject H_0 if $TQ_T^d(\theta) \leq c_{T,1-\alpha}^*\left(\theta\right)$. This gives a test of at most size α.

E.3 An interpretation of the specification tests

The moment equality/inequality conditions that form the basis for the proposed specification tests are linear in the transformed parameter space $\tilde{\theta} = \left\{\phi^{\dagger}, \phi^{\ddagger}, \bar{\Pi}\right\}$, with $\phi^{\dagger} = \mathbf{c}'\mathbf{D}$ and $\phi^{\ddagger} = \mathbf{b}'\mathbf{D}$ in the general case. In particular, in the benchmark application we are considering, with $\mathbf{b} = 1$ and $\mathbf{c} = -1$, the moment conditions (23) and (24) are linear in $\theta = \left\{\mathbf{D}, \bar{\Pi}\right\} \in \mathbf{R}^2$. The upshot is that, as the moment conditions are convex, the identified set Θ^I is fully characterized by its support function $\delta^{\star}\left(q|\Theta^I\right) = \left\{\sup_{\theta\in\Theta^I}\left(q'\theta\right)\ \forall q : ||q|| = 1\right\}$. Specifically $\theta \in \Theta^I$ is equivalent to $q'\theta \leq \delta^{\star}\left(q|\Theta^I\right)$, for all q such that $||q|| = 1$ (see Bontemps et al., 2012; Kaido and Santos, 2014).

This result allows us to provide a geometric interpretation of the specifications test, that illustrates its source of power.[20] Specifically, Bontemps et al. (2012) show that for models defined by a set of linear moment conditions, the identified set is itself convex and its boundary is determined by the hyperplanes that are tangent to it. In our case, as $\Theta^I \in \mathbf{R}^2$ the hyperplanes are lines and the moment conditions (23) and (24) define a diamond shaped region corresponding to the identified set. Suppose there's a single instrument $\mathbf{z}$ satisfying Assumption 2, so that $p = 1$ and, for illustration, that we want to test the moment conditions (25) and (26), corresponding to discretion. Denote $\mathbf{y}_t^1 = \Pi_t - \bar{\Pi}$ and $\mathbf{y}_t^2 = \left(\Pi_t - \bar{\Pi}\right)1\left(\mathbf{S}_{t-1} \leq 0\right)$, and the list of covariates $\mathbf{x}_t^1 = s_t$ and $\mathbf{x}_t^2 = \Delta s_t 1\left(\mathbf{S}_{t-1} \leq 0\right)$. Moreover, to adapt our framework to the set-up in Bontemps et al. (2012), we assume that $\left(\Pi_t - \bar{\Pi}\right)$ has bounded support in $[\Pi_L, \Pi_U] \in \mathbf{R}$. It follows that, under the null hypothesis of discretion, the following linear predictions can be obtained

$$\mathbf{y}_t^1 = \mathbf{D}\mathbf{x}_t^1 + \gamma_1\tilde{\mathbf{z}}_t^1 + e_t^1, \tag{E.17}$$

$$\mathbf{y}_t^2 = \mathbf{D}\mathbf{x}_t^2 + \delta\tau + \gamma_2\tilde{\mathbf{z}}_t^2 + e_t^2, \tag{E.18}$$

where $\tilde{\mathbf{z}}_t^i$ is the residual of the linear prediction of $\mathbf{z}_t$ on $\mathbf{x}_t^i$, $\delta = (\Pi_U - \Pi_L)$ and, following Bontemps et al. (2012), $\tau \in [0, 1]$, is a random variable called a "selection", with the cumulative distribution function $\eta = F\left(\tau \leq \mathtt{T}|\mathbf{y}^1, \mathbf{y}^2, \mathbf{x}^1, \mathbf{x}^2\right)$ an unknown nuisance parameter. The random variable τ is called a "selection" variable because knowledge of η would complete the model and lead to the point identification of $\left\{\bar{\Pi}, \mathbf{D}, \gamma_1, \gamma_2\right\}$ via the conditions $\mathrm{E}\left[\mathbf{x}_t^i e_t^i\right] = 0$ and $\mathrm{E}\left[\tilde{\mathbf{z}}_t^i e_t^i\right] = 0$, for $i = \{1, 2\}$ yielding 4 identification conditions. As η is unknown, we obtain instead partial identification and,

[20] This elegant interpretation was suggested to us by an anonymous referee.

heuristically, the identified set corresponds to the union over the convex domain of the nuisance parameter η of each point identified model.

Finally, when $p \geq 2$ we obtain a surplus of moment conditions. Thus, there are overidentifying restrictions and as stated in Lemma 4 of Bontemps et al. (2012), the identified set may be empty and its non-emptiness requires an additional set of Sargan-like conditions. In particular, the validity of the exclusion conditions $\mathrm{E}[\tilde{\mathbf{z}}_t^i e_t^i] = 0$, for $i = \{1, 2\}$, is satisfied when the restrictions $\gamma_1 = 0$ and $\gamma_2 = 0$ fail to be rejected. It follows that all the moment conditions (either equality or inequality conditions) contribute to the power of the specification test.[21]

[21] This is similar to the framework in Moon and Schorfheide (2009), who use additional moment inequality conditions to provide overidentifying information. Importantly, in our empirical illustration we find that our test is able to reject the null of commitment, while a standard J-test is not.

References

Amato, J. D. and T. Laubach (2003). Estimation and control of an optimization-based model with sticky prices and wages. *Journal of Economic Dynamics and Control 27*(7), 1181–1215.

Andrews, D. W. K. and G. Soares (2010). Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica 78*(1), 119–157.

Baxter, M. (1988). Toward an empirical assessment of game-theoretic models of policymaking: A comment. *Carnegie-Rochester Conference Series on Public Policy 28*, 141–151.

Beresteanu, A. and F. Molinari (2008). Asymptotic properties for a class of partially identified models. *Econometrica 76*(4), 763–814.

Bernanke, B. S. and F. S. Mishkin (1997). Inflation targeting: A new framework for monetary policy? *Journal of Economic Perspectives 11*(2), 97–116.

Blundell, R., M. Browning, and I. Crawford (2008). Best nonparametric bounds on demand responses. *Econometrica 76*(6), 1227–1262.

Bontemps, C., T. Magnac, and E. Maurin (2012). Set identified linear models. *Econometrica 80*(3), 1129–1155.

Bugni, F. A. (2010). Bootstrap inference in partially identified models defined by moment inequalities: Coverage of the identified set. *Econometrica 78*(2), 735–753.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics 12*(3), 383–398.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2000). Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem? *Econometrica 68*(5), 1151–1180.

Chernozhukov, V., H. Hong, and E. Tamer (2007). Estimation and confidence regions for parameter sets in econometric models1. *Econometrica 75*(5), 1243–1284.

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

Ciliberto, F. and E. Tamer (2009). Market structure and multiple equilibria in airline markets. *Econometrica 77*(6), 1791–1828.

Clarida, R., J. Galí, 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*(1), 147–180.

Corts, K. S. (1999). Conduct parameters and the measurement of market power. *Journal of Econometrics 88*(2), 227–250.

Debortoli, D. and A. Lakdawala (2015). How credible is the federal reserve? a structural estimation of policy re-optimizations. *American Economic Journal: Macroeconomics* (forthcoming).

Dennis, R. (2007). Optimal policy in rational expectations models: New solution algorithms. *Macroeconomic Dynamics 11*(1), 31–55.

Erceg, C. J., D. W. Henderson, and A. T. Levin (2000). Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics 46*(2), 281–313.

Galí, J. and M. Gertler (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics 44*(2), 195–222.

Galichon, A. and M. Henry (2011). Set identification in models with multiple equilibria. *Review of Economic Studies*.

Giannoni, M. and M. Woodford (2004). Optimal inflation-targeting rules. In *The Inflation-Targeting Debate*, pp. 93–172. University of Chicago Press.

Giannoni, M. and M. Woodford (2010). Optimal target criteria for stabilization policy. Technical report, National Bureau of Economic Research.

Givens, G. E. (2012). Estimating central bank preferences under commitment and discretion. *Journal of Money, Credit and Banking. 44*(6), 1033–1061.

Golosov, M. and R. E. Lucas, Jr (2007). Menu costs and phillips curves. *Journal of Political Economy 115*, 171–199.

Haile, P. A. and E. Tamer (2003). Inference with an incomplete model of english auctions. *Journal of Political Economy 111*(1), 1–51.

Halunga, A. G., D. R. Osborn, and M. Sensier (2009). Changes in the order of integration of US and UK inflation. *Economic Letters 102*(1), 30–32.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(4), 1029–1054.

Ireland, P. N. (1999). Does the time-consistency problem explain the behavior of inflation in the united states? *Journal of Monetary Economics 44*(2), 279–291.

Kaido, H. and A. Santos (2014). Asymptotically efficient estimation of models defined by convex moment inequalities. *Econometrica 82*(1), 387–413.

Kirsanova, T. and S. D. Le Roux (2013). Commitment vs. discretion in the UK: An empirical investigation of monetary and fiscal policy regime. *International Journal of Central Banking 9*, 99–152.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*(1), 97–126.

Kydland, F. E. and E. C. Prescott (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy 85*(3), 473–491.

Lubik, T. A. and F. Schorfheide (2004). Testing for indeterminacy: An application to U.S. monetary policy. *American Economic Review 94*(1), 190–217.

Manski, C. F. and E. Tamer (2002). Inference on regressions with interval data on a regressor or outcome. *Econometrica 70*(2), 519–546.

Mavroeidis, S. (2005). Identification issues in Forward-Looking models estimated by GMM, with an application to the phillips curve. *Journal of Money, Credit and Banking 37*(3), 421–448.

McCallum, B. T. (1999). Chapter 23 issues in the design of monetary policy rules. In *Handbook of Macroeconomics*, Volume Volume 1, Part C, pp. 1483–1530.

Moon, H. R. and F. Schorfheide (2009). Estimation with overidentifying inequality moment conditions. *Journal of Econometrics 153*(2), 136–154.

Rotemberg, J. and M. Woodford (1997). An optimization-based econometric framework for the evaluation of monetary policy. In *NBER Macroeconomics Annual 1997, Volume 12*, pp. 297–361. MIT Press.

Ruge-Murcia, F. J. (2003). Does the Barro–Gordon model explain the behavior of US inflation? a reexamination of the empirical evidence. *Journal of Monetary Economics 50*(6), 1375–1390.

Sbordone, A. M. (2002). Prices and unit labor costs: a new test of price stickiness. *Journal of Monetary Economics 49*(2), 265–292.

Shi, X. (2015). Model selection tests for moment inequality models. *Journal of Econometrics 187*(1), 1–17.

Svensson, L. E. and M. Woodford (2004). Implementing optimal policy through inflation-forecast targeting. In *The inflation-targeting debate*, pp. 19–92. University of Chicago Press.

Svensson, L. E. O. (1997). Optimal inflation targets, "conservative" central banks, and linear inflation contracts. *American Economic Review 87*(1), 98–114.

Taylor, J. B. (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*(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. *Handbook of Monetary Economics 3*, 723–828.