Equilibrium Under Inflation Targeting in an Infinitely Repeated Game

Alexandre B. Cunha[*†]
Federal University of Rio de Janeiro

December 20, 2023

Abstract Macroeconomic models often have multiple equilibria. Nonetheless, a popular view on inflation targeting posits that this policy helps to coordinate agents' expectations and actions. This paper provides a rationalization for this notion. I study an infinitely repeated game, built on the Barro-Gordon model, in which the central bank incurs a fixed penalty whenever the actual inflation rate differs from the announced target. I assess how changes in the penalty impact the set of subgame perfect equilibrium outcomes supported by trigger strategies that specify reversion to a one-shot Nash equilibrium. The results of this exercise are consistent with the aforementioned view.

Keywords inflation targeting; multiple equilibria; coordination; expectations

JEL classification E31; E52; E58; E61

1 Introduction

Equilibrium multiplicity is a pervasive feature of macroeconomic models. Despite this, several researchers have argued that an inflation targeting policy contributes to anchor the beliefs about future inflation rates and furthers the coordination of the actions of other economic agents with the central bank's goals. Notwithstanding the popularity of this view and the fact that it seems to be consistent with the empirical evidence, it appears that so far no paper has shown, in an intertemporal model, how the introduction of such a policy can lead to equilibrium uniqueness, or at least help players to coordinate on a specific equilibrium.

[*]E-mail: research@alexbcunha.com. Website: https://www.alexbcunha.com.

[†]I thank Luciane Carpena and participants of several conferences and seminars for their comments. Financial support from the National Council for Scientific and Technological Development – CNPq is gratefully acknowledged. Any remaining errors are my sole responsibility.

This paper's main goal consists of filling this gap. Therefore, I study the mechanics of the inflation targeting system in an infinitely repeated game. I adopt the well-known model of Barro and Gordon (1983a and 1983b) as the starting point to construct a suitable stage game. To ease the exposition, I first describe the game in its original form (which does not include an inflation targeting policy). There are two players, the central bank and the general public. At each period, they implement their actions in simultaneous ways. The general public selects its expectations for the inflation rate, while the central bank selects the actual inflation rate. Given those actions, the GDP is determined according to a Lucas-type Phillips curve. The period payoff of the general public is simply the square of its inflation forecast error. Concerning the payoff of the central bank, its preferences are described by a quadratic loss function that depends on the inflation rate and the deviation of output from a level, higher than the natural rate of production, that this player would like to achieve. This stage game has a single Nash equilibrium, in which the inflation rate is positive.

To incorporate the inflation targeting regime, I add two features to the stage game. First, I postulate the existence of an exogenous target for the inflation rate. This target is positive and smaller than the inflation rate in the mentioned Nash equilibrium. Second, I assume that the central bank incurs a fixed penalty (i.e., a payoff loss) whenever it fails to implement the target.[1] The equilibrium set of the modified stage game depends on the penalty. If it is small, the inflation targeting system is irrelevant and the only Nash equilibrium is equal to the one in the stage game without inflation targeting. If the penalty is large, then the system in question is effective to the point that the inflation rate hits the target in the unique Nash equilibrium of the game. If the penalty assumes an intermediate value, then there are two Nash equilibria, each of them being equal to one of the two equilibria just described.[2]

Next, I study the infinitely repeated version (with discount) of this game.[3] As usual, I use trigger strategies that specify reversion to a stage Nash equilibrium to characterize a set of subgame-perfect equilibrium outcomes of the repeated game. Since the set of stage Nash equilibria depends on the value of the penalty, so does the characterization. When the penalty is either small or large, the stage equilibrium is unique. As a consequence, the prescribed reversion must be to the corresponding Nash outcome. If the penalty falls in the intermediate range, then there are two stage equilibria. The general public is indifferent

[1] I further discuss the penalty assumption when reviewing the related literature and in Subsection 2.2.1. For the moment, it suffices to say that this hypothesis has already been used by other authors and that the size of the penalty can be interpreted as measure of the robustness of the institutions that control and supervise the central bank.

[2] In Subsection 2.2.2 I attribute precise meaning to the words *small*, *intermediate*, and *large* by fully characterizing two threshold values so that the penalty is small if it is smaller than the lower threshold, large if it larger than the higher threshold, and intermediate otherwise.

[3] The repeated game I consider is a very standard one. Hence, neither player can commit to future actions.

between the two, while the central bank is worst off in the equilibrium that is equal to the unique stage equilibrium when the penalty is small. Thus, I take this Nash outcome as the reversion threat. Summing up, if the penalty is either small or intermediate, then the trigger strategies specify reversion to a stage Nash equilibrium in which the central bank does not implement the target rate; if the penalty is large, then the reversion is to a stage Nash equilibrium in which the inflation rate is equal to its target.

At this point, I need to clarify a terminology matter. As mentioned in the previous paragraph, I focus on the subgame-perfect equilibrium outcomes of the infinitely repeated game that can be supported by trigger strategies that specify reversion to a stage Nash equilibrium. To avoid some tedious and long repetitions, in the next two paragraphs I will use shorthand expressions as *trigger equilibrium* and *trigger equilibrium outcome.*

In the context of this paper, the inflation targeting regime is fully characterized by two variables: the target rate and the penalty value. Although I treat the target as an exogenous variable, one can see it as an optimal inflation rate that is a function of preferences, technology, tax system, and other features (including institutions) of an economy. On the other hand, the penalty may be interpreted as a concise measure of a society's ability to align the central bank incentives with its desire to achieve a specific inflation rate. Consider now questions of the type "Is there a minimum value for the penalty needed to ensure that the infinitely repeated game has a trigger equilibrium outcome in which the inflation rate is equal to the target at every date?" and "How large does the penalty have to be to ensure that no deviations from the target rate can happen in a trigger equilibrium outcome of the infinitely repeated game?" Questions of this class are particular cases of the more general problem of understanding how changes in the penalty impact the set of trigger equilibrium outcomes of the repeated game, which is precisely the main exercise carried out in this manuscript.

This exercise provides several interesting results. Below I provide a brief five-point summary of them. First, there is a third threshold value for the penalty (smaller than the other two previously mentioned) with the property that there is a trigger equilibrium of the repeated game in which the central bank implements the target rate of inflation at every date if and only if the penalty is greater than or equal to this threshold. Second, an increase in the penalty will not create a new trigger equilibrium in which the central bank never implements the target. Third, if the penalty is already large, then an increase in its value will not create a new trigger equilibrium outcome. Fourth, whenever the penalty is large, the trigger equilibrium in which the central bank implements the target at every day is locally unique (among all trigger equilibria). Fifth, under some additional conditions on the size of the penalty, this uniqueness will be global (among all trigger equilibria). I should point out that none of those findings depends on the players' discount rates.

As previously mentioned, there is a widespread view that the inflation targeting regime has the capacity of helping economic agents to coordinate their expectations and actions. However, to my knowledge so far no paper has provided a theoretical foundation, in an infinite-horizon intertemporal game, for this view. And there is at least one reason for that: equilibrium multiplicity is a prevalent feature of this class of games. Therefore, at first glance, this type of model does not seem to be a viable framework to study how the inflation targeting regime can lead to better coordination among the agents of an economy. Despite that, this paper provides a rationalization for the manner in which an inflation targeting policy can contribute to the coordination of players' actions and lessen the problem of equilibrium multiplicity.

Related literature. With the onset of the rational expectations revolution, the existence of multiple equilibria in macroeconomic models became the focus of several studies. For instance, Sargent and Wallace (1975) concluded that under rational expectations, the price level becomes undetermined in the IS-LM model if the government pegs the nominal interest rate. In a similar vein, Obstfeld and Rogoff (1983) showed that the equilibrium path of the price level is not uniquely determined in a model with infinitely-lived maximizing households with perfect foresight. Also, Sargent (1987, Chapter 8) provided an overview of several results on the indeterminacy of the composition of the public debt. In turn, Woodford (1994) studied the equilibrium indeterminacy issue in an economy with cash and credit goods.

Three aspects of this problem deserve further comments. First, the equilibrium indeterminacy may arise in many distinct forms. For instance, in Sargent and Wallace (1975) and Obstfeld and Rogoff (1983), only the price level is affected by it, while in Woodford (1994) it also impacts real variables. Second, the equilibrium multiplicity became even more ubiquitous when, as in Chari and Kehoe (1990), macroeconomists started to apply the tolls of the theory of infinitely repeated games. Third, Atkeson, Chari and Kehoe (2010) warned against the adoption of the approach that they called unsophisticated implementation when trying to design a macroeconomic policy that uniquely implements an outcome. As they pointed out, this approach fails to check whether the proposed policy also constitutes an equilibrium after a deviation from the equilibrium path. Since I focus on subgame perfect equilibrium, this is not a point of concern here.

The notion that the introduction of an inflation targeting policy may contribute to the coordination of expectations and actions of economic agents is often found in the literature. For instance, Bernanke, Laubach, Mishkin and Posen (1999) stated that the inflation targeting system "provides a focus for the expectations of financial markets and the general public", while Walsh (2009) argued that this type of policy "can align the public's expectations of current and future target rates with the actual goals of the central bank" and may anchor "the public's beliefs about future inflation." Concerning the empirical evidence, Svensson (2011) mentioned that it strongly suggests that "an explicit numerical target for inflation anchors and stabilizes inflation expectations", while Gürkaynak, Levin, and Swanson (2010) stated that their findings "support the view that a transparent and credible inflation target helps to anchor the private sector's perceptions of the distribution of long-run inflation outcomes."

The following picture emerges from the previous paragraphs: equilibrium multiplicity is widely recognized in the broader macroeconomics literature, while the conventional wisdom found in its inflation targeting branch states that the regime in question should help agents to coordinate their expectations and actions. However, from a theoretical point of view, it is not clear how an inflation targeting policy can achieve this. For instance, Smith (1994) showed that multiple competitive equilibria can exist under inflation targeting in an overlapping generations model, while Adão, Correia and Teles (2011) concluded that the same result holds in a cash-in-advance model. This paper is an attempt to conciliate the aforementioned conventional wisdom with the well-known problem of multiple equilibria.

The penalty hypothesis, which plays a central role in this study, has been adopted by other authors. Araujo, Berriel, and Santos (2016) study the matter of equilibrium multiplicity in a game with inflation targeting. Exactly as in this paper, they assume that the central bank incurs a penalty whenever it does not implement the inflation target. They interpret the penalty as a commitment device embedded in the loss function that describes the preferences of the central bank.[4] A similar hypothesis, namely that the central bank faces a fixed cost when the domestic currency is devaluated, can be found in Obstfeld (1995) and Pastine (2000). Likewise, Rebelo and Végh (2008) assumed that an exchange rate devaluation makes the government incur a fixed fiscal cost. In a related vein, Cole and Kehoe (1996 and 2000) postulated that a country that defaults on its foreign debt will face a permanent fall in its total factor productivity.

The seminal papers of Kydland and Prescott (1977), Calvo (1978), and Barro and Gordon (1983a and 1983b) established that, in absence of commitment, even a benevolent central banker can implement an inflation rate higher than the socially optimum one. Rogoff (1985) showed that this problem can be solved by placing an agent whose payoff, when compared to

[4] Although being related, the questions addressed in this paper and by Araujo, Berriel, and Santos (2016) are not exactly the same. They consider a single-period game with imperfect information and evaluate how the ability of the central bank to implement low targets depends on the size of the penalty. They conclude that the lowest target rate consistent with equilibrium uniqueness decreases as the public information becomes noisier. This result in line with the findings of Jia and Wu (2023), who find that ambiguous communication may enhance the credibility of the central bank.

the relevant social welfare function, places a heavier weight on achieving low inflation relative to the weight placed on attaining low unemployment. One can interpret such penalty as a way of achieving Roggoff's solution even if the central banker is not sufficiently averse to inflation.[5]

This paper is also related to the literature on contracts for central bankers started by Person and Tabellini (1993) and Walsh (1995). Those authors, as well as Mishkin and Westelius (2008), argued that an inflation contingent compensation scheme can have effects similar to an inflation targeting policy. Consistent with this view, I show in Subsection 2.2.1 that introducing the penalty is equivalent to assuming that the central banker receives a performance bonus whenever the actual inflation rate is equal to the target. Hence, this paper can be seen as a study of the effects of a specific type of compensation scheme for the central banker.

Organization. The remainder of this paper is organized as follows. Section 2 describes the stage game and characterizes its set of Nash equilibria. Section 3 presents the infinitely repeated game and characterizes the set of equilibrium outcomes that can be supported by trigger strategies that specify reversion to a Nash equilibrium of the stage game. The effects on this set of changes in the penalty are studied in Section 4. Section 5 concludes.

2 Preliminaries: The Stage Game

In this section I present and analyze the stage game underlying the infinitely repeated game that is the focus of this paper. To facilitate the exposition, I first discuss a version of the stage game without inflation targeting. This version is based on the single-period inflation bias models of Kydland and Prescott (1977) and Barro and Gordon (1983a and 1983b). In a second step, I modify the game to incorporate the inflation targeting regime.

2.1 Without Inflation Targeting

There are two players in the game, the central bank (player b) and the general public (player p). Players move simultaneously. Agent p selects a value π^e for the expected inflation rate. Its choice must lie in the set $\Pi = [0, \pi_{\max}]$, where $\pi_{\max}$ is a positive number large enough so that it is never reached in any of the equilibria discussed in this paper. Player b selects a

[5] Although the penalty provides a possible solution to the inflation bias problem, this matter is not the focus of this paper. Instead, its main goal is to study the issue of equilibrium multiplicity under inflation targeting.

value $\pi \in \Pi$ for the actual inflation rate. Given those choices, the value $y \in \mathbb{R}$ of the natural logarithm of GDP is determined according to the Phillips curve

$$y = \bar{y} + \alpha(\pi - \pi^e),$$

where $\bar{y}$ is the natural logarithm of the natural rate of output and α is a positive parameter. The payoff of p is given by

$$V(\pi^e, \pi) = -(\pi - \pi^e)^2.$$

The quadratic loss function

$$w(y, \pi) = -\{\gamma\pi^2 + [y - (1 + \mu)\bar{y}]^2\}, \tag{1}$$

where $\gamma >$ and $\mu > 0$, is the payoff of b.[6]

As usual, one can use the Phillips curve to express the payoff of b as a function of π^e and π. Define the function W so that $W(\pi^e, \pi) = w(\bar{y} + \alpha(\pi - \pi^e), \pi)$. Hence,

$$W(\pi^e, \pi) = -\{\gamma\pi^2 + [\alpha(\pi - \pi^e) - \mu\bar{y}]^2\}. \tag{2}$$

From now on I use the function W to describe the central bank's payoff. This player's problem is

$$\max_{\pi \in \Pi} W(\pi^e, \pi), \tag{3}$$

while player p solves

$$\max_{\pi^e \in \Pi} V(\pi^e, \pi). \tag{4}$$

A *Nash equilibrium* for this game is a vector $(\hat{\pi}^e, \hat{\pi})$ with the properties that: (i) given $\hat{\pi}^e$, $\hat{\pi}$ solves the problem of player b and (ii) given $\hat{\pi}$, $\hat{\pi}^e$ solves the problem of player p.

The game has a unique Nash equilibrium. Denote the partial derivatives $\partial W/\partial \pi^e$ and $\partial W/\partial \pi$ by, respectively, W_1 and W_2. For a fixed π^e, the function $W(\pi^e, \cdot)$ is strictly concave. Therefore, the solution of (3) is characterized by

$$W_2(\pi^e, f(\pi^e)) = 0, \tag{5}$$

[6] Neither Kydland and Prescott (1977) nor Barro and Gordon (1983a and 1983b) spelled out the payoff of the private sector. They just assumed that inflation expectations were formed in a rational manner. The function V is a simple device to incorporate their assumption in the game considered here. Observe that Gibbons (1992) also adopted this payoff specification when discussing the paper of Barro and Gordon (1983b).

where f is the best response function of b. It is a simple exercise to show that

$$f(\pi^e) = \frac{\alpha^2}{\gamma + \alpha^2}\pi^e + \frac{\alpha\mu}{\gamma + \alpha^2}\bar{y}. \tag{6}$$

Optimality by player p requires $\pi^e = \pi$. Thus, $\hat{\pi}^e = \hat{\pi} = f(\hat{\pi})$. As a consequence,

$$\hat{\pi} = \frac{\alpha\mu}{\gamma}\bar{y}. \tag{7}$$

It should be pointed out that $\hat{\pi}$ is the unique fixed point of f.

2.2 With Inflation Targeting

In this subsection I carry out two tasks. First, I modify the previous game to incorporate the inflation targeting regime. Second, I characterize the equilibrium set of the resulting new game. Each of those tasks is carried out in a separate subsection.

2.2.1 Introducing Inflation Targeting

I assume that an outside agent (for instance, the legislature) carries out the task of modifying the game.[7] This agent implements two changes. First, it publicly announces that the central bank has to pursue a target $\pi^* \in (0, \hat{\pi})$ for the inflation rate.

The second change requires a longer discussion. When studying the inflation targeting regime, many authors assume that a term similar to $-\kappa(\pi - \pi^*)^2$, where κ is a positive constant, appears in the objective function of the central bank or in a social welfare function.[8] Following this approach here would require substituting the term $\kappa(\pi - \pi^*)^2$ for $\gamma\pi^2$ in expression (1). However, this procedure implicitly assumes that the external agent that introduces the inflation targeting regime has the ability to modify the function w, which runs contrary to the standard practice of taking preferences as given. Furthermore, once it is accepted that the external agent can change w, the obvious question is why the external agent will not set the parameter μ equal to 0 to ensure that $\pi = \pi^*$ in equilibrium. Therefore, I take a different path. Following Obstfeld (1995), who postulated that a central bank that has pegged its currency will face a fixed penalty if it allows the exchange rate to change, I assume that the external agent assesses a penalty on player b whenever the actual inflation rate differs from the target π^*. This penalty has the property that it leads to a payoff loss

[7] It is possible to assume that the changes are implemented by the central bank itself. I discuss this matter at end of this subsection.

[8] For some examples, see Svensson (1997 and 1999), Drazen (2000), and Capistrán and Ramos-Francia (2010).

equal to $C > 0$. Formally, I define the indicator function I so that

$$I(\pi) = \begin{cases} 1, & \text{if } \pi \neq \pi^* \\ 0, & \text{if } \pi = \pi^* \end{cases}.$$

The function U, which is given by

$$U(\pi^e, \pi) = W(\pi^e, \pi) - I(\pi)C, \tag{8}$$

is the central bank's payoff.

Araujo, Berriel and Santos (2016) adopted this assumption when studying inflation targeting policies in a context of imperfect information. They interpreted the penalty as a commitment technology. The same interpretation is valid here. Furthermore, one can assume there is an increasing relation between the robustness of society's institutions and the value of C. A second interpretation consists of associating the penalty with a performance-based compensation contract for the central banker. Indeed, consider the function $\tilde{U}(\pi^e, \pi) = W(\pi^e, \pi) + [1 - I(\pi)]C$. Suppose now that W describes the preferences of society (or of the government), while the term $[1 - I(\pi)]C$ corresponds to a performance bonus paid to the central banker whenever she/he succeeds in implementing the target rate π^*. If $\tilde{U}$ describes the payoff of the central banker, then this agent has precisely the type of utility function postulated in Walsh (1995). Since $\tilde{U}(\pi^e, \pi) = U(\pi^e, \pi) + C$, for the purposes of this paper, the functions U and $\tilde{U}$ are equivalent. As a consequence, Walsh's interpretation can indeed be applied here.

I close this subsection with a brief remark on the possibility of the above modifications being carried out by the central bank instead of an external agent. The results presented in this paper do not depend on who introduces the changes. However, there are two points to be considered if one assumes that the central bank itself is in charge of implementing the inflation targeting regime. The first is related to the magnitude of C. As shown in the next sections, the value of this parameter impacts the set of equilibrium outcomes: the higher it is, the easier it is to coordinate on an equilibrium in which the inflation rate is equal to the target π^*. Hence, one has to be concerned whether or not the central bank is able to choose a C as high as can be done by an outside institution. The second is related to the selection of the target rate, because if the central bank is in charge of choosing π^*, then it seems reasonable to assume that this agent will attempt to maximize its payoff when carrying out the task in question. Since b's payoff is given by (2) and (8), this player should set $\pi^* = 0$.

2.2.2 Equilibrium

The problem of player p is still given by (4), while player b solves

$$\max_{\pi \in \Pi} U(\pi^e, \pi). \tag{9}$$

A *stage Nash equilibrium with inflation targeting* is a vector (π^e, π) with the properties that: (i) given π^e, π solves the problem of player b; and (ii) given π, π^e solves the problem of player p.

I now turn to the task of characterizing the equilibrium set of the stage game. This requires solving problem (9). Since U is discontinuous at $\pi = \pi^*$, this cannot be done using solely the standard tools. Fortunately, there is a simple procedure that works out in this context: compare the values of $U(\pi^e, f(\pi^e))$ and $U(\pi^e, \pi^*)$ and select the argument $f(\pi^e)$ or π^* that yields the higher payoff. The next lemma formalizes this discussion.[9]

Lemma 1 *The following five statements are true:*
(i) $\max_{\pi \in \Pi} U(\pi^e, \pi) = \max\{U(\pi^e, \pi^*), U(\pi^e, f(\pi^e))\}$*;*
(ii) $\max\{U(\pi^e, \pi^*), U(\pi^e, f(\pi^e))\} = \max\{W(\pi^e, \pi^*), W(\pi^e, f(\pi^e)) - C\}$*;*
(iii) if $U(\pi^e, \pi^*) > U(\pi^e, f(\pi^e))$*, then* π^* *is the unique solution of (9);*
(iv) if $U(\pi^e, \pi^*) < U(\pi^e, f(\pi^e))$*, then* $f(\pi^e)$ *is the unique solution of (9);*
(v) if $U(\pi^e, \pi^*) = U(\pi^e, f(\pi^e))$*, then* π^* *and* $f(\pi^e)$ *are the only two solutions of (9).*

As one should expected, the equilibrium set depends on the value of the penalty C. Suppose that this parameter is small. In this case, the inflation targeting regime is not relevant and the only Nash equilibrium has to be $(\hat{\pi}, \hat{\pi})$, which is the one identified in the game without inflation targeting. On the other hand, if C is large, then the central bank will play π^* regardless of whether $\pi^e = \hat{\pi}$ or $\pi^e = \pi^*$. In anticipation of this, player p will optimally play π^*. Hence, in this case the only Nash equilibrium is (π^*, π^*). Also, an intermediate range for C may exist such that $(\hat{\pi}, \hat{\pi})$ and (π^*, π^*) are the only Nash equilibria.

It turns out that the above conjecture is indeed correct. However, formalizing it requires attributing precise meanings to the notions of small, large and intermediate values of C. Define the parameters k_1 and k_2 according to

$$k_1 = W(\pi^*, f(\pi^*)) - W(\pi^*, \pi^*) \tag{10}$$

and

$$k_2 = W(\hat{\pi}, f(\hat{\pi})) - W(\hat{\pi}, \pi^*). \tag{11}$$

[9] All proofs, as well as two examples, are available in an appendix at the end of the paper.

Since $f(\pi^e)$ uniquely maximizes $W(\pi^e, \cdot)$ and π^* is different from $f(\pi^*)$ and $f(\hat{\pi})$, both k_1 and k_2 are positive. Moreover, k_1 is exactly the value of C for which $U(\pi^*, f(\pi^*)) = U(\pi^*, \pi^*)$. Hence, if $C = k_1$ and player p selects $\pi^e = \pi^*$, then player b will be indifferent between the actions $f(\pi^*)$ and π^*. In similar fashion, $U(\hat{\pi}, f(\hat{\pi})) = U(\hat{\pi}, \pi^*)$ for $C = k_2$. Thus, if the last equality holds, then b can optimally play either $f(\hat{\pi})$ or π^* as a response to $\pi^e = \hat{\pi}$.

Besides playing an important role in this subsection, the next result will be used in the other parts of this paper.

Lemma 2 *The parameters k_1 and k_2 satisfy the inequality $k_1 < k_2$.*

It is now possible to say that the penalty is small when $C < k_1$, large when $C > k_2$, and intermediate when $C \in [k_1, k_2]$. The next proposition, which concludes this subsection, formalizes the intuitive description of the equilibrium set.

Proposition 1 *If $C < k_1$, then $(\hat{\pi}, \hat{\pi})$ is the unique stage Nash equilibrium with inflation targeting. If $C > k_2$, then (π^*, π^*) is the unique stage Nash equilibrium with inflation targeting. And if $C \in [k_1, k_2]$, then $(\hat{\pi}, \hat{\pi})$ and (π^*, π^*) are the only stage Nash equilibria with inflation targeting.*

3 The Infinitely Repeated Game

In this part of the paper I start the study of the game constituted by the infinite repetition of the stage game with inflation targeting of the previous section. In Subsection 3.1, I carry out some basic tasks, such as setting up some notation and presenting a suitable equilibrium concept. In Subsection 3.2, I specify the conditions that characterize the set of all equilibrium outcomes that can be supported by trigger strategies that specify reversion to one of the Nash equilibria of the stage game.

3.1 Structure and Equilibrium Definition

Denote a vector (π_t^e, π_t) of date-t actions by x_t, while ρ and δ are the respective discount factors of players p and b. As usual, these two parameters belong to the interval $(0, 1)$.[10] Given a sequence $\{x_t\}_{t=0}^{\infty}$ of actions, the payoffs of p and b from date t onwards are given,

[10] Two matters concerning the discounting factors should be clarified. First, the results of this paper do not depend on whether or not ρ and δ are different from each other. Second, I did not adopt the standard notation (β) used in the macro literature to emphasize that neither ρ nor δ has to be equal to the discount factor of a typical household of an economy underlying the game studied here.

respectively, by

$$\sum_{r=t}^{\infty} \rho^{r-t} V(\pi_r^e, \pi_r) \tag{12}$$

and

$$\sum_{r=t}^{\infty} \delta^{r-t} U(\pi_r^e, \pi_r). \tag{13}$$

Let h^t be a history $(x_0, x_1, \ldots, x_t)$ of actions. At the beginning of each date t, both players know the history h^{t-1}. Player p implements an action $\pi_t^e \in \Pi$, while b implements an action $\pi_t \in \Pi$. Denote these choices by $s_t(h^{t-1})$ and $\sigma_t(h^{t-1})$. Thus, a strategy for p is a sequence $s = \{s_t\}_{t=0}^{\infty}$, while a strategy for b is a sequence $\sigma = \{\sigma_t\}_{t=0}^{\infty}$. At each date t, given the history h^{t-1} and the strategy σ of player b, p selects a continuation sequence $\{s_r\}_{r=t}^{\infty}$ to maximize (12). This player takes into consideration that the actions of b evolve according to $\pi_r = \sigma_r(h^{r-1})$. In similar fashion, given h^{t-1}, s, and the rule $\pi_r^e = s_r(h^{r-1})$, b selects a continuation sequence $\{\sigma_r\}_{r=t}^{\infty}$ to maximize (13).[11]

An *equilibrium with inflation targeting* is a pair of strategies (s, σ) such that, at every date t and for every history h^{t-1}, $\{s_r\}_{r=t}^{\infty}$ and $\{\sigma_r\}_{r=t}^{\infty}$ solve the problems of the corresponding players. An *equilibrium outcome with inflation targeting* is a sequence $\{x_t\}_{t=0}^{\infty}$ with the property that there is an equilibrium (s, σ) that satisfies $\pi_t^e = s_t(x_0, x_1, \ldots, x_{t-1})$ and $\pi_t = \sigma_t(x_0, x_1, \ldots, x_{t-1})$. Clearly, an equilibrium with inflation targeting is subgame perfect.

3.2 A Set of Equilibrium Outcomes

In this subsection I characterize the set of all equilibrium outcomes that can be supported by trigger strategies that specify reversion to one of the stage Nash equilibria. I denote this set by $\mathcal{T}(C)$, where the $\mathcal{T}$ comes from the word *trigger* and the C is spelled out to emphasize that the set depends on the the value of the penalty.

Given that the set of stage equilibria depends on the value of C, it is necessary to consider separately the cases in which (i) $C < k_1$, (ii) $C \in [k_1, k_2]$, and (iii) $C > k_2$. Since in cases (i) and (iii) there is a single stage equilibrium, the reversion must be to the corresponding equilibrium. Concerning case (ii), observe that player p is indifferent between the outcomes $(\hat{\pi}, \hat{\pi})$ and (π^*, π^*), while $U(\hat{\pi}, \hat{\pi}) < U(\pi^*, \pi^*)$. Therefore, to characterize all equilibrium outcomes that can be supported by reverting to a stage Nash equilibrium, one has to take $(\hat{\pi}, \hat{\pi})$ as the reversion point.

[11] Observe that the game is structured in such a way that no player can commit to future actions.

All that being said, for $C \leq k_2$ the characterizing conditions are

$$\pi_t = \pi_t^e \tag{14}$$

and

$$\sum_{r=t}^{\infty} \delta^{r-t} U(\pi_r^e, \pi_r) \geq \max_{\pi \in \Pi} U(\pi_t^e, \pi) + \delta \frac{U(\hat{\pi}, \hat{\pi})}{1-\delta}. \tag{15}$$

If $C > k_2$, the corresponding expressions are (14) and

$$\sum_{r=t}^{\infty} \delta^{r-t} U(\pi_r^e, \pi_r) \geq \max_{\pi \in \Pi} U(\pi_t^e, \pi) + \delta \frac{U(\pi^*, \pi^*)}{1-\delta}. \tag{16}$$

This discussion is formalized in the next proposition.

Proposition 2 *For $C \leq k_2$, a sequence $\{x_t\}_{t=0}^{\infty}$ belongs to $\mathcal{T}(C)$ if and only if it satisfies (14) and (15). For $C > k_2$, a sequence $\{x_t\}_{t=0}^{\infty}$ belongs to $\mathcal{T}(C)$ if and only if it satisfies (14) and (16).*

The last proposition establishes that expressions (14), (15), and (16) provide sufficient conditions for a sequence X to be an equilibrium outcome with inflation targeting. A natural inquiry is whether every equilibrium outcome must satisfy the conditions in question. As shown in Lemma 3 in the Appendix, the answer is yes for equality (14). Concerning (15) and (16), the question at hand is equivalent to asking whether the worst (under the point of view of player b) equilibrium outcome with inflation targeting is equal to the worst stage Nash equilibrium. It turns out that at least for a patient central bank, the answer is no. Indeed, let $\tilde{\pi}$ be any inflation rate in the interval $(\hat{\pi}, \pi_{\max})$. Regardless of the value of C, for δ sufficiently close to 1 there is an equilibrium outcome in which at date 0 both agents play $\tilde{\pi}$ and in subsequent dates they switch to the worst stage Nash equilibrium. Since $U(\tilde{\pi}, \tilde{\pi})$ is smaller than both $U(\hat{\pi}, \hat{\pi})$ and $U(\pi^*, \pi^*)$, no stage Nash equilibrium can be the worst equilibrium outcome of the repeated game.

4 Equilibrium Outcomes and the Penalty

In this section I study how the set $\mathcal{T}(C)$ is affected by changes in C, as well as under which conditions playing the target π^* on every date is an equilibrium action for both players and some other related issues. I start by introducing some notation. I denote a generic sequence $\{x_t\}_{t=0}^{\infty}$ by X and the sequence in which $x_t = (\pi^*, \pi^*)$ for all t by X^*. Similarly, $\hat{X}$ is the sequence with the property that $x_t = (\hat{\pi}, \hat{\pi})$ for all t. The set of all sequences in $\Pi \times \Pi$ is denoted by $\mathbb{X}$. The subset of $\mathbb{X}$ containing all sequences that have the property that $\pi_t = \pi^*$ for some t is denoted by $\mathbb{X}^*$, while $\tilde{\mathbb{X}}^*$ is the subset of $\mathbb{X}$ containing all sequences with the property that $\pi_t \neq \pi^*$ for every t (i.e., the complement of $\mathbb{X}^*$ with respect to $\mathbb{X}$). I also have to clarify a minor point concerning the usage of the symbols $\subseteq$ and $\subset$. The latter requires the inclusion to be a proper one, while the former allows the two sets under ,analysis to be equal.

4.1 The Sustainability of X^*

In this subsection, I provide a necessary and sufficient condition for X^* to be an element of $\mathcal{T}(C)$. Define the parameter k_0 so that

$$k_0 = (1-\delta)k_1 - \delta[W(\pi^*, \pi^*) - W(\hat{\pi}, \hat{\pi})].$$

Clearly, $k_0 < k_1$. Moreover, $k_0 > 0$ if and only if $\delta < k_1/\{k_1 + [W(\pi^*, \pi^*) - W(\hat{\pi}, \hat{\pi})]\}$.

Proposition 3 *The sequence X^* belongs to $\mathcal{T}(C)$ if and only if $C \geq k_0$.*

If $k_0 \leq 0$, then X^* will be an equilibrium outcome regardless of the value of C (including the limit case of $C = 0$). On the other hand, the implementation of X^* with trigger strategies that specify reversion to a stage Nash equilibrium requires C to be positive whenever $k_0 > 0$.

4.2 The Impacts of Changes in C

In this subsection, I study how the set $\mathcal{T}(C)$ evolves as C increases. The results are split into three propositions.

Suppose that $C \leq k_2$. Then several equilibrium outcomes can be supported by the strategy of reverting to $\hat{X}$ after a deviation. Now, let X' be any sequence such that $\pi_r \neq \pi^*$ for every date r. Consider a generic date t. For dates later than t, the value of C is irrelevant (under the point of view of player b) when comparing the payoffs of X' and $\hat{X}$, since the penalty will be assessed at every date under both sequences. However, by deviating to π^* on date t, player b can avoid the penalty on this date. Therefore, the larger the value of C, the larger will be the incentive of player b to deviate from X' by implementing π^* on date t. Thus, as C increases, sequences belonging to $\tilde{\mathbb{X}}^*$ tend to be dropped out of $\mathcal{T}(C)$.

Proposition 4 *If $C_1 < C_2 \leq k_2$, then $\mathcal{T}(C_2) \cap \tilde{\mathbb{X}}^* \subseteq \mathcal{T}(C_1) \cap \tilde{\mathbb{X}}^*$.*

Mention should be made of two points about this proposition. First, provided that C does not become higher than k_2, if an increase in the penalty adds a sequence X to the set

$\mathcal{T}(C)$, then X must belong to $\mathbb{X}^*$. Second, as illustrated by Example 1 in the Appendix, the sets $\mathcal{T}(C_2) \cap \tilde{\mathbb{X}}^*$ and $\mathcal{T}(C_1) \cap \tilde{\mathbb{X}}^*$ do not have to be equal.

As a consequence of Proposition 3, if X^* is an equilibrium outcome for a given penalty, then it will also be for a higher one. At a first glance, it may appear that this result can be extended to any sequence belonging to $\mathbb{X}^*$. However, an additional condition is required. Indeed, consider two penalties C_1 and C_2 such that $C_1 < C_2 \leq k_2$. Now, let $X \neq X^*$ be an element of $\mathcal{T}(C_1) \cap \mathbb{X}^*$ and t be any date on which $\pi_t \neq \pi^*$. It may happen that an increase in C induces the central bank to deviate to π^* on date t to avoid incurring a higher penalty. However, if on future dates X hits π^* sufficiently often, such a deviation will not be an optimal action for player b. Indeed, consider the inequality

$$\sum_{r=t+1}^{\infty} \delta^{r-t}[1 - I(\pi_r)] \geq I(\pi_t). \tag{17}$$

The sum in the left-hand side is simply a discounted count of the number of times that, after date t, X hits the target π^*. Since $I(\pi^*) = 0$, this inequality will surely hold for $\pi_t = \pi^*$. If $\pi_t \neq \pi^*$, the sum will have to be equal to or greater than 1 for the condition to hold. Denote by $\mathbb{X}^*_\delta$ the subset of $\mathbb{X}^*$ containing all sequences that satisfy (17) for every t. In the next proposition I show that if an element of $\mathbb{X}^*_\delta$ belongs to $\mathcal{T}(C_1)$, then it will also belong to $\mathcal{T}(C_2)$.

Proposition 5 *If $C_1 < C_2 \leq k_2$, then $\mathcal{T}(C_1) \cap \mathbb{X}^*_\delta \subseteq \mathcal{T}(C_2) \cap \mathbb{X}^*_\delta$.*

This proposition implies that provided that the inequality $C \leq k_2$ is not violated, if a sequence is dropped from to set of equilibrium outcomes as a consequence of an increase in the penalty, then it must be the case that X does not hit π^* as often as required by (17). Moreover, Example 2 in the Appendix illustrates that one should not assume that the sets $\mathcal{T}(C_1) \cap \mathbb{X}^*_\delta$ and $\mathcal{T}(C_2) \cap \mathbb{X}^*_\delta$ are equal.

As established in Proposition 2, the conditions that characterize the set of equilibrium outcomes that can be supported by reverting to a Nash equilibria of the stage game depend on whether or not $C > k_2$. Thus, it is important to understand what happens to $\mathcal{T}(C)$ when the value of C changes from k_2 to a higher one. An intuitive analysis of the those conditions suggests that for a penalty C slightly higher than k_2, $\mathcal{T}(C)$ should be a proper subset of $\mathcal{T}(k_2)$. Indeed, for a given penalty, every sequence that satisfies (16) will also satisfy (15). Hence, exactly when the penalty shifts from k_2 to a slightly higher value, every sequence that satisfies (15) but does not satisfy (16) will be dropped from the set of equilibrium outcomes. Furthermore, given a sequence $X \neq X^*$ that respects (16), a further increase in C will impact the inequality in question in such a way that X may fail to satisfy it. Hence, whenever $C_2 > C_1 > k_2$, $\mathcal{T}(C_2)$ should be a subset of $\mathcal{T}(C_1)$. This discussion is formalized in the next proposition.

Proposition 6 *If $C_2 > C_1 > k_2$, then $\mathcal{T}(C_2) \subseteq \mathcal{T}(C_1) \subset \mathcal{T}(k_2)$.*

An obvious implication of this last result is that if C is already larger than k_2, a further increase in the penalty will not add an element to $\mathcal{T}(C)$. It is also worth pointing out that the sets $\mathcal{T}(C_1)$ and $\mathcal{T}(C_2)$ may be equal. Indeed, as a consequence of the forthcoming Proposition 8, both $\mathcal{T}(C_1)$ and $\mathcal{T}(C_2)$ will be equal to $\{X^*\}$ whenever C_1 is sufficiently large.

4.3 The Local and Global Uniqueness of X^*

I present two main results in this subsection. First, in Proposition 7 I show that provided that $C > k_2$, a sequence different from but sufficiently close to X^* cannot be an element of $\mathcal{T}(C)$. Second, in Proposition 8 I prove that X^* is the unique element of $\mathcal{T}(C)$ whenever C is sufficiently large. Furthermore, I provide some complementary findings in Propositions 9 and 10.

To grasp the intuition behind the first of these results, assume that $C > k_2$. Thus, (π^*, π^*) is the stage Nash equilibrium used to support equilibrium outcomes of the repeated game. Now, take a sequence $X \neq X^*$. At any date t on which $\pi_t \neq \pi^*$, a deviation from π_t to π^* will allow b to avoid the penalty C on the date in question. Moreover, if X is such that π_r is sufficiently close to π^* for all $r > t$, the current gain from avoiding the penalty will more than offset any conceivable losses on future dates. Therefore, X cannot belong to $\mathcal{T}(C)$.

Proposition 7 *For every $C > k_2$, there exists a positive number ε (that does not depend on δ) such that if a sequence $X \neq X^*$ has the property that $|\pi_t - \pi^*| < \varepsilon$ for all t, then $X \notin \mathcal{T}(C)$.*

Concerning the second result, its underlying reasoning is relatively simple. As C becomes larger, eventually the central bank's incentives to avoid the penalty will be strong enough to prevent it from ever implementing an inflation rate different from π^*. That being said, the real problem consists of obtaining the desired result without requiring C to be needlessly large.

Since $\hat{X} \in \mathcal{T}(C)$ for all $C \leq k_2$, the aforementioned uniqueness requires $C > k_2$. Next, define k_3 according to

$$k_3 = W(0,0) - W(\pi^*, \pi^*). \tag{18}$$

It is established in the next proposition that $C \geq k_3$ and $C > k_2$ are sufficient conditions for X^* to be the sole element of $\mathcal{T}(C)$. It is worth pointing out that k_3 may be smaller than k_2. Indeed, it is possible to show that $k_2 = (\alpha^2 + \gamma)(\hat{\pi} - \pi^*)^2$ and $k_3 = \gamma(\pi^*)^2$. Thus,

$$k_3 > k_2 \Leftrightarrow \pi^* > \frac{(\alpha^2 + \gamma)^{0.5}}{(\alpha^2 + \gamma)^{0.5} + \gamma^{0.5}}\hat{\pi}.$$

Proposition 8 *If $C > k_2$ and $C \geq k_3$, then $\mathcal{T}(C) = \{X^*\}$.*

Next I investigate whether the converse of the last proposition is true. That is, I study whether the equality $\mathcal{T}(C) = \{X^*\}$ implies that $C > k_2$ and $C \geq k_3$. However, this is equivalent to studying if $C \leq k_2$ or $C < k_3$ implies that $\mathcal{T}(C) \neq \{X^*\}$. Since $\hat{X}$ will belong to $\mathcal{T}(C)$ whenever $C \leq k_2$, it remains to consider what happens when $C \in (k_2, k_3)$.

Suppose that $k_2 < k_3$ and take any C in the interval (k_2, k_3). Hence, (18) implies that $W(0,0) - W(\pi^*, \pi^*) - C > 0$. Now, let π_C be any inflation rate in the interval $(0, \pi^*)$ with the property that

$$W(\pi_C, \pi_C) - W(\pi^*, \pi^*) - C > 0 \tag{19}$$

and $\mathbb{X}(\pi_C)$ be the subset of all sequences in $\mathbb{X}$ such that $\pi_t^e = \pi_t$ and $\pi_t \in [0, \pi_C]$ for all t.

Proposition 9 *Suppose that $k_2 < k_3$. Hence, for every $C \in (k_2, k_3)$ there is a number $\delta_C \in (0,1)$ with the property that if $\delta \in [\delta_C, 1)$, then $\mathbb{X}(\pi_C) \subseteq \mathcal{T}(C)$.*

One may wonder whether the assumption regarding δ can be dispensed with. As shown in the next proposition, the answer is no.

Proposition 10 *Suppose that $k_2 < k_3$. Hence, for every $C \in (k_2, k_3)$ there is a number $\delta_C' \in (0,1)$ with the property that if $\delta \in (0, \delta_C')$, then $\mathcal{T}(C) = \{X^*\}$.*

I close this subsection with two brief comments. First, Proposition 9 makes clear that k_3 is not an unnecessarily high lower bound. Second, the critical discount rates δ_C and δ_C' are uniform over $\mathbb{X}$ (i.e., they do not depend on the sequences).

4.4 Summing Up

I close this section with an assessment of its findings. I have studied under which conditions X^* is an element of $\mathcal{T}(C)$. It turns out that a number k_0 exists with the property that $X^* \in \mathcal{T}(C)$ if and only if $C \geq k_0$. It may happen that $k_0 \leq 0$. If so, then X^* will be an equilibrium outcome with inflation targeting regardless of the value of the penalty (including the limiting case in which $C = 0$). If $k_0 > 0$, then X^* will be an element of $\mathcal{T}(C)$ only if $C > 0$.

I have also investigated how changes in C impact $\mathcal{T}(C)$. Recall that the characterization of this set depends on whether or not $C \le k_2$. I showed that while this inequality holds, an increase in C will not add to $\mathcal{T}(C)$ a sequence X in which $\pi_t \neq \pi^*$ for all t and it will not drop from the set in question a sequence X that hits the target π^* sufficiently often. If $C > k_2$, then $\mathcal{T}(C)$ is proper subset of $\mathcal{T}(k_2)$ and an increase in C will not lead to enlargement of the set $\mathcal{T}(C)$.

Moreover, I have studied under which conditions X^* is the unique element of $\mathcal{T}(C)$. For such a uniqueness to happen, it is necessary that $C > k_2$. When this inequality is satisfied, then X^* has a local uniqueness property. That is, for every $C > k_2$ there is a neighborhood of X^* such that no sequence in this neighborhood will belong to $\mathcal{T}(C)$. Moreover, there is a number $k_3 > 0$ such that if $C \ge k_3$, then X^* is the only element of $\mathcal{T}(C)$.

It is natural to wonder if the conditions required to achieve local or global uniqueness are very restrictive. Despite the simplicity of model considered in this paper, it is still possible to carry out an exploratory analysis of the matter. I start with the local uniqueness. Propositions 1 and 7 imply that local uniqueness around X^* is reached precisely for values of C high enough to break down the $(\hat{\pi}, \hat{\pi})$ stage equilibrium and to ensure that (π^*, π^*) is the unique one-shot equilibrium. This appears to be a high bar to meet.

Concerning global uniqueness, Propositions 8 and 9 together establish that the inequality $C \ge k_3$ must be satisfied to ensure that, regardless of the value of δ, X^* is the only element of $\mathcal{T}(C)$. Suppose that the inequality in question holds. Thus, $W(\pi^*, \pi^*) \ge W(0, 0) - C$, from which follows that $U(\pi^*, \pi^*) \ge U(0, 0)$. Since $U(0, 0) \ge U(\pi, \pi)$ for all $\pi \neq \pi^*$, then the inequality $U(\pi^*, \pi^*) \ge U(\pi, \pi)$ must hold for all π . Therefore, requiring C to be larger than or equal to k_3 entails demanding the penalty to be large enough to ensure that (π^*, π^*) will yield player b at least the same period payoff as any other vector (π, π). That is, the penalty must be sufficiently large to fully offset all incentives the central bank has to deviate from the target. Of course, meeting this requirement is a tall order task.

5 Concluding Remarks

It is well-known that macroeconomic models often display multiple equilibria. Nevertheless, there is a conventional wisdom (which appears to be in line with the empirical evidence) about the inflation targeting regime that claims it may help to coordinate the expectations and actions of economic agents with those of the central bank. However, the ways in which this type of policy can impact the equilibrium set or induce players to coordinate on a particular outcome of an intertemporal model is not yet well understood.

The goal of this paper is to contribute to close this gap. Thus, I analyze how an inflation targeting policy can impact the equilibria of an infinitely repeated game. Its stage game is a variant of the popular Barro-Gordon model. Compared with its parent, this modified version has just two additional features: (i) there is an exogenous target for the inflation rate and (ii) the central bank incurs a fixed penalty (i.e., a payoff loss) whenever it does not implement the target. This penalty can be interpreted as a concise measure of a society's ability to induce the central bank to pursue the target. Hence, a larger penalty can be associated with a more robust policy framework.

I assess how changes in the penalty impact the set of equilibrium outcomes that can be supported by trigger strategies that specify reversion to a Nash equilibrium of the stage game. This exercise establishes that as the penalty increases, outcomes in which the target is never implemented tend to be dropped from the set in question. I also show that if the penalty is larger than a critical value, the equilibrium outcome in which the target is implemented at every date is locally unique with respect to the set, in the sense that there is a neighborhood of this outcome that contains no other element of the set. Furthermore, under some additional conditions the uniqueness is global (i.e., the outcome in question is the sole element of the set). Therefore, the inflation targeting system indeed has, from a theoretical viewpoint, properties consistent with the mentioned conventional wisdom.

Appendix: Proofs and Examples

Proof of Lemma 1. I start with statement (i). Let π^e be any element of Π. Then,

$$\pi \neq \pi^* \Rightarrow U(\pi^e, \pi) = W(\pi^e, \pi) - C \leq W(\pi^e, f(\pi^e)) - C \leq U(\pi^e, f(\pi^e)).$$

Hence,

$$\pi \neq \pi^* \Rightarrow U(\pi^e, \pi) \leq \max\{U(\pi^e, \pi^*), U(\pi^e, f(\pi^e))\}. \quad (20)$$

Since $U(\pi^e, \pi^*) \leq \max\{U(\pi^e, \pi^*), U(\pi^e, f(\pi^e))\}$, the inequality in (20) holds for all $\pi \in \Pi$; therefore, (i) is true. Given that $U(\pi^e, \pi^*) = W(\pi^e, \pi^*)$ and

$$f(\pi^e) \neq \pi^* \Rightarrow U(\pi^e, f(\pi^e)) = W(\pi^e, f(\pi^e)) - C,$$

the equality in (ii) holds if $f(\pi^e) \neq \pi^*$. If $f(\pi^e) = \pi^*$, then both sides of the equality in question will be equal to $W(\pi^e, \pi^*)$. Hence, statement (ii) is true. The last three statements follow directly from (i). ∎

Proof of Lemma 2. Define the function $\psi(\pi^e)$ according to

$$\psi(\pi^e) = W(\pi^e, f(\pi^e)) - W(\pi^e, \pi^*). \tag{21}$$

Thus, $\psi'(\pi^e) = W_1(\pi^e, f(\pi^e)) + W_2(\pi^e, f(\pi^e))f'(\pi^e) - W_1(\pi^e, \pi^*)$. Together, the last equality and (5) imply that $\psi'(\pi^e) = W_1(\pi^e, f(\pi^e)) - W_1(\pi^e, \pi^*)$. Since $W_1(\pi^e, \pi) = 2\alpha[\alpha(\pi - \pi^e) - \mu\bar{y}]$, $\psi'(\pi^e) = 2\alpha^2[f(\pi^e) - \pi^*]$. On the other hand, (6) and (7) imply that

$$f(\pi^*) - \pi^* = \frac{\gamma}{\gamma + \alpha^2}(\hat{\pi} - \pi^*) > 0.$$

Given that f is strictly increasing, $f(\pi^e) - \pi^* \geq f(\pi^*) - \pi^* > 0$ for every $\pi^e \geq \pi^*$. Thus,

$$\psi'(\pi^e) > 0, \ \forall \pi^e \geq \pi^*. \tag{22}$$

Therefore, $\psi(\pi^*) < \psi(\hat{\pi})$. Since $\psi(\pi^*) = k_1$ and $\psi(\hat{\pi}) = k_2$, $k_1 < k_2$. ∎

Proof of Proposition 1. Suppose that $C < k_1$. Combine this inequality with Lemma 2 to conclude that $C < k_2$. Then use (11) to show that $W(\hat{\pi}, \pi^*) < W(\hat{\pi}, f(\hat{\pi})) - C$, which in turn implies that $U(\hat{\pi}, \pi^*) < U(\hat{\pi}, f(\hat{\pi}))$. Thus, if p plays $\pi^e = \hat{\pi}$, then the best action for the central bank consists in playing $\pi = f(\hat{\pi})$. Since $f(\hat{\pi}) = \hat{\pi}$, $(\hat{\pi}, \hat{\pi})$ is a stage Nash equilibrium with inflation targeting. Concerning uniqueness, optimality by player p implies that $(\hat{\pi}, \pi)$ is not an equilibrium for any $\pi \neq \hat{\pi}$. Consider now what happens when p implements an action $\pi^e = \pi' \neq \hat{\pi}$. Suppose that $\pi' = \pi^*$. Then, use the fact that $C < k_1$ and equality (10) to show that $U(\pi^*, \pi^*) < U(\pi^*, f(\pi^*))$. Hence, player b should choose $\pi = f(\pi^*)$. Since $f(\pi^*) \neq \pi^*$, $\pi^e = \pi^*$ is not a best response to $\pi = f(\pi^*)$. For the case in which π' is different from both π^* and $\hat{\pi}$, recall that the optimal choice for b is (i) $\pi = \pi^*$ or (ii) $\pi = f(\pi')$. If (i) is true, then the assumption that $\pi' \neq \pi^*$ implies that such a π' does not solve the problem of player p. If (ii) holds, then the fact that $\hat{\pi}$ is the only fixed point of f implies that $\pi' \neq f(\pi')$. Again, π' cannot be an optimal strategy for p.

Now, assume that $C > k_2$. Apply Lemma 2 again to show that $C > k_1$. Hence, (10) implies that $U(\pi^*, \pi^*) > U(\pi^*, f(\pi^*))$. So, if p implements the action $\pi^e = \pi^*$, the central bank's best response is $\pi = \pi^*$. Thus, (π^*, π^*) is a stage Nash equilibrium with inflation targeting. To show there is no other equilibrium, observe that optimality by p implies that (π^*, π) is not an equilibrium for any $\pi \neq \pi^*$. Next, assume that p plays $\pi^e = \pi' \neq \pi^*$. If $\pi' = \hat{\pi}$, the inequality $C > k_2$ implies that $U(\hat{\pi}, \pi^*) > U(\hat{\pi}, f(\hat{\pi}))$. Thus, b should play π^*, and as a consequence, $\hat{\pi}$ is not an optimal action for player p. If π' is different from both π^* and $\hat{\pi}$, the optimal action for the central bank will be π^* or $f(\pi')$. Again, the facts that $\pi' \neq \pi^*$ and $\pi' \neq f(\pi')$ imply that such a π' is not an optimal strategy for player p.

Finally, consider the case in which $k_1 \leq C \leq k_2$. The inequality $C \geq k_1$ and (10) imply that $U(\pi^*, \pi^*) \geq U(\pi^*, f(\pi^*))$. Thus, if p plays π^*, then b can optimally play π^*. Therefore, (π^*, π^*) is a stage Nash equilibrium with inflation targeting. Similarly, the inequality $C \leq k_2$ and (11) imply that $U(\hat{\pi}, f(\hat{\pi})) \geq U(\hat{\pi}, \pi^*)$. Hence, $(\hat{\pi}, \hat{\pi})$ is also a stage Nash equilibrium with inflation targeting. To establish that there is no other equilibrium, observe that optimality by player p implies that (π^*, π) is not an equilibrium for any $\pi \neq \pi^*$, while $(\hat{\pi}, \pi)$ is not for any $\pi \neq \hat{\pi}$. Moreover, if p plays $\pi' \notin \{\pi^*, \hat{\pi}\}$, then the optimal response for b is π^* or $f(\pi')$. Since $\pi' \neq \pi^*$ and $\pi' \neq f(\pi')$, π' cannot be an equilibrium strategy for p. ∎

Lemma 3 *If $\{x_t\}_{t=0}^{\infty}$ is an equilibrium outcome with inflation targeting, then $\{x_t\}_{t=0}^{\infty}$ satisfies (14).*

Proof. Given a strategy σ for player b, the strategy of setting π_t^e equals to $\sigma_t(h^{t-1})$ will yield player p its maximum attainable payoff (which is 0) at every node of the game. Hence, $V(\pi_t^e, \pi_t) = 0$ in any equilibrium with inflation targeting. Thus, (14) must hold in such an equilibrium. ∎

Proof of Proposition 2. Suppose that $C \leq k_2$. Therefore, $(\hat{\pi}, \hat{\pi})$ is a stage Nash equilibrium with inflation targeting. Hence, any sequence $\{x_t\}_{t=0}^{\infty}$ that satisfies (14) and (15) must belong to $\mathcal{T}(C)$. This establishes the "if part". For the "only if part", let $\{x_t\}_{t=0}^{\infty}$ be any element of $\mathcal{T}(C)$. Apply Lemma 3 to conclude that $\{x_t\}_{t=0}^{\infty}$ satisfies (14). Moreover, the fact that $\{x_t\}_{t=0}^{\infty}$ belongs to $\mathcal{T}(C)$ implies that this sequence must satisfy (15) or (16). However, $U(\hat{\pi}, \hat{\pi}) < U(\pi^*, \pi^*)$. Thus, if $\{x_t\}_{t=0}^{\infty}$ satisfies (16), then it must also satisfy (15). Similar arguments can be applied to the case in which $C > k_2$. ∎

Proof of Proposition 3. I start with the "if part". Suppose that $C \geq k_0$. If $C \geq k_1$, then (π^*, π^*) is a stage Nash equilibrium with inflation targeting, and as a consequence, X^* is an equilibrium outcome of the infinitely repeated game. If $C < k_1$, the fact that $C \geq k_0$ implies that

$$C \geq (1-\delta)k_1 - \delta[W(\pi^*, \pi^*) - W(\hat{\pi}, \hat{\pi})] \Rightarrow$$
$$(1-\delta)C \geq (1-\delta)k_1 - \delta\{W(\pi^*, \pi^*) - [W(\hat{\pi}, \hat{\pi}) - C]\} \Rightarrow$$
$$C \geq k_1 - \frac{\delta}{1-\delta}\{U(\pi^*, \pi^*) - U(\hat{\pi}, \hat{\pi})\}.$$

Combine the last inequality with (10) to conclude that

$$C \geq W(\pi^*, f(\pi^*)) - W(\pi^*, \pi^*) - \frac{\delta}{1-\delta}\{U(\pi^*, \pi^*) - U(\hat{\pi}, \hat{\pi})\} \Rightarrow$$
$$W(\pi^*, \pi^*) + \frac{\delta}{1-\delta}U(\pi^*, \pi^*) \geq W(\pi^*, f(\pi^*)) - C + \frac{\delta}{1-\delta}U(\hat{\pi}, \hat{\pi}).$$

However, $f(\pi^*) \neq \pi^*$. Therefore,

$$U(\pi^*, \pi^*) + \frac{\delta}{1-\delta}U(\pi^*, \pi^*) \geq U(\pi^*, f(\pi^*)) + \frac{\delta}{1-\delta}U(\hat{\pi}, \hat{\pi}). \tag{23}$$

Now, observe that the inequality $C < k_1$ implies that $U(\pi^*, \pi^*) < U(\pi^*, f(\pi^*))$. Thus, it is possible to apply Lemma 1 to show that $\max_{\pi \in \Pi} U(\pi^*, \pi) = U(\pi^*, f(\pi^*))$. Combine this equality with (23) to conclude that X^* satisfies (15). Hence, X^* is an equilibrium outcome with inflation targeting.

Concerning the "only if part", it suffices to show that if $C < k_0$, then $X^* \notin \mathcal{T}(C)$. A reasoning similar to that used to obtain (23) establishes that the reverse inequality holds strictly if $C < k_0$. Thus, player b can enhance its payoff by deviating to $f(\pi^*)$, and as a consequence, $X^* \notin \mathcal{T}(C)$. ∎

Proof of Proposition 4. Let X be any element of $\mathcal{T}(C_2) \cap \tilde{\mathbb{X}}^*$. Thus, X satisfies (14). Furthermore, for any $X \in \tilde{\mathbb{X}}^*$, inequality (15) is equivalent to

$$\sum_{r=t+1}^{\infty} \delta^{r-t}[W(\pi_r^e, \pi_r) - W(\hat{\pi}, \hat{\pi})] \geq$$
$$\max\{W(\pi_t^e, \pi^*) + C, W(\pi_t^e, f(\pi_t^e)) + [1 - I(f(\pi_t^e))]C\} - W(\pi_t^e, \pi_t).$$

Its right-hand side is non-decreasing on C, while the left-hand side does not depend on the variable in question. Given that $C_1 < C_2$ and X satisfies this condition for $C = C_2$, it must be the case that the same is true for $C = C_1$. Therefore, $X \in \mathcal{T}(C_1) \cap \tilde{\mathbb{X}}^*$. ∎

Example 1 The sets $\mathcal{T}(C_1) \cap \tilde{\mathbb{X}}^*$ and $\mathcal{T}(C_2) \cap \tilde{\mathbb{X}}^*$ in Proposition 4 do not need to be equal. Let π' be an inflation rate belonging to the interval $(\pi^*, \hat{\pi})$ and X' be the sequence in which $\pi_t^e = \pi_t = \pi'$ for all t . Since $\pi^* < \pi' < \hat{\pi}$, (10), (11), (21), and (22) imply that

$$k_1 = \psi(\pi^*) < \psi(\pi') < \psi(\hat{\pi}) = k_2.$$

Furthermore, if $C > \psi(\pi')$, then $W(\pi', f(\pi')) - C < W(\pi', \pi^*)$. Hence, for any such C,

$$\max_{\pi\in\Pi}[W(\pi', \pi) - I(\pi)C] = W(\pi', \pi^*)$$

and

$$0 < W(\pi', f(\pi')) - W(\pi', \pi') < W(\pi', \pi^*) - W(\pi', \pi') + C.$$

Therefore,

$$\max_{\pi\in\Pi}[W(\pi', \pi) - I(\pi)C] - [W(\pi', \pi') - C] = W(\pi', \pi^*) - W(\pi', \pi') + C > 0. \tag{24}$$

Now, take a penalty $C_1 \in (\psi(\pi'), k_2)$ and a penalty $C_2 \in (C_1, k_2)$. Thus, the equality in (24) holds for $C = C_1$ and $C = C_2$. Next, define δ' so that

$$\frac{\delta'}{1-\delta'}[W(\pi', \pi') - W(\hat{\pi}, \hat{\pi})] = W(\pi', \pi^*) - W(\pi', \pi') + C_1. \tag{25}$$

Combine the inequality $W(\pi', \pi') > W(\hat{\pi}, \hat{\pi})$ with (24) to conclude that δ' is well defined and lies in the interval $(0, 1)$. Hence, if $\delta = \delta'$ and $C = C_1$, then X' satisfies (15) as an equality, which implies that $X' \in \mathcal{T}(C_1) \cap \tilde{\mathbb{X}}^*$. On the other hand,

$$W(\pi', \pi^*) - W(\pi', \pi') + C_1 < W(\pi', \pi^*) - W(\pi', \pi') + C_2.$$

Combine this inequality with (25) to conclude that X' does not satisfy (15) for $\delta = \delta'$ and $C = C_2$. As a consequence, $X' \notin \mathcal{T}(C_2) \cap \tilde{\mathbb{X}}^*$. ∎

Proof of Proposition 5. Take a sequence X belonging to $\mathcal{T}(C_1) \cap \mathbb{X}^*_\delta$. Clearly, X satisfies (14). Condition (15) can be written as

$$\sum_{r=t+1}^{\infty} \delta^{r-t}[W(\pi^e_r, \pi_r) - W(\hat{\pi}, \hat{\pi})] + \left\{\sum_{r=t+1}^{\infty} \delta^{r-t}[1 - I(\pi_r)] - I(\pi_t)\right\} C \geq$$
$$\max\left\{W(\pi^e_t, \pi^*), W(\pi^e_t, f(\pi^e_t)) - I(f(\pi^e_t))C\right\} - W(\pi^e_t, \pi_t).$$

Since $X \in \mathbb{X}^*_\delta$, (17) implies that the expression inside the large curly brackets is non-negative. Therefore, the left-hand side is non-decreasing in C, while the other side is non-increasing in C. Since X satisfies the above inequality for $C = C_1$, X will also satisfy it for $C = C_2$. Hence, $X \in \mathcal{T}(C_2) \cap \mathbb{X}^*_\delta$. ∎

Example 2 The sets $\mathcal{T}(C_1) \cap X^*_\delta$ and $\mathcal{T}(C_2) \cap X^*_\delta$ in Proposition 5 do not need to be equal. For instance, let X' be the sequence in which $\pi^e_t = \pi_t = \hat{\pi}$ for t even and $\pi^e_t = \pi_t = \pi^*$ for t odd. Assume that $\delta = 0.62$. Therefore, $\delta/(1-\delta^2) \cong 1.0071 > 1$. Thus, (17) holds, which implies that $X' \in X^*_\delta$. Consider the expression

$$\sum_{r=t+1}^{\infty} \delta^{r-t} \left\{ W(\pi^e_r, \pi_r) - W(\hat{\pi}, \hat{\pi}) + [1 - I(\pi_r)]C \right\} \geq \tag{26}$$
$$\max \left\{ W(\pi^e_t, \pi^*), W(\pi^e_t, f(\pi^e_t)) - I(f(\pi^e_t))C \right\} - W(\pi^e_t, \pi_t) + I(\pi_t)C,$$

which is equivalent to (15). The term inside the curly brackets on the left-hand side is positive for r odd and equal to 0 for r even. Therefore, the left-hand side of the inequality is positive. Concerning the right-hand side, the inequality $C \leq k_2$ implies it is equal to 0 for t even. Thus, (26) holds for all even dates. For t odd, the right-hand side is equal to $\max\{0, k_1 - C\}$. Next, assume that $\alpha = 0.5$, $\gamma = \mu = \bar{y} = 1$, and $\pi^* = 0.01$. Hence, $\hat{\pi} = 0.5$ and $k_1 = 0.19208$. The table below contains the results of the numerical evaluation, for odd dates, of both sides of (26) for two distinct values of C.

Numerical evaluation of the left- and
the right-hand sides of expression (26) for t odd

C	left-hand side	right-hand side
0.02	0.16853	0.17208
0.10	0.21849	0.09208

Thus, $X' \notin \mathcal{T}(0.02)$ and $X' \in \mathcal{T}(0.10)$; as a consequence, $\mathcal{T}(0.02) \cap \mathbb{X}^*_\delta \neq \mathcal{T}(0.10) \cap \mathbb{X}^*_\delta$. ∎

Lemma 4 *There exists a real number $\theta > 0$ with the property that $\mathcal{T}(C) \subset \mathcal{T}(k_2)$ for every $C \in (k_2, k_2 + \theta]$.*

Proof. Define θ so that

$$\theta = \frac{\delta}{1-\delta}[W(\pi^*, \pi^*) - W(\hat{\pi}, \hat{\pi}) + k_2]. \tag{27}$$

Now, suppose that C belongs to $(k_2, k_2 + \theta]$ and let X be any element of $\mathcal{T}(C)$. Thus, (16) implies that

$$\sum_{r=t}^{\infty} \delta^{r-t}[W(\pi^e_r, \pi_r) - I(\pi_r)C] \geq \max_{\pi \in \Pi}[W(\pi^e_t, \pi) - I(\pi)C] + \delta \frac{W(\pi^*, \pi^*)}{1-\delta}.$$

As a consequence,

$$\sum_{r=t}^{\infty} \delta^{r-t}[W(\pi_r^e, \pi_r) - I(\pi_r)k_2] \geq \max_{\pi \in \Pi}[W(\pi_t^e, \pi) - I(\pi)(k_2 + \theta)] + \delta \frac{W(\pi^*, \pi^*)}{1-\delta}.$$

Combine this inequality with (27) to conclude that

$$\sum_{r=t}^{\infty} \delta^{r-t}[W(\pi_r^e, \pi_r) - I(\pi_r)k_2] \geq \max_{\pi \in \Pi}[W(\pi_t^e, \pi) - I(\pi)(k_2 + \theta)] + \theta + \delta \frac{W(\hat{\pi}, \hat{\pi}) - k_2}{1-\delta}. \quad (28)$$

On the other hand,

$$\begin{aligned} \max_{\pi \in \Pi}[W(\pi_t^e, \pi) - I(\pi)(k_2 + \theta)] + \theta &= \max_{\pi \in \Pi}\{W(\pi_t^e, \pi) - I(\pi)k_2 + [1 - I(\pi)]\theta\} \Rightarrow \\ \max_{\pi \in \Pi}[W(\pi_t^e, \pi) - I(\pi)(k_2 + \theta)] + \theta &\geq \max_{\pi \in \Pi}[W(\pi_t^e, \pi) - I(\pi)k_2]. \end{aligned}$$

Together, the last inequality and (28) imply that

$$\sum_{r=t}^{\infty} \delta^{r-t}[W(\pi_r^e, \pi_r) - I(\pi_r)k_2] \geq \max_{\pi \in \Pi}[W(\pi_t^e, \pi) - I(\pi)k_2] + \delta \frac{W(\hat{\pi}, \hat{\pi}) - k_2}{1-\delta}.$$

Therefore, X satisfies (15) when the penalty is equal to k_2. Since $X \in \mathcal{T}(C)$, X also satisfies (14). Therefore, $X \in \mathcal{T}(k_2)$, from which it follows that $\mathcal{T}(C) \subseteq \mathcal{T}(k_2)$.

It remains to show that $\mathcal{T}(C) \neq \mathcal{T}(k_2)$. Clearly, $\hat{X} \in \mathcal{T}(k_2)$. Moreover, the inequality $C > k_2$ implies that $W(\hat{\pi}, \hat{\pi}) - C < W(\hat{\pi}, \pi^*)$. Thus,

$$\frac{W(\hat{\pi}, \hat{\pi}) - C}{1-\delta} = W(\hat{\pi}, \hat{\pi}) - C + \delta \frac{W(\hat{\pi}, \hat{\pi}) - C}{1-\delta} < W(\hat{\pi}, \pi^*) + \delta \frac{W(\pi^*, \pi^*)}{1-\delta}.$$

Therefore, $\hat{X}$ does not satisfy (16). As a consequence, $\hat{X} \notin \mathcal{T}(C)$. ∎

For future reference, define the function $R(\pi^e, \pi, C)$ according to

$$R(\pi^e, \pi, C) = \max\{W(\pi^e, \pi^*), W(\pi^e, f(\pi^e)) - C\} - [W(\pi^e, \pi) - I(\pi)C]. \quad (29)$$

Therefore, for sequences in which $\pi_t^e = \pi_t$ for all t, inequality (16) is equivalent to

$$\sum_{r=t+1}^{\infty} \delta^{r-t}\{[W(\pi_r, \pi_r) - I(\pi_r)C] - W(\pi^*, \pi^*)\} \geq R(\pi_t, \pi_t, C). \quad (30)$$

Lemma 5 *Suppose that $C > k_2$. Thus, $R(\pi^*, \pi^*, C) = 0$, $R(\pi, \pi, C) > 0$ for $\pi \neq \pi^*$, and $R(\pi, \pi, C)$ is non-decreasing in C.*

Proof. Take any $C > k_2$. Since $C > k_1$, $W(\pi^*, \pi^*) > W(\pi^*, f(\pi^*)) - C$. Combine this inequality with (29) to conclude that $R(\pi^*, \pi^*, C) = 0$. Next, observe that

$$\pi \neq \pi^* \Rightarrow R(\pi, \pi, C) = \max\{W(\pi, \pi^*) - W(\pi, \pi) + C, W(\pi, f(\pi)) - W(\pi, \pi)\}. \tag{31}$$

Since $W(\hat{\pi}, \pi^*) - W(\hat{\pi}, \hat{\pi}) + C = C - k_2$, $R(\hat{\pi}, \hat{\pi}, C) \geq C - k_2 > 0$. Therefore, $R(\pi, \pi, C) > 0$ if $\pi = \hat{\pi}$. On the other hand, if π is different from both $\hat{\pi}$ and π^*, then $f(\pi) \neq \pi$ and, as a consequence, $W(\pi, f(\pi)) - W(\pi, \pi) > 0$. I conclude again that $R(\pi, \pi, C) > 0$. Finally, an inspection of the equality in (31) establishes that $R(\pi, \pi, C)$ is non-decreasing in C whenever $\pi \neq \pi^*$. Since $R(\pi^*, \pi^*, C) = 0$ for all C, it must be the case that $R(\pi, \pi, C)$ is non-decreasing in C for all π. ∎

Proof of Proposition 6. Take two penalties C_1 and C_2 such that $C_2 > C_1 > k_2$. Let X be an element of $\mathcal{T}(C_2)$. Thus, X satisfies (30) for $C = C_2$. Now, observe that the left-hand side of (30) is non-increasing in C, while Lemma 5 implies that its right-hand side is non-decreasing in C. Therefore, X also satisfies (30) for $C = C_1$. Hence, $X \in \mathcal{T}(C_1)$, and as a consequence, $\mathcal{T}(C_2) \subseteq \mathcal{T}(C_1)$.

It remains to show that $\mathcal{T}(C_1) \subset \mathcal{T}(k_2)$. The previous conclusion implies that the inclusion $\mathcal{T}(C_1) \subseteq \mathcal{T}(C_0)$ holds for every $C_0 \in (k_2, C_1)$. By making C_0 sufficiently close to k_2, it is possible to apply Lemma 4 to conclude that $\mathcal{T}(C_0) \subset \mathcal{T}(k_2)$. Thus, $\mathcal{T}(C_1) \subset \mathcal{T}(k_2)$. ∎

Proof of Proposition 7. Take penalty $C > k_2$. The continuity of W implies that there exists a $\varepsilon > 0$ such that

$$|\pi - \pi^*| < \varepsilon \Rightarrow |W(\pi, \pi) - W(\pi^*, \pi^*)| < C. \tag{32}$$

Clearly, ε does not depend on δ. Now, observe that

$$\begin{aligned} |W(\pi, \pi) - W(\pi^*, \pi^*)| < C \Rightarrow W(\pi, \pi) - C - W(\pi^*, \pi^*) < 0 \Rightarrow \\ [W(\pi, \pi) - I(\pi)C] - W(\pi^*, \pi^*) \leq 0. \end{aligned} \tag{33}$$

Next, take a sequence $X \neq X^*$ with the property $|\pi_r - \pi^*| < \varepsilon$ for all r and let t be the first date on which $\pi_t \neq \pi^*$. Inequality (33) implies the sum in the left-hand side of (30) is smaller than or equal to 0. On the other hand, the fact that $\pi_t \neq \pi^*$ combined with Lemma 5 implies that $R(\pi_t, \pi_t, C) > 0$. Thus, X does not satisfy (30), so $X \notin \mathcal{T}(C)$. ∎

Proof of Proposition 8. Take any C that satisfies the stated conditions. Since $C > k_2$, $X^* \in \mathcal{T}(C)$. Next, take a sequence $X \neq X^*$ that satisfies (14). It suffices to show that X does not satisfy (30). Let t be the first date on which $\pi_t \neq \pi^*$. Apply Lemma 5 to conclude that

$R(\pi_t, \pi_t, C) > 0$. Consider now the left-hand side of (30). If $\pi_r = \pi^*$, then the term inside the curly brackets is equal to 0. Suppose now that $\pi_r \neq \pi^*$. Since $W(0,0) \geq W(\pi_r, \pi_r)$,

$$W(\pi_r, \pi_r) - W(\pi^*, \pi^*) \leq W(0,0) - W(\pi^*, \pi^*) = k_3 \Rightarrow W(\pi_r, \pi_r) - W(\pi^*, \pi^*) \leq C.$$

Hence, the term inside the curly brackets is nonpositive if $\pi_r \neq \pi^*$. As a consequence, the sum on the left-hand side of (30) is smaller than or equal to 0. Since $R(\pi_t, \pi_t, C) > 0$, it follows that X does not satisfy (30). ∎

Proof of Proposition 9. Define $\bar{R}$ according to

$$\bar{R} = \max_{\pi \in \Pi} \left[\max\{W(\pi, \pi^*) - W(\pi, \pi) + C, W(\pi, f(\pi)) - W(\pi, \pi)\} \right]. \tag{34}$$

Since the objective function is continuous and Π is compact, $\bar{R}$ is well defined. Furthermore, $W(\pi^*, f(\pi^*)) - W(\pi^*, \pi^*) > 0$, so $\bar{R} > 0$. Next, define δ_C so that

$$\delta_C = \frac{\bar{R}}{[W(\pi_C, \pi_C) - C - W(\pi^*, \pi^*)] + \bar{R}}.$$

Combine (19) with the fact $\bar{R} > 0$ to conclude that $\delta_C \in (0,1)$. Now, take an $X \in \mathbb{X}(\pi_C)$ and let t be any date. Since $\pi_r \leq \pi_C < \pi^*$, $W(\pi_r, \pi_r) \geq W(\pi_C, \pi_C)$ and $I(\pi_r)C = C$. Hence,

$$W(\pi_r, \pi_r) - I(\pi_r)C - W(\pi^*, \pi^*) \geq W(\pi_C, \pi_C) - C - W(\pi^*, \pi^*).$$

Thus, if $\delta \in [\delta_C, 1)$, then

$$\sum_{r=t+1}^{\infty} \delta^{r-t}[W(\pi_r, \pi_r) - I(\pi_r)C - W(\pi^*, \pi^*)] \geq \frac{\delta_C}{1-\delta_C}[W(\pi_C, \pi_C) - C - W(\pi^*, \pi^*)] = \bar{R}. \tag{35}$$

Now, compare the equality in (31) with the objective function in (34). Since $\pi_t \neq \pi^*$, it must be the case that $\bar{R} \geq R(\pi_t, \pi_t, C)$. Together with (35), this last inequality implies that X satisfies (30), and as a consequence, $X \in \mathcal{T}(C)$. ∎

Proof of Proposition 10. Given a penalty $C \in (k_2, k_3)$, take any ε as in the statement of Proposition 7 and let Π_ε be the set $\{\pi \in \Pi : |\pi - \pi^*| \geq \varepsilon\}$. Since $C < k_3$, (18) and (32) together imply that $0 \in \Pi_\varepsilon$. Therefore, Π_ε is not empty. Next, consider the problem of selecting $\pi \in \Pi_\varepsilon$ to minimize $R(\pi, \pi, C)$. Since $\pi^* \notin \Pi_\varepsilon$, the objective function is continuous in Π_ε. Thus, the compactness of this set implies that there is a solution π_ε. Clearly, $\pi_\varepsilon \neq \pi^*$;

hence, Lemma 5 implies that $R(\pi_\varepsilon, \pi_\varepsilon, C) > 0$. Now, define δ'_C according to

$$\delta'_C = \frac{R(\pi_\varepsilon, \pi_\varepsilon, C)}{(k_3 - C) + R(\pi_\varepsilon, \pi_\varepsilon, C)}.$$

Observe that $\delta'_C \in (0,1)$. Now, take a sequence $X \neq X^*$ that satisfies (14). Apply Proposition 7 to conclude that if $|\pi_t - \pi^*| < \varepsilon$ for all t, then $X \notin \mathcal{T}(C)$. Suppose now that $|\pi_t - \pi^*| \geq \varepsilon$ for some date t. Since $W(\pi_r, \pi_r) \leq W(0,0)$,

$$W(\pi_r, \pi_r) - C - W(\pi^*, \pi^*) \leq W(0,0) - C - W(\pi^*, \pi^*) = k_3 - C \Rightarrow$$
$$W(\pi_r, \pi_r) - I(\pi_r)C - W(\pi^*, \pi^*) \leq k_3 - C.$$

Hence, if $\delta \in (0, \delta'_C)$, then

$$\sum_{r=t+1}^{\infty} \delta^{r-t}[W(\pi_r, \pi_r) - I(\pi_r)C - W(\pi^*, \pi^*)] < \frac{\delta'_C}{1 - \delta'_C}(k_3 - C) = R(\pi_\varepsilon, \pi_\varepsilon, C).$$

Furthermore, $\pi_t \in \Pi_\varepsilon$, so $R(\pi_\varepsilon, \pi_\varepsilon, C) \leq R(\pi_t, \pi_t, C)$. Therefore, X does not satisfy (30) and this implies that $X \notin \mathcal{T}(C)$. ∎

References

Adão, Bernardino, Isabel Correia, and Pedro Teles (2011). Unique Monetary Equilibria with Interest Rate Rules. *Review of Economic Dynamics* 14, 432–442.

Araujo, Aloisio, Tiago Berriel, and Rafael Santos (2016). Inflation Targeting with Imperfect Information. *International Economic Review* 57, 255–269.

Atkeson, Andrew, V.V. Chari, and Patrick J. Kehoe (2010). Sophisticated Monetary Policies. *Quarterly Journal of Economics* 125, 47–89.

Barro, Robert J., and David B. Gordon (1983a). A Positive Theory of Monetary Policy in a Natural Rate Model. *Journal of Political Economy* 91, 589–610.

Barro, Robert J., and David B. Gordon (1983b). Rules, Discretion and Reputation in a Model of Monetary Policy. *Journal of Monetary Economics* 12, 101–121.

Bernanke, Ben S., Thomas Laubach, Frederic S. Mishkin, and Adam S. Posen (1999). *Inflation Targeting: Lessons from the International Experience.* Princeton, Princeton University Press.

Capistrán, Carlos, and Manuel Ramos-Francia (2010). Does Inflation Targeting Affect the Dispersion of Inflation Expectations?. *Journal of Money, Credit and Banking* 42, 113–134.

Calvo, Guillermo A. On the Time Consistency of Optimal Policy in a Monetary Economy. *Econometrica* 46, 1411–1428.

Chari, V.V., and Patrick J. Kehoe (1990). Sustainable plans. *Journal of Political Economy* 98, 783–802.

Cole, Harold L., and Kehoe, Timothy J. (1996). A Self-Fulfilling Model of Mexico's 1994-1995 Debt Crisis. *Journal of International Economics* 41, 309–330.

Cole, Harold L., and Kehoe, Timothy J. (2000). Self-Fulfilling Debt Crises. *Review of Economic Studies* 67, 91–116.

Drazen, Allan (2000). *Political Economy in Macroeconomics.* Princeton, Princeton University Press.

Gibbons, Robert (1992). *Game Theory for Applied Economists.* Princeton, Princeton University Press.

Gürkaynak, Refet S., Andrew Levin, and Eric Swanson (2010). Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from the U.S., UK, and Sweden. *Journal of the European Economic Association* 8, 1208–1242.

Jia, Chengcheng, and Jing Cynthia Wu (2023). Average Inflation Targeting: Time Inconsistency and Ambiguous Communication. *Journal of Monetary Economics* 138, 69–86.

Kydland, Finn E., and Edward C. Prescott (1977). Rules Rather than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy* 85, 473–492.

Mishkin, Frederic S., and Niklas J. Westelius (2008). Inflation Band Targeting and Optimal Inflation Contracts. *Journal of Money, Credit and Banking* 40, 557–582.

Obstfeld, Maurice (1995). The Logic of Currency Crises. In: Eichengreen, Barry, Jeffry Frieden, and Jürgen Hagen (eds.) *Monetary and Fiscal Policy in an Integrated Europe.* Heidelberg, Springer.

Obstfeld, Maurice, and Kenneth Rogoff (1983). Speculative Hyperinflations in Maximizing Models: Can We Rule Them Out? *Journal of Political Economy* 91, 675–687.

Pastine, Ivan (2000). Devaluation of Fixed Exchange Rates: Optimal Strategy in the Presence of Speculation. *Economic Theory* 15, 631–661.

Persson, Torsten, and Guido Tabellini (1993). Designing Institutions for Monetary Stability. *Carnegie-Rochester Conference Series on Public Policy* 39, 53–84.

Rebelo, Sergio, and Carlos A. Végh (2008). When Is It Optimal to Abandon a Fixed Exchange Rate?. *Review of Economic Studies* 75, 929–955.

Rogoff, Kenneth (1985). The Optimal Degree of Commitment to an Intermediate Monetary Target. *Quarterly Journal of Economics* 100, 1169–1189.

Sargent, Thomas J. (1987). *Dynamic Macroeconomic Theory.* Cambridge, Harvard University Press.

Sargent, Thomas J., and Neil Wallace (1975). "Rational" Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule. *Journal of Political Economy* 83, 241–254.

Smith, Bruce D. (1994) Efficiency and Determinacy of Equilibrium Under Inflation Targeting. *Economic Theory* 4, 327–344.

Svensson, Lars E.O. (1997). Optimal Inflation Targets, "Conservative" Central Banks, and Linear Inflation Contracts. *American Economic Review* 87, 98–114.

Svensson, Lars E.O. (1999). Inflation Targeting as a Monetary Policy Rule. *Journal of Monetary Economics* 50, 607–654.

Svensson, Lars E.O. (2011). Inflation Targeting. In: Friedman, Benjamin M., and Michael Woodford (eds.) *Handbook of Monetary Economics*, Volume 3B. Amsterdam, Elsevier.

Walsh, Carl E. (1995). Optimal Contracts for Central Bankers. *American Economic Review* 85, 150–167.

Walsh, Carl E. (2009). Inflation Targeting: What Have We Learned? *International Finance* 12, 195–233.

Woodford, Michael (1994). Monetary Policy and Price Level Determinacy in a Cash-in-Advance Economy. *Economic Theory* 4, 345–380.