*Macroeconomic Dynamics* (2023), **27**, pp. 1866–1892
doi:10.1017/S1365100522000499

ARTICLE

# Implementing the commitment solution via discretionary policy-making*

Volker Hahn

University of Konstanz
Email: volker.hahn@uni-konstanz.de

**Abstract**
This paper demonstrates that, in a large class of linear-quadratic models with rational expectations, losses due to time-inconsistency problems can be avoided, as the commitment solution can be implemented by a policy-maker who acts under discretion. We focus on two approaches. First, we show that a non-Markovian, reputational equilibrium that implements the commitment solution always exists. Second, we show how delegation to a policy-maker with an additional objective for the policy instrument can be used to implement the commitment solution via a standard discretionary Markov equilibrium. Implementation is facilitated by the fact that the commitment outcome can be attained irrespective of the weight that the policy-maker assigns to the additional target. Using the standard new Keynesian model as an example, we study the dynamics of the economy under optimal additional output targets as well as optimal interest-rate targets for central banks.

**Keywords:** Discretion; Commitment; Reputation; Optimal Delegation

## 1. Introduction

In August 2020, the Federal Reserve revised its monetary policy strategy. Since then, the Fed seeks to achieve its inflation objective of 2% on average over time.[1] Notably, the Fed intends to pursue a history-dependent policy as, after longer periods of inflation below 2%, "monetary policy will likely aim to achieve inflation moderately above 2% for some time." The revised strategy can be seen in light of the well-known theoretical finding that a central bank that could commit to a specific policy in the future would choose a policy that is history-dependent (Clarida et al. 1999). A history-dependent policy has the advantage that it may enable the central bank to influence expectations in a desirable way (Woodford, 2003).[2] However, it is not obvious how a central bank that can re-optimize its policy in every period can credibly pursue such a history-dependent policy. Why should the Fed find it optimal ex post to achieve higher inflation rates after periods of unusually low inflation? This paper demonstrates that a policy-maker that acts purely under discretion, that is, a policy-maker that cannot commit to a specific future behavior but chooses its instrument optimally every period, taking its own future behavior as given, can often implement the optimal solution under commitment. Thus, the Fed's revised strategy may well lead to improved monetary policy.

In particular, the present paper shows the implementability of the commitment solution by a discretionary policy-maker for the large class of linear models with rational expectations considered by Blanchard and Kahn (1980) and a quadratic objective function.[3,4] Hence, the social

*I would like to thank Hans Gersbach, Fabian Hoch, Stefan Niemann, Jochen Güntner, seminar participants in Linz, two anonymous referees, and an associate editor for many valuable comments and suggestions.

losses stemming from the time-inconsistency problem in models where current economic variables depend on expectations about future variables can be avoided. The class of models to which our approach can be applied includes the canonical new Keynesian model with a purely forward-looking Phillips curve (Clarida et al. 1999), models with lagged endogenous variables in the Phillips curve [see Steinsson (2003), among others], and even complex DSGE models like Christiano et al. (2005) and Smets and Wouters (2007).

We follow two different approaches to implementing the commitment solution: reputational equilibria and optimal delegation.[5] In our first approach, we focus on discretionary equilibria that are not Markov-perfect but where the optimal behavior of all agents including the policy-maker is dependent on payoff-irrelevant histories.[6] We prove that such reputational equilibria enable the exact implementation of the commitment solution by a discretionary policy-maker in a large class of linear-quadratic models with backward-looking and forward-looking variables (see Propositions 1 and 3).

As it is well known, when Markov perfection is imposed, a discretionary policy-maker typically does not find it optimal to set non-predetermined variables to the values implied by the commitment solution because, even if the private sector believed that the policy-maker would do so, the policy-maker would later find it optimal to deviate. In the discretionary equilibrium that implements the commitment solution, the policy-maker's current choice of instrument affects future non-predetermined variables but not current non-predetermined variables. As a consequence, the policy-maker can ensure in every period that the non-predetermined variables in the next period are expected to be equal to the values implied by the commitment solution. It is rational for the private sector to have such expectations because the policy-maker cannot influence the non-predetermined variables later.

It is crucial to note that the inability to influence current non-predetermined variables is not imposed by assumption. By contrast, it is a property of the discretionary equilibrium that implements the commitment solution and, in particular, a consequence of the way how the rational expectations evolve in equilibrium. In equilibrium, it is impossible to influence current non-predetermined variables because the direct effect of a change in the instrument on the current non-predetermined variables is exactly offset by an indirect effect. This indirect effect arises because changes in the instrument lead to changes in expectations about future non-predetermined variables.

In our analysis of the second approach, we prove that delegation to a policy-maker whose preferences differ from those of society can always ensure the exact implementation of the commitment solution via a standard discretionary Markov equilibrium (see Propositions 2 and 4). In particular, optimal delegation can be achieved if the policy-maker cares not only about welfare but also about a specific time-varying objective for the policy instrument, which itself depends on past choices of the instrument. This approach is related to Woodford (2003), and a more recent paper by Bilbiie (2014). Woodford (2003) shows that delegating monetary policy to a central bank that pursues an interest-rate smoothing objective can be socially desirable even in a model where interest-rate smoothing is not socially desirable *per se*. Bilbiie (2014) sharpens this result by proving that an appropriately designed loss function enables exact implementation of the commitment solution by a discretionary central bank in the standard new Keynesian model. Our approach extends these previous findings in different directions.

First and most importantly, our result about delegation is applicable to a large class of linear models with forward-looking variables, whereas Woodford (2003) as well as Bilbiie (2014) consider only the standard new Keynesian model. It may be interesting to note that, with regard to models such as Christiano et al. (2005) and Smets and Wouters (2007), Bilbiie (2014, p. 75) writes that "it is likely that in more empirically realistic models . . .exact implementation of the timeless-optimal commitment solution will be impossible. . ." The present paper proves that the commitment solution can be implemented in these models as well.

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

Second, in contrast with existing approaches like Bilbiie (2014) or Woodford (2003), our approach to optimal delegation involves that the weight that the central bank puts on the additional objective can be chosen arbitrarily and, in particular, can be very small. In practice, it may be difficult to determine the weights that the policy-maker assigns to conflicting objectives with high precision. Thus, this property of our approach would facilitate implementation. At the same time, the additional target that the policy-maker has to pursue may be comparably complex. However, all the information regarding the additional objective can be condensed into a single index. This index is straightforward to compute and could be given an interpretation as a measure of "dynamic policy performance."

How can it be ensured that the policy-maker, for example, a central bank, cares about the additional targets for the instrument or the performance index in addition to its existing objectives? One possibility are official statements that the central bank intends to take an additional target into account when conducting monetary policy. As a result, deviations from this target may involve some, possibly small, costs in the future. These costs may be psychological, as the communication of an additional target can be interpreted as implying promises about future behavior, and individuals, all else being equal, may prefer to keep their promises to some degree. Moreover, deviations from an official monetary policy strategy may lead to a loss in prestige. Gersbach and Hahn (2011) discuss anecdotal evidence for costs of deviating from statements about one's own future decisions. For example, according to Svensson (2009), members of the Swedish monetary policy committee agreed not to signal likely future interest-rate decisions, as such signals might "pre-commit" committee members. An alternative path to optimal delegation is to use incentive contracts that make the policy-maker's pay dependent on deviations of the instrument from the target.[7] Importantly, as has been mentioned before, even very small additional incentives are sufficient to guarantee that a policy-maker acting under discretion can implement the commitment solution via a Markov equilibrium.

The paper is related to several other strands of literature. Potential gains from commitment have been identified in the classic literature on the inflation bias [Kydland and Prescott (1977) and Barro and Gordon (1983)]. Kydland and Prescott (1977) and Stokey (1989) demonstrate the time inconsistency of optimal government policies in other fields like taxation or patent protection. In the standard new Keynesian model, another time-inconsistency problem arises, the so-called stabilization bias [Clarida et al. (1999) and Woodford (1999)]. Papers that aim to quantify the gains from commitment for central banks find that they are potentially sizable [Dennis and Söderström (2006) and Levine et al. (2008)]. As a consequence, it is important to answer the question how the gains from commitment can be achieved.

A recent paper by Debortoli et al. (2018) assumes that the central bank can implement the commitment solution for a given loss function that captures a specific mandate. Debortoli et al. (2014) and Debortoli and Lakdawala (2016) find empirical evidence that the Federal Reserve operates with a high degree of commitment. These papers leave open the question how central banks can implement the commitment solution. The present paper lays out such a mechanism.

A related paper by Hahn (2021) shows the existence of non-Markov-perfect discretionary equilibria and points out the potential for them to be welfare enhancing. In contrast with the present paper, Hahn (2021) focuses exclusively on the canonical new Keynesian model. Moreover, for the class of equilibria considered by Hahn (2021), a discretionary central bank can never implement the optimal commitment solution. By contrast, the present paper shows that exact implementation of the commitment solution is possible under discretionary policy-making in a large class of models. Hahn (2021) also does not consider optimal delegation to an independent policy-maker.

Reputation can also be modeled with the help of the sustainable equilibrium concept (Chari and Kehoe, 1990, 1993), which, like our approach, relies on strategies that violate the Markov property and thus depend on payoff-irrelevant histories.[8] Compared to the sustainable equilibrium concept, the discretionary equilibrium allows only for one-shot deviations by the policy-maker in each period rather than deviations which specify policies for all possible future histories.

One advantage of our approach is that it guarantees that the commitment solution can be implemented in a large class of models. By contrast, there is no comparably general result that ensures that the commitment solution (or Ramsey policy) can be supported by a sustainable equilibrium.

The game-theoretic literature has considered various variants of the Folk theorem, which implies that for infinitely repeated games all individually rational payoff combinations can be achieved for sufficiently high discount factors [see (Fudenberg and Maskin, 1986), for one particular variant of the Folk theorem].[9] In contrast with the present paper, these analyses do not allow for the possibility that the different periods are connected via forward-looking and backward-looking variables. Reputational equilibria have also been considered in macroeconomic models [see, e.g., Barro and Gordon (1983), Loisel (2008), Levine et al. (2008) and Abreu et al. (1990)].

Our approach differs from existing analyses of reputational equilibria in two respects. First, existing analyses typically rely on trigger strategies to construct reputational equilibria. While mathematically convenient, trigger strategies may often involve an implausibly harsh punishment in response to minor deviations [al Nowaihi and Levine (1994)]. By contrast, the reputational equilibria constructed in this paper have the advantage that small policy errors or small measurement errors for economic variables do not lead to large consequences in future periods. In particular, we show that these equilibria are compatible with "proportional punishment," where small deviations only lead to small and temporary increases in losses in future periods. Second, existing approaches to modeling reputational equilibria typically only guarantee the implementability of the socially optimal solution when the discount factor is sufficiently large. By contrast, our finding about the implementability of the commitment solution holds for arbitrary discount factors. This is a consequence of the fact that deviations by the policy-maker are not punished by a reversal to the standard discretionary equilibrium but by proportional punishments, which lead to outcomes slightly worse than the equilibrium outcome.[10]

It is convenient to formulate the history-dependent behaviors in our reputational equilibria with the help of additional state variables that, on the equilibrium path, are identical to the Lagrange multipliers associated with the constraints for the non-predetermined (or forward-looking) variables in the corresponding problem under commitment. Thus, our approach may be reminiscent of Marcet and Marimon (2019), who show how the solution to the planner's problem in nonlinear models with forward-looking constraints can be obtained via a recursive saddle-point functional equation that involves co-state variables that are introduced recursively as functions of the Lagrange multipliers associated with the forward-looking constraints.[11] The aim of their paper is completely different from ours. They develop an elegant, computationally advantageous method to determine the solution to social planning problems when the optimal policy is not time-consistent, whereas we demonstrate how the commitment solution can be implemented by policy-makers acting under discretion. At the danger of oversimplification, one could say that Marcet and Marimon show how time-inconsistent policies can be computed, while the present paper shows how time-inconsistent policies can be turned into time-consistent ones (e.g., by delegation to a policy-maker with an additional objective for the instrument).[12] Another key difference between our approach and theirs is that the endogenous additional state variables introduced by our approach effectively turn non-predetermined variables into predetermined ones, which will be explained in more detail in Section 2.3.

This paper is organized as follows. In Section 2, we illustrate the main mechanism behind our results with the help of the canonical new Keynesian model. In particular, Section 2.6 shows how the commitment solution can be implemented by delegation of monetary policy to a central bank that has an additional time-varying target for the output gap. Section 3 generalizes our results to a large class of linear-quadratic models with rational expectations. As an application of our approach, we show that the optimal commitment policy in the new Keynesian model can also be implemented by delegation to a discretionary policy-maker that faces an additional interest-rate objective. Section 4 concludes.

## 2. Simple new Keynesian model

### 2.1 Framework

To illustrate the main mechanisms behind our results, we now use the simple new Keynesian model as an example [Clarida et al. (1999)]. The general results for a broad class of models will be derived in Section 3.

In every period $t = 0, 1, 2, \ldots$, the private sector's behavior is summarized by:

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

$$\xi_{t+1} = \rho \xi_t + \varepsilon_{t+1}, \tag{2}$$

where $\beta \in (0, 1)$ is a common discount factor, $\kappa > 0$, $\rho \in (0, 1)$, and the $\varepsilon_t$'s are i.i.d. shocks that are drawn from a normal distribution with mean zero. $\pi_t$ is the inflation rate and $u_t$ is the output gap. Equation (1) is the new Keynesian Phillips curve and (2) describes the evolution of markup shocks. The initial value of $\xi_t$, $\xi_0$, is exogenously given. Taking (1) and (2) into account, the central bank minimizes the expected discounted sum of losses:

$$l(\pi_t, u_t) = \frac{1}{2}\pi_t^2 + \frac{a}{2}u_t^2 \tag{3}$$

with $a > 0$. In our setup, $u_t$ constitutes the central bank's instrument. $\pi_t$ is a forward-looking or non-predetermined variable. $\xi_t$ is a predetermined variable.[13] The IS curve is omitted to simplify the exposition.

### 2.2 Optimal commitment

We first construct the optimal commitment solution, which enables us to show later how a discretionary policy-maker can achieve this solution. The optimal commitment solution can be obtained by setting up the Lagrangian:

$$\mathcal{L}_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[ \frac{1}{2}\pi_t^2 + \frac{a}{2}u_t^2 + \lambda_{t+1} \left( \beta^{-1}\pi_t - \pi_{t+1} - \kappa \beta^{-1} u_t - \beta^{-1}\xi_t \right) \right], \tag{4}$$

where $\lambda_t$ $(t = 1, 2, \ldots)$ are the multipliers associated with the new Keynesian Phillips curve.

As it is well known [Clarida et al. (1999) and Woodford (1999)], the commitment solution can be characterized by:

$$u_{t+1} - u_t = -\frac{\kappa}{a}\pi_{t+1} \qquad \textit{for t=0,1,2, ...} \tag{5}$$

$$u_0 = -\frac{\kappa}{a}\pi_0 \tag{6}$$

together with (1) and (2).

In line with Backus and Driffill (1986) and Söderlind (1999), the commitment solution can be described by equations that specify the joint evolution of $\xi_t$ and $\lambda_t$ as well as by equations that state how the non-predetermined variable $\pi_t$ and the instrument $u_t$ depend on the current values of $\xi_t$ and $\lambda_t$. With the help of the optimal commitment solution in Clarida et al. (1999), it is straightforward to show that, for the simple new Keynesian model under consideration, these equations are

$$\begin{pmatrix} \xi_{t+1} \\ \lambda_{t+1} \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ -\frac{\beta\delta}{1-\delta\beta\rho} & \delta \end{pmatrix} \begin{pmatrix} \xi_t \\ \lambda_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{t+1} \\ 0 \end{pmatrix}, \tag{7}$$

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

where

$$\delta = \frac{1+\beta+\frac{\kappa^2}{a}-\sqrt{(1-\beta)^2+2(1+\beta)\frac{\kappa^2}{a}+\frac{\kappa^4}{a^2}}}{2\beta} \in (0,1), \tag{8}$$

as well as

$$\begin{pmatrix} \pi_t \\ u_t \end{pmatrix} = \begin{pmatrix} \frac{\delta}{1-\delta\beta\rho} & \frac{1-\delta}{\beta} \\ -\frac{\kappa\delta}{a(1-\delta\beta\rho)} & \frac{\kappa\delta}{a\beta} \end{pmatrix} \begin{pmatrix} \xi_t \\ \lambda_t \end{pmatrix}. \tag{9}$$

For exogenous values of $\xi_0$ and $\lambda_0 = 0$, (7) and (9) describe the entire dynamics of the system. Loosely speaking, equation (7) can be interpreted as introducing a new predetermined state variable in addition to $\xi_t$: the Lagrange multiplier on the forward-looking constraint, $\lambda_t$.

### 2.3 Discretionary policy-making

Abandoning the restriction to Markov-perfect strategies for the time being, we now construct a discretionary equilibrium that implements the commitment solution. This equilibrium involves an additional state variable. A general law of motion for the additional state variable $s_t$ is

$$s_{t+1} = \phi_\xi \xi_t + \phi_s s_t + \phi_u u_t, \tag{10}$$

with a given initial value $s_0$. It is clear that, in every period $t$, $s_t$ can be written exclusively as a function of past shocks and past values of the instrument. Thus, decisions made by the policy-maker or other agents that depend on $s_t$ are influenced by past events that are not payoff-relevant. This is a typical feature of reputational equilibria such as those considered by Chari and Kehoe (1990, 1993). In classic analyses of reputational equilibria in the inflation bias literature [Barro and Gordon (1983)], inflation expectations are specified to be functions of past deviations of inflation from some optimal level.

A discretionary equilibrium for the economy under consideration can be defined as follows:

**Definition 1.** *Consider a central bank with instrument $u_t$, loss function (3), and discount factor $\beta$, which faces the constraints (1), (2), as well as the law of motion (10) for the payoff-irrelevant state variable $s_t$, where $\phi_u$, $\phi_\xi$, $\phi_s$, and $s_0$ are given. Then*

$$\begin{pmatrix} \pi_t \\ u_t \end{pmatrix} = M \begin{pmatrix} \xi_t \\ s_t \end{pmatrix} \tag{11}$$

*with $2 \times 2$-dimensional matrix $M$ describes a discretionary equilibrium if the following properties hold:*

1. *The path of $\{\pi_t, u_t, \xi_t\}_{t=0}^{\infty}$ implied by (2), (10), and (11) satisfies (1).*
2. *In each period $t$, no profitable deviation exists for the central bank, that is, the choice of $u_t$ implied by (11) is optimal given (1), where the central bank takes the process by which rational inflation expectations $\mathbb{E}_t\pi_{t+1}$ are formed as given.*
3. *Rational expectations $\mathbb{E}_t\pi_{t+1}$ are formed in line with (2), (10), and (11).*

This definition is identical to the one used in the literature [see e.g., Backus and Driffill (1986)] with the only modification that the additional state variable $s_t$, which is not payoff-relevant, is introduced.[14]

For payoff-relevant state variables, the coefficients in the law of motion are typically exogenously given. By contrast, $s_t$ is an auxiliary variable that is used to describe how decisions and

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

expectations are affected by payoff-irrelevant past events. Note that in our paper as in other analyses of reputational mechanisms, there are multiple equilibria.[15] Thus, the $\phi$'s in the law of motion for $s_t$, (10), are coefficients that will be pinned down in order to obtain the equilibrium of interest, namely the one implementing the commitment outcome. Other choices of the coefficients would lead to other reputational equilibria or may not result in reputational equilibria at all. The flexibility with which the additional state variable can be introduced may be reminiscent of sunspots.[16] A major difference between the two concepts is the fact that sunspot variables are exogenous stochastic processes, whereas the state variable that we introduce is endogenous and can be influenced by the policy-maker, in particular.

We make use of the flexibility with which $s_t$ can be introduced and set $s_0 = 0$ as well as

$$\phi_u = -\frac{\kappa}{1-\delta}, \tag{12}$$

$$\phi_\xi = -\frac{1}{1-\delta\beta\rho}, \tag{13}$$

$$\phi_s = \frac{1}{\beta}. \tag{14}$$

It may be instructive to give an intuition for how the parameters for the law of motion for $s_t$ are chosen and why the central bank cannot profitably deviate in the discretionary equilibrium implementing the commitment solution that we are constructing.

In the Markov-perfect equilibrium of the standard new Keynesian model where the central bank minimizes social losses, there is no endogenous state variable, and inflation in all periods is only a function of the current markup shock. Hence, inflation expectations $\mathbb{E}_t\pi_{t+1}$ cannot be influenced by the central bank's choice of $u_t$, as the future markup shock $\xi_{t+1}$ is exogenous to monetary policy. Together with this observation, the Phillips curve (1) implies that an increase in $u_t$ affects current inflation only via the traditional marginal-cost channel, where an increase of $u_t$ by $\Delta u$ entails an increase in inflation by $\kappa\Delta u$.

In a discretionary equilibrium that violates the Markov property, changes in the instrument $u_t$ may lead to changes in inflation expectations because changes in $u_t$ influence $s_{t+1}$, which affects expectations about inflation in the future. These inflation expectations, in turn, influence current inflation, according to the new Keynesian Phillips curve. Thus, the central bank can affect current inflation $\pi_t$ not only via the traditional marginal-cost channel, which we have described above, but also via an expectations channel. As will be shown formally later, the particular choice of $\phi_u$ in (12) guarantees that the effects of a change in $u_t$ on inflation via the marginal-cost channel and the expectations channel exactly offset each other. As a consequence, inflation $\pi_t$ effectively becomes a predetermined variable, which cannot be influenced by the central bank's choice in period $t$.

This observation is key to understanding how the discretionary equilibrium we are constructing can implement the commitment solution. The central bank can choose its current instrument in a way such that inflation is expected to correspond to the value compatible with the commitment solution in the next period because, in the next period, it will not be possible for the central bank to influence inflation in that period. By contrast, a central bank in the discretionary equilibrium without an additional payoff-irrelevant state variable can always influence the current value of inflation and thereby implement a policy such that inflation differs from the value implied by the commitment solution.[17]

Finally, we discuss how the remaining parameters are set. Obviously, $s_0 = 0$ ensures $s_0 = \lambda_0$.[18] The parameter choices for $\phi_\xi$ and $\phi_s$ were made in a way such that, for the value of $\phi_u$ given in (12) and $s_0 = 0$, the evolution of the variables $\pi_t$, $u_t$, and $s_t$ equals the evolution of the corresponding variables $\pi_t$, $u_t$, and $\lambda_t$ in the commitment solution. Why the specific value of $\phi_s$ is required to implement the commitment solution is discussed in more detail in Section 2.4. In Appendix A, we

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

comment on a potential concern that, together with (10), $\phi_s > 1$ could involve explosive dynamics for $s_t$.

The following proposition shows that the commitment solution of the standard new Keynesian model can be implemented via a discretionary equilibrium:

**Proposition 1.** *Consider a central bank with instrument $u_t$, loss function (3), and discount factor $\beta$, which faces the constraints (1), (2), as well as the following law of motion for the payoff-irrelevant state variable $s_t$:*[19]

$$s_{t+1} = \frac{1}{\beta}s_t - \frac{1}{1-\delta\beta\rho}\xi_t - \frac{\kappa}{1-\delta}u_t, \qquad \text{with } s_0 = 0. \tag{15}$$

*Then*

$$\begin{pmatrix} \pi_t \\ u_t \end{pmatrix} = \begin{pmatrix} \frac{\delta}{1-\delta\beta\rho} & \frac{1-\delta}{\beta} \\ -\frac{\kappa\delta}{a(1-\delta\beta\rho)} & \frac{\kappa\delta}{a\beta} \end{pmatrix} \begin{pmatrix} \xi_t \\ s_t \end{pmatrix} \tag{16}$$

*describes a discretionary equilibrium. This equilibrium implements the commitment solution.*

While it is easy to confirm the statement of the proposition for specific parameter combinations with standard software routines that can be used to compute discretionary equilibria (Söderlind, 1999), it is instructive to also consider a formal proof.

*Proof.* As a preliminary step, we note that the dynamics of $\pi_t$ and $u_t$ equal those in the commitment solution. This can be seen by comparing (15) and (16) to (7) and (9). In particular, plugging the expression for $u_t$ given in (16) into (15) yields an expression for $s_{t+1}$ as a function of $\xi_t$ and $s_t$. With the help of (8) as well as $s_0 = \lambda_0 = 0$, it is straightforward to show that this equation for $s_{t+1}$ yields dynamics for the additional state variable $s_t$ that are identical to the dynamics for the Lagrange multiplier in the commitment solution, which are described by (7).

To prove that (16) describes a discretionary equilibrium, we show that, first, together with (2) and (15), it is compatible with the Phillips curve (1) and that, second, the central bank's choice of $u_t$ in every period $t$ is optimal subject to the new Keynesian Phillips curve in this period $t$, given the rational inflation expectations $\mathbb{E}_t\pi_{t+1}$ that are formed in line with (2), (15), and (16).

The candidate equilibrium is compatible with the new Keynesian Phillips curve, because (i) the economy evolves as in the commitment solution and (ii) the commitment solution satisfies the new Keynesian Phillips curve. To prove that the central bank behaves optimally requires a few steps. First, we calculate inflation expectations, as they enter the new Keynesian Phillips curve, which represents a constraint for the central bank. Using (2), (15), and (16), we obtain the following expression:

$$\mathbb{E}_t\pi_{t+1} = \frac{1}{1-\delta\beta\rho}\left(\delta\rho - \frac{1-\delta}{\beta}\right)\xi_t + \frac{1-\delta}{\beta^2}s_t - \frac{\kappa}{\beta}u_t. \tag{17}$$

As we have argued before, it is crucial that inflation expectations are a function of $u_t$. This property stems from the fact that the central bank's choice of $u_t$ affects $s_{t+1}$, which in turn affects inflation in period $t+1$.

As a next step, we plug the inflation expectations (17) into the new Keynesian Phillips curve (1) and obtain

$$\pi_t = \frac{\delta}{1-\delta\beta\rho}\xi_t + \frac{1-\delta}{\beta}s_t. \tag{18}$$

It is noteworthy that the central bank's choice of $u_t$ has no influence on inflation in the same period. As has been explained before, this is a property of the particular equilibrium under consideration and is a consequence of the specific choice of $\phi_u$. Section 2.5 studies a case where,

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

due to small deviations from rational expectations, the influence of $u_t$ on $\pi_t$ is not zero but small. In this case, outcomes close to the commitment outcome can still be implemented.

Consider optimal central bank behavior in a particular period $t$. For this purpose, we set up the corresponding Bellman equation:

$$W(\xi_t, s_t) = \min_{u_t} \left\{ \frac{1}{2}\pi_t^2 + \frac{a}{2}u_t^2 + \beta \mathbb{E}_t W(\xi_{t+1}, s_{t+1}) \right\}$$
$$\textit{subject to (2), (15), (18).} \tag{19}$$

We obtain the following first-order condition as well as a condition that results from the envelope theorem:

$$0 = au_t - \frac{\kappa}{1-\delta}\beta \mathbb{E}_t W_s(\xi_{t+1}, s_{t+1}), \tag{20}$$

$$W_s(\xi_t, s_t) = \frac{1-\delta}{\beta}\pi_t + \mathbb{E}_t W_s(\xi_{t+1}, s_{t+1}), \tag{21}$$

where the subscript $s$ stands for the respective partial derivative. Equations (20) and (21) can be combined to

$$\mathbb{E}_t u_{t+1} - u_t = -\frac{\kappa}{a}\mathbb{E}_t \pi_{t+1} \qquad \textit{for } t\textit{=0,1,2,...} \tag{22}$$

This condition characterizes optimal central bank behavior in the candidate discretionary equilibrium. As the paths of $\pi_t$ and $u_t$ satisfy the condition for optimal central bank behavior under commitment, (5), they also satisfy (22). *Q.E.D.*

The additional state variable $s_t$ can be interpreted as the burden of past promises. To make this more precise, note that (16) implies that the central bank's instrument $u_t$ is an increasing function of $s_t$ in equilibrium. Thus, a positive value of $s_t$ represents a promise to choose a rather expansionary monetary policy, while a negative value represents a promise to choose a comparably contractionary policy. As (15) entails that $s_{t+1}$ is a decreasing function of $u_t$, current expansionary monetary policy tends to entail the promise of contractionary monetary policy in the future. As a consequence, the central bank behaves as in the commitment solution. In the commitment solution, a negative markup shock calls for expansionary policy in the same period but contractionary policy in the next period.

### *2.4 Dynamic response to policy errors*

We continue our discussion of the standard new Keynesian model by illustrating the dynamics of the economy if the central bank deviates from its optimal choice of $u_t$ in period 0 but not in future periods. We set the parameters as follows: $\kappa = 0.3$, $\beta = 0.99$, $\rho = 0.9$, $a = 0.05$, and $\xi_0 = 0$, in addition to $s_0 = 0$. Moreover, assume that the central bank chooses $u_0 = 1$ rather than $u_0 = 0$, which would be optimal.

Figure 1 illustrates the dynamics of the system in the absence of markup shocks $\xi_t$, that is, for $\varepsilon_t = 0$ in $t = 1, 2, 3, \ldots$. The deviation of the central bank cannot influence inflation (displayed as a solid line) in period 0, as the discretionary equilibrium under consideration involves that inflation is effectively predetermined.[20] According to the dash-dotted line, which displays $s_t$, the suboptimally high level of $u_0$ drives $s_1$ below zero (compare (10) and (12)). Because $s_{t+1}$ depends positively on $s_t$ on the equilibrium path ($s_{t+1} = \delta s_t$), $s_t$ stays negative in the consecutive periods $t = 2, 3, \ldots$ As inflation and output are increasing functions of $s_t$ (compare (9) for $s_t = \lambda_t$), both output and inflation are negative from period 1 onward.

The figure shows that deviations from the equilibrium behavior in a particular period lead to changes in the additional state variable $s$ in future periods and thereby to increases in losses in

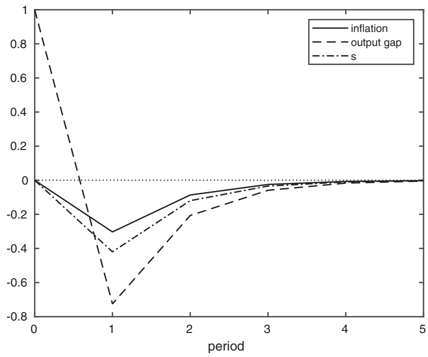


**Figure 1.** Impulse responses of inflation (solid line), the output gap (dashed line), and the additional state variable $s_t$ (dash-dotted line) in response to a one-time deviation of the output gap. Markup shocks have been set to zero in all periods.

these periods. It may be interesting to contrast the consequences of a deviation in this paper and the respective consequences in papers that use trigger strategies to overcome time-inconsistency problems [see, e.g., Loisel (2008) and Levine et al. (2008)].[21] Compared to equilibria with trigger strategies, a small deviation of the central bank only has small, transitory consequences for the economy. As can be seen easily, the response of the economy after a deviation is always proportional to the magnitude of the deviation.[22]

It may also be instructive to consider the situation where $s_t$ differs from zero in one period $t$, there are no markup shocks, that is $\xi_t = 0$, and the central bank ignores the burden of past "promises," $s_t$, by setting $u_t = 0$ in this period. In this case, $s_{t+1} = \beta^{-1} s_t$ would hold. Thus, the burden of past "promises" would grow by a factor $\beta^{-1}$. This mild increase in $s_t$ deters deviations in the first place. It might also be noteworthy that the factor $\beta^{-1}$ decreases with $\beta$. This is plausible, as larger discount factors require smaller punishments in the future.

### *2.5 Deviations from rational expectations*

In a sense, the discretionary equilibrium implementing the commitment solution represents a knife-edge case in which the effects of changes in the central bank's instrument on inflation via the marginal-cost channel and the expectations channel cancel each other exactly. This may raise questions about the robustness of our approach.

To study this issue, we consider small deviations from rational expectations. In particular, we assume that inflation expectations are not given by (17) but by

$$\mathcal{E}_t \pi_{t+1} = \frac{1}{1-\delta\beta\rho}\left(\delta\rho - \frac{1-\delta}{\beta}\right)\xi_t + \frac{1-\delta}{\beta^2} s_t - \left(\frac{\kappa}{\beta} + \epsilon\right) u_t, \tag{23}$$

where we use $\mathcal{E}_t$ to denote expectations that may not be fully rational, and $\epsilon$ is a parameter that describes the magnitude of deviations from rationality. Obviously, for $\epsilon = 0$, (23) is identical to (17). Under the assumption that private sector expectations are governed by (23), the Phillips curve (1) implies that inflation is given by:

$$\pi_t = \frac{\delta}{1-\delta\beta\rho}\xi_t + \frac{1-\delta}{\beta} s_t - \beta\epsilon u_t, \tag{24}$$

which is a generalization to (18). Importantly, for $\epsilon \neq 0$, it is no longer true that the central bank cannot influence current inflation by changing its instrument $u_t$. However, for small values of $\epsilon$, the effect of changes in the instrument $u_t$ on $\pi_t$ is small.

As a next step, we analyze the dynamic consequences of this departure from rational expectations. To complete the description of the variant of the model, we assume that the central bank

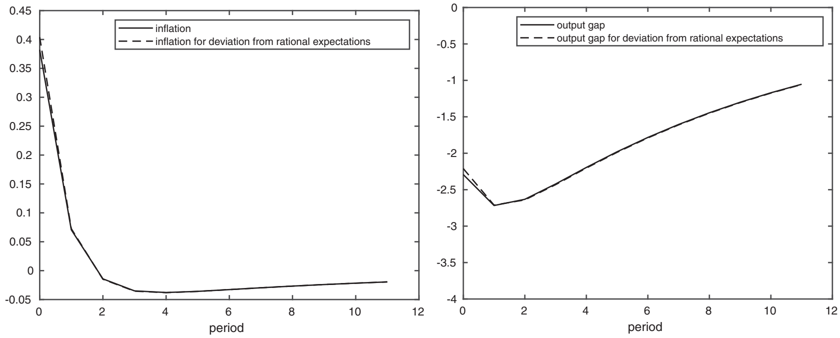


**Figure 2.** Impulse responses of inflation (left panel) and the output gap (right) panel for a markup shock with $\varepsilon_0 = 1$. The solid lines represent the impulse responses when inflation expectations are perfectly rational ($\epsilon = 0$). The dashed lines show the impulse responses when inflation expectations not perfectly rational, that is, they are given by (23) with $\epsilon = 0.01$.

knows that expectations are formed in line with (23) and that $s_t$ still evolves according to (15). Figure 2 shows the impulse responses of inflation (left panel) and the output gap (right panel) to a markup shock. The solid lines stand for $\epsilon = 0$, that is, rational expectations, and the dashed lines for small deviations from rational expectations where we have set $\epsilon = 0.01$.

The figure shows that small deviations from rational expectations only have a small effect on the dynamics of the economy. The reason for this finding is that, while changes in the instrument affect current inflation for small but nonzero values of $\epsilon$ and thus inflation cannot be interpreted as an effectively predetermined variable, non-negligible changes in inflation require very large changes in the instrument, which tend to be costly for the central bank. Hence, small deviations from rational expectations allow the implementation of an equilibrium with social losses that are close to the losses under commitment.

While we have only studied a very specific deviation from the equilibrium implementing the commitment outcome, the conclusions from this analysis hold more generally: Small deviations from the equilibrium implementing the commitment solution lead to equilibria that involve social losses very close to the ones implied by the commitment solution. Consider, for example, a discretionary equilibrium where the central bank responds to the state variable $s_t$ but where $s_t$ follows a law of motion (10) with coefficients $\phi_u$, $\phi_\xi$, and $\phi_s$ slightly different from the ones specified in (12)–(14). Note that inflation and output in equilibrium can always be expressed as functions of current and lagged markup shocks and that the coefficients in front of these realizations of markup shocks are continuous functions of the $\phi$'s. As a consequence, small changes in the $\phi$'s only have small effects on the dynamic response of the economy to shocks and therefore only small effects on social losses.

### *2.6 Implementation through delegation*

As a next step, we show that the commitment solution can also be implemented by a standard discretionary Markov equilibrium when monetary policy is delegated to a central bank with a loss function that is different from the social loss function.[23] Our approach is related to the result by Woodford (2003) that the delegation of monetary policy to a central bank with an interest-rate smoothing objective can be socially desirable even if interest-rate smoothing is not socially desirable *per se*. It extends this finding in two dimensions. First, our approach implements the commitment solution and thus allows the central bank to reap all of the welfare gains that are theoretically possible. By contrast, interest-rate smoothing typically allows the central bank to achieve only a part of the possible welfare gains. Second, our approach is more general as it can be applied to the broad class of linear models considered in Section 3 and is not restricted to the standard new Keynesian model. Optimal interest-rate targets that facilitate the implementation of the commitment solution in the new Keynesian model are derived in Section 3.4.

The following proposition is shown in Appendix C:

**Proposition 2.** *Suppose that monetary policy-making is delegated to a central bank whose instantaneous losses are*

$$l_{CB}(\pi_t, u_t, s_t) = \frac{1}{2}\pi_t^2 + \frac{a}{2}u_t^2 + \frac{b}{2}(u_t - u_t^*)^2, \tag{25}$$

*where*

$$u_t^* = -\frac{\kappa\delta}{a(1-\delta\beta\rho)}\xi_t + \frac{\kappa\delta}{a\beta}s_t \tag{26}$$

*and b is an arbitrary parameter with $b > 0$. $s_t$ is given by (15). The central bank discounts its future losses with the same discount factor $\beta$ as society. Then the commitment solution associated with the social loss function (3) can be implemented by a discretionary Markov equilibrium.*

Compared to the social loss function, the central bank loss function (25) includes the additional term $b(u_t - u_t^*)^2/2$, which penalizes deviations of the central bank's instrument from the time-varying target $u_t^*$. $u_t^*$ mimics the evolution of $u_t$ in the discretionary equilibrium implementing the commitment solution (compare (16)). Thus, on the equilibrium path, $u_t^*$ equals the central bank's choice of instrument for the commitment solution. In this sense, the central bank's modified loss function punishes deviations from the commitment solution.[24]

While it is perhaps unsurprising that large costs of deviations from the commitment solution ensure that the central bank implements this solution, the proposition implies that even small incentives, that is, arbitrarily small values of $b$, are sufficient. Loosely speaking, the reason why small values of $b$ are sufficient to ensure that the central bank behaves as under optimal commitment is that the term $b(u_t - u_t^*)^2/2$ does not alter the incentives of the central bank compared to the equilibrium discussed in Proposition 1, as the central bank already chooses $u_t = u_t^*$ in this equilibrium. The purpose of the additional term in the central bank's loss function is to add an additional state variable to the central bank's optimization problem.

It may also be noteworthy that it is possible to completely eliminate $s_t$ from the central bank's optimization problem, as it is straightforward to formulate $u_t^*$ as a function of the current value of $\xi_t$ as well as lagged values of $u_t^*$ and $u_t$. More specifically, we can write

$$u_t^* = \frac{1}{\beta}u_{t-1}^* - \frac{\kappa\delta}{a(1-\delta\beta\rho)}\xi_t - \frac{\kappa^2\delta}{a\beta(1-\delta)}u_{t-1}. \tag{27}$$

One might think that the coefficient $\frac{1}{\beta}$ implies explosive dynamics for the output-gap target $u_t^*$. However, it is easy to see that this is not the case. For this purpose, we rewrite (27) as:

$$u_t^* = \frac{1}{\beta}\left[1 - \frac{\kappa^2\delta}{a(1-\delta)}\right]u_{t-1}^* - \frac{\kappa\delta}{a(1-\delta\beta\rho)}\xi_t - \frac{\kappa^2\delta}{a\beta(1-\delta)}(u_{t-1} - u_{t-1}^*) \tag{28}$$

and note that, in equilibrium, $u_t = u_t^*$ holds for all periods and that the coefficient in front of $u_{t-1}^*$ is smaller than 1 (and positive) for all admissible parameter values.

For the parameter values, we have used before ($\kappa = 0.3$, $\beta = 0.99$, $\rho = 0.9$, and $a = 0.05$), we obtain

$$u_t^* = 0.2851u_{t-1}^* - 2.2929\xi_t - 0.7250(u_{t-1} - u_{t-1}^*). \tag{29}$$

The property that the target value $u_t^*$ depends on lagged values of the central bank's instrument is crucial for the implementability of the commitment solution. In particular, it is responsible for the key feature of our approach that non-predetermined economic variables are effectively turned into predetermined ones.

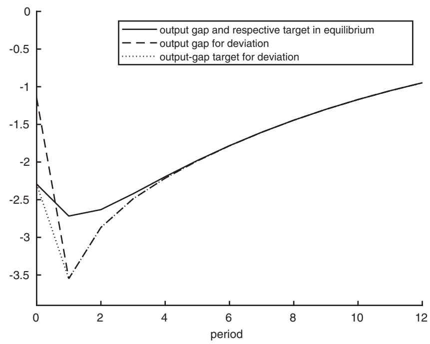


**Figure 3.** Impulse responses of the equilibrium output gap and the corresponding target in response to a markup shock with $\varepsilon_0 = 1$ (solid line); the output gap in response to the markup shock with a deviation in period 0 (dashed line), and the output-gap target for the deviation (dotted line).

Figure 3 illustrates the dynamics of the output-gap target $u_t^*$ for a markup shock with $\varepsilon_0 = 1$ and $\varepsilon_t = 0$ for $t = 1, 2, 3, \ldots$.[25] The evolution of the output gap under commitment is identical to the evolution of the output-gap target as well as the output gap chosen by a discretionary central bank under optimal delegation (solid line). In addition, we show the dynamics of the output-gap target (dotted line) if the central bank deviates in period 0 and chooses only 50% of the optimal value for the output gap but implements the optimal policy thereafter. As a consequence of this deviation, the output-gap target drops in period 1 but quickly returns toward the equilibrium path. The output gap (dashed line) differs from its target in period 0 by construction but corresponds to the target in all remaining periods. We have confirmed that, in line with our theoretical results, the value of $b$, which describes the intensity of the additional incentives to align output with the additional target $u_t^*$, does not affect the equilibrium outcomes under delegation.

At this point, it is instructive to discuss the relationship to the interest-rate smoothing objective considered by Woodford (2003) in more detail. Woodford augments the central bank's loss function by terms that are proportional to $(i_t)^2$ and $(i_t - i_{t-1})^2$, where the central bank's instrument $i_t$ is the deviation of the nominal interest rate from its long-run level. These additional terms in the loss function can be interpreted as introducing an additional objective for the central bank's instrument, where the target value $i_t^*$ is proportional to $i_{t-1}$. This is obviously related to our approach, where the target $u_t^*$ depends on lagged values of the instrument $u_t$ as well. By adding an IS curve to the new Keynesian model studied in this section, one can easily specify an optimal interest-rate target $i_t^*$ that allows for the commitment solution to be implemented. This approach is pursued in Section 3.4.

How could one ensure that the central bank cares about deviations from an additional output target, that is, that $b(u_t - u_t^*)^2/2$ enters its loss function? First, following Walsh (1995), incentive contracts could be used and central banker's pay could depend on deviations of their policies from the targets $u_t^*$. Importantly, the monetary incentives can be arbitrarily small. Second, as has been discussed in the Introduction, it is plausible that a public announcement that the central bank intends to set its instrument in line with $u_t^*$ will entail small direct costs for the central bank if it deviates from this announcement at a later stage.[26]

## 3. General results

### 3.1 Setup

In the following, we generalize the results for the simple new Keynesian model to a broad class of linear-quadratic models with rational expectations. There are $n_x$ predetermined variables, contained in the column vector $x_t$, and $n_y$ non-predetermined variables, contained in the

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

vector $y_t$. Let $z_t$ be the $(n_x + n_y)$-dimensional vector that contains all predetermined and all non-predetermined variables, where the predetermined variables come first. There is also a $k$-dimensional vector of instruments $u_t$.

Social losses are described by:

$$L = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left( z_t' Q z_t + 2 z_t' U u_t + u_t' R u_t \right), \tag{30}$$

where $\beta \in (0, 1)$. $Q$ is an $(n_x + n_y) \times (n_x + n_y)$-dimensional matrix, $U$ is an $(n_x + n_y) \times k$-dimensional matrix, and $R$ is a $(k \times k)$-dimensional matrix. Without loss of generality, we assume $Q$ and $R$ to be symmetric. For the time being, we assume that the policy-maker's loss function is identical to the social loss function. In Section 3.4, we will examine the case where the policy-maker's losses include an additional term that involves target values for the instruments $u_t$.

The predetermined and non-predetermined variables evolve according to

$$x_{t+1} = A_{xx} x_t + A_{xy} y_t + B_x u_t + \varepsilon_{x,t+1}, \tag{31}$$

$$\mathbb{E}_t y_{t+1} = A_{yx} x_t + A_{yy} y_t + B_y u_t, \tag{32}$$

where $A_{xx}$, $A_{xy}$, $A_{yx}$, $A_{yy}$, $B_x$, and $B_y$ are given matrices whose coefficients have been obtained from log-linearized equations describing the private sector equilibrium, for example. The $n_x$ components of $\varepsilon_{x,t+1}$ describe the innovations to the predetermined variables $x_{t+1}$. They have zero mean and covariance matrix $\Sigma$.

As is well known, models with lagged variables and expectations more than one period ahead can also be cast in the form considered here. The same is also be true for some models with lagged expectations of present and future variables [see Blanchard and Kahn (1980)].

### 3.2 Commitment

The commitment solution can be obtained by setting up the Lagrangian [Backus and Driffill (1986) and Söderlind (1999)]:

$$\mathcal{L}_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[ z_t' Q z_t + 2 z_t' U u_t + u_t' R u_t + 2 \rho_{t+1}' \left( A z_t + B u_t - z_{t+1} \right) \right], \tag{33}$$

where

$$A = \begin{pmatrix} A_{xx} & A_{xy} \\ A_{yx} & A_{yy} \end{pmatrix}, \; B = \begin{pmatrix} B_x \\ B_y \end{pmatrix}, \tag{34}$$

and $\rho_t$ is an $(n_x + n_y)$-dimensional vector of Lagrange multipliers ($t = 1, 2, 3, \ldots$). The first-order conditions with respect to $z_t$ and $u_t$ are

$$\beta A' \mathbb{E}_t \rho_{t+1} = -\beta Q z_t - \beta U u_t + \rho_t, \tag{35}$$

$$-B' \mathbb{E}_t \rho_{t+1} = U' z_t + R u_t. \tag{36}$$

We assume that the commitment solution exists and involves unique paths of $z_t$ and $u_t$. As shown by Backus and Driffill (1986), the commitment solution can then be described by an equation that specifies the evolution of the predetermined variables $x_t$ and the multipliers $\rho_{y,t}$ associated with the non-predetermined variables:

$$\begin{pmatrix} x_{t+1} \\ \rho_{y,t+1} \end{pmatrix} = H \begin{pmatrix} x_t \\ \rho_{y,t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t+1} \\ 0 \end{pmatrix}, \tag{37}$$

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

where $H$ is an $(n_x + n_y) \times (n_x + n_y)$-dimensional matrix. The initial values of the predetermined variables $x_t$ are exogenously given. The initial values of the elements in $\rho_{y,t}$ are zero, because the non-predetermined variables can be chosen freely in $t = 0$. The non-predetermined variables $y_t$, the Lagrange multipliers $\rho_{x,t}$ associated with the predetermined variables, and the policy-makers' instruments $u_t$ can be expressed as functions of the current values of $x_t$ and $\rho_{y,t}$. In particular, $y_t$ can be written as:

$$y_t = C \begin{pmatrix} x_t \\ \rho_{y,t} \end{pmatrix}. \tag{38}$$

$C$ is an $n_y \times (n_x + n_y)$-dimensional matrix that can be decomposed as $C = \left(C_x, C_{\rho_y}\right)$, where $C_x$ has dimensions $n_y \times n_x$ and $C_{\rho_y}$ has size $n_y \times n_y$.

### 3.3 Discretionary policy-making

We are now in a position to formulate a generalization of Proposition 1 to a large class of models:[27]

**Proposition 3.** *Consider the unique commitment solution of the model characterized by (31), (32) and loss function (30). Let $C = (C_x, C_{\rho_y})$ be the matrix that describes how $y_t$ depends on $x_t$ as well as $\rho_{y,t}$ (see (38)). Then $C_{\rho_y}$ is invertible. For the discretionary policy-maker, introduce an $n_y$-dimensional vector of additional state variables, $s_t$. In the initial period, set $s_0 = 0_{n_y}$.[28] For $t = 0, 1, 2, \ldots$, assume that $s_t$ evolves according to*

$$s_{t+1} = A_{sx}x_t + A_{ss}s_t + B_s u_t, \tag{39}$$

*where the matrices $B_s$, $A_{sx}$, and $A_{ss}$ are given by:*

$$B_s = C_{\rho_y}^{-1} \left(B_y - C_x B_x\right), \tag{40}$$

$$A_{sx} = C_{\rho_y}^{-1} \left[\left(A_{yy} - C_x A_{xy}\right) C_x - C_x A_{xx} + A_{yx}\right], \tag{41}$$

$$A_{ss} = C_{\rho_y}^{-1} \left(A_{yy} - C_x A_{xy}\right) C_{\rho_y}. \tag{42}$$

*Then a discretionary equilibrium for (31), (32), (39), and loss function (30) exists that implements the commitment solution.*

The proof is given in Appendix D.

The main idea for the proof of the general result is similar to the one for the new Keynesian model. First, the fact that the commitment solution can be described by the joint dynamics of $x_t$ and $\rho_{y,t}$ suggests that $n_y$ additional state variables should be added to the discretionary policy-maker's problem in addition to the payoff-relevant state variables $x_t$. These additional state variables are the components of $s_t$. Second, to ensure that the policy-maker does not renege on past promises, the dynamics of $s_t$ are specified in a way such that, in every period $t$, the policy-maker cannot influence current non-predetermined variables (as opposed to future non-predetermined variables). This is achieved by (40), which is a straightforward generalization to (12). Third, (41) and (42) specify the matrices $A_{ss}$ and $A_{sx}$ in a manner such that the dynamics of $s_t$ on the equilibrium path of the discretionary equilibrium equal those of $\rho_{y,t}$ in the commitment solution. Finally, one needs to show that the discretionary policy-maker does not wish to deviate. At this point of the proof, it is helpful that the value function of the discretionary policy-maker can be constructed from the commitment solution under the assumption that both approaches lead to identical dynamics.[29]

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

Proposition 3 shows that policy-makers that lack a means of commitment can nevertheless implement socially desirable policies that could not be supported in a Markov-perfect equilibrium. Thus, our approach clarifies that it is possible to implement history-dependent strategies like average inflation targeting, which was discussed at the beginning of the Introduction.

It may also be interesting to relate Proposition 3 to a theorem in Backus and Driffill (1986) that examines whether the commitment solution can be supported by trigger strategies.[30] In particular, they consider the case where deviations from the commitment solution are punished by a grim trigger, that is, a permanent switch to the standard Markov-perfect discretionary equilibrium. According to their theorem, the commitment solution can be sustained (i) if the support of shocks is bounded and (ii) if the discount factor is sufficiently large. By contrast, the discretionary equilibria constructed in this paper always allow for the implementation of the commitment solution, irrespective of the magnitude of the discount factor and for an arbitrary support of shocks.[31] This result stems from the fact that, loosely speaking, the punishment for deviations from the commitment solution is not restricted to a switch to the standard discretionary equilibrium.

### 3.4 Implementation through delegation

Having shown that the commitment solution can be implemented by a discretionary non-Markovian equilibrium, we proceed by showing that the finding of Section 2.6 about optimal delegation and the implementation of the commitment solution by a discretionary Markov equilibrium can be generalized to the broad class of models specified in Section 3.1. For this purpose, we consider a policy-maker whose loss function is identical to the social loss function, except for an additional term that penalizes deviations of the instrument from an additional target for the instrument.

In Appendix E, we show:

**Proposition 4.** *Suppose that policy-making is delegated to a policy-maker whose instantaneous loss is*

$$L_{del} = z_t' Q z_t + 2 z_t' U u_t + u_t' R u_t + (u_t - u_t^*)' \mathcal{B} (u_t - u_t^*), \tag{43}$$

*where* $u_t^* = -F \begin{pmatrix} x_t \\ s_t \end{pmatrix}$ *and* $s_t$ *is given by (39). F is the matrix that describes how the instrument* $u_t$ *is set as a function of* $x_t$ *and* $\rho_{y,t}$ *in the commitment solution, that is, it satisfies* $u_t = -F \begin{pmatrix} x_t \\ \rho_{y,t} \end{pmatrix}$.[32] $\mathcal{B}$ *is an arbitrary* $k \times k$*-dimensional positive definite matrix. The policy-maker discounts its future losses with the same discount factor* $\beta$ *as society. Then the commitment solution associated with the social loss function can be implemented by a discretionary Markov equilibrium.*

For the standard new Keynesian model, we have shown that the additional incentives for the policy-maker that are necessary to ensure the implementation of the commitment solution can be very small. Importantly, this finding extends to the general setup considered here. Formally, this is an implication of the fact that $\mathcal{B}$ is an arbitrary positive definite matrix. As a consequence, it is possible to choose a matrix $\mathcal{B}$ with very small positive eigenvalues.[33]

In the following, we explain in more detail how delegation of monetary policy to a central bank with an optimally designed interest-rate target can implement the commitment solution in the new Keynesian model. For this purpose, we augment the standard new Keynesian model discussed in Section 2 by an IS curve:

$$y_t = -\sigma^{-1} \left( i_t - \mathbb{E}_t[\pi_{t+1}] \right) + \mathbb{E}_t[y_{t+1}], \tag{44}$$

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

where $i_t$, the central bank's instrument, is the nominal interest rate net of the natural real rate of interest, $\sigma$ is a positive parameter, and we use $y_t$ rather than $u_t$ for the output gap, as the output gap does not correspond to the instrument in this variant of the model. Henceforth, we will call $i_t$ the nominal interest rate for simplicity.

We select $\sigma = 1$ and set the remaining parameter values to the levels used before. As there is only one instrument, $\mathcal{B}$ corresponds to an arbitrary positive number, whose value does not affect our results. It is straightforward to describe the dynamics of the two auxiliary variables $s_t$ with the help of (39), where, to compute $A_{ss}$, $A_{sx}$, and $B_s$, we have to use matrices $A$ and $B$ that take the IS curve into account. The additional interest-rate target is then given by $i_t^* = -F({x_t}', {s_t}')'$, where $F$ is defined as in Proposition 4. The loss function under optimal delegation incorporates the additional interest-rate target in the following way:

$$l_{CB,i}(\pi_t, u_t, i_t^*) = \frac{1}{2}\pi_t^2 + \frac{a}{2}y_t^2 + \frac{b}{2}(i_t - i_t^*)^2, \tag{45}$$

where $b$ is an arbitrary positive number.

It is easy to show that the central bank's optimization problem can be formulated without the auxiliary variables $s_t$. More specifically, the interest-rate target $i_t^*$ can be written as $i_t^* = I_t^* + \Delta_t$ with the properties that $I_t^* = i_t$ and $\Delta_t = 0$ on the equilibrium path and where, for our parameter values and an arbitrary positive value of $b$, the dynamics of $I_t^*$ and $\Delta_t$ are governed by:

$$I_{t+1}^* = 0.353645\, \xi_{t+1} + 0.490248\, \xi_t + 0.285081\, I_t^* - 1.443798\, (i_t - I_t^*), \tag{46}$$

$$\Delta_{t+1} = 0.584252\, (\Delta_t + I_t^* - i_t), \tag{47}$$

with initial values $I_0^* = 0.353645\, \xi_0$ and $\Delta_0 = 0$.

It is noteworthy that $\Delta_{t+1} = 0$ whenever the central bank set its instrument in line with its overall target $i_t^* = I_t^* + \Delta_t$ in the previous period. Thus, optimal delegation involves that the central bank faces an interest-rate target $i_t^*$ that can be decomposed into a long-term component $I_t^*$, which evolves recursively and is a decreasing function of the lagged instrument, as well as a short-term component $\Delta_t$, which affects the central bank's overall target only in periods following deviations of the central bank's instrument $i_t$ from the overall target $i_t^* = I_t^* + \Delta_t$. In particular, whenever the central bank's choice of instrument $i_t$ is too high compared to the target $i_t^*$, then $\Delta_{t+1} < 0$, that is, the target for period $t + 1$ is pushed below $I_{t+1}^*$.

The dynamics of the economy after a markup shock are illustrated in Figure 4.[34] The left panel shows inflation (black solid line) and output (gray solid line) under optimal delegation, which correspond to the respective values under commitment. The corresponding dashed lines show the dynamics of output and inflation after a deviation in period 0 where the central bank selects only 50% of the optimal value in that period and implements the optimal policy in subsequent periods. As has been mentioned before, our construction turns non-predetermined variables (inflation and output in this case), into predetermined ones. This explains why output and inflation correspond to their optimal levels in the period of the deviation. From period 1 onward, they differ from the values under optimal commitment. However, these differences become very small after a few periods.

The right panel of the figure illustrates the dynamics of the interest rate in more detail. The solid line stands for the interest rate under optimal delegation and optimal commitment. The interest-rate path for the one-time deviation described above corresponds to the dashed line. $\Delta_t$ differs from zero only in period 1 (dotted line). In particular, it is negative due to the fact that the interest-rate set by the central bank in period 0 was too high under the deviation. The overall interest-rate target $i_t^*$ is the sum of $\Delta_t$ (dotted line) and $I_t^*$ (the dash-dotted line). One can see that, following the deviation in period 0, the interest rate and the overall interest-rate target $i_t^*$ tend to move back to their equilibrium values over time.

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

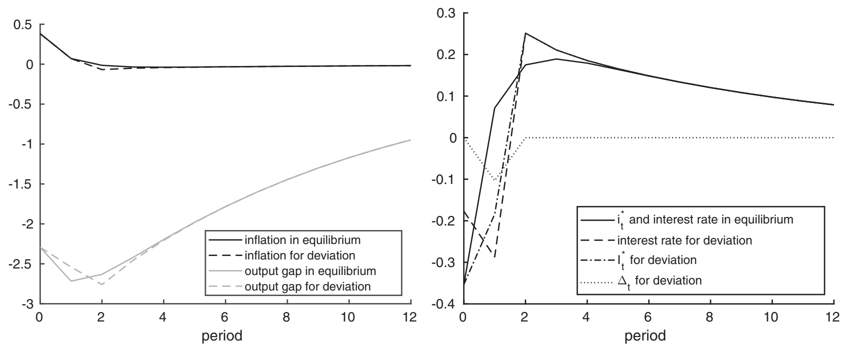


**Figure 4.** Dynamic responses to a markup shock under optimal delegation to a central bank with an interest-rate target. Left panel: inflation (solid black line) and output (solid gray line) under optimal delegation as well as optimal commitment; inflation (dashed black line) and output (dashed gray line) under delegation when the central bank chooses a level of $i_0$ that corresponds to only 50% of the optimal level. Right panel: the interest-rate target $i_t^* = I_t^* + \Delta_t$ and the interest rate $i_t$ in equilibrium (solid line), the nominal interest rate under the deviation in the first period (dashed line), the component $I_t^*$ of the interest-rate target (dash-dotted line) and $\Delta_t$, the second component of the interest-rate target (dotted line).

### *3.5 Optimal delegation in practice*

In general, one may be concerned that the additional targets for the central bank may be complex while other approaches to optimal delegation involve objectives that may be easier to interpret. For example, in Woodford (2003), the central bank is assigned four objectives: (1) an inflation stabilization objective, (2) an output stabilization objective, (3) an interest-rate stabilization objective, and (4) an objective to stabilize changes in interest rates (see equation 4.1 in his paper). While these objectives have straightforward interpretations, optimal delegation involves that one has to set the relative weights that the central bank attaches to these four targets to specific values.

In practice, hiring a central banker who attaches the right weights to these different objectives may be challenging. As a consequence, one may design incentive contracts that specify how central bankers' pay depends on the degree to which the different targets are met. However, even such an approach may be problematic as central bankers are plausible to be intrinsically motivated as well and, as an additional complication, the intrinsic motivation may interact with the extrinsic incentives in non-obvious ways. To sum up, it appears challenging to ensure that, in practice, a central bank acts in line with a loss function with specific, optimally selected weights on different targets.

By contrast, the approach proposed in this paper does not involve this problem as the weight that the central bank attaches to the additional target compared to its other objectives is irrelevant. The specification of the additional target may be complex, but measuring this target and the central bank's performance with respect to meeting this target is straightforward.

Even for a comparably rich model like the one by Smets and Wouters (2007), the term $(u_t - u_t^*)'\mathcal{B}(u_t - u_t^*)$ in Proposition 4 can be condensed into a one-dimensional index that can be calculated in a straightforward manner. This index could be given an interpretation as a measure of "dynamic central bank performance" and could be published on a regular basis. This performance index would take a value of zero when the central bank acts as in the equilibrium implementing the commitment solution and would take higher values otherwise.

The use of quality indices, which condense complex information in order to make it easier to digest, is not uncommon in other areas. For example, in the EU, energy labels on an A-G scale are used to provide summary information on the energy efficiency of appliances such as fridges.[35] Moreover, many central banks publish inflation forecasts, which also condense various pieces of information. These arguments suggest that the publication of an additional performance index, which can be obtained by calculating $(u_t - u_t^*)'\mathcal{B}(u_t - u_t^*)$, may be a way of implementing our approach in practice.

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

## 4. Conclusions

This paper has shown that the losses associated with time-inconsistency problems can often be avoided. In particular, we have demonstrated the existence of discretionary equilibria implementing the commitment solution, where the policy-maker responds to lagged variables that do not enter the social loss function directly. Thus, our approach explains how history-dependent strategies such as average inflation targeting can be implemented by a central bank that cannot commit to a particular strategy.

An important task for the policy-maker is to help coordinate economic agents on the discretionary equilibrium that facilitates the commitment outcome.[36] A promising way in this regard is the communication of a strategy that specifies how the policy-maker intends to behave in the future and which economic variables are relevant for the conduct of the policy. The public is plausible to believe that the policy-maker will behave in this way because, in this case, such behavior will be optimal for the policy-maker in every future period. The implementation of the commitment outcome does not require any direct costs of deviating from the communicated strategy.

To further ensure that the optimal commitment solution can be implemented, incentive contracts can be used, which stipulate that the decision-maker's pay depends on deviations of the instruments from specific targets. Importantly, the weights that the policy-maker assigns to these additional targets compared to its other objects are irrelevant for the implementability of the commitment solution. Even small direct losses for the policy-maker caused by deviations from the target values ensure that all of the theoretically possible welfare gains can be achieved by a discretionary policy-maker in a standard Markov equilibrium. This property represents an advantage over other approaches to delegation that typically require that the weights on additional objectives have to be set optimally, which is arguably difficult to achieve in practice.

We have argued that, even for complex models, all relevant information about the additional targets can be condensed into a single index, which is straightforward to calculate and can be interpreted as an index of "dynamic policy performance." The publication of this index on a regular basis can ensure the implementation of the commitment outcome by a central bank even if the extent to which the central bank cares about this index compared to its other objectives may be small.

Finally, one might wonder whether our approach might be applicable also to nonlinear models, for example, the new Keynesian model with a lower bound on nominal interest rates. While it appears plausible that there are nonlinear models in which the commitment solution can be implemented by a discretionary policy-maker as well, it would be difficult to obtain a general theoretical result. The proofs of our propositions for general linear-quadratic rational expectations models rely on the explicit solution to the commitment problem [Backus and Driffill (1986)]. An analytical solution for the commitment problem typically does not exist for nonlinear models. However, numerical analyses of the implementability of the commitment solution by discretionary policy-makers in nonlinear models would be a promising avenue for future research.[37]

## Notes

**1** See the "Statement on Longer-Run Goals and Monetary Policy Strategy," amended effective August 27, 2020, available from https://www.federalreserve.gov.

**2** The commitment solution of the standard new Keynesian model (Clarida et al. 1999) has the property that periods of below-target inflation that are caused by a negative markup shock should be followed by periods with inflation rates that are higher than the ones chosen by a discretionary policy-maker. The optimal commitment solution in a liquidity trap is analyzed by Eggertsson and Woodford (2003). It also implies that expectations about future inflation should be raised when the economy experiences a low rate of inflation.

**3** The discretionary solution and the commitment solution have been analyzed in Oudiz and Sachs (1985), Backus and Driffill (1986), Currie and Levine (1993), and Söderlind (1999). Dennis (2007) presents solution algorithms.

**4** Waki et al. (2018) consider the optimal degree of discretion in a framework where society imposes dynamic constraints on the policy-maker.
**5** The delegation of monetary policy to a central bank whose preferences differ from the preferences of society has been studied by Rogoff (1985), Woodford (2003), Walsh (2003), and Vestin (2006), among others.
**6** Responding to payoff-irrelevant variables can be optimal in linear-quadratic models of discretionary policy-making if current economic variables depend on expectations about future economic variables and if the payoff-irrelevant variables are defined recursively [Hahn (2021)].
**7** Incentive contracts for central bankers are introduced in Walsh (1995). While he derives the optimal contract for a neoclassical one-period model, our analysis generates optimal contracts for a general class of dynamic linear-quadratic models.
**8** In the reputational equilibria considered in Section IV of Barro and Gordon (1983), the behaviors of private agents and the central bank depend on past deviations of inflation from the socially optimal rate.
**9** Abreu et al. (1986, 1990) introduce a method to study reputation in infinitely repeated games that focuses on sets of continuation values as opposed to strategies.
**10** In the linear-quadratic models considered in the present paper, it is always possible to achieve outcomes that are slightly worse than the equilibrium outcome, because the set of admissible payoffs is not bounded (and thus not compact). We consider this to be a plausible assumption in many macroeconomic applications because, for example, no maximum value for the inflation rate exists. Appendix B considers the case of bounded shocks and shows that, in this case, endogenous economic variables remain bounded as well, even if the policy-maker deviates to the standard discretionary equilibrium in an arbitrary period.
**11** Cole and Kubler (2012) extend Marcet and Marimon's method to weakly concave Pareto sets. Gerali and Lippi (2003) present a toolkit for solving the commitment solution in linear forward-looking economies.
**12** Obviously, for every dynamic optimization problem involving forward-looking constraints, it is possible to specify some optimization problem such that the discretionary solution to this optimization problem would be the commitment solution to the original problem. For example, one could follow Marcet and Marimon's approach to obtain a saddle-point functional equation with additional state variables. In principle, a discretionary policy-maker who does not minimize the loss function subject to the given constraints but solves the saddle-point functional equation in every period will automatically implement the commitment solution. However, this approach does not provide an alternative solution to the main problem considered in this paper, which is how a discretionary policy-maker who minimizes a given loss function subject to given constraints can implement the commitment solution.
**13** The definitions of predetermined variables and non-predetermined variables are in line with Backus and Driffill (1986) and Söderlind (1999).
**14** On p. 14, Backus and Driffill (1986) write that a discretionary policy-maker "cannot take as given either the non-predetermined variables . . .or the artificial variables $p_{y,t}$ which enabled it to sustain the commitments needed for the optimal program." One might have the impression that this quote suggests that the equilibrium concept used in this paper differs from the one used in Backus and Driffill (1986). This, however, is not the case (except for the obvious difference that we explicitly allow for choices to depend on payoff-irrelevant variables). According to our approach, the central bank does not take the Lagrange multiplier (which correspond to $p_{y,t}$ in the above quote) from the commitment solution as given but a predetermined, payoff-irrelevant variable $s_t$ with a specific law of motion.
**15** See Walsh (2017, Sec. 6.3.1) for an analysis of multiple reputational equilibria in a classic model of the inflation bias.
**16** See Cass and Shell (1983) for a seminal contribution and Farmer and Zabczyk (2022) for a recent application.
**17** In Section 2.5, we show that, even if the effects of changes in the instrument on inflation via the marginal-cost channel and the expectations channel do not cancel each other perfectly, it is still possible to achieve a discretionary equilibrium that results in an outcome close to the one under commitment.
**18** One might wonder whether a discretionary policy-maker could also implement the policy that is optimal from a timeless perspective (Woodford, 1999). The timeless-perspective solution is identical to the commitment solution but imposes (5) rather than (6) also in the initial period. A discretionary policy-maker could easily implement this solution for the appropriate initial value $s_0$.
**19** Please recall our previous discussion of the size of the coefficient in front of $s_t$.
**20** In principle, this property suggests a simple empirical test of whether the central bank is able to implement the commitment solution. If the central bank is successful in this regard, monetary policy shocks will affect inflation only with a lag.
**21** Abreu et al. (1990) show that, under certain conditions, efficient sequential equilibria of infinitely repeated games have the so-called bang-bang property, which involves that only extreme payoff combinations occur for all possible histories.
**22** In a related analysis, Appendix B examines the case where the exogenous shocks are drawn from a distribution with finite support.
**23** In a seminal contribution, Rogoff (1985) advocates the delegation of monetary policy to a conservative central banker, that is, an individual whose preferences differ from those of society.

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

**24** It may be noteworthy that the standard discretionary Markov equilibrium for a central bank minimizing the social loss function, which involves $u_t = -\{\kappa/[\kappa^2 + a(1-\beta\rho)]\}\xi_t$, does not correspond to a discretionary equilibrium in the case where monetary policy is delegated to a central bank with loss function (25).
**25** Matlab code for the simulations can be downloaded from https://www.wiwi.uni-konstanz.de/hahn/. The code utilizes the routines provided by Söderlind (1999) to demonstrate that the equilibrium we have constructed does indeed correspond to a discretionary equilibrium.
**26** Gersbach and Hahn (2011) examine a model where the central bank's prestige depends on the precision of its forecasts to some extent.
**27** The definition of a discretionary equilibrium is standard and a straightforward extension to Definition 1. Loosely speaking, the policy-maker chooses its set of instruments optimally in every period, taking its own behavior in future periods and the process by which the public forms its rational expectations as given.
**28** We use the definition that $0_{n_y}$ is an $n_y$-dimensional column vector of zeros.
**29** The equilibrium for the discretionary policy-maker's problem specified in Proposition 3 is typically not unique (see also our related discussion in Section 2.3). In all cases where the economy admits a Markov-perfect discretionary equilibrium, this equilibrium will also correspond to an equilibrium for the economy where the additional state variables $s_t$ have been added via (39).
**30** Currie and Levine (1993) find an analogous result for continuous-time models. A related finding for the new Keynesian model is due to Kurozumi (2008).
**31** Appendix B studies the case of bounded shocks.
**32** Backus and Driffill (1986) show that $u_t$ can be written as $u_t = -F({x_t}', {\rho_{y,t}}')'$ under optimal commitment. The equation is a straightforward generalization to *(9)*.
**33** One may wonder whether the equilibrium implementing the commitment solution is the only discretionary Markov equilibrium when the policy-maker's loss function is (43). As shown in Blake and Kirsanova (2012), linear-quadratic models with endogenous predetermined state variables may have multiple discretionary Markov equilibria. Together with the observation that $u_t^*$ (or $s_t$) correspond to such endogenous predetermined state variables, this implies that the uniqueness of the Markov equilibrium under optimal delegation cannot be guaranteed. However, Blake and Kirsanova's results typically imply local uniqueness of the Markov equilibrium implementing the commitment outcome.
**34** Our experiments suggest that, for arbitrary starting values, the routines presented in Söderlind (1999) converge to the discretionary solution under optimal delegation with an output target. For an interest-rate target, these routines typically do not converge if arbitrary starting values are used. However, by using the correct starting values, which can be constructed from the commitment solution, one can utilize these routines to confirm the existence of the discretionary equilibrium under optimal delegation with an interest-rate target and the finding that it implements the commitment solution. The fact that the algorithm that computes discretionary equilibria is not guaranteed to converge is noted also by Söderlind. Matlab code for the simulations can be downloaded from https://www.wiwi.uni-konstanz.de/hahn/.
**35** See https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1575536811417&uri=CELEX:32019R2014.
**36** Armenter (2016) shows how a specific interest-rate policy allows the central bank to ensure coordination on the socially best equilibrium in a standard neoclassical model of the inflation bias.
**37** It might also be interesting to examine whether solutions under intermediate degrees of commitment [Schaumburg and Tambalotti (2007), Debortoli and Nunes (2010), Debortoli et al. (2014), and Guo and Krause (2014)], where the policy-maker may renege on past promises with a constant probability every period, can be implemented by a fully discretionary policy-maker as well.
**38** It might be worth noting that, in the absence of shocks, the new Keynesian Phillips curve can be written as $\pi_{t+1} = \beta^{-1}\pi_t - \beta^{-1}\kappa u_t$. Thus, the coefficient in front of $\pi_t$ is larger than one just like the coefficient in front of $s_t$ in (10).
**39** The standard Markovian discretionary solution involves $u_t = -\{\kappa/[\kappa^2 + a(1-\beta\rho)]\}\xi_t$.
**40** There appears to be a small mistake in Backus and Driffill (1986), as they omit the discount factor in the relationship between $\rho_t$ and $Vz_t$. That (54) is correct can be confirmed by comparing (36) and the first-order condition for optimization problem (53), which is $U'z_t + Ru_t + \beta B'V\mathbb{E}_t z_{t+1} = 0$.

## References


Abreu, D., D. Pearce and E. Stacchetti. (1986) Optimal cartel equilibria with imperfect monitoring. *Journal of Economic Theory* 39(1), 251–269.

Abreu, D., D. Pearce and E. Stacchetti. (1990) Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* 58(5), 1041–1063.

al Nowaihi, A.and P. L. Levine. (1994) Can reputation resolve the monetary policy credibility problem? *Journal of Monetary Economics* 33(2), 355–380.

Armenter, R. (2016) Sustainable monetary policy and inflation expectations. *The BE Journal of Macroeconomics* 16(2), 301–334.

Backus, D. and J. Driffill. (1986) *The Consistency of Optimal Policy in Stochastic Rational Expectations Models*, CEPR Discussion Paper No. 124,

Barro, R. and D. B. Gordon. (1983) A positive theory of monetary policy in a natural rate model. *Journal of Political Economy* 91(4), 589–610.
Bilbiie, F. O. (2014) Delegating optimal monetary policy inertia. *Journal of Economic Dynamics and Control* 48(9), 63–78.
Blake, A. P. and T. Kirsanova. (2012) Discretionary policy and multiple equilibria in LQ RE models. *Review of Economic Studies* 79, 1309–1339.
Blanchard, O. J. and C. M. Kahn. (1980) The solution of linear difference models under rational expectations. *Econometrica* 48(5), 1305–1311.
Cass, D. and K. Shell. (1983) Do sunspots matter? *Journal of Political Economy* 91(2), 193–227.
Chari, V. V. and P. J. Kehoe. (1990) Sustainable plans. *Journal of Political Economy* 98(4), 783–802.
Chari, V. V. and P. J. Kehoe. (1993) Sustainable plans and mutual default. *Review of Economic Studies* 60(1), 175–195.
Christiano, L. J., M. Eichenbaum and C. L. Evans. (2005) Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113(1), 1–45.
Clarida, R., J. Galí and M. Gertler. (1999) The science of monetary policy: A new Keynesian perspective. *Journal of Economic Literature* 37(4), 1661–1707.
Cole, H. and F. Kubler. (2012) Recursive contracts, lotteries and weakly concave Pareto sets. *Review of Economic Dynamics* 15(4), 479–500.
Currie, D. and P. Levine. (1993) *Rules, Reputation and Macroeconomic Policy Coordination*. Cambridge: Cambridge University Press.
Debortoli, D. and A. Lakdawala. (2016) How credible is the Federal Reserve? A structural estimation of policy re-optimizations. *American Economic Journal: Macroeconomics* 8(3), 42–76.
Debortoli, D. and R. Nunes. (2010) Fiscal policy under loose commitment. *Journal of Economic Theory* 145(3), 1005–1032.
Debortoli, D., J. Maih and R. Nunes. (2014) Loose commitment in medium-scale macroeconomic models: Theory and applications. *Macroeconomic Dynamics* 18(1), 175–198.
Debortoli, D., J. Kim, J. Lindé and R. Nunes. (2018) Designing a simple loss function for central banks: Does a dual mandate make sense? *Economic Journal* 129(621), 2010–2038.
Dennis, R. (2007) Optimal policy in rational expectations models: New solution algorithms. *Macroeconomic Dynamics* 11(1), 31–55.
Dennis, R. and U. Söderström. (2006) How important is precommitment for monetary policy? *Journal of Money, Credit, and Banking* 38(4), 847–872.
Eggertsson, G. B. and M. Woodford. (2003) The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity* 139(1), 2003–2211.
Farmer, R. E. A. and P. Zabczyk. (2022) A sunspot-based theory of unconventional monetary policy. *Macroeconomic Dynamics* 26(1), 186–217.
Fudenberg, D. and E. Maskin. (1986) The Folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54(3), 533–554.
Gerali, A. and F. Lippi. (2003) Optimal Control and Filtering in Linear Forward-looking Economies: A Toolkit, *CEPR Discussion Paper No. 3706*.
Gersbach, H. and V. Hahn. (2011) Monetary policy inclinations. *Journal of Money, Credit, and Banking* 43(8), 1707–1717.
Guo, J.-T. and A. Krause. (2014) Optimal dynamic nonlinear income taxation under loose commitment. *Macroeconomic Dynamics* 18(6), 1403–1427.
Hahn, V. (2021) Discretionary policy and multiple equilibria in a new Keynesian model. *Oxford Economic Papers* 73(1), 423–445.
Kurozumi, T. (2008) Optimal sustainable monetary policy. *Journal of Monetary Economics* 55(7), 1277–1289.
Kydland, F. E. and E. C. Prescott. (1977) Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85(3), 473–492.
Levine, P., P. McAdam and J. Pearlman. (2008) Quantifying and sustaining welfare gains from monetary commitment. *Journal of Monetary Economics* 55(7), 1253–1276.
Loisel, O. (2008) Central bank reputation in a forward-looking model. *Journal of Economic Dynamics and Control* 32(11), 3718–3742.
Marcet, A. and R. Marimon. (2019) Recursive contracts. *Econometrica* 87(5), 1589–1631.
Oudiz, G. and J. D. Sachs. (1985) International policy coordination in dynamic macroeconomic models. In: Oudiz, G. and J. D. Sachs.(eds.), *International Economic Policy Coordination,* NBER Books, pp. 274–330. Cambridge: Cambridge University Press.
Rogoff, K. (1985) The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics* 100(4), 1169–1189.
Schaumburg, E. and A. Tambalotti. (2007) An investigation of the gains from commitment in monetary policy. *Journal of Monetary Economics* 54(2), 302–324.
Smets, F. and R. Wouters. (2007) Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review* 97(3), 586–606.

Söderlind, P. (1999) Solution and estimation of RE macromodels with optimal policy. *European Economic Review* 43(4-6), 813–823.
Steinsson, J. (2003) Optimal monetary policy in an economy with inflation persistence. *Journal of Monetary Economics* 50(7), 1425–1456.
Stokey, N. L. (1989) Reputation and time consistency. *American Economic Review (Papers and Proceedings)* 79, 134–139.
Svensson, L. E. O. (2009) Transparency under flexible inflation targeting: Experiences and challenges. *Sveriges Riksbank Economic Review* 5, 2009–2044.
Vestin, D. (2006) Price-level versus inflation targeting. *Journal of Monetary Economics* 53(7), 1361–1376.
Waki, Y., R. Dennis and I. Fujiwara. (2018) The optimal degree of monetary discretion in a new Keynesian model with private information. *Theoretical Economics* 13(3), 1319–1367.
Walsh, C. E. (2003) Speed limit policies: The output gap and optimal monetary policy. *American Economic Review* 93(1), 265–278.
Walsh, C. E. (1995) Optimal contracts for central bankers. *American Economic Review* 85, 150–167.
Walsh, C. E. (2017) *Monetary Theory and Policy*. 4th edition, Cambridge, MA: MIT Press.
Woodford, M. (1999) Commentary: How should monetary policy be conducted in an era of price stability?, *New Challenges for Monetary Policy*, pp. 277–316. Kansas City: Federal Reserve Bank of Kansas City.
Woodford, M. (2003) Optimal interest-rate smoothing. *Review of Economic Studies* 70(4), 861–886.

## Appendix A. Potential concern that $\phi_s > 1$

This appendix discusses a potential concern that, together with (10), $\phi_s > 1$ might involve explosive dynamics for $s_t$ if, for example, the central bank set $u_t = 0$ in every period.

First, we would like to stress that, while there are typically economic conditions, for example, transversality conditions from the underlying nonlinear model, that rule out explosions in real variables, there is no such condition that rules out explosions in the additional state variable $s_t$. Second, it can be easily confirmed that, in the equilibrium implementing the commitment solution, which is described in the following analysis, the central bank sets its policy in a way such that explosive dynamics of $s_t$ as well as $u_t$ and $\pi_t$ do not occur.[38]

## Appendix B. Bounded shocks

As has been highlighted in Section 2.4, one advantage of our approach is that the punishment after a deviation is always in proportion to the deviation itself. To see this from a slightly different angle, assume that the shocks $\varepsilon_t$ are drawn from a distribution with bounded support $[-\overline{\varepsilon}, +\overline{\varepsilon}]$. It is then easy to see that the state variables $\xi_t$ and $s_t$ would remain bounded in equilibrium. In particular, $\xi_t$ would lie in the interval $[-\overline{\varepsilon}/(1-\rho), +\overline{\varepsilon}/(1-\rho)]$ in all periods, provided that this is true for the initial value $\xi_0$. As a consequence, the policy-maker's choices of instrument $u_t$ would also be bounded in all periods. As is straightforward to show, this would also be true for all future periods if the policy-maker deviated to the standard (Markovian) discretionary policy in a particular period.[39] Moreover, inflation would remain bounded as well. The boundedness of $\pi_t$ and $u_t$ is an immediate consequence of the observation that the state variable $s_t$ remains in a bounded set in all periods following a deviation to the choice that would be optimal in the standard Markovian discretionary equilibrium. In this sense, our approach always relies only on moderate punishments. *Q.E.D.*

## Appendix C. Proof of Proposition 2

We focus on a particular period $t$ and assume that, in all other periods $t' \neq t$, the central bank chooses its instrument in line with the commitment solution, that is, it chooses $u_{t'} = u^*_{t'}$, where $u^*_{t'}$ is given by (26). Then we show that a central bank acting under discretion does not deviate from the commitment solution in period $t$. As a consequence, we can conclude that the policy of a central bank that minimizes (25) under discretion is compatible with the commitment solution in every period.

The central bank's decision problem in period $t$ can be formulated as:

$$\widetilde{W}(\xi_t, s_t) = \min_{u_t} \left\{ \frac{1}{2}\pi_t^2 + \frac{a}{2}u_t^2 + \frac{b}{2}(u_t - u_t^*)^2 + \beta \mathbb{E}_t \widetilde{W}(\xi_{t+1}, s_{t+1}) \right\}$$

$$\textit{subject to (2), (15), (18), (26), } \xi_t \textit{ as well as } s_t \textit{ given,} \tag{48}$$

where we have used the fact that the new Keynesian Phillips curve (1) can be written as (18), as the central bank is expected to behave as in the commitment solution in period $t+1$.

We obtain the following first-order condition as well as a condition that results from the envelope theorem:

$$0 = au_t + b(u_t - u_t^*) - \frac{\kappa}{1-\delta}\beta \mathbb{E}_t \widetilde{W}_s(\xi_{t+1}, s_{t+1}), \tag{49}$$

$$\widetilde{W}_s(\xi_t, s_t) = \frac{1-\delta}{\beta}\pi_t - b\frac{\kappa\delta}{a\beta}(u_t - u_t^*) + \mathbb{E}_t \widetilde{W}_s(\xi_{t+1}, s_{t+1}), \tag{50}$$

where the subscript $s$ stands for the respective partial derivative. Equations (49) and (50) can be combined to

$$\mathbb{E}_t u_{t+1} - u_t - \frac{b}{a}(u_t - u_t^*) = -\frac{\kappa}{a}\mathbb{E}_t \pi_{t+1} - \frac{b}{a}\left(1 - \frac{\kappa^2}{a}\frac{\delta}{1-\delta}\right)\mathbb{E}_t[u_{t+1} - u_{t+1}^*] \qquad \textit{for t=0,1,2,} \ldots \tag{51}$$

This condition characterizes optimal central bank behavior. The finding that $u_t = u_t^*$ is an optimal policy for $t = 0, 1, 2, \ldots$ follows from the fact that, for $u_t = u_t^*$ and $\mathbb{E}_t u_{t+1} = \mathbb{E}_t u_{t+1}^*$, (51) collapses to (22), which holds for a path that implements the commitment solution (see (5)).

Thus we have shown the claim of the proposition. *Q.E.D.*

## Appendix D. Proof of Proposition 3

It will be useful to note that the commitment solution can also be obtained by an alternative approach outlined in Backus and Driffill (1986). They formulate (31) and (32) jointly as

$$z_{t+1} = Az_t + Bu_t + \varepsilon_{t+1}, \tag{52}$$

where the first $n_x$ elements of $\varepsilon_{t+1}$ are the exogenous disturbances $\varepsilon_{x,t+1}$ and the remaining components are given by the endogenous expectational errors $\varepsilon_{y,t+1}$, which have to satisfy the requirement that $\mathbb{E}_t \varepsilon_{y,t+1} = 0_{n_y}$ for $t = 0, 1, 2, \ldots$.

In a first step of the alternative approach, the policy-maker takes $z_t$ as given when it makes its choice regarding $u_t$ for $t = 0, 1, 2, \ldots$. In a second step, it chooses the initial value of the non-predetermined variable, $y_0$, as well as the expectational errors $\varepsilon_{y,t}$ for $t = 1, 2, 3, \ldots$.

Thus, in the first step, the policy-maker selects $u_t$ to minimize

$$\min_{u_t} \left\{ z_t' Q z_t + 2z_t' U u_t + u_t' R u_t + \beta \mathbb{E}_t \left[ z_{t+1}' V z_{t+1} + k \right] \right\},$$

$$\textit{subject to (52), } z_t \textit{ given,} \tag{53}$$

where $z_{t+1}' V z_{t+1} + k$ is the cost-to-go at $t+1$ with a $(n_x + n_y) \times (n_x + n_y)$-dimensional symmetric matrix $V$ and a constant $k$. Importantly, $V$ is related to the multipliers from the Lagrangian approach via[40]

$$\rho_t = \beta V z_t. \tag{54}$$

We construct matrices $V_{xx}$, $V_{xy}$, $V_{yx}$, and $V_{yy}$ by partitioning $V$ conformably with $x_t$ and $y_t$.

https://doi.org/10.1017/S1365100522000499 Published online by Cambridge University Press

In the second step, the policy-maker selects $y_0$ as well as the endogenous forecast errors $\varepsilon_{y,t+1}$. The latter choice is immaterial for our purposes and therefore omitted. The former choice is obtained as a result of minimizing $z_0' V z_0 = x_0' V_{xx} x_0 + 2x_0' V_{xy} y_0 + y_0' V_{yy} y_0$ with respect to $y_0$ for given $x_0$. The corresponding first-order condition is

$$V_{yx} x_0 + V_{yy} y_0 = 0. \tag{55}$$

Since we have assumed that the commitment solution involves a unique path of all economic variables, the optimal choice of $y_0$ is unique. Thus, we can conclude that $V_{yy}$ is invertible.

Comparing (38) and $\rho_{y,t} = \beta V_{yx} x_t + \beta V_{yy} y_t$ (see (54)) leads to the implication that $C_{\rho_y}$ is invertible as well and that

$$V_{yx} = -\beta^{-1} C_{\rho_y}^{-1} C_x, \tag{56}$$

$$V_{yy} = \beta^{-1} C_{\rho_y}^{-1}. \tag{57}$$

It will be useful to introduce matrix $T$ as:

$$T = \begin{pmatrix} I_{n_x} & 0_{n_x \times n_y} \\ \beta V_{yx} & \beta V_{yy} \end{pmatrix}, \tag{58}$$

where $I_{n_x}$ is the $n_x \times n_x$-dimensional identity matrix and $0_{n_x \times n_y}$ is an $n_x \times n_y$ matrix of zeros. Matrix $T$ allows us to transform $z_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$ into $\begin{pmatrix} x_t \\ \rho_{y,t} \end{pmatrix}$:

$$\begin{pmatrix} x_t \\ \rho_{y,t} \end{pmatrix} = T \begin{pmatrix} x_t \\ y_t \end{pmatrix} \tag{59}$$

We note that $T$ is invertible with

$$T^{-1} = \begin{pmatrix} I_{n_x} & 0_{n_x \times n_y} \\ -V_{yy}^{-1} V_{yx} & \beta^{-1} V_{yy}^{-1} \end{pmatrix}. \tag{60}$$

After these preliminary steps, we now begin to formulate the optimization problem under discretion, assuming that all variables $z_t$ and $u_t$ evolve as in the commitment solution and introducing a vector of additional state variables $s_t$, which is always identical to $\rho_{y,t}$ on the equilibrium path. We postulate that, in period $t+1$, $y_{t+1}$, will depend on $x_{t+1}$ and $s_{t+1}$ in the same way that $y_{t+1}$ depends on $x_{t+1}$ and $\rho_{y,t+1}$ in the commitment solution (compare (38)). In this case, (32) can be formulated as:

$$\mathbb{E}_t C \begin{pmatrix} x_{t+1} \\ s_{t+1} \end{pmatrix} = A_{yx} x_t + A_{yy} y_t + B_y u_t. \tag{61}$$

Combining (31), (39), and (61) yields

$$\left(A_{yy} - C_x A_{xy}\right) y_t = \left(C_x A_{xx} + C_{\rho_y} A_{sx} - A_{yx}\right) x_t + C_{\rho_y} A_{ss} s_t + \left(C_x B_x + C_{\rho_y} B_s - B_y\right) u_t. \tag{62}$$

We choose $B_s$ such that $C_x B_x + C_{\rho_y} B_s - B_y = 0$, which is equivalent to (40). This entails that the policy-maker cannot affect $y_t$ by changing $u_t$:

$$\left(A_{yy} - C_x A_{xy}\right) y_t = \left(C_x A_{xx} + C_{\rho_y} A_{sx} - A_{yx}\right) x_t + C_{\rho_y} A_{ss} s_t \tag{63}$$

Comparing with (38) yields

$$\left(A_{yy}-C_xA_{xy}\right)C_x=C_xA_{xx}+C_{\rho_y}A_{sx}-A_{yx}, \tag{64}$$

$$\left(A_{yy}-C_xA_{xy}\right)C_{\rho_y}=C_{\rho_y}A_{ss}. \tag{65}$$

These equations are equivalent to (41) and (42). They pin down $A_{sx}$ and $A_{ss}$.

We introduce a symmetric $(n_x+n_y)\times(n_x+n_y)$-dimensional matrix $W$ such that, conditional on optimal behavior by the policy-maker in every period, the present value of discounted losses is $(x_t', s_t')W(x_t{}', s_t{}')'$, up to a constant term. The discretionary policy-maker's optimization problem can then be stated as:

$$\min_{u_t}\left\{2\left(x_t'\ y_t'\right)Uu_t+u_t'Ru_t+\beta\mathbb{E}_t\left[(x_{t+1}', s_{t+1}')W\begin{pmatrix}x_{t+1}\\ s_{t+1}\end{pmatrix}\right]\right\}$$

$$\textit{subject to (31), (39), } x_t,\ y_t, \textit{ and } s_t \textit{ given.} \tag{66}$$

It is noteworthy that the term $z_t'Qz_t$ from the loss function can be ignored in the minimization problem, because $x_t$ is predetermined in $t$ and $y_t$, due to our choice of $B_s$, is effectively predetermined as well.

As a next step, we note that the matrix $W$ is related to $V$ via the relation

$$W=\left(T^{-1}\right)'VT^{-1} \tag{67}$$

because the candidate discretionary equilibrium we are constructing involves the same cost-to-go as the commitment solution.

With the help of (67), we can state the first-order condition for optimization problem (66) as:

$$U'z_t+Ru_t+\beta\left(B_x'\ B_s'\right)\left(T^{-1}\right)'VT^{-1}\mathbb{E}_t\begin{pmatrix}x_{t+1}\\ s_{t+1}\end{pmatrix}=0. \tag{68}$$

Because of

$$T^{-1}\begin{pmatrix}x_{t+1}\\ s_{t+1}\end{pmatrix}=\begin{pmatrix}x_{t+1}\\ y_{t+1}\end{pmatrix}=z_{t+1}, \tag{69}$$

$$\beta V\mathbb{E}_t z_{t+1}=\mathbb{E}_t\rho_{t+1}, \tag{70}$$

and

$$\left(B_x'\ B_s'\right)\left(T^{-1}\right)'=\left(B_x'\ B_y'\right)=B', \tag{71}$$

the condition for optimal behavior by the discretionary policy-maker, (68), is equivalent to (36), which holds for the commitment solution. Thus, we have constructed a discretionary equilibrium that implements the commitment solution. *Q.E.D.*

## Appendix E. Proof of Proposition 4

As in Appendix C we focus on a particular period $t$ and assume that the policy-maker chooses a policy compatible with the commitment policy in all other periods $t'\neq t$, that is, $u_{t'}=u_{t'}^*=-F\begin{pmatrix}x_{t'}\\ s_{t'}\end{pmatrix}$. Then we show that the policy-maker finds it optimal to choose a policy compatible with commitment in period $t$. This allows us to conclude that the claim of the proposition holds.

If the optimal commitment solution is implemented in periods $t' \neq t$ and therefore $u_{t'} = u^*_{t'}$, then the cost-to-go in period $t+1$ will be identical to the one for the discretionary policy-making problem considered in Appendix D. Moreover, we would like to remind the reader that the coefficients in the law of motion (39) have been chosen in a way such that $y_t$ is effectively exogenously given in this case.

Then, as a straightforward extension to (66), the policy-maker's problem in period $t$ with the modified loss function (43) can be stated as:

$$\min_{u_t} \left\{ 2({x_t}', {y_t}')Uu_t + u_t'Ru_t + (u_t - u_t^*)'\mathcal{B}(u_t - u_t^*) + \beta \mathbb{E}_t \left[ (x_{t+1}', {s_{t+1}}')W \begin{pmatrix} x_{t+1} \\ s_{t+1} \end{pmatrix} \right] \right\}$$

$$\textit{subject to (31), (39), } u_t^* = -F \begin{pmatrix} x_t \\ s_t \end{pmatrix}, x_t, y_t, \textit{ and } s_t \textit{ given,} \tag{72}$$

where $W$, which describes the cost-to-go, is identical to the respective expression in Appendix D.

In Appendix D, we have already shown that $u_t = u_t^*$ solves this optimization problem if the additional term $(u_t - u_t^*)'\mathcal{B}(u_t - u_t^*)$ is absent. As the additional term $(u_t - u_t^*)'\mathcal{B}(u_t - u_t^*)$ is obviously minimal for $u_t = u_t^*$, we can conclude that $u_t = u_t^*$ solves the minimization problem (72) as well. *Q.E.D.*

**Cite this article:** Hahn V (2023). "Implementing the commitment solution via discretionary policy-making." *Macroeconomic Dynamics* **27**, 1866–1892. https://doi.org/10.1017/S1365100522000499