SciencesPo
DEPARTMENT OF ECONOMICS

Discussion paper 2017-01

CREDIBILITY AND MONETARY POLICY

Jean Barthélemy
Eric Mengus

Sciences Po Economics Discussion Papers

Credibility and Monetary Policy*

Jean Barthélemy   Eric Mengus

May 2, 2017
First draft: October 2016

PRELIMINARY

Abstract

This paper revisits the ability of central banks to manage private sector's expectations depending on its credibility and how this affects the use of interest rate rules and pegs to achieve monetary policy objectives. When private agents can only provide limited incentives for the central bank to follow a policy, we show that resulting limited credibility allows a central bank to prevents the inflation from diverging by defaulting on past promises if necessary. As a result, the Taylor rule, when expected, anchors inflation expectations on a unique equilibrium path as long as the Taylor principle is satisfied. Finally, we also show that limited credibility restricts the impact of long-term interest rate pegs, so as to make current conditions less dependent on future policy changes.

Keywords: Taylor principle, credibility, Forward Guidance.
JEL Classification: E31, E52, E65.

*Barthélemy: Sciences-Po, Department of Economics, 28 rue des Saints Pères 75007 Paris, France and Banque de France, 31 rue Croix des Petits Champs, 75001 Paris cedex, France. Email: jean.barthelemy@sciencespo.fr. Mengus: HEC Paris, Economics and Decision Sciences Department, 1 rue de la Libération, 78350 Jouy-en-Josas, France. Email: mengus@hec.fr. We thank Marco Bassetto, Gaetano Gaballo, Guillaume Plantin, Franck Portier, Eduardo Perez and Xavier Ragot for helpful comments as well as seminar participants at Sciences Po and 2017 T2M (Lisbon). All remaining errors are ours.

1 Introduction

To stabilize the economy and guarantee price stability, the central bank needs to be credible about its future moves to steer private sector's expectations (Barro and Gordon, 1983) and also to anchor these expectations on a single path. To this purpose, central bank may adhere to interest-rate rules as argued by Kydland and Prescott (1977) and Taylor (1993). Such rules however allow self-fulfilling diverging inflation paths as noted by Woodford (2003) among others, and therefore fail to anchor private sector's expectations.[1] In addition, more sophisticated rules that restore the anchoring of private expectations may not necessarily be credible (Cochrane, 2011).[2]

In this paper, we investigate the implementation of monetary policy when the central bank's credibility to follow a policy is potentially limited. In our framework, the central bank's credibility stems from the incentives resulting from private agents' reactions in case of a policy deviation, as, for example, in Chari and Kehoe (1990). As a result, depending on private agents' punishment scheme, the central bank's credibility may be limited.[3] Our key finding is that, when private agents' reactions only leads to limited credibility, the central bank ultimately prefers inflation stabilization over its previously announced policy plan if inflation diverges. In particular, if inflation diverges, the central bank is always better off by deviating thus forcing inflation to be locally bounded in equilibrium. As a result, we obtain that, paradoxically, limited credibility ensures the anchoring of inflation expectations in the case where the central bank is expected to follow the Taylor rule satisfying the Taylor principle. However, limited credibility also allows for alternative equilibria, where the central banker is not expected to follow a Taylor rule, thus suggesting that the central banker cannot credibly rule out the multiplicity of equilibria simply by promising to follow a policy.

In addition, we show that limited credibility has implications on unconventional monetary policy as it also constrains the outcomes of an interest rate peg – as the ones advocated by forward guidance policy – on current inflation. We show that under limited credibility only short-term and permanent pegs are credible, the first result echoing Friedman (1968). This result implies that long-term but not permanent announced policy changes, being not

[1] This happens even when the Taylor principle is satisfied (Taylor, 1999; Clarida et al., 1999).

[2] See Loisel (2009) or Atkeson et al. (2010) among others for examples of more sophisticated rules or policies. Cochrane (2011) argues that these policies usually rely on the time-inconsistent commitment to "blow up the economy".

[3] This parallels the literature on sovereign debt that has analyzed the sustainability of sovereign debt depending on punishment schemes (see Eaton and Gersovitz, 1981; Bulow and Rogoff, 1989; Kletzer and Wright, 2000).

credible, do not have a contemporaneous effect on macroeconomic variables. We also show that there is no "neo-Fisherian paradox" under limited credibility as there is no continuity between short-run inflationary interest rate pegs as used by forward guidance and potentially deflationary permanent pegs.[4]

Importantly, these two effects of limited credibility arise because of the restrictions implied by limited credibility on out-of-equilibrium outcomes[5] and our results hold even when limited credibility does not constrain the optimal allocation. In contrast, when central bank is perfectly credible, the Taylor rule cannot anchor inflation expectations –there are multiple equilibria even under the Taylor principle– and the "neo-Fisherian paradox" arises for long-term interest rate pegs.

But when is credibility limited and how is it connected to policy decisions? We identify that credibility is limited when the cost of losing credibility on its ability to manage future expectations – which results from the deviation from past policy promises – is smaller than the cost of hyperinflation – that is the cost of losing the credibility of money itself.[6] In that case, when inflation is costly from the central bank's point of view, there is an incentive for the central bank to deviate from the inflationary policy, even at the cost of losing reputation. As both the inability of the Taylor rule to anchor inflation expectations and the long term inflationary effect of long interest rate pegs both rely on diverging inflation paths, there exists in both cases a point after which inflation becomes sufficiently large so that this incentive triggers a policy deviation. In addition, anticipating this future policy move, agents adjust their expectations, forcing the central bank to deviate from its policy plan in advance.[7]

We first illustrate these mechanisms in a simple flexible price model satisfying a Fisher equation in the case of the Taylor rule (Section 2). We assume that a time-inconsistent central bank cannot commit to adhere to a policy plan but has reputation concerns. We then characterize the set of sustainable equilibria as in Chari and Kehoe (1990) and study the sustainability of Taylor rule-type policies. This allows us to study how limited credibility affects the implementation of a particular allocation using a policy plan, by verifying that off-

[4]When extended infinitely, a commitment to low interest rate may push current inflation to explode. This result has been found to be paradoxical as an infinite peg may also have the effect to decrease inflation at lower but bounded level (the "neo-Fisherian effect") and, in addition, this also leads to the counter-intuitive outcome that current inflation is highly sensitive to policy changes in the very far future.

[5]From that perspective, our work echoes Bassetto (2005), where he studies how physical constraints restrict policy strategies and modify the implementation of a particular policy.

[6]In terms of empirical counterparts, periods of high inflation as the US Great Inflation can be interpreted as periods of lost credibility on monetary policy promises but they were not periods of hyperinflation.

[7]This is consistent with Cochrane (2011)'s comment that deviations are anticipated by private agents and that there cannot be two different policies implemented at the same time.

equilibrium reactions that guarantee such an implementation are credible. We show that in this model diverging inflation path are not sustainable and there is a unique paths of private agents' beliefs that are consistent with expecting a Taylor rule.

We then extend our reasoning to a standard New Keynesian model (Section 3). We start our investigations in perfect foresight. Our main finding is that the worst sustainable equilibrium leads to an infinite loss for the central banker, thus making the central banker fully credible to implement any feasible allocation. We then confirm that full credibility makes Taylor rules satisfying the Taylor principle insufficient to ensure well-anchored inflation expectations, as there exist multiple equilibria. But, this also means that any off-equilibrium actions that can curb diverging inflation expectations are also sustainable,[8] thus potentially guaranteeing well-anchored inflation expectations. Finally, full credibility has not bite on interest rate pegs that may have infinite effects on current inflation. This leads the neo-Fisherian paradox to hold and there is no restrictions on the contemporaneous effects of future policy changes.

This motivates to turn to situations where the central bank's credibility is limited (Section 4). To this purpose, we focus on equilibria where the private sector coordinates on milder punishment schemes that leads to a bounded welfare loss.[9] As a result, we first show that inflation remains bounded across all equilibria. The main reason is that unbounded inflation would lead to infinite costs, whereas the central bank can deviate and implement a better outcome by optimally setting its policy at each point in time.

We then apply this result to study the set of equilibria where the central bank follows a Taylor rule. When the Taylor principle holds, this set is only composed of one equilibrium – the zero-inflation equilibrium – in which the Taylor rule is *always* followed. Furthermore, we show that there are no equilibria where the Taylor rule is followed for some periods before shifting to a discretionary policy when inflation starts to rise: the value for the central banker to follow a rule disappears when private agents anticipate the shift to another policy in the future.

Yet, we also show that there exist alternative equilibria, starting with the time-consistent

[8] See Loisel (2009) for such a strategy using solely the interest rate instrument. See Cochrane (2011) for a description and discussion of alternative solutions external to the New-Keynesian model.

[9] In our benchmark case, we focus on the time-invariant discretionary equilibrium. This assumption about the punishment is widespread in the literature: see Kurozumi (2008), Loisel (2008), Nakata (2014) among others. In Appendix A, we provide a game-theoretic foundation of such a punishment scheme. Our results are robust to considering punishments where agents coordinate on the permanent liquidity trap equilibrium described by Benhabib et al. (2001). See also Kletzer and Wright (2000) for the investigation of private agents' coordination on punishment schemes in the context of sovereign debt.

equilibrium in which the central bank re-optimizes its decisions at every date. Such a multiplicity does not stem for the Taylor rule itself: we show that, in none of these alternative equilibria, private agents expect the central bank to follow a Taylor rule and that these equilibria would arise with any other rule. This simply results from the combination of rational expectation by private agents and sequential rationality for the central banker: whatever the intended rule, if private agents anticipate, for example, that the central banker discretionarily re-optimizes its policy at every date, there is no value for the central banker to follow the intended rule and he is forced to discretion.[10]

Turning to forward guidance and interest rate pegs, we show that limited credibility imposes an upper bound on the maximal duration of the peg, except if the peg is implemented infinitely, in which case, inflation adjusts downward following the "neo-Fisherian effect". This finding has two implications: marginal policy changes in the future do not have large contemporaneous effects, as they are simply not credible and there is no "neo-Fisherian" paradox as short-term and infinite pegs are disjointed sets of equilibria.

We then investigate the consequences of limited credibility in the presence of fundamental or non-fundamental shocks.

First, in Section 5.1, we show that limited credibility still allow for many equilibria when the Taylor principle is not satisfied. Yet, we show that limited credibility restricts this set of equilibria and, thus, limits the potential fluctuations of inflation. More precisely, we show that fluctuations boundaries are connected to the time-consistency of the central bank's preferences, thus bridging the identification of non-fundamental shocks in Clarida et al. (2000) with the standard view on inflation bias and time-consistency as in Barro and Gordon (1983).

These results lead to a reinterpretation of the Taylor principle: when the Taylor principle holds, it forces any out-of-equilibrium expectations to trigger a change in policy by the central bank. In our model, this takes the form of an immediate shift to discretion. In contrast, when the Taylor principle is not satisfied, multiple expectations can emerge without modifying the incentive to follow through on the Taylor rule, thus allowing for multiplicity.

Second, in Section 5.2, we extend our reasoning to a stochastic environment where the economy is hit by fundamental shocks. The presence of shocks do not change our conclusions, in particular when the effects of shocks are stabilized by the Taylor rule or when shocks are bounded.[11] The only difference is that, when the Taylor rule is suboptimal or do not perfectly

[10]For example, in Atkeson et al. (2010), the central bank can commit to off-equilibrium deviations (See Section 5.4 for a more detailed discussion).

[11]Our conclusions are also robust to standard forms of unbounded shocks as standard autoregressive processes.

stabilize the effects of shocks, the equilibrium policy may feature shifts to discretionary policy. We show that these shifts are in fact stabilizing: a relaxed version of the Taylor principle holds in that case.

Third, we discuss how are results are affected by the presence of a Zero Lower Bound (ZLB) (Section 5.3), in which case the private agents can coordinate on the permanent liquidity trap equilibrium as described by Benhabib et al. (2001) to punish deviations.

In the end, we argue that limited credibility is key to understand monetary policy ability to manage inflation. Taylor rules are not *per se* producing multiplicity, but multiplicity is the outcome of the agents' perception of future central bank's policies: if agents believe in the implementation of the Taylor rule, they will coordinate their expectations on a unique path of inflation. However, what ensures the anchoring of inflation expectations is only partially Taylor rules but also the expectation that the central bank eventually cares more about stabilizing the economy than following its rule. From the central bank's perspective, this then requires not to blindly follow a Taylor rule,[12] but to build up the reputation to be willing to stabilize inflation. In terms of central banks' communication, this then implies that the central bank emphasizes this willingness to stabilize inflation in addition to their desire to stick to a Taylor rule, in line with the communication of many central banks that have associated the use of rules with the objective of price stability. This also means that credible monetary policy is sufficient to ensure well-anchored inflation expectations, even in the absence of fiscal backing and that more complex rules or monetary procedures than the Taylor rule are not necessary to ensure price stability (see Section 5.4 for further discussions).

Literature review Our paper is connected to several strands of the literature.

The existence of multiple equilibria consistent with the Taylor rule and the Taylor principle is well known (Woodford, 2003; Cochrane, 2011). The literature has answered to this problem in a number of ways.

First, Obstfeld and Rogoff (1983), Obstfeld and Rogoff (1986) or Atkeson et al. (2010) propose to switch from an interest-rate rule to another monetary policy (money growth rule, commodity standard, and so on) in case of excessive inflation due to the interest-rate rule. However, as noticed by Cochrane (2011), this switch is not necessarily time-consistent. In

[12] This contrasts with the policy recommendation to legally constrain monetary policy to follow a Taylor rule. It is then consistent with the recent legislation entitled *Requirements for Policy Rules of the Federal Open Market Committee*, Section 2 of the Fed Oversight Reform and Modernization Act (H.R. 3189), where the central bank can change its strategy when required by circumstances, in our case, diverging inflation expectations.

contrast, in our approach, we do not assume any form of commitment on central bank's strategies.

Second, one of the major alternative for active monetary policy is the fiscal theory of the price level (Leeper, 1991; Sims, 1994; Woodford, 1995; Sims, 2013), where the value of money is determined by future fiscal surpluses. Yet, to ensure determinacy, fiscal policy has to be active, thus requiring full credibility of future surpluses. Instead, our results emphasize that active monetary policy can be sufficient to ensure inflation determinacy, even in the absence of a commitment technology.

Third, Loisel (2009), Adao et al. (2011) or Hall and Reis (2016) have investigated alternative rules that respond to expectations. By responding to private agents' expectations, the central banker can ensure the existence of a unique allocation that does not depend on expectations. In the case of Hall and Reis (2016), the response to expectation goes through conducting the interest payments on central bank's reserves. However, these rules are not necessarily credible and will not be optimally followed through by the central banker, for example when economic agents expect the discretionary outcome. Thus, they are comparable to the Taylor rule under limited credibility without having the simplicity and transparency of the Taylor rule.

Fourth, Bullard and Mitra (2002) and McCallum (2009) show that the Taylor principle delivers a unique learnable equilibrium based on the concept of E-stability. This interpretation of the Taylor principle based on learnability of the bounded equilibrium has been challenged by Cochrane (2009).

Our paper is more generally connected to the literature on implementation and, in particular, to Bassetto (2005) (see also Bassetto, 2002). Using Bassetto (2005)'s wording, the Taylor rule is a strategy that the central bank can commit to following a "Schelling-timing". However, such a strategy is insufficient to ensure the unique implementation and we then consider the "Schelling-timing" in the absence of a commitment technology.

The "neo-Fisherian paradox" of the forward guidance is stated in García-Schmidt and Woodford (2015) or Cochrane (2016). The solutions to these proposals have been either to depart from rational expectations (Gabaix, 2016; García-Schmidt and Woodford, 2015) or use a different equilibrium selection (Cochrane, 2013). We instead show that future policy that triggers large suboptimal deviation in the short run are not credible under limited credibility and does not belong to any properly defined equilibrium.

As mentioned, our paper also parallels the literature on sovereign debt sustainability. Indeed, we identify that monetary policy implementation crucially depends on the ability by the private sector to punish policy deviations in a similar spirit as the sovereign debt literature identified the impact of punishment schemes for the sustainability of sovereign debt as in Eaton and Gersovitz (1981), Bulow and Rogoff (1989) or Kletzer and Wright (2000) for the impact of the time-consistency and the coordination of lenders in implementing punishment schemes.

The role of reputation for monetary policy was first introduced by Barro and Gordon (1983) to study the inflation bias stemming from the desire to maintain the unemployment rate below its natural rate. Other sources of time-inconsistency of monetary policy– that gives value to maintain the central bank's reputation – have been investigated as the stabilization bias. In addition, the inflation bias has been showed to arise due to multiple frictions: the steady-state distortion of monopolistic competition in the New Keynesian model and, among more recent contributions, the desire to redistribute wealth in the presence of nominal contracts in Nuno and Thomas (2017) or the inefficiently low steady state level of employment in the absence of private insurance as in Challe (2017).

This role of reputation has been investigated by Chang (1998) and Ireland (1997) who study the sustainability of the Friedman rule in monetary models.[13] Recent research has focused on the sustainability of the Ramsey allocation in the New-Keynesian model (see Kurozumi, 2008; Loisel, 2008; Sunakawa, 2015, among others) or Nakata (2014) in the presence of the ZLB. Their main result is that the Ramsey allocation is usually sustainable under plausible calibrations. Our main contribution with respect to this literature is that we focus on the ability of the central bank to coordinate private agents on a particular allocation using a policy, which requires to investigate the sustainability of the whole policy strategy and not only the sustainability of the desired allocation. Importantly, we show that limited credibility has an impact on the implementation of monetary policy, even in the case where limited credibility does not constrain the optimal allocation.

2 A simple model

In this section, we highlight the two key ingredients that underline most of this paper's results in the context of a simple flexible price model. First, limited credibility rules out infinite levels of inflation, second, the anticipation of future deviations makes immediate deviation desirable. The model features a Fisher equation linking private agents expectations and nominal interest rate and a (time-inconsistent) objective function for the central bank.

[13]See also Abreu (1988) or Abreu et al. (1990) on repeated games and Phelan and Stacchetti (2001) for the application to the Ramsey tax model. In our context, we are able to determine the worst sustainable equilibrium without investigating the full set of equilibria as in Abreu et al. (1990).

The environment We consider an infinite model of inflation determination with an optimizing central bank and private agents forming rational expectations. Each period is denoted by $t \in \{0, 1, ...\}$.

The central bank is in charge of maintaining price stability and the anchoring of long term inflation expectations and is free to choose the level of inflation at each date. At date t, given private agents' expectations about future inflation, the central bank seeks to maximize the objective function W_t over current inflation rate π_t:

$$W_t = -(\pi_t - \bar{\pi})^2 - \sum_{k>t} \beta^{k-t} (E_t \pi_k)^2,$$

where β is the central banker's discount factor and $\bar{\pi} > 0$ measures the intensity of the temptation to raise inflation in the short run. This objective functions is *ad hoc* but captures in a simple way the short-run trade-off between maintaining anchored inflation expectations at zero in the medium run and raising current inflation at higher level ($\bar{\pi}$).[14]

The private sector forms rational expectations about future inflation rates, hence *in equilibrium* $\pi_{t+1} = E_t \pi_{t+1}$. We also assume that, *in equilibrium*, a Fisher equation holds:

$$i_t = E_t \pi_{t+1} + r, \tag{1}$$

where i_t denotes the date-t nominal interest, $E_t \pi_{t+1}$ the date-t expectations formed by private agents about future inflation and r the risk-free real return.[15]

In the following, we are looking for situations where the central bank adheres to a Taylor rule of the form:

$$i_t = r + \phi \pi_t, \tag{2}$$

where ϕ is a positive scalar measuring the response of the nominal interest rate to inflation.

With a commitment technology As a benchmark, let us assume that the central bank has a commitment technology and decides at date 0 to commit to follow a Taylor rule forever. In this case, combining equations (1) and (2) yields the following condition:

$$\phi \pi_t = E_t \pi_{t+1}. \tag{3}$$

[14]In the standard New-Keynesian model, this inflation bias derives from the steady-state distortions of monopolistic competition. See also Nuno and Thomas (2017) or Challe (2017) for alternative micro-foundations of this bias in the presence of incomplete markets. Alternatively, the stabilization bias due to cost-push shocks may produce the same kind of time-inconsistency.

[15]This equation can be micro-funded by a no-arbitrage condition when, at each date, risk-neutral households can choose to hold government bonds yielding a risk-free nominal return i_t or real assets yielding a risk-free real return r. Without loss of generality, we assume that the return of the real asset is constant.

There is a continuum of solutions to that equation in perfect foresight indexed by the date-0 inflation rate π_0:

$$\pi_t = \phi^t \pi_0. \tag{4}$$

This means that multiple equilibria emerge from following a Taylor rule when this rule is always credible. In particular, the rule produces either zero inflation at any date or leads to diverging inflation paths when initial inflation is different from zero.

Equilibrium definition Let us now assume that the central bank does not have a commitment technology and that following a Taylor rule can only be the outcome of an equilibrium. In particular, let us define an equilibrium in this model as a sequence of current and expected inflation rates and nominal interest rates such that:

(i) At each date, given past histories and given private sector expectations, the central bank maximizes over its objective function.

(ii) Given past histories, private agents form rational expectations and equation (1) holds.

The constant inflation rate $\pi_t = \bar{\pi}$ is an equilibrium that corresponds to the so-called discretionary equilibrium in which the central bank maximizes its objective without internalizing its impact on private agents' expectations. In such an equilibrium, the Fisher equation implies that $i_t = \bar{\pi} + r$.

Limited credibility As agents' behaviors can be function of past histories, the potential set of equilibria can include more equilibria than the time-consistent discretionary one, due to the possibility of trigger strategies as in Chari and Kehoe (1990), among others. Let us focus on the trigger strategy, where agents stick to a policy and then shift to the discretionary equilibrium if there was any deviation in the past.[16]

As a result, the incentive to deviate at date t from a particular policy leading to an inflation path $\{\pi_k\}_{k\geq t}$ is, taking into account rational expectations:

$$(\pi_t - \bar{\pi})^2 + \sum_{k>t} \beta^{k-t} (\pi_k)^2 \leq \underbrace{0}_{\text{Deviation}} + \underbrace{\frac{\beta}{1-\beta}\bar{\pi}^2}_{\text{Shift to discretion}} . \tag{5}$$

The right hand side corresponds to the discounted sum of losses when deviating. In the period of the deviation, the central bank can freely set inflation and, thus, sets it at $\pi_t = \bar{\pi}$. In any

[16]In general focusing on a particular trigger strategy is only a sufficient condition to be an equilibrium. This is also a necessary condition if the "punishment" equilibrium is the worst sustainable equilibrium (see Abreu et al., 1990), which is the case of the discretionary equilibrium in this simple context.

future periods, the private sector expects the central bank to set inflation under discretion, leading inflation expectations to always equal $\bar{\pi}$.

Importantly, this discounted sum is bounded so that the incentive constraint for not deviating from a policy leads date-t loss – the left hand term of inequality (5) – to be bounded as well.

The incentives to follow a Taylor rule Let us now investigate how the credibility constraint (5) affects the implementation of a Taylor rule. To do so, let us substitute in (5), the values of inflation implied by following a Taylor rule:

$$(\phi^t \pi_0 - \bar{\pi})^2 + \sum_{k>t} \beta^{k-t} (\phi^k \pi_0)^2 \leq \frac{\beta}{1-\beta} \bar{\pi}^2. \tag{6}$$

To start with, in the case where initial inflation equals $\pi_0 = 0$, this inequality boils down to:

$$\bar{\pi}^2 \leq \frac{\beta}{1-\beta} \bar{\pi}^2, \tag{7}$$

which is satisfied when $\beta > 1/2$, which we assume from now on. As a result, there exists an equilibrium in which the Taylor rule is followed at every date and where inflation always equals 0.

But what about situations where inflation is initially different from 0? In that case, equation (6) is as follows:

$$(\phi^t \pi_0 - \bar{\pi})^2 + \sum_{k>t} \beta^{k-t} (\phi^k \pi_0)^2 = \infty, \text{ if } \beta\phi^2 \geq 1; \tag{8}$$

$$(\phi^t \pi_0 - \bar{\pi})^2 + \sum_{k>t} \beta^{k-t} (\phi^k \pi_0)^2 = (\phi^t \pi_0 - \bar{\pi})^2 + \phi^{2t} \frac{\beta\phi^2 \pi_0^2}{1-\beta\phi^2}, \text{ if } \beta\phi^2 < 1. \tag{9}$$

As a result, there are two situations to consider depending on the value of $\beta\phi^2$.

When $\beta\phi^2 \geq 1$, equation (5) is never satisfied as the right hand side of (5), $\frac{\beta}{1-\beta}\bar{\pi}^2$, is bounded, thus implying that there does not exist an equilibrium where the Taylor rule is followed and $\pi_0 \neq 0$.

When $\beta\phi^2 \leq 1$, equation (5) may be satisfied at least in the short run. However, as time goes by and t increases, inflation increases as well at the rate ϕ. When taking the limit when t goes to infinity, one obtains, when $\pi_0 \neq 0$:

$$\lim_{t\to\infty} (\phi^t \pi_0 - \bar{\pi})^2 + \phi^{2t} \frac{\beta\phi^2 \pi_0^2}{1-\beta\phi^2} = \infty. \tag{10}$$

Ultimately the loss to stick to the Taylor rule becomes large and exceeds the bounded loss obtained from deviating and sticking to discretion so that the inequality (5) does not hold in the long run and the central bank has an incentive to deviate at some point in the future.

In the end, when $\pi_0 \neq 0$ and under the Taylor rule, there always exists a date at which the central bank is better off deviating and, hence, the only equilibrium where the Taylor rule is always followed is when the inflation rate is perfectly stabilized at zero.

Expectations and switch If diverging inflation expectations cannot arise due to limited credibility constraints – the central bank preferring to deviate from its policy –, does this necessarily rule out multiple equilibria where the Taylor rule is followed, at least in the short run? Indeed, in the case where $\beta\phi^2 < 1$, the central bank does not immediately deviate from the Taylor rule but only in the long run when the inflation rate has become sufficiently large. One may then wonder whether there exist equilibria where the central bank follows the Taylor rule in the short run before switching to discretion?

To answer these questions, let us elaborate further about the off-equilibrium case where inflation becomes sufficiently large at some date t to trigger a deviation. In equilibrium, this deviation "has to be" anticipated by private agents, who adjust their expectations at previous dates. Let us illustrate how this affects the incentives to follow the Taylor rule at any previous period.

When deviating at period t, the central bank sets inflation at $\pi_t = \bar{\pi}$. In period $t-1$, under rational expectations, private agents perfectly anticipate that the level of inflation at date t will be $\bar{\pi}$. As, in equilibrium, the Fisher equation holds, the date-$t-1$ nominal interest rate i_{t-1} has to satisfy $i_{t-1} = \bar{\pi} + r$. Following the Taylor rule at that date would imply an inflation level $\phi\pi_{t-1} = \bar{\pi}$. The central bank then compares this welfare outcome resulting from sticking at date-$t-1$ to the Taylor rule with what it can do otherwise by re-maximizing over π_{t-1}, given that future inflation expectations are set anyway at $\bar{\pi}$:

$$(\bar{\pi}/\phi - \bar{\pi})^2 + \frac{\beta}{1-\beta}\bar{\pi}^2 \geq 0 + \frac{\beta}{1-\beta}\bar{\pi}^2, \tag{11}$$

when the Taylor principle ($\phi > 1$) holds. As a result, the central bank has no incentives to follow the Taylor rule in period $t-1$ as well: this would lead to an inflation rate of $\pi_{t-1} = \bar{\pi}/\phi$, while the best discretionary policy is $\pi_{t-1} = \bar{\pi}$ and there is no cost of deviating from the Taylor rule since private agents expectations are already anchored at $\bar{\pi}$ for next periods.

As a result, when expecting the central bank's future deviation, private agents reduce the current incentives for the central bank to adhere to a Taylor rule. This then implies that there does not exist an equilibrium outcome where the central bank follows a Taylor rule in the short run before switching to discretion when inflation increases too much. This comes from the very reason that central banks adhere to rules because they have a long term value.

In the end, there is only one equilibrium where the Taylor rule is always followed and this is when inflation is perfectly stabilized at zero. Note, however, that this does not rule out multiple equilibria – the time-consistent discretionary outcome is also an equilibrium– but in this alternative equilibria, another policy that the Taylor rule is expected.

This simple model illustrates the two important consequences of limited credibility. First, diverging inflation paths incentivize the central bank to deviate from its previous policy. Second, the expectation of deviations modifies the incentives to adhere to a rule in any previous periods. These two ingredients lead to restrict the set of equilibria where the Taylor rule is expected to be followed.

3 New-Keynesian model

In this section, we first describe the model, we then derive optimal monetary policies under commitment and discretion.

3.1 Model

The private sector behavior is captured by two equations:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa y_t, \tag{NKPC}$$

$$y_t = E_t y_{t+1} - \frac{1}{\sigma} \left(i_t - E_t \pi_{t+1} \right), \tag{EE}$$

where π_t denotes date-t inflation rate, y_t the output gap and i_t the nominal interest rate. Parameters σ, β and κ are positive constants. The operator E_t is the time-t conditional expectation operator. The New Keynesian Phillips curve (NKPC) describes the dynamics of prices. The Euler equation (EE) sums up the consumer's inter temporal consumption choice. We introduce shocks in Section 5.2.

Welfare We suppose that date-t social welfare, W_t, can be measured by the following quadratic loss function, L_t:[17]

$$W_t = -L_t \text{ and } L_t = \frac{1}{2} E_t \sum_{s=t}^{\infty} \beta^{s-t} \left[\pi_s^2 + \lambda \left(y_s - y^* \right)^2 \right], \tag{12}$$

where the scalar y^* is the desired level of the output gap and is referred to as the inflation bias and the scalar λ measures the weight of output gap fluctuations in the loss function.

[17] Microfoundation of such a loss can be found in Woodford (2003).

We crucially assume that the steady state of the economy is inefficient –the scalar y^* in the loss function measures this inefficiency– such that the optimal monetary policy is time-inconsistent. Otherwise, rules would not be better than discretion and therefore central banker would not need to use rules and monetary policy would be essentially trivial.

3.2 Optimal policies under commitment and discretion

In this subsection, we investigate the optimal policies when the (benevolent) central banker can commit and when he acts under discretion.

Commitment The optimal date-t monetary policy is to maximize welfare with respect to the entire sequence of nominal interest rate, $\{i_\tau\}_{\tau\geq t}$, given that the allocation $\{\pi_\tau, y_t\}_{\tau\geq t}$ solves equations (NKPC) and (EE).

Problem 1 (Ramsey problem)**.**

$$\max_{i_\tau, \tau\geq t} W_t$$

under the constraints (NKPC) and (EE).

The Lagrangian of this problem can be written as:

$$\mathcal{L} = E_t \sum_{\tau=t}^{\infty} \beta^\tau \left[\frac{1}{2}(\pi_\tau^2 + \lambda\left(y_\tau - y^*\right)^2) + \mu_\tau(\pi_\tau - \beta E_\tau \pi_{\tau+1} - \kappa y_\tau)\right.$$
$$\left. + \nu_t(y_\tau - E_\tau y_{\tau+1} + \frac{1}{\sigma}\left(i_\tau - E_\tau \pi_{\tau+1} - r_\tau^n\right))\right].$$

First order conditions lead to, for any $\tau \geq t$,

$$\pi_\tau + \mu_\tau - \mu_{\tau-1} = 0, \tag{13}$$

$$\lambda(y_\tau - y^*) - \kappa\mu_\tau = 0, \tag{14}$$

where μ_τ is the Lagrange multiplier associated to the constraints (NKPC) and the initial multiplier $\mu_{t-1} = 0$. The presence of the lagged Lagrange multiplier, $\mu_{\tau-1}$ proves that the optimal monetary policy is backward-looking and therefore time-inconsistent as the economy is purely forward-looking.

Zero-inflation allocation The above allocation is history-dependent. However, inflation and the output gap converge to zero. We call the asymptotic allocation the zero-inflation allocation. Woodford (2003) refers to it as the timeless-perspective allocation. Such an allocation is important since Taylor rules are supposed to implement it.

Discretion A discretionary allocation is an allocation that, at each period t, solves the following problem:

Problem 2 (Discretion)**.**

$$\max_{i_t} W_t$$

under the constraints (NKPC) and (EE).

First order conditions lead to:

$$\pi_t + \mu_t = 0, \tag{15}$$

$$\lambda(y_t - y^*) - \kappa\mu_t = 0, \tag{16}$$

where μ_t is, again, the Lagrange multiplier associated to the constraint (NKPC). The time-invariant discretionary allocation is the pair of constants $\{\pi, y\}$ solving the problem 2. The other discretionary allocations involve inflation paths that diverge toward $+\infty$ or $-\infty$ depending on an arbitrary initial condition π_0:

$$\pi_t - \pi = \left[\frac{1 + \frac{\kappa^2}{\lambda}}{\beta}\right]^t (\pi_0 - \pi) \tag{17}$$

3.3 Sustainable equilibrium

In this section, we first lay down some elements to define the set of sustainable equilibria. We then define the policy strategy that we call 'following a Taylor rule'. Finally, we characterize the set of sustainable equilibria in the New-Keynesian model defined in section 3 and discuss the implications for monetary policy.

Definition of sustainable equilibrium We assume that the central banker is benevolent and chooses sequentially his best policy to minimize the loss function, while private agents form their expectations according to rational expectations. The actions and the timing of the game are as follows: first, private agents form their expectations for all future periods rationally, then, the central bank optimally sets the nominal interest rate, finally, private agents choose the current level of consumption and inflation competitively, which results into equations (NKPC) and (EE).[18]

[18]An underlying assumption is that private agents are all identical, anonymous and decide both prices –as shareholders of firms– and quantities –as households.

More precisely, we denote by the sequence $h_{t-1} = \{\pi_\tau, y_\tau, i_\tau, \tau < t\}$ the history prior to period t. For any expectations rule $f = (f_0, f_1...)$, we denote by $f^t = (f_t, f_{t+1}...)$ the continuation of the expectations rule at time t, where $f_t(h_{t-1})$ denotes private agents expectations $\{E_t\pi_{t+1}, E_t y_{t+1}\}$.[19] Similarly, we denote by σ the policy plan of the central bank and by σ^t its continuation. The policy rule, σ_t depends on the history, h_{t-1} but also on private agents expectations, $\{E_t\pi_{t+1}, E_t y_{t+1}\}$ that are supposed observable at the time of the decision. Thus, $i_t = \sigma_t(h_{t-1}, E_t\pi_{t+1}, E_t y_{t+1})$. Finally, economic agents choose the level of the output gap and inflation at date t as a function of economic agents expectations $\{E_t\pi_{t+1}, E_t y_{t+1}\}$ and current interest rate i_t.[20]

We can now describe how each decision is determined given others' decision rule, before presenting a formal definition of a sustainable equilibrium. We present these actions by starting from the end of the period t.

The history induced by the continuation of the policy plan σ^t and continuation of expectations rule f^t, after an history h_{t-1}, follows:

$$\begin{aligned} h_t =& \{h_{t-1}, \dots \\ & \dots \pi_t(f_t(h_{t-1}), \sigma_t(h_{t-1}, f_t(h_{t-1})), y_t(f_t(h_{t-1}), \sigma_t(h_{t-1}, f_t(h_{t-1}))), \sigma_t(h_{t-1}, f_t(h_{t-1}))\}, \end{aligned} \tag{18}$$

where inflation and the output gap are uniquely pinned down by equations (NKPC) and (EE).

We focus on rational expectations and therefore we suppose that expectations of agents are always correct. Given an history h_{t-1} and a policy strategy σ, the continuation of the expectations rule satisfy rational expectations:

$$f_t(h_{t-1}) = [\pi_{t+1}(f_{t+1}(h_t), \sigma_{t+1}(h_t, f_{t+1}(h_t)), y_{t+1}(f_{t+1}(h_t), \sigma_{t+1}(h_t, f_{t+1}(h_t)), \tag{19}$$

where h_t is induced by h_{t-1} following equation (18).

Finally, the policymaker chooses the continuation of the policy plan σ^t given an history h_{t-1} and the continuation of expectations f^t to maximize social welfare:

$$\begin{aligned} \max_{\sigma_\tau, \tau \le t} - \sum_{\tau=t}^{\infty} \beta^{\tau-t} \left[\pi_\tau(f_\tau(h_{\tau-1}), \sigma_\tau(h_{\tau-1}, f_\tau(h_{\tau-1})))^2 + \dots \right. \\ \left. \dots \lambda(y_\tau(f_\tau(h_{\tau-1}), \sigma_\tau(h_{\tau-1}, f_\tau(h_{\tau-1}))) - y^*)^2 \right] \end{aligned} \tag{20}$$

where h_τ is induced by $h_{\tau-1}$ following equation (18).

This allows us to define a sustainable equilibrium as follows:

[19] Expectations at time t only depend on variables at time $t-1$ because of the timing of the game and the absence of shocks. In Section 5.2, these expectations also include current shocks.

[20] Since the first order conditions of private agents are forward-looking, inflation and the output gap will depend on h_{t-1} at the equilibrium only through expectations and monetary policy.

Definition 1. *A sustainable equilibrium is a policy plan σ and the private expectations f such that for every history h_{t-1}:*

(i) [Optimal policy] Given the private expectations f, the continuation of the policy plan σ^t solves central bank's problem (20)*;*

(ii) [Competitive equilibrium and Rational expectations] Given the policy plan σ, the continuation of expectations f^t satisfies equation (19)*.*

In particular, note that any discretionary allocation is a sustainable equilibrium. Indeed, if agents expect the discretionary equilibrium, the central banker will optimally choose to stick to it. This thus satisfies item (i) of the definition. In addition, private agents expectations are correct and they fulfill equations (NKPC) and (EE). Item (ii) is thus also satisfied.

Characterizing the set of sustainable equilibria We now characterize the set of sustainable equilibria along the lines of Abreu (1988) and Chari and Kehoe (1990), where any sustainable equilibrium should deliver a better outcome than the trigger strategy consisting of the optimal deviation from the policy plan given private agents expectations followed by the worst equilibrium.

Before stating the result, let us define the best deviation policy. Given private agents expectations f_t and an history h_{t-1}, the best deviation policy, $\sigma_t^d(h_{t-1}, f_t(h_{t-1}))$ is given by maximizing the instantaneous social welfare with respect to i_t:

$$U_t^d = \max_{i_t} -\left[\pi_t^2 + \lambda(y_t - y^*)^2\right] \tag{21}$$

under the constraints (NKPC) and (EE) for given expectations, $f_t(h_{t-1}) = \{E_t\pi_{t+1}, E_t y_{t+1}\}$.

As pointed out in Section 3.2, allocations under discretion can lead to diverging inflation when $\lambda > 0$, in which case they imply an infinite loss. These allocations are also sustainable, implying that the worst sustainable equilibrium necessarily triggers an infinite social loss. Therefore, we get the trivial result that any feasible allocation is sustainable:

Proposition 1. *When $\lambda > 0$, any feasible allocation satisfying* (EE) *and* (NKPC) *is the outcome of a sustainable equilibrium.*

When $\lambda = 0$, the unique sustainable allocation is the zero-inflation allocation ($\pi_t = 0$ for all t).

Proof. See Appendix C.1. □

It is only when the central bank does not weight output stabilization ($\lambda = 0$) as Rogoff (1985)'s conservative central banker that the worst equilibrium yields a bounded loss. But, at the same time, there is only one sustainable equilibrium that corresponds with the first best ($\pi_t = 0$ at each date t).

In the end, there is no limits to credibility that would restrict the actions of the central bank.

3.4 Taylor rules

In what follows and to make things simple, we consider a simplified version of the Taylor rule (Taylor, 1993):

$$i_t = \phi \pi_t, \tag{TR}$$

where the positive scalar ϕ measures the reactiveness of the nominal interest rate to inflation. In this context, the Taylor principle corresponds to ϕ being larger than 1.[21]

To make such a rule consistent with the timing of the decisions defined above, we rewrite the Taylor rule as a function of private agents' expectations using the equations (NKPC) and (EE):

$$i_t = \phi \frac{\kappa E_t y_{t+1} + (\beta + \kappa/\sigma) E_t \pi_{t+1}}{1 + \phi\kappa/\sigma}. \tag{22}$$

This latest equation defines a policy strategy that consists of adhering to the Taylor rule. We say that an allocation $\{\pi_t, y_t, i_t\}_t$ is implemented by the Taylor rule when this allocation satisfies (NKPC), (EE) and (TR) for any t.

Proposition 2. *If $\lambda > 0$ and the Taylor principle is satisfied, then there exist multiple sustainable equilibria implemented by the Taylor rule.*

If $\lambda = 0$, the unique sustainable equilibrium implemented by the Taylor rule corresponds to the zero-inflation allocation.

Proof. See Appendix C.2. □

Proposition 2 is a direct corollary of Proposition 1. This latter proposition proves that, when $\lambda > 0$, any strategy that implies to blow up the economy is sustainable as the worst equilibrium leads to an infinite loss. Therefore, following through on the Taylor rule is always credible even if it means blowing up the economy! The Taylor rule is thus consistent with many sustainable equilibria that depend on arbitrary initial expectations. They all exhibit diverging inflation paths and large social losses.

[21] See Woodford (2003) for a proof.

On the other hand, as any central bank's strategy is sustainable, the strategies that the central bank may find useful to use to avoid the multiplicity resulting from the Taylor rule are also sustainable. One example would be to follow the Taylor rule if expectations are not inflationary and shift to Loisel (2009)'s bubble-free rule if inflation expectations are diverging.[22]

3.5 Neo-Fisherian paradox of forward guidance

Let us now investigate the neo-Fisherian paradox of forward guidance is affected by sustainability.

More precisely, consider an allocation $\{\pi_t, y_t, i_t\}_{t\geq 0}$ satisfying equations (NKPC) and (EE) at each date t and such that i_t follows an interest rate peg to a low interest rate $i_{low} < 0$ from date 0 to date $T-1$ and then verifies a Taylor rule satisfying the Taylor principle afterward. We restrict attention to situations where, at date T, economic agents coordinate on the zero inflation equilibrium[23]. For exposition purpose, we do not consider any deflationary shock that could have motivated such policy strategy to focus on the impact of the peg. We refer to each of these allocations as the *T-period peg allocation*. When the peg holds indefinitely, the corresponding allocation is the ∞*-period peg allocation*

The neo-Fisherian paradox of forward guidance is that the initial level of inflation in the T-period peg allocation, π_0, monotonously diverges to $+\infty$ when T tends to $+\infty$, while the infinite peg may yield a bounded level of inflation, $\pi_{\infty-peg} = i_{low} < 0$.

Proposition 3. *For any $T < \infty$ and for $T = \infty$, the T-period peg allocation is an outcome of a sustainable equilibrium if and if $\lambda > 0 <$.*

Proof. See Appendix C.3. □

When the punishment is potentially infinite, any allocation is sustainable and, thus sustainability has no bite on the neo-Fisherian effect and does not help much to clarify the neo-Fisherian paradox of forward guidance.

In the end, when agents can infinitely punish central banks, any feasible allocation is sustainable. As a result, the Taylor rule is not sufficient to ensure determinacy – but, any policy that may restore inflation determinacy is also sustainable – and forward guidance

[22]In richer environments, such strategies may include alternative tools as fiscal ones, under the assumption that the preferences of the authority in charge of these other tools are aligned with the central bank's ones.

[23]This restriction is standard in the literature even if other allocations exist. When introducing limited credibility, such a restriction will become unnecessary.

policies face the neo-Fisherian paradox. Such a result highly relies on the punishment scheme that agents are coordinating on and also has the consequence to inflict them an infinite loss. In the following, we investigate how milder punishment schemes that provide limited incentives – and, thus, limited credibility – for the central bank to follow through on past promises affect these two benchmark results.

4 Limited credibility

In this section, we first introduce limited credibility by assuming that private agents coordinate on a milder punishment scheme. We then characterize the set of sustainable equilibrium and show that limited credibility implies that inflation remains bounded. We then apply this result to investigate the set of sustainable equilibria with Taylor rules – as we did in Section 2 in the case of flexible prices – and to the equilibrium outcomes when the central bank follows an interest rate peg.

4.1 Milder punishment schemes

The generic outcome that any feasible equilibrium is sustainable directly derives from the infinite loss in the worst equilibrium. If, instead, one thinks that the central banker's credibility is limited, the punishment by the private agents, in case of a deviation, should be milder and bounded. In particular, this can be obtained when agents coordinate on the time-invariant discretionary equilibrium, as such an equilibrium yields a positive but finite loss.

Assumption 1 (Limited credibility). *In case of a deviation, agents coordinate on the time-invariant discretionary equilibrium to punish the central banker.*

In the appendix, we provide a micro-foundation for such a coordination, based on the idea that private agents' punishment in case of a deviation should not hurt too much private agents themselves i.e. should be consistent with potential renegotiation of private agents.[24] There are, of course, alternative ways to produce limited credibility, e.g. punishment with finite length as (Loisel, 2008).

[24] This assumption actually parallels the literature in sovereign debt where in addition to lack of commitment on the borrower's side as in Eaton and Gersovitz (1981) or Bulow and Rogoff (1989), one can consider the problem of commitment of lenders to follow through on punishments as in Kletzer and Wright (2000). Note that renegotiation on punishment schemes in monetary models has been evoked by Nakata (2014).

This assumption on the punishment scheme leads to define Limited Credibility equilibria (or, to make it short LC-equilibria).

Definition 2. *An allocation, i.e. a sequence* $\{\pi_t, y_t, i_t\}_{t\geq 0}$*, is the outcome of a LC-equilibrium if and only if (i) it solves equations* (NKPC) *and* (EE) *where expectations are replaced by realized values (ii) for every t, the following inequality holds:*

$$-\sum_{s=t}^{\infty} \beta^{s-t} \left[\pi_s^2 + \lambda \left(y_s - y^*\right)^2\right] \geq U_t^d + \beta W^D. \tag{SUST}$$

In equation (SUST), the left-hand-side infinite sum corresponds to the date-t welfare evaluated along the allocation. The first right-hand-side member corresponds to the best instantaneous welfare given expectations consistent with the date-t+1 allocation and the second member corresponds to the welfare in the time-invariant discretionary equilibrium.

In the case of a conservative central banker ($\lambda = 0$), the set of sustainable equilibria as previously defined and the set of LC-equilibria coincide and boil down to a unique allocation ($\pi_t = 0$, for all period t).

4.2 Some properties of LC-equilibria

Let us now investigate how the constraint (SUST) restricts the set of allocations. In particular, let us show that limited credibility rules out diverging nominal paths:

Proposition 4. *In a LC-equilibrium, inflation is bounded.*

Proof. See Appendix C.4. □

The main, and simple, intuition of proposition 4's result is that unbounded inflation leads to very low levels of welfare, which ultimately fall well below the level of welfare in a discretionary equilibrium. As a result, the central bank always optimally deviates from sticking to a diverging-inflation allocation. In the end, proposition 4 implies that inflation should remain bounded. This result is important since the multiplicity of allocations consistent with a Taylor rule that we found in Section 3.4 is based on allocations that exhibit diverging inflation path.

Zero-inflation allocation The loss associated with the zero-inflation allocation is $y^*/(1-\beta)$. This loss has to be compared with the best deviation which is given by equation (21) and the loss under the time-invariant discretionary allocation. This leads to the following proposition:

Proposition 5. *The zero-inflation allocation is the outcome of a LC-equilibrium if and only if the following condition is satisfied:*

$$\frac{1}{1-\beta} \leq \frac{\kappa^2}{\kappa^2+\lambda} + \kappa^2 \frac{\beta}{1-\beta} \frac{\lambda+\kappa^2}{(\kappa^2+\lambda(1-\beta))^2}. \tag{SUST-TP}$$

Proof. This Proposition is a direct application of definition 2. □

This condition is trivially satisfied when the weight on the output gap is 0 ($\lambda = 0$). If there is no inflation-output gap trade-off, the central banker finds always desirable to fully stabilize inflation and the zero-inflation allocation is the outcome of a LC-equilibrium. Conversely, when the weight λ is infinite, as output gap stabilization around the target, y^*, has infinitely more weight than inflation stabilization, the zero-inflation allocation is not a LC-equilibrium.

In the rest of the paper, we make the following assumption:

Assumption 2. *Condition* (SUST-TP) *is satisfied.*

Given our previous discussion, this corresponds to a range of intermediate values for the weight λ that is put on the output gap stabilization objective. Assumption 2 involves no loss of generality: when this assumption is not satisfied, the allocation implemented by the Taylor rule is simply not a LC-equilibrium. This assumption is also motivated by the literature on the impact of limited credibility on the optimal allocation. Kurozumi (2008) and Loisel (2008) have shown that, under plausible calibrations and using the discretionary equilibrium as a punishment, the optimal allocation is sustainable in normal times and Nakata (2014) extended this result to the case of liquidity traps and forward guidance.

4.3 Taylor rules

Let us now investigate whether the Taylor rule is sufficient to determine the inflation rate when credibility is limited. To this purpose, we are looking for equilibria where the policy instrument i_t is set following a Taylor rule, at least for some periods. Our main finding is that the Taylor principle - nominal interest rate has to respond more than one for one to changes in inflation - is sufficient to ensure inflation determinacy. We examine to what extent limited credibility restrict the set of equilibria outside the Taylor principle in subsection 5.1.

We are interested not only in equilibria in which the Taylor rule is always followed but also in those consistent with the Taylor rule in the short run but not in the long run. We denote by the integer τ the date of a switch from the Taylor rule to the discretionary policy.

The following proposition states this subsection's main result:

Proposition 6. *The Taylor principle ensures that there exist only two LC-equilibria consistent with a Taylor rule strategy under limited credibility:*

(i) the discretionary equilibrium ($\tau = 0$),

(ii) the zero-inflation allocation ($\tau = \infty$).

Proof. See Appendix C.5. □

First, Proposition 6 shows that there exists a unique LC-sustainable equilibrium implemented by the Taylor rule, and this is the zero-inflation allcoation. The proof goes as follows. Suppose that there exists another allocation implemented by the Taylor rule. In particular, such an allocation features some inflation different from 0 in some period. Without loss of generality, such inflation rate can be assumed to be positive. Then, because of the Taylor principle, this implies that inflation should rise in the future and, more specifically, the inflation rate should even diverge to infinity. In turn, this implies that the central banker is better off deviating from the Taylor rule at some finite date in the future. Such an allocation is thus not sustainable.

Second, Proposition 6 proves that there does not exist equilibria featuring Taylor rule only in the short run under limited credibility. When the Taylor principle is satisfied, inflation is non-zero only when some inflation is expected in the future, eventually leading the central bank to switch from the Taylor rule to the discretionary policy at some point.

Yet, when such a move is anticipated, the central banker has any incentives to switch earlier. As a result, by backward induction, the central bank never follows the Taylor rule. In the end, there exist only two equilibria: in the first one the Taylor rule is always followed by the central bank, in the second one the central bank always adopts the discretionary monetary policy.

Proposition 6 shows that if there is multiplicity of equilibria it is not because the Taylor rule leads to multiple equilibria but instead because economic agents can coordinate their expectations on the discretionary equilibrium.

Cochrane (2011) argues that policy switch in the future in case of diverging inflation path does not prevent the equilibrium to exist. We instead argue that the threat to switch to more optimal monetary policy in case of diverging inflation path can prevent the formation of such a trajectory. Indeed, the expectations of a policy deviation makes desirable to switch immediately, and therefore, if agents believe in the Taylor rule in the first place they should not coordinate on non-zero inflation trajectories.

In the end, if the central banker's credibility is limited, thus preventing costly off-equilibrium reaction by the central banker, the diverging inflation paths consistent with the Taylor rule are also not sustainable and, thus, there exists a unique equilibrium in which the Taylor rule is always followed by the central banker.

4.4 Neo-Fisherian paradox of forward guidance

We now reconsider the paradox explained in section 3.5 under limited credibility. Indeed, Proposition 4 shows that explosive trajectories due to too-low-for-too-long interest rates do not exist because it is not credible to follow through on such a promise.

Proposition 7. *There exists a maximal duration of forward guidance, $\bar{T}$, compatible with a LC-equilibrium (for finite T).*

If $T < \bar{T}$, the T-period peg allocation is the outcome of a LC-equilibrium and the initial level of inflation π_0 increases with T but is below some thresholds $\bar{\pi}$.

If $T = \infty$, the low inflation rate allocation $\pi = i_{low}$ is the outcome of a LC-equilibrium if and only if i_{low} is sufficiently close to 0.

Proof. See Appendix C.6. □

This proposition shows that the puzzling discontinuity between long forward guidance that triggers infinite inflation while a permanent interest rate shift is consistent with stable inflation disappears when introducing limited credibility as, ultimately, forward guidance is not credible anymore. This leads to a clear dichotomy: either the central banker engages in a one-for-all monetary policy change (an infinite peg) or a transitory forward guidance with a limited duration. In addition, according to Proposition 7, there is a point after which any further extension of an already long forward guidance has no effect on contemporaneous inflation. Third, a central bank cannot set a policy rate lower than $\bar{i}$ forever.

5 Extensions

In this section, we investigate the consequences of limited credibility when enriching our environment. First, we consider how limited credibility affects the multiplicity of equilibrium when the Taylor principle is not satisfied. Second, we extend our results to the presence of shocks. Third, we investigate how the zero lower bound may affect our results.

5.1 Limited credibility outside the Taylor principle

In this subsection, we investigate how limited credibility modifies the equilibrium outcome when the central bank is expected to follow a Taylor rule that does not satisfy the Taylor principle. In this case, it is well known that the inflation rate is not determinate. Yet, we show that limited credibility limits inflation.

Proposition 8. *When the Taylor principle is not satisfied ($\phi < 1$), there exist many sustainable equilibria consistent with a Taylor rule strategy under limited credibility:*

(i) the discretionary equilibrium ($\tau = 0$),

(ii) the zero-inflation allocation ($\tau = \infty$),

(iii) self-fulfilling prophecies equilibrium ($\tau = \infty$) described as follows:

$$\pi_t = \lambda_1^t \pi_0, \quad \textit{and} \quad y_t = \lambda_1^t \omega \pi_0,$$

where the initial level of inflation π_0 is any scalar in $[\underline{\pi}, \bar{\pi}]$, the scalar λ_1 is the stable eigenvalue of the three-equation system and the scalar ω denotes the relative weight of the associated eigenvector to the output gap (see Appendix C.7 for formal definitions).

Furthermore, the upper (lower) bound for inflation $\bar{\pi}$ ($\underline{\pi}$) increases (resp. decreases) with the inflation bias y^*.

Proof. See Appendix C.7. □

First, let us suppose that economic agents anticipate that the central banker follows the Taylor rule, in this case we know that there exist many equilibria that depend on a free parameter. We index each equilibrium by the initial level of inflation π_0. Limited credibility rules out two types of equilibria, those with a diverging inflation path –in which inflation and the output gap belong to the unstable eigenvector– and those with a too high inflation initial level –with a too large π_0. The zero-inflation allocation belongs to the set of sustainable equilibria, but because of the failure of the Taylor principle many other equilibria coexist.

Second, there is no equilibria featuring a Taylor rule in the short run but not in the long run. Interestingly, this result does not depend on the Taylor principle in a deterministic environment. Suppose that the Taylor principle is not satisfied. If private agents form expectations that lead to inflation consistent with a sustainable level of inflation $\bar{\pi}_0 \in [\underline{\pi}, \bar{\pi}]$ in period 0, then inflation converges toward zero. Along this trajectory, the central bank adheres to the Taylor rule and has no incentives to deviate from it. However, if they form expectations outside this admissible bracket or if they put a weight on the explosive eigenvalue, then the central bank optimally deviates immediately at date 0 and never stick to the Taylor rule.

Third, the range of fluctuations of inflation depends on the inflation bias. Indeed, the social welfare in the worst equilibrium increases with this bias, so the higher the bias, the more the bias the more costly the abandon from the Taylor rule is, making higher expectations sustainable. As far as we know, this is the first time such a connection is established. This result suggests that adhering to the Taylor principle is more crucial when the policymaker suffers from a high level of time-inconsistency, but is less crucial when the policymaker is more conservative.

If we reinterpret the results by Clarida et al. (2000) and Lubik and Schorfheide (2004) on the Great Inflation in the US through the lens of this result, the large volatility of inflation in the 70s may have reflected both a departure from the Taylor principle and a significant inflation bias. Without inflation bias, the Fed would have departed from the Taylor rule to discretion and inflation would have been more stable.

5.2 Stochastic environment

This section extends our results to a stochastic environment and we investigate whether shocks can modify the results of our benchmark analysis. Our main result is that, under various conditions (bounded shocks or certain unbounded shocks), our results go through as the distance between the punishment equilibrium and any other equilibria has to remain bounded, thus ruling out self-fulfilling paths of diverging inflation.

We incorporate shocks into the baseline model as follows:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa y_t + u_t, \tag{NKPC}$$

$$y_t = E_t y_{t+1} - \frac{1}{\sigma} \left(i_t - E_t \pi_{t+1} - r_t^n \right), \tag{EE}$$

where the random processes u_t and r_t^n denote cost-push shocks and natural interest rate shocks respectively. Besides, we assume that the central banker can arbitrary deviate from the Taylor rule as follows:

$$i_t = \alpha \pi_t + \epsilon_t, \tag{23}$$

where ϵ_t is a monetary policy shock.

LC-equilibrium in a stochastic environment The timing is modified to take into account the stochastic property of the game. We now assume that the current shocks are known by private agents and by the central bank at each stage of the game. Thus expectations that they form as well as the policy plans implemented by the central bank are functions of these shocks. We also assume that private agents form expectations that are correct in average. More formally, we define a sustainable equilibrium in such an environment by including shocks in the state space, $h_{t-1} = \{\pi_\tau, y_\tau, i_\tau, u_\tau, r^n_\tau, \epsilon_\tau, \tau < t\}$ and assume that the right-hand-side member of equation (19) is replaced by an integral over all future shocks. We now define LC-equilibrium in such an environment.

The presence of shocks implies fluctuations in the time-invariant discretionary equilibrium. If shocks are bounded then the punishment equilibrium is bounded as well and so is any LC-equilibrium, thus implying that all our previous results still hold. We extend this result to certain situations where shocks are unbounded:

Proposition 9. *The distance between inflation in any LC-equilibrium and inflation in the time-invariant discretionary equilibrium remains bounded if:*

(i) Shocks are bounded.

(ii) Shocks follow a first-order autoregressive process with second-order process innovation.[25]

Proof. See Appendix C.8. □

In the following, we apply this result to validate the robustness of our benchmark results.

Taylor principle and the 'divine coincidence' If the natural interest rate shock, r^n_t, is the only shock affecting the economy and is incorporated in the Taylor rule (as suggested by Woodford, 2003; Gali, 2008) then propositions 6 is not modified. The Taylor principle guarantees the uniqueness of the equilibrium as long as economic agents believe in the Taylor rule in the first place. Indeed, if the interest rate reacts by one-for-one to the natural interest rate shock, this latter shock disappears from the economy and no more affects inflation and the output gap.[26]

[25]Formally, this embeds any autoregressive process such that: $x_t = \rho x_{t-1} + \nu_t$ with $\rho < 1$, $E_t\nu_{t+1} = 0$ and $E_t|\nu_{t+1}|^2 < \infty$.

[26]If the Taylor principle is not satisfied, sunspot shocks can hit the economy:

$$[\pi_t \quad y_t]' = V_1 w_{1,t}, \texttt{ where } w_{1,t+1} = \lambda_1 w_{1,t} + \eta_{t+1}, \tag{24}$$

where $w_{1,0}$ corresponds to an (arbitrary) initial condition and the stochastic process η_{t+1} is any bounded zero-mean process, sometimes referred to as sunspot shocks. In the existing literature, the variance and the support of this shock is unbounded. Limited credibility imposes the boundedness of sunspot shocks (as in Proposition ??).

Suboptimal Taylor rules If the Taylor rule does not respond to the natural rate of interest r_t^n, if the economy is hit by a cost-push shock u_t or if monetary policy deviates from the Taylor rule, then large shocks can generate an incentive for policy switching, the Taylor rule being suboptimal compared to the discretionary equilibrium (or the best deviation policy).

For simplicity, we suppose all shocks to be i.i.d. and zero-mean. We also assume that the variance of shocks is small compared to the inflation bias y^*. This latter assumption guarantees that the Taylor rule delivers a higher expected welfare relative to the one delivered by the time-invariant discretionary equilibrium. If this assumption is not verified, the Taylor rule will never be implemented in a LC-equilibrium.

Proposition 10. *If shocks (r^n, u, ϵ) are not too large, the Taylor principle guarantees that there exists a unique LC-equilibrium implemented by the Taylor rule.*

Proof. See Appendix C.9. □

If shocks are not too large, central banker always prefers sticking to the Taylor rule and the Taylor rule implements a LC-equilibrium. Proposition 10 does not depend on the nature of shocks as the three kind of shocks we consider lead to some inefficiency of the Taylor rule.

On the contrary, when shocks are too large, then deviating from the Taylor rule policy triggers a short-term gain that dominates the long term gains of adhering to it and achieving the zero-inflation allocation (compared to the time-invariant discretionary allocation). The presence of large shocks may lead to allocations that are implemented by a Taylor rule in the short run but not in the long run. We now investigate the impact of such large shocks ex ante.

To simplify the exposition, we only assume large monetary policy shocks. Consider an economy in which monetary policy follows a Taylor rule (as long as it is optimal) that can be affected by a very large (positive or negative) monetary policy shock. This is a shortcut for...

We assume that the monetary policy shock is either 0 with probability $(1-p)$ or ϵ $(-\epsilon)$ with probability $p/2$ $(p/2)$. We also assume that the size of the shock ϵ is so large that it always leads the central banker to switch to discretion.

Proposition 11. *There exists a threshold $\bar{p}$, such that:*

If $p < \bar{p}$, there exists a unique LC-equilibrium implemented by the Taylor rule (as long as the economy is not affected by the shock) if and only if the adjusted taylor principle $\phi > 1-p$ is verified. When the monetary policy shock is zero, inflation and the output gap increase with p because of the anticipation of the inflation bias in the discretionary equilibrium.

If $p > \bar{p}$, the Taylor rule is never followed by the central banker.

Proof. See Appendix C.10. □

The higher the probability of a large shock, the less incentives the central banker has to stick to the Taylor rule. There exists a thresholds, $\bar{p}$, such that when the probability of large shocks is above the threshold $\bar{p}$, the central banker never follows the Taylor rule. There are two reasons: first, inflation and the output gap are closer and closer to the discretionary allocation when p increases, second the long-term gains of the Taylor rule decreases with p (it is as if the discount factor of the central banker is lower).

5.3 Taylor rule and the zero lower bound

Introducing a zero lower bound on nominal interest rate potentially introduces more equilibria in the game that we have studied. In particular, the permanent liquidity trap situation as discussed by Benhabib et al. (2001) may be a sustainable equilibrium. Importantly, this equilibrium corresponds to a finite loss (see Appendix B), and so, our results on the time-invariant discretionary equilibrium can be extended to the case where private agents punish the central banker by coordinating their expectations on those compatible with the permanent liquidity trap.

5.4 More complex rules

In this subsection, we discuss two types of policies put forward in the literature to anchor inflation expectations at a unique level: first, bubble-free rules that incorporate private agents' expectations to disconnect these expectations and the current level of the economy Loisel (2009),[27] second, sophisticated rules that integrate a policy switch in case of inflation diverging (Woodford, 2003, for a short review).[28] Under full credibility, these more complex rules succeed in implementing a unique equilibrium by committing to blow up the economy when economic agents form diverging expectations. However, they fail when credibility is limited.

Under limited credibility, off-equilibrium outcomes of such policies are not necessarily credible and this may lead to multiple equilibria. In particular, when economic agents expect that inflation and the output gap will be at their time-invariant discretionary counterparts,

[27] More precisely, Loisel (2009) put forward rules like: $i_t = r_t^n + E_t\pi_{t+1} + \psi\pi_t + 1/\sigma E_t(y_{t+1} - y_t)$, where the parameter ψ is positive while Adao et al. (2011) advocate for rules like: $R_t = \frac{\xi_t}{E_t \frac{\beta u_C(t+1)}{P_{t+1}}}$, where ξ_t is an exogenous variable and u_C is the marginal utility of consumption in a non-linear setting. For a related proposal, see also Hall and Reis (2016).

[28] Atkeson et al. (2010) for instance argue that if inflation diverges from the target, the central bank can switch to a money-based rule to stabilize inflation.

then the central banker has any incentives not to follow the complex rule (either the bubble-free or the sophisticated rule) but rather to immediately choose the discretionary policy. As a result these more complex rules suffer from the same problem that the Taylor rule as identified in this paper and therefore there is no gains to adhere to these more complex rules compared to standard Taylor rule, which is – at least – simple to define and verify.

6 Conclusion

In this paper, we have revisited the implementation of monetary policy in the new Keynesian model. Our main result is that limited credibility paradoxically allows to ensure the anchoring of inflation expectations when a Taylor rule is expected to be implemented. In addition, limited credibility reduces the effectiveness of forward guidance as it restricts the possible interest rate pegs that can be implemented – thus solving the neo-Fisherian paradox.

References

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383–396.

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1063.

Adao, B., I. Correia, and P. Teles (2011): "Unique Monetary Equilibria with Interest Rate Rules," *Review of Economic Dynamics*, 14, 432–442.

Atkeson, A., V. V. Chari, and P. J. Kehoe (2010): "Sophisticated Monetary Policies," *The Quarterly Journal of Economics*, 125, 47–89.

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

Bassetto, M. (2002): "A Game-Theoretic View of the Fiscal Theory of the Price Level," *Econometrica*, 70, 2167–2195.

——— (2005): "Equilibrium and government commitment," *Journal of Economic Theory*, 124, 79–105.

Benhabib, J., S. Schmitt-Grohe, and M. Uribe (2001): "The Perils of Taylor Rules," *Journal of Economic Theory*, 96, 40–69.

Bullard, J. and K. Mitra (2002): "Learning about monetary policy rules," *Journal of Monetary Economics*, 49, 1105–1129.

Bulow, J. and K. Rogoff (1989): "Sovereign Debt: Is to Forgive to Forget?" *American Economic Review*, 79, 43–50.

Challe, E. (2017): "Uninsured unemployment risk and optimal monetary policy," Mimeo.

Chang, R. (1998): "Credible Monetary Policy in an Infinite Horizon Model: Recursive Approaches," *Journal of Economic Theory*, 81, 431–461.

Chari, V. V. and P. J. Kehoe (1990): "Sustainable Plans," *Journal of Political Economy*, 98, 783–802.

Clarida, R., J. Galí, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, 115, 147–180.

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

Cochrane, J. H. (2009): "Can learnability save new-Keynesian models?" *Journal of Monetary Economics*, 56, 1109–1113.

——— (2011): "Determinacy and Identification with Taylor Rules," *Journal of Political Economy*, 119, 565–615.

——— (2013): "The New-Keynesian Liquidity Trap," NBER Working Papers 19476, National Bureau of Economic Research, Inc.

——— (2016): "Michelson-Morley, Occam and Fisher: The Radical Implications of Stable Inflation at Near-Zero Interest Rates," Mimeo.

Eaton, J. and M. Gersovitz (1981): "Debt with Potential Repudiation: Theoretical and Empirical Analysis," *Review of Economic Studies*, 48, 289–309.

Friedman, M. (1968): "The Role of Monetary Policy," *American Economic Review*, 58.

Gabaix, X. (2016): "A Behavioral New Keynesian Model," NBER Working Papers 22954, National Bureau of Economic Research, Inc.

Gali, J. (2008): *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press.

García-Schmidt, M. and M. Woodford (2015): "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis," NBER Working Papers 21614, National Bureau of Economic Research, Inc.

Hall, R. E. and R. Reis (2016): "Achieving Price Stability by Manipulating the Central Bank?s Payment on Reserves," NBER Working Papers 22761, National Bureau of Economic Research, Inc.

Ireland, P. N. (1997): "Sustainable monetary policies," *Journal of Economic Dynamics and Control*, 22, 87–108.

Kletzer, K. M. and B. D. Wright (2000): "Sovereign Debt as Intertemporal Barter," *American Economic Review*, 90, 621–639.

Kurozumi, T. (2008): "Optimal sustainable monetary policy," *Journal of Monetary Economics*, 55, 1277–1289.

Kydland, F. E. and E. C. Prescott (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85, 473–491.

Leeper, E. M. (1991): "Equilibria under 'active' and 'passive' monetary and fiscal policies," *Journal of Monetary Economics*, 27, 129–147.

Loisel, O. (2008): "Central bank reputation in a forward-looking model," *Journal of Economic Dynamics and Control*, 32, 3718–3742.

——— (2009): "Bubble-free policy feedback rules," *Journal of Economic Theory*, 144, 1521–1559.

Lubik, T. A. and F. Schorfheide (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94, 190–217.

McCallum, B. T. (2009): "The Role of "Determinacy" in Monetary Policy Analysis," IMES Discussion Paper Series 09-E-17, Institute for Monetary and Economic Studies, Bank of Japan.

Nakata, T. (2014): "Reputation and Liquidity Traps," Finance and Economics Discussion Series 2014-50, Board of Governors of the Federal Reserve System (U.S.).

Nuno, G. and C. Thomas (2017): "Optimal Monetary Policy and Heterogeneous Agents," Mimeo.

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

——— (1986): "Ruling out divergent speculative bubbles," *Journal of Monetary Economics*, 17, 349–362.

Phelan, C. and E. Stacchetti (2001): "Sequential Equilibria in a Ramsey Tax Model," *Econometrica*, 69, 1491–1518.

Rogoff, K. (1985): "The Optimal Degree of Commitment to an Intermediate Monetary Target," *The Quarterly Journal of Economics*, 100, 1169–89.

Sims, C. A. (1994): "A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy," *Economic Theory*, 4, 381–399.

——— (2013): "Paper Money," *American Economic Review*, 103, 563–84.

Sunakawa, T. (2015): "A quantitative analysis of optimal sustainable monetary policies," *Journal of Economic Dynamics and Control*, 52, 119–135.

Taylor, J. B. (1993): "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

——— (1999): *Monetary Policy Rules*, no. tayl99-1 in NBER Books, National Bureau of Economic Research, Inc.

Woodford, M. (1995): "Price-level determinacy without control of a monetary aggregate," *Carnegie-Rochester Conference Series on Public Policy*, 43, 1–46.

——— (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

A Renegociation

Proposition 1 stems from the ability of private agents to commit to punish harshly the central banker in case of a deviation. Yet, such a punishment comes at their own expense. Conversely, there exist equilibria where private agents do not trust the central banker but where the society welfare remains bounded (e.g. the time-invariant discretionary equilibrium). As a result, taking into account private agents' own time-inconsistency and their desire to renegotiate (in a similar spirit as Kletzer and Wright, 2000, for the case of sovereign debt) is very likely to alter Proposition 1. In this section, we allow such a renegotiation to take place and show that it provides a micro-foundation for the LC-equilibrium definition (definition 2).

Setting. We enrich the model so as to let households coordinate their expectations once the central banker has implemented his policy. We however restrict the set of renegotiation to forbid coordination on a Pareto dominated equilibrium or an equilibrium involving time-inconsistency. The new timing of the game is as follows:

(i) In the beginning of the period, given the policy plan, σ, private agents form expectations that satisfy rational expectations, $f_t(\sigma, h_{t-1})$;

(ii) The central banker implements optimal monetary policy by setting i_t taking f_t as given;

(iii) If the central banker deviates from the initial plan, private agents coordinate their expectations on the worst equilibrium, $\tilde{f}_{t+1} = f^W$;

(iv) Private agents choose whether to renegotiate. Either they stick to their past expectations (f_{t+1} or $\tilde{f}_{t+1}$) or they decide to coordinate their expectations on the best time-consistent allocation $\bar{f}_{t+1}$.[29]

Steps (i) to (iii) describe the standard game supporting the definition of a sustainable equilibrium. We add a renegotiation step at the end of the period. Step (iv) is a relatively conservative departure from the standard case without renegotiation. Suppose that in step (iii), economic agents coordinate on the worst equilibrium. Then it can be Pareto improving to renegotiate to avoid inflecting themselves too harsh punishment. We restrict the renegotiation set to allocations that prevent the central banker to deviate in the future. This restriction can be interpreted in two different ways.

First, agents internalize that if they choose an allocation that will lead the central banker to re-optimize at each period, they prevent themselves from punishing the central banker in case of deviation, indirectly pushing him to renege on past promises too often. Second, this restriction captures the idea that it is unlikely that (i) private agents accept to negotiate with the central banker if this latter has already defaulting on his promises (ii) the central banker accepts to negotiate if private agents change their expectations.[30]

More formally, we denote by the set $\mathcal{S}$ the set of renegotiation. $\mathcal{S}$ consists of all sustainable equilibria satisfying:

$$-E_t \sum_{s=t}^{\infty} \beta^{s-t} \left[\pi_s^2 + \lambda \left(y_s - y^*\right)^2\right] = U_t^d + \beta E_t W_{t+1}, \tag{25}$$

[29]Note that this is equivalent to having some kind of Walrasian auctioneer that will make an offer to the rest of the private agents, who will always agree as they are homogeneous.

[30]Such an assumption is equivalent to the central banker punishing the private sector if they reset their expectations.

where W_{t+1} denotes the continuation value. This condition can be simplified as follows:

$$-\left[\pi_t^2 + \lambda\left(y_t - y^*\right)^2\right] = U_t^d. \tag{26}$$

All the discretionary equilibria and the constant deflation equilibrium are in the set $\mathcal{S}$. In particular, as among the discretionary equilibria, there exists an equilibrium implying an infinite loss, this means that the $\mathcal{S}$ contains an equilibrium achieving the worst outcome.

This results into the following definition:

Definition 3 (Renegotiation-proof equilibrium)**.** *A renegotiation-proof equilibrium is a policy rule σ and private agents expectations f that satisfy conditions (i) to (iv).*

Characterization of the set of renegotiation-proof equilibria. Let us characterize the set of renegotiation-proof equilibria:

Proposition 12. *The worst renegotiation-proof equilibrium is the time-invariant discretionary equilibrium.*

A sequence $\{\pi_t, y_t, i_t\}_{t\geq 0}$ is the outcome of a sustainable equilibrium if and only if (i) it satisfies rational expectations (ii) for every t, the following inequality holds:

$$-E_t \sum_{s=t}^{\infty} \beta^{s-t}\left[\pi_s^2 + \lambda\left(y_s - y^*\right)^2\right] \geq U_t^d + \beta E_t W_{t+1}^d. \tag{27}$$

Proof. Let us first consider the worst renegotiation-proof equilibrium.

Suppose that there exists a worse equilibrium than the time-invariant discretionary one. In this case private agents should renegotiate in step (iv) and coordinate on the time-invariant discretionary equilibrium. Knowing this ex ante, private agents should immediately choose expectations consistent with the time-invariant discretionary equilibrium. Besides, this equilibrium is renegotiation-proof (no reason to deviate for the central banker in step (iii) as well as for private agents in step (iv)). Therefore the discretionary equilibrium is the worst renegotiation-proof equilibrium.

Since the worst equilibrium is the discretionary one, if the central banker deviates the maximal punishment is the discretionary one. Thus, step (iv) can be introduced in step (iii), by assuming that private agents coordinate on the discretionary equilibrium in case of central banker default. Therefore, the sustainability constraint is (27). □

Note that, the deflationary equilibrium and the deflation spiral are not renegotiation-proof equilibrium. In other words, the zero lower bound cannot be self-fulfilling as long as economic agents can renegotiate.

B Loss in the permanent liquidity trap.

We now suppose that monetary policy faces a lower bound: $i_t \geq -\bar{i}$, with $\bar{i} > 0$. The permanent liquidity trap is a pair of steady-state output gap and inflation (y, π) solving the following model:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa y_t \text{ and } y_t = E_t y_{t+1} - \frac{1}{\sigma}\left(-\bar{i} - E_t \pi_{t+1}\right) \tag{28}$$

The resulting per period loss is:

$$\bar{i}^2 + \left(\frac{1-\beta}{\kappa}\bar{i} + y^*\right)^2 < \infty. \tag{29}$$

As this loss is constant over time, the total welfare loss in a permanent liquidity trap is also bounded:

$$\frac{1}{1-\beta}\left(\bar{i}^2 + \left(\frac{1-\beta}{\kappa}\bar{i} + y^*\right)^2\right) < \infty. \tag{30}$$

C Proofs

C.1 Proof of Proposition 1.

Case $\lambda > 0$. Let us first show that there exists an allocation satisfying (EE), (NKPC) and the first-order conditions of the discretionary problem (15) and (16) that yields an infinite loss. As the allocation solves the policy problem under discretion, it is also time-consistent and sustainable.

By manipulating the equations, we can write y_t and π_t as functions of $E_t\pi_{t+1}$ and of the time-invariant discretionary allocation (y, π):

$$y_t - y = -\frac{\lambda\beta}{\lambda+\kappa^2}E_t(\pi_{t+1} - \pi), \tag{31}$$

$$\pi_t - \pi = \frac{\kappa\beta}{\lambda+\kappa^2}E_t(\pi_{t+1} - \pi). \tag{32}$$

Thus, there exist a continuum of allocations solving these equations that can be indexed by date-0 inflation π_0:

$$\pi_t - \pi = \left(\frac{\lambda+\kappa^2}{\lambda\beta}\right)^t(\pi_0 - \pi). \tag{33}$$

We can then write date-0 loss:

$$W_0 = (1+\frac{\kappa^2}{\lambda^2})\sum_t \beta^t\pi_t^2 = (1+\frac{\kappa^2}{\lambda^2})\sum_t \beta^t\left[(\pi_t-\pi)^2 + 2\pi(\pi_t-\pi) + \pi^2\right]. \tag{34}$$

Since $\frac{(\lambda+\kappa^2)^2}{\lambda^2\beta} > 1$, the first member of the sum diverges toward ∞ and dominates the other members. Finally, one can obtain the Proposition's result by considering the strategy where agents coordinate on one of the diverging discretionary equilibrium in case of a policy deviation.

Case $\lambda = 0$. In this case, the discretionary outcome leads to $\pi_t = 0$ at each date 0, which is the first-best allocation. Let us now show that there are no other sustainable equilibria.

Let us consider an allocation that yields a different loss from 0 and let us assume that it is the outcome of a sustainable equilibrium. This implies that there exists a period τ where $\pi_\tau \neq 0$. Then, given the same private sector expectation, there exists a policy strategy that yields a strictly better outcome by setting inflation at 0 at every date, thus contradicting that the allocation is consistent with a sustainable equilibrium.

C.2 Proof of Proposition 2.

We denote by $A(\phi)$ the 2-matrix that links current and expected variables when the central banker follows a Taylor rule:

$$\begin{bmatrix} E_t\pi_{t+1} \\ E_t y_{t+1} \end{bmatrix} = A(\phi) \begin{bmatrix} \pi_t \\ y_t \end{bmatrix}, \tag{35}$$

where $A(\phi)$ is defined as follows:

$$A(\phi) = \begin{bmatrix} 1/\beta & -\kappa/\beta \\ \phi/\sigma - 1/(\sigma\beta) & \kappa/(\sigma\beta) + 1 \end{bmatrix}, \tag{36}$$

When the central bank follows a Taylor rule satisfying the Taylor principle, the matrix $A(\phi)$ has two unstable eigenvalues. There exists a unique stable allocation $\pi_t = 0$ and many unstable allocations that can be all expressed as a function of initial conditions (y_0, π_0) (Woodford, 2003). All these allocations are sustainable if and only if $\lambda > 0$ because of Proposition 1.

If $\lambda = 0$, then the unique sustainable allocation is the zero-inflation allocation. This allocation does not contradict the Taylor rule in equilibrium; thus the unique sustainable equilibrium implemented by the Taylor rule corresponds to the zero-inflation allocation.

C.3 Proof of Proposition 3.

Proposition 3 is a direct application of Proposition 1. For any finite $T < \infty$, the loss associated with the T-period peg allocation is finite and feasible, therefore it is the outcome of a sustainable equilibrium if $\lambda > 0$. However, if $\lambda = 0$, such an allocation is not the outcome of a sustainable equilibrium. Similarly, the ∞-period peg allocation is the outcome of a sustainable equilibrium if and if $\lambda > 0$.

C.4 Proof of Proposition 4.

Let assume that there exists a LC-equilibrium for which inflation is not bounded, i.e. for a certain date $\tau+1$ inflation $\pi_{\tau+1}$ is above a scalar $M > 0$ that can be as large as we want. We now consider the sustainability condition at date τ:

$$L_\tau > \pi_\tau^2 + \lambda(y_\tau - y*)^2 + \beta(\pi_{\tau+1}^2 + \lambda(y_{\tau+1} - y*)^2), \tag{37}$$

where L_τ is the social loss at period τ induced by the sustainable equilibrium with unbounded inflation path. On the right hand side of the inequality, the sum of the first two terms is greater than U_τ^d, the loss given the best deviation at period τ. Furthermore, since $\pi_{\tau+1}$ is very large the last two terms of the inequality may lead to a very high loss, and especially above the one generated by the time-invariant discretionary equilibrium. Thus:

$$L_\tau > U_\tau^d + \beta W_{\tau+1}^d. \tag{38}$$

Finally, this proves that the unbounded inflation path is inconsistent with a sustainable equilibrium.

Intuition: When the central banker anticipates a very high level of inflation next period, he is better off by deviating immediately since the worst equilibrium triggers a finite loss.

C.5 Proof of Proposition 6.

First, let us note that the zero-inflation allocation where $\pi_t = y_t = 0$ at each date t satisfies (NKPC), (EE) and (TR) which is sustainable by Assumption 2.

Now, consider another allocation consistent with (NKPC), (EE) and (TR). This means that there exists a date t at which the inflation rate differs from 0 (there is no loss of generality to focus on π_t as we can make the same reasoning with y_t and note that either π_t or π_{t+1} should then differs from 0). We can easily rewrite all variables after this date as a product of the power of an unstable matrix times the level of inflation and output gap at time t:

$$\begin{bmatrix} E_t\pi_{t+k} \\ E_t y_{t+k} \end{bmatrix} = A(\phi)^k \begin{bmatrix} \pi_t \\ y_t \end{bmatrix},$$

where $A(\phi)$ is the matrix defined in equation (36). As the Taylor principle is satisfied, all eigenvalues of $A(\phi)$ are unstable, and therefore a non zero inflation level at time t leads to diverging trajectory for the output gap and inflation. All trajectories that lead to an infinite level of inflation are not sustainable. Thus such allocations are not the outcome of a sustainable equilibrium.

Short-run Taylor rules. Suppose that there exists a date τ at which the central bank switches from the Taylor rule to the discretionary policy. As a result, $E_{\tau-1}\pi_\tau$ and $E_{\tau-1}y_\tau$ are equal to the values under discretion and central bank have any incentives to switch to the discretionary equilibrium at date $\tau - 1$. By backward induction down to the initial date, the central bank switches to discretion at date 0. Therefore, there does not exist any LC-equilibrium implemented by the Taylor rule in the short run but not in the long run.

C.6 Proof of Proposition 7.

We denote by T, the duration of the peg. For all $t \leq T$, the central banker sets the policy instrument at a low level, $i_T = i_{low} < 0$, and for $t > T$, the central banker follows a Taylor rule (satisfying the Taylor principle).

If T is finite, the unique allocation from a $T+1$ perspective that is an outcome of a sustainable equilibrium is the zero-inflation allocation (see Proposition 6). Therefore for any date t from 0 to T, the economy is described by the following dynamic system:

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = A(0)^{-1} \begin{bmatrix} E_t\pi_{t+k} \\ E_t y_{t+k} \end{bmatrix} + Di_{low},$$

Thus,

$$\begin{bmatrix} \pi_0 \\ y_0 \end{bmatrix} = (I - A(0)^{-1})^{-1}(I - [A(0)^{-1}]^{T+1})Di_{low}. \tag{39}$$

Because the nominal interest rate does not respond to inflation during the peg the matrix $A(0)^{-1}$ has one eigenvalue above and the other below 1. Therefore, π_0 diverges to $+\infty$ when T tends to $+\infty$. Indeed, $[A(0)^{-1}]^{T+1}$ explodes with T and all the coefficients of matrix $A(0)^{-1}$ are positive while those of D are negative. So there exists an upper bound on T, $\bar{T}$ such that if the duration of the peg is above this threshold the allocation is not the outcome of a LC-equilibrium.

However, it is easy to see that the allocation $\pi_t = i_{low}$ for any t is the outcome of a LC-equilibrium if the peg lasts forever and if i_{low} is sufficiently close to zero (otherwise inflation is too low and the central banker prefers deviating from this policy plan).

C.7 Proof of Proposition 8.

Let suppose that the Taylor rule is always satisfied. Then, following previous subsection, we can write the model as follows:

$$\begin{bmatrix} E_t\pi_{t+1} \\ E_t y_{t+1} \end{bmatrix} = A(\phi) \begin{bmatrix} \pi_t \\ y_t \end{bmatrix}.$$

When the Taylor principle is not verified there is one eigenvalue lower than one and one above one, we call them $\lambda_1 < 0$ (associated with the vector V_1) and $\lambda_2 > 0$ (associated with the vector V_2) respectively. We thus can decompose $[\pi_t \quad y_t]' = V_1 w_{1,t} + V_2 w_{2,t}$ where:

$$E_t w_{i,t+1} = \lambda_i w_{i,t}.$$

Therefore, using the argument developed in the proof C.5, we see that to be a LC-equilibrium one might impose $w_{2,t} = 0$. Therefore, all the LC-equilibria consistent with a Taylor rule which does not satisfy the Taylor principle satisfy:

$$[\pi_t \quad y_t]' = V_1 w_{1,t}, \texttt{ where } w_{1,t+1} = \lambda_1 w_{1,t}. \tag{40}$$

We can index each equilibrium by the initial (arbitrary) condition $\pi_0 = V_{11}\pi_0$, where V_{11} is the first member of the column vector V_1. Thus,

$$\pi_t = \lambda^t \pi_0.$$

In addition, since $y_t = V_{12} w_{1,t} = V_{12}\lambda_1^t w_{1,0} = \frac{V_{12}}{V_{11}}\lambda_1^t \pi_0$, i.e.

$$y_t = \lambda^t \omega \pi_0, \texttt{ with } \omega = \frac{V_{12}}{V_{11}}.$$

However, such an allocation is the outcome of a LC-equilibrium if and only if the initial condition π_0 is not too large (see Proposition 4).

The role of the inflation bias. Let suppose that the initial level of inflation is π_0 and the level of output gap is $y_0 = \frac{V_{12}}{V_{11}}\pi_0$. This is a situation in which the economy converges to the zero-inflation allocation. The maximal and minimal values for inflation correspond to levels of inflation such that the sustainability constraint (equation (SUST)) is exactly binding.

$$\sum_{t=0}^{\infty} \beta^t[\pi_t^2 + \lambda(y_t - y^*)^2] = [\frac{\lambda\kappa}{\lambda+\kappa^2}y^* + \frac{\lambda\beta}{\lambda+\kappa^2}\pi_1]^2 + \lambda[\frac{\lambda}{\lambda+\kappa^2}y^* - \frac{\kappa\beta}{\lambda+\kappa^2}\pi_1 - y^*]^2 + \frac{\beta}{1-\beta}W^d,$$

where W^d denotes the loss under discretion. One can replace $\pi_t = \lambda_1^t \pi_0$ in this expression and find a quadratic equation in y^* and π_0:

$$a\pi_0^2 + b\pi_0 y^* + c{y^*}^2 = 0,$$

If the inflation bias is absent, then π_0 is necessarily zero as well and the minimum and maximum levels of inflation are zero as well. Otherwise, we can divide this quadratic equation by ${y^*}^2$ and we find a quadratic polynomial in π_0/y^*. First, Assumption 2 guarantees that if π_0 is zero the path is sustainable, this translates into $c < 0$. Second, after some computations we can prove that a is always positive (this is consistent with Proposition 4 that shows that inflation is always bounded in a sustainable equilibrium). Indeed, coefficient a is as follows:

$$a = \frac{1}{1-\beta\lambda_1^2}\left[1 + \lambda\left(\frac{V_{12}}{V_{11}}\right)^2\right] - \lambda_1^2\lambda\beta,$$

which is positive. Thus, π_0/y^* satisfies a quadratic polynomial which is negative in 0 and tends to $+\infty$ for very large and negative values. The quadratic polynomial has two roots of opposite signs that we denote by $x_1 < 0$ and $x_2 > 0$. Consequently, the lower and upper bounds ($\underline{\pi}$ and $\bar{\pi}$) satisfy:

$$\underline{\pi} = x_1 y^* \text{ and } \bar{\pi} = x_2 y^*$$

Therefore, the upper and the lower bounds for inflation are directly proportional to the inflation bias.

C.8 Proof of Proposition 9.

Let first show that if shocks are bounded, inflation is bounded in any LC-equilibrium. Suppose that there exists an allocation with a very large level of inflation at date t. This high level of inflation is either associated with a high output gap or high inflation expectations.

First case: high inflation expectations. Since the shocks are bounded, the loss under the time-invariant discretionary allocation is bounded as well. If inflation expectations are really high they thus imply a loss that dominates the loss under the time-invariant discretionary allocation which is impossible since it is the worst LC-equilibrium.

Second case: high output gap and low inflation expectations. If inflation expectations are low, the best deviation policy triggers a mild loss while the current loss is very high; the central banker has thus any incentives to deviate (even if the continuation value of the loss is zero).

In the two cases, the allocation with unbounded inflation path is not the outcome of a LC-equilibrium.

We now deal with the AR(1) with L-2 innovations shocks. It is easy to see that under this assumption the loss under the time-invariant discretionary policy is bounded. Therefore, we can apply the exact same reasoning than before.

C.9 Proof of Proposition 10.

Because shocks are i.i.d. and zero mean expected inflation and output gap are zero if the central bank follows a Taylor rule forever. Consequently, we get

$$\pi_t^{TR} = \frac{1}{1+\alpha\kappa/\sigma}(u_t - \kappa/\sigma(\epsilon_t - r_t^n)), \tag{41}$$

$$y_t^{TR} = \frac{1}{1+\alpha\kappa/\sigma}(-\alpha/\sigma u_t - 1/\sigma(\epsilon_t - r_t^n)). \tag{42}$$

$$\tag{43}$$

While inflation and the output gap under the best deviation policy are:

$$\pi_t^{BD} = \frac{\lambda\kappa}{\kappa^2+\lambda}(y^* + u_t/\kappa), \tag{44}$$

$$y_t^{BD} - y^* = -\kappa/\sigma\pi_t. \tag{45}$$

$$\tag{46}$$

Furthermore, an allocation implemented by the Taylor rule is a LC-equilibrium if and only if:

$$[\pi_t^{TR}]^2 + \lambda(y_t^{TR} - y^*)^2 + \beta L^{TR} \leq [\pi_t^{BD}]^2 + \lambda(y_t^{BD} - y^*)^2 + \beta L^{BD}, \tag{47}$$

where L^{TR} and L^{BD} are the losses associated with the allocation implemented by the Taylor rule policy while and the time-invariant discretionary equilibrium respectively. Since shocks are i.i.d. these losses are independent of realized shocks at date t.

Finally to prove Proposition 10, we simply need to prove that (47) is verified for small shocks. This can be proved by continuity. When the variance of shocks tends to zero and the size of the current shock tends to zero as well this inequality converges toward the inequality ensuring that the zero-inflation allocation is a LC-equilibrium.

One can remark that as if we suppose small variance of shocks, then this inequality is violated as soon as one of the three shocks is too large. For the monetary policy shock and the natural rate of interest shock this is trivial since the right-hand-side of the inequality is unaffected by these shocks while the right-hand-side first two members increase with the size of these shocks. For the cost-push shock, one can use a Taylor expansion to show that the left-hand-side members diverge more quickly than the right-hand-side toward $+\infty$.

C.10 Proof of Proposition 11.

When the monetary policy shock is zero, equations (NKPC) and (EE) are modified as follows:

$$\pi_t = \beta(1-p)\pi_{t+1} + \beta p\bar{\pi}^D + \kappa y_t, \tag{48}$$

$$y_t = (1-p)y_{t+1} + (1-p)\bar{y}^D - 1/\sigma(\phi\pi_t - (1-p)\pi_{t+1} - p\bar{\pi}^D), \tag{49}$$

where variables π_t and y_t denote the sequence of inflation and the output gap in the absence of shocks at date t, scalars $\bar{\pi}^D$ and $\bar{y}^D$ denote the average inflation and output-gap in the discretionary allocation in case of non-zero monetary policy shock.

We can remark that all forward-looking variables are multiplied by $(1-p)$ so all the eigenvalues of the dynamic system are multiplied by $(1-p)$. The dynamic system is unstable if and only if the lowest eigenvalue is largest than 1 thus the economy diverges for any non-zero initial conditions (absent of shocks) if and only if $\phi > 1 - p$.

In addition, the unique stable solution is given by:

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = [I - A(1-p)]^{-1} Ap \begin{bmatrix} \kappa \\ 1 \end{bmatrix} \frac{\lambda y^*}{\lambda + \kappa^2} \tag{50}$$

We can observe that an increase in the probability of large shock increases inflation and the output gap (even in absence of shocks) through private agents' expectations. It also reduces long-term gains to follow the Taylor rule because the probability of following it decreases with p.

Therefore, there exists a probability $\bar{p}$ such that following a Taylor rule –even in the absence of contemporaneous shock– is always suboptimal. The central banker never follows the Taylor rule when the probability p is above this threshold.

The higher the probability of a large shock, the less incentives the central banker has to stick to the Taylor rule. There exists a thresholds, $\bar{p}$, such that when the probability of large shocks is above the threshold $\bar{p}$, the central banker never follows the Taylor rule. There are two reasons: first, inflation and the output gap are closer and closer to the discretionary allocation when p increases, second the long-term gains of the Taylor rule decreases with p (it is as if the discount factor of the central banker is lower).