Testing for optimal monetary policy via moment inequalities

Laura Coroneo, Valentina Corradi and Paulo Santos Monteiro

No 985

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF
WARWICK

Testing for optimal monetary policy via moment inequalities

Laura Coroneo*, Valentina Corradi and Paulo Santos Monteiro†

March 29, 2012

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.*

*University of Manchester, Economics - School of Social Sciences
†University of Warwick, Department of Economics.

We thank James Banks, Herman Bierens, Jean-Marie Dufour, Alastair Hall, Lynda Khalaf, Maral Kichian, Tatiana Kirsanova, Karel Mertens, Micheal McMahon, Sushama Murty, Milan Nedeljkovic, Carlos Noton, Denise Osborn, Adam Rosen for useful comments. We also thank seminar participants at the University of Chile, the Central Bank of Chile, the University of Exeter, the University of Manchester, the University of Warwick, the 2009 Brussels Waseda Seminar, the 2009 Econometric Society European Meeting, the 2009 North American Summer Meetings of the Econometric Society, the 2010 EC^2 conference, the 2011 Time Series Conference at CIREQ, the 2011 Society for Advancement in Economic Theory Conference, the 2011 International Conference on Computing in Economics and Finance, the Bank od Serbia, the 2012 North America Winter Meeting of the Econometric Society and the Federal Reserve Board.

1 Introduction

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

In a discretionary regime the inflation rate on average exceeds the level that would be optimal if commitment to history dependent policy rules was feasible. This is the celebrated Barro and Gordon (1983) inflationary bias result, arising in the presence of distortions that imply a suboptimal natural output rate. In addition to this deterministic inflation-bias, under discretion there is a state-contingent inflation bias resulting from the fact that the monetary authority sets policy independently of the history of shocks. The upshot of this state-contingent bias is that when the output gap is negative, the inflation rate under discretion in the following period is higher than what it would be if the monetary authority was able to commit to history-dependent plans. This state-contingent inflationary bias allows for the derivation of an inflation lower-bound (obtained under commitment) and an upper-bound (obtained under discretion), based on the first order conditions that characterize optimal monetary policy under each policy regime. These bounds are compatible with a continuum of monetary policy rules characterized by differing degrees of commitment, and in which full commitment and discretion are the two extreme cases. In particular, they include the quasi-commitment model proposed by Schaumburg and Tambalotti (2007), where the monetary authority deviates from commitment-based optimal plans with a fixed, exogenous probability, known to the public.

Our framework relies on the state-contingent inflation bounds, which we use to derive moment inequality conditions associated with optimal monetary policy, and to identify the set of structural parameters for which the moment inequalities hold, i.e. the identified set. We estimate the identified set implied by optimal monetary policy, and construct confidence regions that cover the identified set with a pre-specified probability, using inference methods developed in Chernozhukov, Hong, and Tamer (2007). We then test whether the moment restrictions implied by a specific policy regime are satisfied. Assuming either discretion or commitment allows for point identification of the underlying structural parameters. Hence, parameters can be consistently estimated and it is possible to perform standard tests of overidentifying restrictions (Hansen, 1982). However, if our objective is to test for discretion or commitment under the maintained assumption of optimal monetary policy, the

standard Hansen's J-test does not make use of all the available information. Instead, we propose a test for discretion and a test for commitment 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, Hong, and Tamer (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.

We apply our testing procedure to investigate whether the time series of inflation and output gap in the United States are consistent with the New Keynesian model of optimal monetary policy that has been widely used in recent studies of monetary policy, following Rotemberg and Woodford (1997), Clarida, Galí and Gertler (1999), and Woodford (2003). Using the sample period running from 1983:1 to 2008:3, we find evidence in favor of discretionary optimal monetary policy, and against 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.

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

Using a full-information maximum-likelihood approach, Givens (2010) estimates a New Keynesian model for the US economy in which the monetary authority conducts optimal monetary policy. The model is estimated separately under the two alternatives of commitment and discretion, using quarterly data over the Volcker–Greenspan–Bernanke era; a comparison of the log-likelihood of the two alternative models based on a Bayesian information criterion (to overcome the fact that the two models are non-nested) strongly favors discretion over commitment. A similar Bayesian approach has been used by Kirsanova and le Roux (2011), who also find evidence in favor of discretion for both monetary and fiscal policy in the UK. The partial identification framework that we propose in this paper permits, instead, a general econometric specification that nests commitment and discretion as two special cases. Unlike full-information methods, our approach does not require a complete representation of the economy, nor strong assumptions about the nature of the forcing variables.

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

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

This paper also contributes to a growing literature proposing partial identification methods to overcome lack of information about the economic environment. For instance, Manski and Tamer (2002) examine inference on regressions with interval outcomes. In the industrial organization literature, Haile and Tamer (2003) use partial identification to construct bounds on valuation distributions in second price auctions. Blundell, Browning and Crawford (2008) derive bounds that allow to set-identify predicted demand responses in the study of consumer behavior. Ciliberto and Tamer (2009) study inference in entry games without requiring equilibrium selection assumptions. Galichon and Henry (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 theoretical economy and characterizes optimal monetary policy. Section 3 derives the bounds for inflation implied by the structural model of optimal monetary policy. Section 4 outlines the inference procedure. Section 5 describes the proposed test for optimal monetary policy under discretion and under commitment. Finally, Section 6 reports the empirical findings and Section 7 concludes. Appendix A contains details about the theoretical model and Appendix B collects all proofs.

2 Optimal Monetary Policy

The structural framework corresponds to the new-Keynesian model with staggered prices and monopolistic competition that has become widely used to study optimal monetary policy. 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 cost and expected inflation

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

often called the *New Keynesian Phillips Curve*.[2] Here, ψ and β are positive parameters related to technology and preferences, $\mathbb{E}_t$ denotes the expectations formed by the economic agents at t, π_t is the rate of change of a general index of goods prices and s_t is the real marginal cost in deviation from the flexible-price steady state.

The welfare-theoretical 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, taking the form

$$\mathcal{W} = \mathbb{E}_0 \left\{ -\frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left[\phi \, \pi_t^2 + \psi \left(s_t + u_t \right)^2 \right] \right\}, \tag{2}$$

where ϕ is a positive parameter that relates to technology and preferences. The variable $u_t \leq 0$ is an exogenous stochastic shock resulting from time-varying markups and other distortions.[3]

The model of optimal monetary policy under commitment is based on the assumption that the monetary authority maximizes (2) subject to the constraint imposed by the Phillips curve equation (1). If the monetary authority is able to commit to a state contingent path for inflation and the output gap, the first order conditions solving the monetary authority's problem at some given period τ are

$$\begin{aligned} s_t + u_t + \lambda_t &= 0, \qquad t = \tau, \tau + 1, \ldots \\ \phi \pi_t + \lambda_{t-1} - \lambda_t &= 0, \qquad t = \tau, \tau + 1, \ldots \end{aligned}$$

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

$$\phi \pi_t = - \left(s_t + u_t \right) + \left(s_{t-1} + u_{t-1} \right). \tag{3}$$

However, the commitment solution is time inconsistent in the Kydland and Prescott (1977) sense: each period t, the monetary authority is tempted to behave as if $\lambda_{t-1} = 0$, ignoring the impact of its current actions on the private sector expectations. When the monetary authority lacks a commitment technology, it must set policy sequentially.

Under discretion, optimal monetary policy satisfies the Markov property in the sense that the policy is chosen independently of past choices of the monetary authority. Thus, the policymaker acts as if $\lambda_{t-1} = 0$ and the resulting joint path for inflation and output gap is

$$\phi \pi_t = - \left(s_t + u_t \right). \tag{4}$$

Let ϕ_0 denote the "true" value of ϕ. We define $\pi_t^c \left(\phi_0 \right)$ as the inflation in period t consistent with the first order conditions for optimal policy under commitment, given knowledge of s_t and u_t, and the structural parameter ϕ_0. In the same way, $\pi_t^d \left(\phi_0 \right)$ is the inflation in period t consistent with the first order conditions for optimal policy under discretion. Thus, $\pi_t^c \left(\phi_0 \right)$ and $\pi_t^d \left(\phi_0 \right)$ are,

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

[3] It is common to refer to fluctuations of u_t around its steady-state as cost-push shocks (Clarida et al., 1999). For more details, see Appendix A.

respectively, given by

$$\pi_t^c(\phi_0) = -\phi_0^{-1}(s_t + u_t) + \phi_0^{-1}(s_{t-1} + u_{t-1}), \quad (5)$$

$$\pi_t^d(\phi_0) = -\phi_0^{-1}(s_t + u_t). \quad (6)$$

To model optimal monetary policy requires a decision about whether the first order conditions of the policy maker are represented by (5) or, instead, by (6). But how does one decide whether the behavior of the monetary authority should be classified as discretionary or commitment-like? We propose a general characterization of optimal monetary policy that nests 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 (5) or (6) is satisfied. This characterization also allows for arbitrarily frequent switches between commitment and discretion and is, therefore, consistent with the quasi-commitment model proposed by Schaumburg and Tambalotti (2007).[4]

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 (5) or (6). 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 *Optimal monetary policy implies that* $\Pr\left(\pi_t^c(\phi_0) \leq \pi_t(\phi_0) \leq \pi_t^d(\phi_0) \middle| s_{t-1} \leq 0\right) = 1$, *where* $\pi_t(\phi_0)$ *is the actual inflation rate in period* t.

The bounds for inflation in Lemma 1 are derived from equations (5) and (6). Recalling that u_t is a random variable with support in $\mathbb{R}^-$, it follows that

$$\Pr\left(s_{t-1} + u_{t-1} \leq 0 \middle| s_{t-1} \leq 0\right) = 1,$$

which implies that $\pi_t^d(\phi_0) \geq \pi_t^c(\phi_0)$, whenever $s_{t-1} \leq 0$.

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

Assumption 1 *Let* $\pi_t(\phi_0)$ *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 zero and variance* σ_v^2.

[4] In the Schaumburg and Tambalotti (2007) model, the monetary policy is delegated to a sequence of policymakers with tenures of random duration. Each policymaker is assumed to formulate optimal commitment plans that satisfy (5). However, when a new policymaker takes office it is optimal for her to deviate from the preexisting plan and, therefore, the inflation rate at the inception of a new policymaker's tenure satisfies (6).

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

$$\Pr\left(\pi_t^c(\phi_0) + v_t \leq \Pi_t \leq \pi_t^d(\phi_0) + v_t \middle| s_{t-1} \leq 0\right) = 1, \tag{7}$$

which establishes a lower bound and an upper bound for the observed inflation rate, Π_t. Denoting by $1(s_{t-1} \leq 0)$ an indicator function taking value 1 if $s_{t-1} \leq 0$, we derive moment inequalities that are implied by optimal monetary policy, as follows

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

$$-\mathrm{E}\left[(\phi_0\Pi_t + s_t + u_t - \phi_0 v_t)\, 1(s_{t-1} \leq 0)\right] \geq 0, \tag{8}$$

$$\mathrm{E}\left[(\phi_0\Pi_t + \Delta s_t + \Delta u_t - \phi_0 v_t)\, 1(s_{t-1} \leq 0)\right] \geq 0, \tag{9}$$

are implied by optimal monetary policy under either commitment or discretion, where $\phi_0 > 0$ denotes the "true" structural parameter and E *is the unconditional expectation operator.*

Proposition 1 follows immediately from (7), noticing that the bounds on inflation are valid any time $s_{t-1} \leq 0$ and, therefore, they also hold when multiplied by $1(s_{t-1} \leq 0)$.

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. *$\mathrm{E}\left[v_t\, 1(s_{t-1} \leq 0)\; Z_t\right] = 0$;*
3. *$\mathrm{E}\left[u_{t-r}\, 1(s_{t-1} \leq 0)\, Z_t\right] = \mathrm{E}\left[u_{t-r}\, 1(s_{t-1} \leq 0)\right]\mathrm{E}\left[Z_t\right]$, for $r = 0, 1$;*
4. *$\mathrm{E}\left[\Pi_t\, 1(s_{t-1} \leq 0)\, Z_t\right] \neq 0$, $\mathrm{E}\left[s_t\, 1(s_{t-1} \leq 0)\, Z_t\right] \neq 0$ and $\mathrm{E}\left[\Delta s_t\, 1(s_{t-1} \leq 0)_t\, Z_t\right] \neq 0$.*

Assumption 2.1. guarantees that (without loss of generality) the vector of instruments can be restricted to have positive support. The instrumental variables are assumed to be uncorrelated with the measurement error v_t and with $u_{t-r}\, 1(s_{t-1} \leq 0)$, for $r = 0, 1$. In particular, Assumption 2.3. implies that

$$\mathrm{E}\left[(\Delta u_t)\; 1(s_{t-1} \leq 0)\, Z_t\right] = \mathrm{E}\left[(\Delta u_t)\; 1(s_{t-1} \leq 0)\right]\mathrm{E}\left[Z_t\right].$$

In what follows, we assume that $\mathrm{E}\left[u_{t-1}\, 1(s_{t-1} \leq 0)\right] \approx \mathrm{E}\left[u_t\, 1(s_{t-1} \leq 0)\right] = \overline{u}_0$, so that the term $E\left[(\Delta u_t)\; 1(s_{t-1} \leq 0)\right] \approx 0$.[5] Finally, Assumption 2.4 requires that the instruments are relevant.

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

$$\mathrm{E}\Big[m_{d,t}(\phi_0, \overline{u}_0)\Big] \equiv \mathrm{E}\Big[-\left((\phi_0\Pi_t + s_t)\; 1(s_{t-1} \leq 0) + \overline{u}_0\right) Z_t\Big] \geq 0, \tag{10}$$

$$\mathrm{E}\Big[m_{c,t}(\phi_0)\Big] \equiv \mathrm{E}\Big[(\phi_0\Pi_t + \Delta s_t)\; 1(s_{t-1} \leq 0)\, Z_t\Big] \geq 0. \tag{11}$$

[5] This approximation is accurate provided $cov\left(\Delta u_t, 1(s_{t-1} \leq 0)\right) \approx 0$. If we do not make this approximation there is an additional nuisance parameter to be estimated, corresponding to $\mathrm{E}\left[u_{t-1}\, 1(s_{t-1} \leq 0)\right]$.

Notice that $\overline{u}_0$ is a nuisance parameter to be estimated along with ϕ_0, the structural parameter of interest. We use $\theta = (\phi,\ \overline{u}) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^-$ 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 (10) and (11). Thus, we define the identified set consistent with optimal monetary policy as follows

Definition 1 *Let $\theta = (\phi,\ \overline{u}) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^-$. 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

$$\mathrm{E}\left[m_t\left(\theta\right)\right] \equiv \begin{cases} \mathrm{E}\Big[m_{d,t}\left(\phi, \overline{u}\right)\Big] = \mathrm{E}\Big[-\left(\left(\phi\Pi_t + s_t\right)\ 1\left(s_{t-1} \leq 0\right) + \overline{u}\right) Z_t\Big], \\ \mathrm{E}\Big[m_{c,t}\left(\phi\right)\Big] = \mathrm{E}\Big[\left(\phi\Pi_t + \Delta s_t\right)\ 1\left(s_{t-1} \leq 0\right) Z_t\Big]. \end{cases} \tag{12}$$

Under optimal monetary policy, Θ^I is never empty. In fact, the moment inequality conditions in (12) can be written as

$$\mathrm{E}\Big[m_{d,t}(\phi, \overline{u})\Big] = \mathrm{E}\Big[m_{d,t}(\phi_0, \overline{u}_0)\Big] - \mathrm{E}\Big[\left(\phi - \phi_0\right)\Pi_t\, 1\left(s_{t-1} \leq 0\right) Z_t + \left(\overline{u} - \overline{u}_0\right) Z_t\Big] \geq 0 \tag{13}$$

$$\mathrm{E}\Big[m_{c,t}(\phi)\Big] = \mathrm{E}\Big[m_{c,t}(\phi_0)\Big] + \mathrm{E}\Big[\left(\phi - \phi_0\right)\Pi_t\, 1\left(s_{t-1} \leq 0\right) Z_t\Big] \geq 0, \tag{14}$$

where the first terms on the RHS of equations (13) and (14) are non-negative because of (10) and (11). Hence, by construction $(\phi_0, \overline{u}_0) \in \Theta^I$. On the other hand, Θ^I may be non-empty even if (7) does not hold. In fact, violation of (7) does not necessary imply a violation of (10) and/or (11). Thus, $(\phi_0, \overline{u}_0)$ 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.

Note that the set of values ϕ for which the inequality condition associated with (13) is satisfied increases linearly in $-\overline{u}$. In fact, since $\overline{u} \leq 0$, the smaller $\overline{u}$, the larger the set of values of ϕ satisfying the inequality constraint. Heuristically, the higher the level of distortions, the higher the level of inflation under discretion and, hence, the larger the range of inflation rates consistent with optimal monetary policy. This is illustrated graphically in Figure 1, which represents the linear relation between ϕ and $\overline{u}$ under discretion and for $Z_t = 1$. The area below the line represents, for each value of $\overline{u}$, the set of values of ϕ compatible with optimal monetary policy.

Although our moment inequalities are linear in the parameters, 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 sharp 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. 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 will estimate the identified set

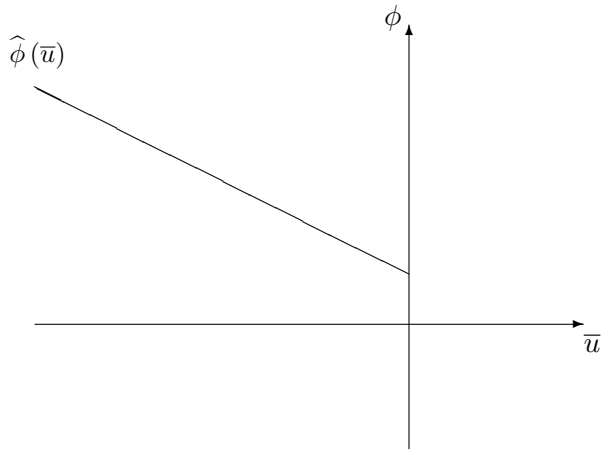


Figure 1: Discretion inequality condition in the $\overline{u}$ and ϕ space

according to the criterion function of Chernozhukov, Hong and Tamer (2007).

Before proceeding, notice that one may be tempted to reduce the two moment inequalities in (13)–(14) into a single moment equality condition, given by

$$\mathrm{E}\Big\{ \left[\phi\Pi_t + (\Delta s_t + \Delta u_t) + \varphi_t \left(s_{t-1} + u_{t-1}\right)\right] 1\left(s_{t-1} \leq 0\right) Z_t \Big\} = 0,$$

where $\varphi_t \in \{0, 1\}$ is a random variable taking value 0 in the case of commitment and 1 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 sometimes used in the industrial organization literature. However, if we allow for periodic switches between commitment and discretion, φ_t cannot be in general treated as a fixed parameter and this approach is no longer implementable (see Corts, 1999 and Rosen, 2007). For example, if the monetary policy authority sticks to a commitment plan unless s_{t-1} gets "too negative", then φ_t depends on s_{t-1} and/or $1\left(s_{t-1} < 0\right)$ and cannot be consistently estimated. Instead, the inflation bounds that we derive are compatible with a monetary authority that deviates from commitment-based optimal plans with some probability, not necessarily exogenous.

4 Set Identification

In this section we describe how to estimate the identified set Θ^I using a partial identification approach. The basic idea underlying the estimation strategy is to use the bounds for the observed inflation rate derived from the theoretical model to generate a family of moment inequality conditions that are consistent with optimal policy. These moment inequality conditions are then used to construct a criterion function whose set of minimizers is the estimated identified set. Provided that the estimated identified set is non-empty, we then proceed to construct the corresponding confidence region.

We define the following $2p$ vector of moment conditions associated with (12)

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

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

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

where $m_t^i(\theta)$ is the i-th element of $m_t(\theta)$. Let $V(\theta)$ be the asymptotic variance of $\sqrt{T}m_T(\theta)$ and $\widehat{V}_T(\theta)$ the corresponding heteroscedasticity and autocorrelation consistent (HAC) estimator.[6] The criterion function we use for the inferential procedure is

$$Q_T(\theta) = \sum_{i=1}^{2p} \frac{\left[m_T^i(\theta)\right]_-^2}{\hat{v}_T^{i,i}(\theta)}, \tag{15}$$

where $[x]_- = x\,1(x\leq 0)$, and $\hat{v}_T^{i,i}(\theta)$ is the $i-$th element of the diagonal of $\widehat{V}_T(\theta)$. The probability limit of $Q_T(\theta)$ is given by $Q(\theta) = \mathrm{p}\lim_{\mathrm{T}\to\infty} Q_T(\theta)$. 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.

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

$$\widehat{\Theta}_T^I = \left\{\theta\in\Theta \text{ s.t. } TQ_T(\theta)\leq d_T^2\right\}, \tag{16}$$

where d_T satisfies the conditions in Proposition 2 below.

Assumption 3 *The following conditions are satisfied*

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

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

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

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

4. $\sup_{\theta\in\Theta}\left|\nabla_\theta m_T(\theta) - D(\theta)\right| \xrightarrow{pr} 0$ *uniformly for all* θ *in* Θ*, where* $D(\theta)$ *is full rank.*

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

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

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

and $\rho_H\left(\widehat{\Theta}_T^I, \Theta^I\right) = O_p\left(\frac{d_T}{\sqrt{T}}\right)$.[7]

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

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

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

where

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

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

To compute the critical value $c_{\alpha,T}$ of the distribution of $\sup_{\theta\in\Theta^I} TQ_T(\theta)$, we replace the unknown set Θ^I by its consistent estimator $\widehat{\Theta}_T^I$, as shown in Proposition 2, and we use bootstrap critical values.[8] In order to reproduce the serial correlation of the moment conditions, we rely on block-bootstrap. In particular, let $T = bl$, where b denotes the number of blocks and l denotes the block length, and let $W_t^* = (\Pi_t^*, s_t^*, Z_t^*)$ denote the re-sampled observations. For each $\theta\in\widehat{\Theta}_T^I$, we construct

$$TQ_T^*(\theta) = \sum_{i=1}^{2p}\left(\sqrt{T}\left[\frac{m_T^{i*}(\theta) - m_T^i(\theta)}{\sqrt{\widehat{v}^{i,i*}(\theta)}}\right]_- 1\left[m_T^i(\theta) \le \sqrt{\widehat{v}^{i,i}(\theta)}\sqrt{2\ln\ln T/T}\right]\right)^2, \qquad (17)$$

where $m_T^{i*}(\theta)$ is the bootstrap analog of the sample moment conditions $m_T^i(\theta)$, constructed using the bootstrapped data (Π_t^*, s_t^*, Z_t^*), and $\widehat{v}^{i,i*}(\theta)$ is the $i-$th element in the diagonal of the bootstrap analog of the variance of the moment conditions V_θ^*. The indicator function in (17) implements the Generalized Moment Selection (GMS) procedure introduced by Andrews and Soares (2010), using information about the slackness of the sample moment conditions to infer which population

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

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

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 construct the $(1-\alpha)$-percentile $c^*_{\alpha,T,B}$. The following proposition can be established:

Proposition 3 *Given Assumptions 1–3, and given $\widehat{\Theta}_T^I$ defined as in (16), as $T, B \to \infty$, $l \to \infty$, $l^2/T \to 0$*

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

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

5 Testing for the policy regime

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 a θ, then we have evidence in favor of discretion (commitment). Thus, we test for either policy regimes, each time maintaining the assumption of optimal monetary policy. The test consists of a two-step procedure that requires in the first step estimating the structural parameter under the null hypothesis, and in the second step testing if the parameter is in the identified set implied by optimal monetary policy. Under discretion, the moment inequalities in (10) hold as equalities and point identify $\theta_0 \equiv (\phi_0, \overline{u}_0)$. Instead under commitment, the moment inequalities in (11) hold as equalities point-identifying ϕ_0, while $\overline{u}_0$ can only be set-identified by the upper bound implied by optimal monetary policy.

5.1 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 is given by

$$\mathrm{E}\Big[m_{d,t}(\phi_0, \overline{u}_0)\Big] = \mathrm{E}\Big[-\left(\left(\phi_0\Pi_t + s_t\right) 1\left(s_{t-1} \leq 0\right) + \overline{u}_0\right) Z_t\Big] \quad = \quad 0, \tag{18}$$

$$\mathrm{E}\Big[m_{c,t}(\phi_0)\Big] = \mathrm{E}\Big[\left(\phi_0\Pi_t + \Delta s_t\right) 1\left(s_{t-1} \leq 0\right) Z_t\Big] \quad \geq \quad 0, \tag{19}$$

where the moment equality conditions in (18) follow from the assumption of discretion and the moment inequality conditions (19) impose a lower bound to the observed inflation rate as implied by optimal monetary policy. As already mentioned, conditions (18) point identify both ϕ_0 and the nuisance parameter $\overline{u}_0$, 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 3 *Let $\theta_0 \equiv (\phi_0,\ \overline{u}_0) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^-$. We define the null hypothesis of discretion and optimal monetary policy as,*

$$H_0^d :\ \theta_0 \textit{ satisfies conditions } (18)\text{–}(19).$$

against the alternative

$$H_1^d: \ \theta_0 \text{ does not satisfy conditions } (18)\text{–}(19).$$

To test the null hypothesis of discretion we follow a two-step procedure. Under the null hypothesis, the "true" structural parameter θ_0 is point-identified and it can be consistently estimated via the optimal GMM estimator using the moment conditions (18).[9] Thus, to test the null hypothesis of discretion we first obtain an estimate for the structural parameter vector using the optimal GMM estimator, denoted $\widehat{\theta}_T^d$. In the second step, we construct the following test statistic

$$TQ_T^d\left(\widehat{\theta}_T^d\right) = T\left(\sum_{i=1}^{p}\frac{m_{d,T}^i\left(\widehat{\theta}_T^d\right)^2}{\hat{v}_T^{i,i}\left(\widehat{\theta}_T^d\right)} + \sum_{i=p+1}^{2p}\frac{\left[m_{c,T}^i\left(\widehat{\theta}_T^d\right)\right]_-^2}{\hat{v}_T^{i,i}\left(\widehat{\theta}_T^d\right)}\right), \tag{20}$$

where $\hat{v}_T^{i,i}\left(\widehat{\theta}_T^d\right)$ is the i-th diagonal element of $\widehat{V}_T\left(\widehat{\theta}_T^d\right)$, the HAC estimator of the asymptotic variance of $\sqrt{T}\left[m_{d,T}\left(\widehat{\theta}_T^d\right), m_{c,T}\left(\widehat{\theta}_T^d\right)\right]$, which takes into account the estimation error in $\widehat{\theta}_T^d$.[10] Notice that since the first p moment conditions hold with equality, they all contribute to the asymptotic distribution of $TQ_T^d\left(\widehat{\theta}_T^d\right)$. Thus, we apply the GMS procedure only to the inequality moment conditions.

Andrews and Soares (2010) study the limiting distribution of the statistic in (20) 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. Furthermore, we need to 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 4 *Let Assumptions 1–3 hold. Let $c_{T,B,\alpha}^{*d}$ be the $(1-\alpha)$ percentile of the empirical distribution of $TQ_T^{*d}\left(\widehat{\theta}_T^{*d}\right)$, the bootstrap counterpart of $TQ_T^d\left(\widehat{\theta}_T^d\right)$. Then, as $T, B \to \infty$, $l \to \infty$, $l^2/T \to 0$:*

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

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

where B denotes the number of bootstrap replications.

[9] If we assume that θ_0 satisfies (18)–(19), 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 $\theta_0 \in \Theta^I$ satisfying (18)–(19).

[10] See Appendix B for the definition of $\widehat{V}_T\left(\widehat{\theta}_T^d\right)$.

5.2 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

$$\mathrm{E}\Big[m_{d,t}\left(\phi_0, \overline{u}_0\right)\Big] = \mathrm{E}\Big[-\left(\left(\phi_0 \Pi_t + s_t\right) 1\left(s_{t-1} \leq 0\right) + \overline{u}_0\right) Z_t\Big] \quad \geq \quad 0, \tag{21}$$

$$\mathrm{E}\Big[m_{c,t}\left(\phi_0\right)\Big] = \mathrm{E}\Big[\left(\phi_0 \Pi_t + \Delta s_t\right) 1\left(s_{t-1} \leq 0\right) Z_t\Big] \quad = \quad 0, \tag{22}$$

where the moment equality condition (22) follows from the assumption of commitment and the moment inequality condition (21) imposes an upper bound to the observed inflation rate, as implied by optimal monetary policy. Notice that conditions (21)–(22) only set-identify $\overline{u}_0$. Intuitively, this happens because under commitment the average level of distortions is irrelevant for the conduct of optimal monetary policy, since there is no inflationary bias.[11] We define the following test for optimal monetary policy under commitment.

Definition 4 *Let $\theta_0 \equiv (\phi_0,\ \overline{u}_0) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^-$. We define the null hypothesis of commitment and optimal monetary policy as,*

$$H_0^c:\ \theta_0 \text{ satisfies conditions } (21)\text{–}(22).$$

against the alternative

$$H_1^c:\ \theta_0 \text{ does not satisfy conditions } (21)\text{–}(22).$$

Although the null hypothesis of optimal monetary policy under commitment has the same structure as the null hypothesis under discretion, the implementation of the test is different because the conditions (22) do not point-identify $\overline{u}_0$. Therefore, we must implement the test for a sequence of values for $\overline{u}$, and select the most conservative test statistic. Instead, ϕ_0 is point identified and it can be estimated via optimal GMM using the moment equalities in (22). Let $\widehat{\phi}_T^c$ denote the estimated parameter. For a fixed $\overline{u}$, we construct the test statistic

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

and compute the corresponding critical value $c_{T,B,\alpha}^{*c}\left(\overline{u}\right)$, as discussed in Section 5.1. In practice, we construct the test statistic over a dense grid of values for $\overline{u}$ and report the test statistic yielding the highest p-value.

Proposition 5 *Let Assumptions 1–3 hold. Let $c_{T,B,\alpha}^{*c}\left(\overline{u}\right)$ be the $(1-\alpha)$ percentile of the empirical distribution of $TQ_T^{*c}\left(\widehat{\phi}_T^{*c}, \overline{u}\right)$, the bootstrap counterpart of $TQ_T^c\left(\widehat{\phi}_T^c, \overline{u}\right)$. Then, as $T, B \to \infty$,*

[11] In the commitment case, as $\overline{u}_0$ cannot be point identified, the estimation procedure of Moon and Schorfheide (2009) cannot be directly implemented, even if we were willing to assume that (21) and (22) hold.

$l \to \infty$, $l^2/T \to 0$:

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

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

where B *denotes the number of bootstrap replications.*

6 Empirical Application

In the previous section, we proposed a model specification test based on the criteria functions (20) and (23) to test for, respectively, discretion and commitment, under the maintained assumption of optimal monetary policy. We now apply this framework to study the monetary policy in the United States.

6.1 Data

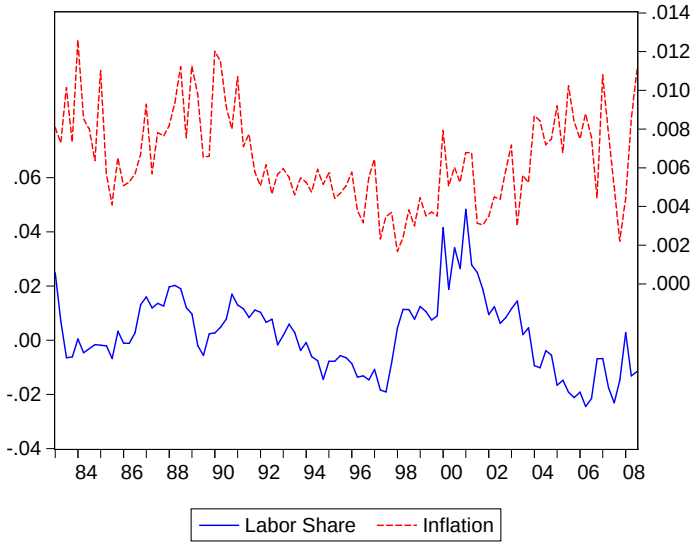
We use quarterly time-series for the US economy for the sample period 1983:1 to 2008:3. Following Sbordone (2002) and Gali and Gertler (1999), we exploit the relationship between the economy-wide real marginal cost s_t and the labor income share (equivalently, real unit labor costs). In the theoretical economy, the real marginal cost is proportional to the unit labor cost (see Appendix A). Hence, we use the linearly detrended labor income share in the non-farm business sector to measure s_t. Our measure of inflation is the annualized percentage change in the GDP deflator.

The econometric framework developed in this paper relies on a stationarity assumption (see Assumption 3). 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, Gali and Gertler (2000), we have decided to study the sample starting from 1983:1, that removes the first three years of the Volcker era. Clarida, Gali and Gertler (2000) offer two reasons for doing this. First, this period was characterized by a sudden and permanent disinflation episode bringing inflation down from about 10 percent to 4 percent. Second, over the period 1979:4-1982:4, the operating procedures of the Federal Reserve involved targeting non-borrowed reserves as opposed to the Federal Funds rate. Thus, our empirical analysis focuses on the sample period 1983:1 to 2008:3, which spans the period starting after the disinflation and monetary policy shifts that occurred in the early 1980s and extends until the period when the interest rate zero lower bound becomes a binding constraint.[12] Figure 2 plots the time series of the US labor income share, s_t, and inflation for the sample period 1983:1 to 2008:3.

It is frequently assumed that movements in military purchases are exogenous; moreover, fluctuations in military spending account for much of the variation in government purchases (see

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

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



Hall 2009). Thus, we use as instrument the variable *'Military Spending'*, given by the log of real government expenditure in national defense detrended using the HP-Filter.[13] As a second instrument, we use the variable *'Oil Price Change'*, given by the log difference of the spot oil price. The instrumental variables are adjusted using the following transformation guaranteeing positiveness: $Z_+ = Z - min(Z)$. The complete instrument list includes the variables *'Military Spending'*, *'Oil Price Change'*, and the unit vector, yielding $p = 3$ instruments and 6 moment conditions overall.

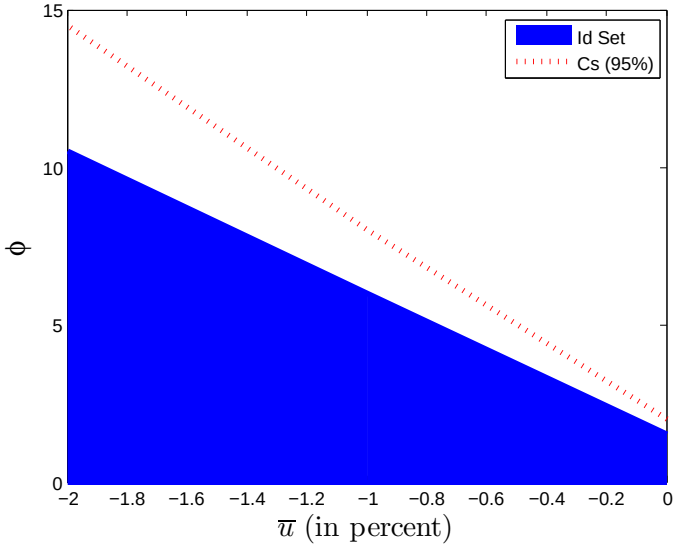
6.2 Model specification tests

We begin by providing the estimation results for the identified set. For the implementation of the estimator, we specify a truncation parameter for the computation of the HAC variance estimator equal to four. The confidence sets are constructed using 1,000 block-bootstrap replications, also with block size equal to four. In Figure 3, we show the estimated identified set implied by optimal monetary policy $\widehat{\Theta}_T^I$, and the corresponding confidence region at 95% confidence level, $\widehat{C}_T^{0.95}$. As discussed earlier, the upper-bound of the identified set grows linearly in the level of distortions given by $-\overline{u}$. Moreover, the fact that the identified set is not empty provides some evidence in favor of a path for inflation and output gap consistent with optimal monetary policy.

We next examine the formal test statistics developed in sections 5.1 and 5.2 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 (18); next, using the estimated vector of parameters $\widehat{\theta}_T^d$ we construct the test statistic for discretion $TQ_T^d\left(\widehat{\theta}_T^d\right)$ and compute the bootstrap critical value. To

[13] Benigno and Woodford (2005) show that in the case of a distorted steady state (i.e., when $\overline{u} \neq 0$) the optimal response to a variation in government purchases involves changes in the inflation rate and the output gap (our endogenous regressors).

Figure 3: Identified Set and Confidence Set under Optimal Monetary Policy



Note: The shaded area corresponds to the estimated identified set. The dashed line denotes the confidence set at 95% confidence level.

test commitment, we proceed in the same way, except that the moment condition (22) only permits identification of the parameter ϕ^c. Thus, after estimating $\widehat{\phi}^c$ by optimal GMM, we implement the test for a sequence of values for $\overline{u}$. For each $\overline{u}$, we construct the test statistic $TQ_T^c\left(\widehat{\phi}^c;\overline{u}\right)$ and compute the bootstrap critical value. Therefore, to test commitment we consider the test statistic over a grid of values for $\overline{u}$ and select the most conservative test statistic.

Given that we are using a sufficient number of instrumental variables for overidentification, we start by reporting results from the standard Hansen test statistic for overidentifying restrictions. The upper panel of Table 1 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 (18) and (22). Under the null of commitment or discretion, we can also derive the moment conditions consistent with commitment or discretion directly from the first order conditions in equations (6) and (5). These are given by, respectively,

$$\mathrm{E}\Big[-\left(\phi_0\Pi_t + s_t + u_t\right)Z_t\Big] = 0 \tag{24}$$

$$\mathrm{E}\Big[\left(\phi_0\Pi_t + \Delta s_t\right)Z_t\Big] = 0 \tag{25}$$

In this case, the moment conditions does not include the indicator function taking value 1 when $s_{t-1} \leq 0$ (which we used to derive the state-contingent inflation bounds) and, thus, uses the full sample of data to estimate the parameters. The corresponding test statistics for overidentifying restrictions, $J-$test*, are reported in the middle panel of Table 1. As can be seen from the Table,

Table 1: Model Specification Tests

	Discretion	Commitment
$J-$test	2.675	3.304
p-val	(0.176)	(0.261)
$J-$test*	0.162	1.235
p-val	(0.636)	(0.562)
TQ_T	11.177	63.998
p-val	(0.244)	(0.015)

Note: The upper and middle panels report the Hansen test of overidentifying restriction based on the GMM model under discretion or commitment. The $J-$test reported in the top panel refer to the moment conditions (18) and (22), respectively. The $J-$test* reported in the middle panel refer to the moment conditions given by (24) and (25), respectively. The p values for the $J-$tests are computed using 1,000 block-bootstrap replications with block-size 4. The bottom panel reports the test statistic TQ_T, defined by equations (20) and (23), for commitment and discretion. The p values for the the test statistics TQ_T are computed via the bootstrapping procedure described in Sections 5.1 and 5.2.

in both cases the standard test for overidentifying restrictions fails to reject either model.[14] Thus, by not making use of the full set of implications of optimal monetary policy, we are unable to discriminate between the two alternative policy regimes.

However, using the additional information implied by the maintained assumption of optimal monetary policy, we can test the composite null hypothesis of optimal monetary policy and a specific policy regime—discretion or commitment. The test statistic is based on equation (20) for the case of discretion, and equation (23) for the case of commitment. The results are shown in the bottom panel of Table 1. For the case of discretion, the p-value associated with the test statistic is 24.4 percent and, therefore, we fail to reject the null hypothesis of discretion at all conventional levels. For the case of commitment, the parameter $\overline{u}$ is not identified by GMM and, therefore, needs to be fixed. We consider a dense grid of values for $\overline{u}$, and the resulting p-values associated with the test statistic (23) range between 1.1 percent and 1.5 percent. Hence, even choosing the most conservative case, a p-value of 1.5 percent allows for rejection of the null hypothesis of optimal policy under commitment at the 5% confidence level.

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

7 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 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 and yield set identification of the economy's structural parameters. We propose a two-step model specification test that makes use of the set of moment inequality and equality conditions implied by optimal monetary policy under a specific policy regime. We test the null hypotheses of discretionary optimal monetary policy and of optimal monetary policy under commitment.

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.

By extending the Generalized Moment Selection method of Andrews and Soares (2010) to take into account the contribution of parameter estimation error on the relevant covariance matrix, our two-step testing procedure can be used more generally to test the validity of models combining moment equalities and inequality conditions, when the parameters of the model can be consistently estimated under the null hypothesis.

Appendix

A The Structural Model

The framework is that of the New Keynesian forward-looking model with monopolistic competition and Calvo price-setting exposed in Woodford (2011). The representative household seeks to maximize the following utility function

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ u\left(C_t; \xi_t\right) - \int_0^1 v\left(H_t\left(j\right); \xi_t\right) dj \right\},$$

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

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

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

$$\frac{v_h\left(H_t\left(j\right); \xi_t\right)}{u_c\left(C_t; \xi_t\right)} \chi_t = w_t\left(j\right), \tag{26}$$

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

The aggregate resource constraint is given by

$$C_t + G_t \leq \left[\int_0^1 y_t\left(i\right)^{(\vartheta-1)/\vartheta} di \right]^{\vartheta/(\vartheta-1)} \equiv Y_t \tag{27}$$

where $y_t\left(i\right)$ is the quantity produced of commodity i and G_t (which is included in the vector of exogenous disturbances ξ_t) represents the government's expenditure on an exogenously given quantity of the same basket of commodities that is purchased by the households. In equilibrium condition (27) holds with equality. Except for the fact that labor is immobile across industries, there is a common technology for the production of all goods: the commodity i is produced according to the production function

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

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

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

where the real marginal cost function is defined by

$$S\left(y, Y; \tilde{\xi}\right) = \chi \frac{v_h\left(f^{-1}(y); \xi\right)}{u_c(Y - G; \xi)} \left[f'\left(f^{-1}(y)\right)\right]^{-1}, \tag{28}$$

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

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

where the aggregate price index P_t is defined by

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

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

$$\frac{1}{\mu} = S\left(Y_t^n, Y_t^n; \tilde{\xi}_t\right), \tag{29}$$

where Y_t^n is the natural rate of output, which corresponds to the equilibrium level of output under flexible prices. Nonetheless, it is assumed prices are sticky, so that the output of each commodity i differs. A log-linear approximation to the marginal cost function (28) around the deterministic equilibrium under flexible prices yields the condition

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

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

$$s_t = \left(\omega + \sigma^{-1}\right)\left(\widehat{Y}_t - \widehat{Y}_t^n\right), \tag{30}$$

where s_t denotes the economy-wide real marginal cost in log deviation from steady state. Thus the output gap is related to variations in the economy-wide real marginal cost.[15]

[15] The output gap $\widehat{Y}_t - \widehat{Y}_t^n$ is not directly observable and there is no reason to believe that it can be proxied by the deviation of output from a smooth statistical trend. Sbordone (2002) and Gali and Gertler (1999) notice that the most direct way to measure time variation in the output gap is based on the variation in the production costs, as implied by equation (30).

To model the behavior of prices, we employ the probabilistic model due to Calvo (1983). In any period each industry j has a fixed probability $1-\alpha$ that it may adjust its price during that period. Else, with probability α, the producers in that industry must keep their price unchanged. A producer that changes its price in period t, chooses the new price to maximize the discounted flow of profits

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

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

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

and the profit function is given by

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

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

$$P_t=\left[\alpha P_{t-1}^{1-\vartheta}+\left(1-\alpha\right)p_t^{\star 1-\vartheta}\right]^{1/(1-\vartheta)}. \tag{31}$$

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

$$\mathbb{E}_t\left[\sum_{s=t}^{\infty}\left(\alpha\beta\right)^{s-t}\Gamma\left(\frac{p_t^\star}{P_s},Y_s,\tilde{\xi_s}\right)\right]=0, \tag{32}$$

where the function Γ is given by

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

By log-linearizing equations (31) and (32) around the deterministic steady state (under zero inflation), we obtain the following two conditions

$$\log P_t=\alpha\log P_{t-1}+\left(1-\alpha\right)\log p_t^\star \tag{33}$$

and

$$\sum_{s=t}^{\infty}(\alpha\beta)^{s-t}\,\mathbb{E}_t\left[\log p_t^{\star}-\log P_s-\zeta\left(\widehat{Y}_s-\widehat{Y}_s^n\right)\right]=0 \tag{34}$$

where $\zeta=\left(\omega+\sigma^{-1}\right)/\left(1+\omega\vartheta\right)$. Combining the equations (33) and (34) yields the New Keynesian Phillips curve, given by

$$\pi_t=\beta\mathbb{E}_t\pi_{t+1}+\kappa x_t,$$

where $\kappa=(1-\alpha)(1-\alpha\beta)(\zeta/\alpha)$, and the variable $x_t\equiv\left(\widehat{Y}_t-\widehat{Y}_t^n\right)$ is the output gap, which is proportional to the average real marginal cost. Thus, making use of equation (30) we obtain the expression (1) in the main text, with $\psi\equiv\kappa/\left(\omega+\sigma^{-1}\right)$.

The efficient level of output satisfies the condition

$$S\left(Y_t^{\star},Y_t^{\star},\tilde{\xi}_t\right)=\chi_t, \tag{35}$$

and corresponds to the level of output under flexible prices and without distortions resulting from firm's market power and the labor wedge. Thus, from equations (29) and (35), we derive the following relationship

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

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

$$\mathcal{W}=\mathbb{E}_0-\frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left\{\pi_t^2+\frac{\kappa}{\vartheta}\left[x_t+\left(\omega+\sigma^{-1}\right)^{-1}u_t\right]^2\right\} \tag{36}$$

$$\propto\mathbb{E}_0\left\{-\frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left[\phi\pi_t^2+\psi\left(s_t+u_t\right)^2\right]\right\},$$

where $\phi\equiv\left(\omega+\sigma^{-1}\right)\vartheta$. The derivation of (36) is given in Benigno and Woodford (2005).

B Proofs

Proof of Lemma 1: Immediate from the definition of $\pi_t^c(\phi_0)$ and $\pi_t^d(\phi_0)$, Equations (5)–(6).

Proof of Proposition 1: Given Assumption 1, it follows straightforwardly from Equation (7).

Proof of Proposition 2: The statement follows from Theorem 3.1 in Chernozhukov, Hong and Tamer (2007), with $\widehat{c}=d_T^2$, $a_T=T$, $\gamma=2$, once we show that Assumptions 1–3 imply the satisfaction of their Conditions 1 and 2. Condition 1(a) holds as $\theta=(\phi,\overline{u})$ lies in a compact subset of $\mathbb{R}^+\times\mathbb{R}^-$. Assumptions 1–2 allow to state the sample objective function as $Q_T(\theta)$ in (15). 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 (15) satisfies Condition 2.

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

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

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

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

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

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

with $D(\theta) = diag(V(\theta))$ and $V(\theta) = p\lim_{T\to\infty}\widehat{V}_T(\theta)$, as defined in footnote 6. Finally, $h(\theta) = (h_1(\theta), ..., h_{2p}(\theta))'$ is a vector measuring the slackness of the moment conditions, given by

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

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

$$\sup_{\theta\in\widehat{\Theta}_T^I} TQ_T(\theta) \xrightarrow{d} \sup_{\theta\in\Theta^I}\sum_{i=1}^{2p}\left(\left[\sum_{j=1}^{2p}\omega_{i,j}(\theta)\mathcal{Z}_i + h_i(\theta)\right]_-\right)^2. \tag{37}$$

We need to show that the $(1-\alpha)$-percentile of the right-hand side of (37), $c_{\alpha,T}$, 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 $i = 1, ..., 2p$, we have that, almost surely,

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

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

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

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

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

Hence, as T gets large, only the moment conditions holding with equality contribute to the bootstrap limiting distribution, and the probability of eliminating a non-slack moment condition approaches zero. Further, given the block resampling scheme, for all i, $\mathrm{E}^*\left(\sqrt{T}\left(m^*_{i,T}(\theta)-m_{i,T}(\theta)\right)\right)=O_p\left(l/\sqrt{T}\right)$ and $\mathrm{var}^*\left(\sqrt{T}\left(m^*_{i,T}(\theta)\right)\right)=\widehat{v}^{i,i}(\theta)+O_p\left(l/\sqrt{T}\right)$, where E^* and var^* denote the mean and variance operator under the probability law governing the resampling scheme. Since $l=o\left(\sqrt{T}\right)$, as $T\to\infty$, conditional on the sample,

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

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

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

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

where $\widetilde{\theta}^d_T=\arg\min_\theta m_{d,T}(\theta)'\,m_{d,T}(\theta)$, and $\widehat{\Omega}_{dd,T}\left(\widetilde{\theta}^d_T\right)$ is the HAC estimator of the variance of $\sqrt{T}m_{d,T}(\theta_0)$. If we knew $\theta_0=(\phi_0,\overline{u}_0)$, the statement would follow by a similar argument as in the proof of Proposition 3, 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, $\widehat{\theta}^d_T$. Thus, the parameter estimation error term, $\sqrt{T}\left(\widehat{\theta}^d_T-\theta_0\right)$, 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,

$$\sqrt{T}m_{d,T}\left(\widehat{\theta}^d_T\right)=\sqrt{T}m_{d,T}(\theta_0)+D_{d,T}\left(\overline{\theta}^d_T\right)\sqrt{T}\left(\widehat{\theta}^d_T-\theta_0\right) \tag{38}$$

$$\sqrt{T}m_{c,T}\left(\widehat{\theta}^d_T\right)=\sqrt{T}m_{c,T}(\theta_0)+D_{c,T}\left(\overline{\theta}^d_T\right)\sqrt{T}\left(\widehat{\theta}^d_T-\theta_0\right) \tag{39}$$

with $\overline{\theta}_T^d \in \left(\widehat{\theta}_T^d, \theta_0\right)$, $D_{d,T}(\theta) = \nabla_\theta m_{d,T}(\theta)$ and $D_{c,T}(\theta) = \nabla_\theta m_{c,T}(\theta)$. From (38) it follows that

$$\begin{aligned}\operatorname{avar}\left(\sqrt{T}m_{d,T}\left(\widehat{\theta}_T^d\right)\right) &= \operatorname{avar}\left(\sqrt{T}m_{d,T}(\theta_0)\right) + \operatorname{avar}\left(D_{d,T}\left(\widehat{\theta}_T^d\right)\sqrt{T}\left(\widehat{\theta}_T^d - \theta_0\right)\right) + \\ &\quad +2\operatorname{acov}\left(\sqrt{T}m_{d,T}(\theta_0), D_{d,T}\left(\widehat{\theta}_T^d\right)\sqrt{T}\left(\widehat{\theta}_T^d - \theta_0\right)\right). \end{aligned} \tag{40}$$

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

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

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

$$\sqrt{T}\left(\widehat{\theta}_T^d - \theta_0\right) = -\widehat{B}_{d,T}D_{d,T}\left(\widehat{\theta}_T^d\right)'\widehat{\Omega}_{dd,T}\left(\widehat{\theta}_T^d\right)^{-1}\sqrt{T}m_{d,T}(\theta_0), \tag{41}$$

with

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

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

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

where the first entry can be computed using (40) and and (41), i.e.

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

Also,

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

Note that, for the computation of 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 $\widehat{V}_{dc,T}\left(\widehat{\theta}_T^d\right)$. Let

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

It is easy to see that $V_{dd}\left(\theta_0\right)$ is of rank $p-2$, while $V_{cc}\left(\theta_0\right)$ is of full rank p, hence the asymptotic variance covariance matrix $V\left(\theta_0\right)$ is of rank $2p-2$. However, this is not a problem, as we are only concerned with the elements along the main diagonal.

We now outline how to construct bootstrap critical values. The bootstrap counterpart of $TQ_T^d\left(\widehat{\theta}_T^d\right)$ writes as:

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

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

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

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

$$\widehat{\Omega}_{dd,T}^*\left(\widetilde{\theta}_T^{*d}\right) = \frac{1}{T}\sum_{k=1}^{b}\sum_{j=1}^{l}\sum_{i=1}^{l}\left(m_{d,I_k+i}\left(\widetilde{\theta}_T^{*d}\right)-m_{d,T}\left(\widehat{\theta}_T^d\right)\right)\left(m_{d,I_k+j}\left(\widetilde{\theta}_T^{*d}\right)-m_{d,T}\left(\widehat{\theta}_T^d\right)\right)', \quad (42)$$

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

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

As for the computation of the bootstrap critical values, we need only the elements along the main

diagonal, below we report only the expressions for $\widehat{V}^*_{dd,T}\left(\widehat{\theta}^{*d}_T\right)$ and $\widehat{V}^*_{cc,T}\left(\widehat{\theta}^{*d}_T\right)$, which are

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

where

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

with $\widehat{D}^*_{d,T}\left(\widehat{\theta}^{*d}_T\right) = \nabla_\theta m^*_{d,T}\left(\widehat{\theta}^{*d}_T\right)$ and where $\widehat{\Omega}^*_{dd,T}\left(\widehat{\theta}^{*d}_T\right)$ is defined as in (42), but with $\widetilde{\theta}^{*d}_T$ replaced by $\widehat{\theta}^{*d}_T$, also

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

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

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

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

We compute B bootstrap replication of $TQ^{*d}_T\left(\widehat{\theta}^{*d}_T\right)$, say $TQ^{*d}_{T,1}\left(\widehat{\theta}^{*d}_T\right), ..., TQ^{*d}_{T,B}\left(\widehat{\theta}^{*d}_T\right)$, and compute the $(1-\alpha)-$th percentile of its empirical distribution, $c^{*d}_{T,B,\alpha}\left(\widehat{\theta}^{*d}_T\right)$. 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^d_0 whenever $TQ^d_T\left(\widehat{\theta}^d_T\right)$ is larger than (smaller than or equal to) $c^{*d}_{T,B,\alpha}\left(\widehat{\theta}^{*d}_T\right)$ provides a test with asymptotic size α and unit asymptotic power. To this end, we show that, under H^d_0, $TQ^{*d}_T\left(\widehat{\theta}^{*d}_T\right)$ has the same limiting distribution as $TQ^d_T\left(\widehat{\theta}^d_T\right)$, conditionally on the sample, and for all samples except a set of probability measure approaching zero. On the other hand, under H^d_1, $TQ^{*d}_T\left(\widehat{\theta}^{*d}_T\right)$ has a well defined limiting distribution, while $TQ^d_T\left(\widehat{\theta}^d_T\right)$ diverges to infinity.

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

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

and

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

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

$$\mathrm{E}^{*}\left(\sqrt{T}\left(m_{d,T}^{*}\left(\widehat{\theta}_{T}^{*d}\right)-m_{d,T}\left(\widehat{\theta}_{T}^{d}\right)\right)\right)=O_{p}\left(\frac{l}{\sqrt{T}}\right)=o_{p}(1),$$

$$\mathrm{var}^{*}\left(\sqrt{T}\left(m_{d,T}^{*}\left(\widehat{\theta}_{T}^{*d}\right)-m_{d,T}\left(\widehat{\theta}_{T}^{d}\right)\right)\right)=\widehat{V}_{T}\left(\widehat{\theta}_{T}^{d}\right)+O_{p}\left(\frac{l}{\sqrt{T}}\right)=\widehat{V}_{T}\left(\widehat{\theta}_{T}^{d}\right)+o_{p}(1),$$

and

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

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

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

The statement in the Proposition then follows.

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

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

where $\widetilde{\phi}_{T}^{c}=\arg\min_{\phi}m_{c,T}(\phi)'\,m_{c,T}(\phi)$ and $\widehat{\Omega}_{cc,T}\left(\widetilde{\phi}_{T}^{c}\right)$ is the HAC estimator of the variance of $\sqrt{T}m_{c,T}(\phi_{0})$. Via mean value expansion

$$\begin{aligned}\sqrt{T}m_{d,T}\left(\widehat{\phi}_{T}^{c},\overline{u}\right) &= \sqrt{T}m_{d,T}(\phi_{0},\overline{u})+D_{d,T}(\overline{\phi},\overline{u})\sqrt{T}\left(\widehat{\phi}_{T}^{c}-\phi_{0}\right)\\ \sqrt{T}m_{c,T}\left(\widehat{\phi}_{T}^{c}\right) &= \sqrt{T}m_{c,T}(\phi_{0})+D_{c,T}(\overline{\phi})\sqrt{T}\left(\widehat{\phi}_{T}^{c}-\phi_{0}\right)\end{aligned}$$

where $\overline{\phi}\in(\widehat{\phi}_{T}^{c},\phi_{0})$, $D_{d,T}(\phi,\overline{u})=\nabla_{\phi}m_{d,T}(\phi,\overline{u})$ and $D_{c,T}(\phi)=\nabla_{\phi}m_{c,T}(\phi)$. Expanding the GMM first order condition around ϕ_{0}

$$\sqrt{T}\left(\widehat{\phi}_{T}^{c}-\phi_{0}\right)=-\widehat{B}_{c,T}D_{c,T}\left(\widehat{\phi}_{T}^{c}\right)'\widehat{\Omega}_{cc,T}\left(\widehat{\phi}_{T}^{c}\right)^{-1}\sqrt{T}m_{c,T}(\phi_{0})$$

where

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

The asymptotic variance of the moment conditions $\sqrt{T} m_T (\phi_0, \overline{u})$ can be estimated by

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

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

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

where

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

Also,

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

Let

$$V_{cc} (\phi_0) = \text{plim}_{T \to \infty} \widehat{V}_{cc,T} \left(\widehat{\phi}^c_T \right), \quad V_{dd} (\phi_0, \overline{u}) = \text{plim}_{T \to \infty} \widehat{V}_{dd,T} \left(\widehat{\phi}^c_T, \overline{u} \right).$$

Again, it is easy to see that $V_{cc} (\phi_0)$ is of rank $p-1$, while $V_{dd} (\phi_0, \overline{u})$ is of full rank p, hence the asymptotic variance covariance matrix $\widehat{V}_T \left(\widehat{\phi}^c_T, \overline{u} \right)$ is of rank $2p-1$. The bootstrap counterpart of $\widehat{V}_T \left(\widehat{\phi}^c_T, \overline{u} \right)$ is given by

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

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

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

where

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

and where

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

and

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

with

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

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

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

References

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

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

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

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

Benigno P. and M. Woodford (2005). Inflation Stabilization and Welfare: The Case of Distorted Steady State. *Journal of the European Economic Association*, 3(6), 1185-1236.

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

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

Bontemps, C., T. Magnac and E. Maurin (2012). Set Identified Linear Models. *Econometrica*, forthcoming.

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

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

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

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

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

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

Corts, K.S. (1999). Conduct Parameters and the Measurement of Market Power. *Journal of Econometrics*, 88, 227–250.

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

Galí, J. and M. Gertler (1999). Inflation Dynamics: A Structural Econometric Analysis. *Journal of Monetary Economics*, 44, 195-222.

Galichon, A. and M. Henry (2011). Set Identification in Models with Multiple Equilibria. *Review of Economic Studies*, 78 (4), 1264-1298.

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

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

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

Halunga A., D. Osborn and M. Sensier (2009). Changes in the Order of Integration of US and UK Inflation. *Economic Letters*, 102, 30–32.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50, 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, 279-291.

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

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

Lubik, T. A. and F. Schorfheide (2004). Testing for Indeterminacy: An Application to U.S. Monetary Policy. *The 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, 519-547.

McCallum, B. (1999). Issues in the design of monetary policy rules. *Handbook of Macroeconomics*, in: J. B. Taylor and M. Woodford (ed.), Handbook of Macroeconomics, Ed. 1, Vol. 1, Ch. 23, pp. 1483-1530.

Moon, R. and F. Schorfheide (2009). Estimation with Overidentifying Inequality Moment Conditions. *Journal of Econometrics*, 153 (2), 136-154.

Rosen, A. (2007). Identification and Estimation of Firms' Marginal Cost Functions with Incomplete Knowledge of Strategic Behavior. CeMMAP working paper CWP03/07.

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

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

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

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

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

Woodford, M. (2001). The Taylor Rule and Optimal Monetary Policy. *American Economic Review*, 91(2), 232-237.

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

Woodford, M. (2011). Optimal Monetary Stabilization Policy, in: B.M. Friedman and M. Woodford, eds., Handbook of Monetary Economics, vol. 3B.

Woodford, M. and M. P. Giannoni (2010). Optimal Target Criteria for Stabilization Policy. NBER Working Paper 15757.