Universidade Federal do Rio Grande do Sul
Faculdade de Ciências Econômicas
Departamento de Economia e Relações Internacionais
Trabalho de Conclusão de Curso

Time Inconsistency in Monetary Policy: A Discussion of Solutions

Trabalho de Conclusão de Curso

Pedro Henrique Zecchin Costa

Porto Alegre, RS
2021

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
DEPARTAMENTO DE ECONOMIA E RELAÇÕES INTERNACIONAIS
TRABALHO DE CONCLUSÃO DE CURSO

TIME INCONSISTENCY IN MONETARY POLICY: A DISCUSSION OF SOLUTIONS

PEDRO HENRIQUE ZECCHIN COSTA

Work presented in partial fulfillment of the requirements for the degree of Bachelor in Economics.

Advisor:
Prof. Ronald Otto Hillbrecht

PORTO ALEGRE, RS
2021

Costa, Pedro Henrique.
Time Inconsistency in Monetary Policy:
A Discussion of Solutions / Pedro Henrique Zecchin
Costa. -- 2021.
62 f.

Orientador: Ronald Otto Hillbrecht

(Trabalho de Conclusão de Curso) - Universidade
Federal do Rio Grande do Sul, Faculdade de Ciências
Econômicas, Departamento de Economia e Relações
Internacionais, Porto Alegre, BR-RS, 2021.

Rules versus Discretion, Time Inconsistency,
Inflation Bias, Monetary Policy I. Hilbrecht, Ronald.,
orient.

PEDRO HENRIQUE ZECCHIN COSTA

TIME INCONSISTENCY IN MONETARY POLICY: A DISCUSSION OF SOLUTIONS

Work presented in partial fulfillment of the requirements for the degree of Bachelor in Economics.

Advisor:
Prof. Ronald Otto Hillbrecht

Aprovado em: Porto Alegre, ________de ________ de 2021

BANCA EXAMINADORA:

____

Prof. Dr. Ronald Otto Hillbrecht – Orientador
UFRGS

____

Prof. Dr. Marcelo de C. Griebeler
UFRGS

____

Prof. Dr. Marcelo Savino Portugal
UFRGS

ACKNOWLEDGMENTS

This work wouldn't be possible, quite literally, without the people that brought me to life, Beatriz and João. For this gift, there are no words that can sum up my gratitude. I need to try, though. So: thank you.

I want to dedicate this work to my family, friends, and professors, without whom I wouldn't be who I am today. Special thanks to Dani, Tom, and Túlio for our longstanding friendship. Shout-out to Kingira: Bê, Guerra, Rafa, Rodney, and Tuio. UFRGS certainly wouldn't have been such fun without you guys. Finally, I want to thank my favorite sister: Julia. Thank you all for transforming me into a better economist, person, and friend.

"All men are physico-chemically equal"

— Aldous Huxley, Brave New World

ABSTRACT

Kydland and Prescott (1977) demonstrate that the inability of a discretionary policymaker to commit to a future plan of action may lead to a suboptimal result. A promise to follow a better outcome is time-inconsistent: the policymaker has incentives to ex-post deviate from the announced policy. This study discusses how rules, rather than discretion, can mitigate the time-inconsistency problem in monetary policy.

The second chapter builds a canonical model to demonstrate that the discretionary policy leads to inflation bias. The model is then modified to examine solutions suggested by the relevant literature. The third chapter discusses the trade-offs associated with institutional solutions, in which the policymaker alters the behavior of the monetary authority. More specifically, I consider three rule-based solutions: a zero-inflation rule, an independent central bank, and a state-contingent rule. Finally, I investigate how reputation can alleviate the time consistency problem even with a discretionary policymaker.

Keywords: Rules versus Discretion, Time Inconsistency, Inflation Bias, Monetary Policy

RESUMO

Kydland and Prescott (1977) mostram que a incapacidade de um *policymaker* discricionário em se comprometer a um plano de ação futuro pode levar a um resultado subótimo. Prometer seguir um resultado melhor é temporalmente inconsistente: o *policymaker* tem incentivos para desviar, ex-post, da política anunciada. Esse estudo discute como regras, em vez de discrição, podem atenuar o problema de inconsistência temporal em política monetária.

O segundo capítulo constrói um modelo canônico para demonstrar como a política discricionária leva a um viés inflacionário. O modelo é subsequentemente modificado para examinar soluções sugeridas pela literatura relevante. O terceiro capítulo discute os *trade-offs* associados a soluções institucionais, nas quais o *policymaker* altera o comportamento da autoridade monetária. Mais especificamente, considero três soluções: uma regra de inflação zero, um banco central independente, e uma regra condicional. Por último, investigo como reputação pode diminuir o problema de inconsistência temporal mesmo com uma política discricionária.

Palavras-chave: Rules versus Discretion, Time Inconsistency, Inflation Bias, Monetary Policy

List of Symbols

y_t	Real output at period t
y_n	Natural output
π_t	Inflation at period t
α	Output sensibility to excess inflation
π_t^e	Expected inflation at period t
σ^2	Offer shock's variance
ε_t	Offer shock at period t
y^*	Socially optimal real output
κ	Gap between optimal and real output
φ	Relative preference parameter of society
$\tilde{\varphi}$	Relative preference parameter of the monetary authority
τ	Cost for the policymaker to overrule the monetary authority decision
δ	Intertemporal discount rate
$\bar{\pi}$	Announced future inflation

List of Figures

Contents

1. INTRODUCTION

Monetary policy is a primary instrument to manage aggregate demand in the short run. Given its importance, economists have debated for decades (and still do) the best ways to conduct monetary policymaking. Embedded into this discussion is the debate of rules versus discretion. Should the monetary authority be bounded to a predetermined plan of action? Or should it have discretionary power to decide the best policy at the moment, ad-hoc? The debate soon evolved to other related issues among economists and political scientists. Should the central bank be independent? What's the optimal governance structure for a monetary authority? How accountable should the central bank be? How to shield policymaking from short-sighted voters-seekers politicians?

Kydland and Prescott (1977) pioneered the debate by showing that discretionary policy often produces an inefficient equilibrium. The inefficiency arises from the incapacity of the discretionary policymaker to ex-ante credibly commit to the socially optimal policy. A promise to pursue this optimal policy is time-inconsistent: the policymaker has ex-post incentives to deviate from the announcement. In monetary policy, the inefficient outcome emerges as higher than optimal inflation with no output benefit. This excess inflation is called inflation bias. There are many possible theoretical explanations for the existence of an inflation bias, and I will briefly mention some, but my objective here is specific. I want to focus on the problem of dynamic time inconsistency.

This study guides the reader throughout the rules versus discretion debate, developing the relevant models. I intend to do this in a didactic approach, using a standard notation across models, and stressing the relevance and robustness of the assumptions. I also organized an extensive mathematical appendix, explaining most of the mathematical steps necessary to arrive at the exposed equations. With the help of this appendix, a reader with basic mathematical skills should be able to reach the results themselves.

First, I explain why the time inconsistency problem emerges in general and subsequently develop a canonical model to demonstrate it in monetary policy. Then I start addressing solutions in which the policymaker alters the behavior of the monetary authority. I build three models exploring these solutions: a zero-inflation rule, an independent conservative central bank, and a state-contingent rule. Next, I consider how reputation might alleviate the problem. I will develop two game theory models that build around the idea of reputation. Finally, I discuss the relevance of time inconsistency today.

2. Time Inconsistency under the Discretionary Policy

This chapter is divided into two sections. The first introduces the reader to the time inconsistency problem, explaining under which conditions it might emerge. The second builds a canonical model that demonstrates how it appears under discretionary policymaking in monetary policy.

2.1 Why discretion is suboptimal

Kydland and Prescott (1977) famously show that even a benevolent policymaker may lead to an inefficient outcome by behaving discretionarily. It's worthwhile to precise what discretion means in this discussion: it means that the policymaker can choose (at his discretion) a policy given the current circumstances. It's from this discretion that emerges the difficulty of the policymaker to credibly commit to a plan of action. The commitment may be time-inconsistent: when the time to fulfill the action comes, it may be optimal to deviate from the promised policy, given the circumstances.

One could look at this result and not be impressed, "politicians are always lying to us," some could say. But the policymaker is assumed to be benevolent. There shouldn't be a conflict of interests. Or so it seems. Even a policymaker aligned with the social preferences might not act as promised. Why would a policymaker do such a thing? The short answer is that he would do it for the benefit of society. After all, he is benevolent.

The problem with this tendency to look for social welfare is that it can backfire. By attempting to do good, the discretionary policy will, instead, do harm. Why? So far, I have talked about one agent: the policymaker. But there is another essential one: rational private agents. Rationality doesn't imply that they have perfect eyesight and can predict the future accurately. It only means they observe how the policymaker acts, forming expectations endogenously. And, as with most economic models, their expectations matter. They matter because it's their interaction with the policymaker's actions that will produce the equilibrium outcomes. It's from this interplay between the rational expectations and the discretionary policy that emerges the inefficient result.

Why would private agents form expectations that would decrease social welfare? We will see formally how this happens, but the intuition is that rational agents don't like being deceived, even for their own good.

Because the policymaker is discretionary, he always sets the optimal policy given the already-formed private expectations. The problem is that the discretionary policy doesn't consider its effects on expectations. The private agents know how the policymaker behaves, so they form expectations such that the policymaker hasn't incentives to deceive. The final result is inefficient: there's an equilibrium in which everyone would be better off. But this Pareto superior equilibrium isn't attainable through discretionary policymaking.

The broad picture is: in a dynamic system with rational agents, the policymaker's behavior affects the expectations. By not considering its effects on expectations, the discretionary policy might worsen the results.

Kydland and Prescott (1977) suggest how to solve the problem: rules. Time inconsistency only exists because the policymaker has discretionary power to decide, ad-hoc, which policy to follow. The inefficiency derives from the policymaker's inability to commit credibly to a future plan of action. If he obeyed a binding rule, from which he couldn't deviate, then there's no time inconsistency. That's why their seminal paper is considered the precursor of the rules versus discretion debate.

That said, it's important to mark that rule-based policy wasn't a new idea. In the field of monetary theory, we can trace back the defense of policy rules back to Fisher (1919) or even Wicksell (1907). Friedman (1948) famously advocates a constant growth rate for the money supply. The difference relies on why Kydland and Prescott (1977) defend rules: avoiding a time inconsistency problem in a dynamic setup with rational agents.

In practice, there are several reasons to support rules rather than discretion besides time inconsistency. Taylor (2017), in a review of rule-based policy, lists some additional reasons, such as less short-run political pressure, reduction of uncertainty, and greater accountability. For an extensive survey on the technical detail for rules, the reader can see McCallum (1999) and Taylor and Williams (2010). My focus here will be around the time inconsistency.

I want to stress that the time inconsistency problem isn't exclusive to monetary policy. It applies to a broad range of real-world problems. It can occur whenever an agent (not necessarily a policymaker) cannot credibly commit themselves to future behavior. I will comment on some examples in the chapter prior to the conclusion.

2.2 DISCRETIONARY MONETARY POLICY

The latter section laid the intuition behind the time inconsistency problem under the discretionary policy. This section formalizes the discussion around a canonical model showing how this problem arises in monetary policymaking. This model is the foundation from which the future models will develop.

The model presented is very similar to the one developed by Barro and Gordon (1983a). They weren't the first to apply the time inconsistency problem to monetary policy: Kydland and Prescott (1977) and also Calvo (1978) expose similar models.

Real output follows a typical expectations-augmented supply curve:

$$y_t = y_n + \alpha(\pi_t - \pi_t^e) + \varepsilon_t \tag{2.2.1}$$

Where ε_t are offer shocks that follow a normal distribution with zero mean and variance σ^2. As usual, inflation is expansionary: an inflation rate above the expected level will push output above its natural level y_n. α is a positive parameter that measures how sensitive output is to excess inflation.

This specification, together with rational expectations, implies that, in equilibrium, the expected output must equal the natural level. This must be true because rational expectations require that $\pi_t^e = \mathbb{E}[\pi_t]$. Such that:

$$\begin{aligned} \mathbb{E}[y_t] &= \mathbb{E}\left[y_n + \alpha(\pi_t - \pi_t^e) + \varepsilon_t\right] \\ &= y_n + \alpha\mathbb{E}(\pi_t - \pi_t^e) + \mathbb{E}[\varepsilon_t] \\ &= y_n \end{aligned} \tag{2.2.2}$$

There are only two agents: a policymaker that controls monetary policy and private agents that form inflation expectation. As mentioned earlier, the policymaker is benevolent, meaning that its objective is to maximize social welfare. To build a social welfare function for our models, it's common to postulate that people dislike inflation and would like the output to be around the socially optimal level y^*. The functional form of the social loss function (such that the higher the value of the loss function, the lower the social welfare) is:

$$\mathcal{L} \equiv \pi_t^2 + \varphi(y_t - y^*)^2 \tag{2.2.3}$$

Where φ represents the relative preference that society gives to income stabilization around y^* in detriment of inflation-fighting. For example, a society with a low value of φ doesn't care much about output stabilization and prefers lower inflation. The implicit inflation bliss-point is zero, but this can be relaxed by changing the term π_t^2 with $(\pi_t - \pi^*)^2$, where π^* is the inflation bliss-point [1]. Naturally, this specification of the loss function is a simplification, and it's possible to take different forms. Barro and Gordon (1983a) actually work with a linear welfare function, instead of a quadratic one [2]. Also, in real life, the private agents probably have different preferences. For example, someone in debt would like a bit more inflation, such that the real value of the debt is smaller. The use of a social welfare function hides this heterogeneity. I mention this to remind the reader to be skeptical of the "everyone is better off" argument. The Pareto superiority of the outcomes is a consequence of using the welfare function that abstracts from complexity.

An essential assumption is that the socially optimal output is higher than the natural output, that is, $y_n < y^*$. There are several reasons to justify this assumption. Barro and Gordon (1983a) mention that inefficiencies caused by taxation make the natural level to be lower than the efficient level. Clarida et al. (1999) and other New-Keynesian models work under an imperfect competition framework, which also creates similar inefficiencies. Likewise, Cukierman (1992) argues that unions keep real wages above the market-clearing level. To simplify the notation, I denote this gap between the optimal level and the natural one as $\kappa \equiv y^* - y_n > 0$.

The timing of events:

1. Private agents form π_t^e
2. ε_t is realized
3. Policymaker sets π_t

The timing structure is crucial for two main reasons. First, when the

[1] Society could want a marginally positive rate of inflation, see Schmitt-Grohé and Uribe (2010) for a discussion of the optimal inflation rate.

[2] See Cukierman (1992) for a discussion on the different types of welfare functions.

policymaker decides which level of inflation to set, the private agents have already formed the expectations. The policymaker, therefore, considers inflation expectation as given. Second, because the policymaker observes the offer shock, he can mitigate the effects on output.

As an alternative specification, it's possible to work with wage formation instead of inflation expectation. Private agents sign one-period wage contracts rationally, hence they can't rapidly react to offer shocks, but the policymaker can. This approach is interesting because it makes it explicit that this timing implies a one-period nominal rigidity. I made this digression to explain that it's from this timing that derives the non-neutrality of monetary policy in the short-term, even though we have rational agents.

The objective of the policymaker is to minimize social loss (or, equivalently, maximize social welfare). He can manipulate the inflation level to achieve this goal. I assume that the policymaker can set whatever rate of inflation perfectly. It would be possible to add instrument uncertainty[3], but I will stick with the simplification. So the policymaker solves the following problem:

$$\min_{\pi_t} \quad \pi_t^2 + \varphi(y_t - y^*)^2 \tag{2.2.4}$$

We can substitute y_t with the supply curve (2.2.1) and solve the optimization problem. As mentioned in the introduction, the appendix contains detailed explanations for most of the mathematical steps. This way, the main text is not polluted with mathematical development, and the curious reader can also be satisfied. Look for a clickable link after the equation to go to the appendix. The first-order condition of the optimization problem (2.2.4) yields:

$$\pi_t = \frac{\varphi\alpha}{1 + \varphi\alpha^2}\left[\kappa + \alpha\pi_t^e - \varepsilon_t\right] \tag{2.2.5}$$

Appendix A.1.1

We can interpret this expression as the reaction function of the policymaker, given the parameters and the observed variables. For example, a negative offer shock makes the policymaker react by increasing inflation, such that the effect of the shock on output is dampened.

[3] For example, the monetary authority uses a policy instrument (the growth of the money supply), but demand shocks affect the final equilibrium. See Walsh (2017) for such an approach.

The private agents know how the policymaker acts, and form expectations accordingly. The assumption of rational expectations implicitly set their loss function as:

$$\mathcal{L}^p \equiv (\pi_t - \pi_t^e)^2 \tag{2.2.6}$$

In practice it means that the agents will take the expected value of the monetary authority reaction function. Remember that they don't observe the offer shock. Solving the problem of the private agents we obtain:

$$\pi_t^e = \varphi\alpha\kappa \tag{2.2.7}$$

Appendix: A.1.2

Now that we have an expression for the inflation expectation, we can put it in the monetary authority reaction function to find equilibrium inflation. We can do the same for the supply curve to find equilibrium output:

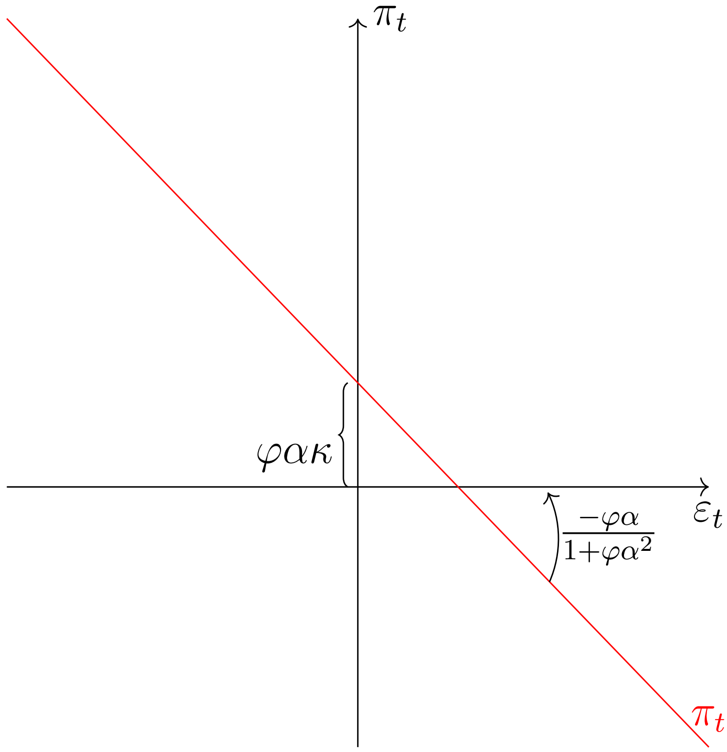
$$\begin{aligned} \pi_t &= \varphi\alpha\kappa - \frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t \\ y_t &= y_n + \frac{1}{1+\varphi\alpha^2}\varepsilon_t \end{aligned} \tag{2.2.8}$$

Here we can already see one of the main results highlighted by Kydland and Prescott (1977). The equilibrium inflation has a part ($\varphi\alpha\kappa$) fully predicted by the private agents. Consequently, this part has no expansionary effect. This excess inflation that has no benefit is the inflation bias.

Observe what parameters affect the inflation bias ($\varphi\alpha\kappa$). If society puts a higher weight on output stabilization ($\uparrow \varphi$), it leads to a higher inflation bias. The higher the sensibility of output to surprise inflation ($\uparrow \alpha$), the higher the inflation bias. Also, a larger gap between the natural and the optimal output ($\uparrow \kappa$) leads to more inflation bias.

It's easy to note that society is better off if the monetary authority would set $\pi_t = -\frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t$. As I will comment afterward, this is the socially optimal rule. So why doesn't the policymaker do that? And remember: the policymaker is benevolent. To demonstrate this, let's assume that people believed that the expected inflation would

FIGURE 2.1. Inflation under the discretionary policymaker



be zero. But then the policymaker could improve social welfare by increasing inflation to boost output. Given $\pi_t^e = 0$, the optimal reaction of the policymaker is to set:

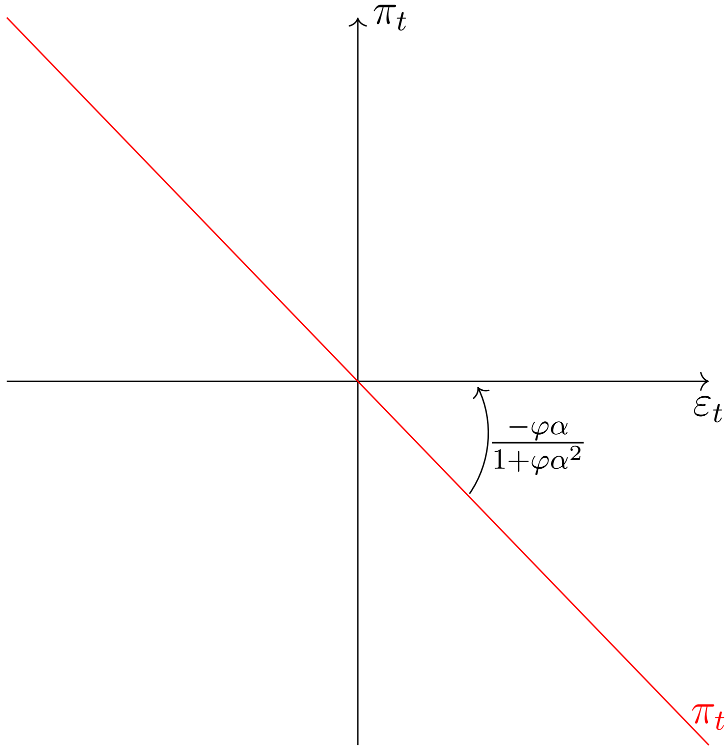
$$\pi_t = \frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa - \varepsilon_t\right] \tag{2.2.9}$$

The policymaker has incentives to produce surprise inflation. But this can't be an equilibrium because agents are rational. The private agents know the monetary authority would want to set higher inflation if their expected inflation is too low. So they raise the inflation they expect until the policymaker has no incentive to deceive the private agents. That's why Barro and Gordon (1983a) consider discretionary monetary policymaking inefficient. The average inflation produced is higher than it should be, without any advantage.

Before moving to potential solutions for the inflation bias, it's worthwhile to find a benchmark. What is the socially optimal level of inflation? In this case, the monetary authority would consider the expected inflation as endogenous. The socially optimal policy provides the same stabilization of offer shocks, but without the inflation bias:

$$\begin{aligned}\pi_t &= -\frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t \\ \pi_t^e &= 0 \\ y_t &= y_n + \frac{1}{1+\varphi\alpha^2}\varepsilon_t\end{aligned} \tag{2.2.10}$$

FIGURE 2.2. Inflation under the socially optimal rule



It's easy to see that this result is better than the discretionary equilibrium, but we can see it more concretely by computing the expected welfare loss of each scenario. We need to input the equilibrium outcomes into the loss function and take the expected value of it. The expected welfare loss associated with discretionary monetary policy yields:

$$\mathbb{E}[\mathcal{L}] = \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa^2\left(1+\varphi\alpha^2\right)^2 + \sigma^2\right] \tag{2.2.11}$$

Appendix: A.1.3

And the expected welfare loss associated with the socially optimal equilibrium yields:

$$\mathbb{E}[\mathcal{L}] = \varphi\kappa^2 + \frac{\varphi}{1+\varphi\alpha^2}\sigma^2 \tag{2.2.12}$$

Appendix: A.1.4

This socially optimal rule cannot be the equilibrium because of the incentives faced by the discretionary monetary authority. Nonetheless, it's useful to have it as a benchmark.

The results exposed in Barro and Gordon (1983a) suggest that discretionary monetary policy is suboptimal, but the authors depend on many assumptions to arrive at this conclusion. One of the crucial assumptions is that the monetary authority wants to set output above the natural level. Blinder (1999), ex-Vice Chairman of the FED, argues that this doesn't happen in practice. That said, with a different model framework it's possible to have an inflation bias even without this assumption. Cukierman and Gerlach (2003) achieve this by assuming that the policymaker is uncertain about the economy's state and has asymmetrical preferences (recessions are worse than booms). I will further discuss the empirical evidence after addressing the independent central bank.

3. INSTITUTIONAL SOLUTIONS

This chapter discusses how the policymaker can alter the behavior of the monetary authority to diminish the time inconsistency problem. I build three different models exploring these institutional solutions: a zero-inflation rule, an independent central bank, and a state-contingent rule.

3.1 ZERO-INFLATION RULE

Since Kydland and Prescott (1977) exposed the time inconsistency problem, economists have sought solutions to solve it. As we will see, there are many (potential) ways to mitigate this problem. I start with the one that Kydland and Prescott (1977) suggested: a simple rule.

I need to introduce another agent: the central bank. In the previous models, I used the terms policymaker and monetary authority (central bank) interchangeably. I could do so because there was no distinction in the objectives of these two agents. Now they are two separate agents. The policymaker is still benevolent, but he can change the behavior of the central bank. The monetary authority, in turn, sets inflation seeking to pursue the objective that the policymaker has determined.

The timing of events:

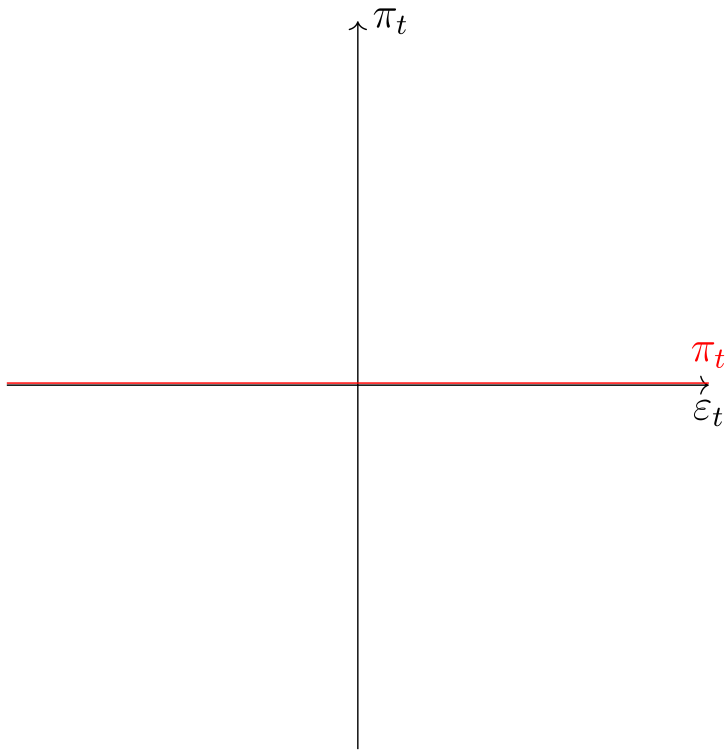
1. Policymaker instructs the monetary authority to set inflation $\pi_t = 0 \; \forall t$
2. Private agents form π_t^e
3. ε_t is realized
4. Monetary authority sets π_t

For this model, the policymaker will instruct the central bank to set inflation at zero[4], always. It's a simple rule that private agents can easily verify. The equilibrium outcomes for this scenario would be:

[4] The social bliss-point of inflation is zero, but if it was $\pi^* > 0$ the rule could be $\pi_t = \pi^* \; \forall t$.

$$\begin{aligned} \pi_t &= 0 \\ \pi_t^e &= 0 \\ y_t &= y_n + \varepsilon_t \end{aligned} \tag{3.1.1}$$

FIGURE 3.1. Inflation under the zero-inflation rule



There's no expected inflation, so there's no inflation bias. The problem is solved! However, mind the catch. The offer shock affects output fully. Compare this with the discretionary equilibrium (2.2.8). There the inflation response has a stabilization effect.

Therefore, by following a strict zero inflation rule, the monetary policy loses its stabilization effect. But it manages to diminish inflation. This rule accomplishes something that the people want (lower inflation) at the cost of something that people don't want (a more volatile output). The result shouldn't shock any economist: there is no such thing as a free lunch, only trade-offs. But the policymaker wonders: which option is better? It's not hard to compute the expected welfare loss associated with the simple rule:

$$\mathbb{E}[\mathcal{L}] = \varphi\left[\kappa^2 + \sigma^2\right] \tag{3.1.2}$$

Appendix A.2.1

The policymaker can compare the expected welfare loss of discretionary policy to the one of a simple rule and see in which situation one trumps the other. A simple zero-inflation rule is better than discretionary policy if and only if:

$$\sigma^2 < \kappa^2(1 + \varphi\alpha^2) \tag{3.1.3}$$

Appendix A.2.2

When the offer shock's variance is below a certain threshold, the gains stemming from lower inflation outweigh the stabilization costs. This result makes intuitive sense. By following a simple rule we have no stabilization policy, but if the variance of shocks is already low, then stabilization isn't that crucial.

Here we compare the outcomes from two extremes (zero-inflation against the discretionary policy), but can't we seek middle ground? In the next section, I build a model that better explores this trade-off between inflation-fighting and stabilization policy.

3.2 INDEPENDENT CENTRAL BANK

In the previous section, we discovered a trade-off between discretionary policy and a zero-inflation rule. Under certain circumstances, one is better than the other. But what if we could seek an intermediary option between these policies? After all, if there is a trade-off, then there is an optimal level.

That's what Rogoff (1985) proposes. In the model, the policymaker chooses the inflation-fighting relative preference of the monetary authority[5]. By changing this parameter, the policymaker affects the central bank's behavior (and, indirectly, the private agent's). In effect, it means that he can move between the outcomes of the discretionary policy and the zero-inflation rule. The (somewhat) shocking result is that it's always optimal for society to have a central bank that weighs inflation-fighting more vigorously than society itself does.

[5]Alesina and Grilli (1991) develop a model in which it's the median voter that chooses this parameter, the results are similar.

The timing of events:

1. Policymaker determines the monetary authority's preference parameter $\tilde{\varphi}$
2. Private agents form π_t^e
3. ε_t is realized
4. Monetary authority sets π_t

The policymaker chooses the relative preference of the central bank $\tilde{\varphi}$. The monetary authority then solves the familiar optimization problem:

$$\min_{\pi_t} \ \pi_t^2 + \tilde{\varphi}(y_t - y^*)^2 \tag{3.2.1}$$

Notice that the only difference so far is that I substituted φ, the true relative preference of society, with $\tilde{\varphi}$, the one the policymaker assigned to the central bank. This assignment is common knowledge, such that the private agents know it. The equilibrium outcomes will be similar (only changing φ with $\tilde{\varphi}$) to the discretionary equilibrium:

$$\begin{aligned} \pi_t &= \tilde{\varphi}\alpha\kappa - \frac{\tilde{\varphi}\alpha}{1+\tilde{\varphi}\alpha^2}\varepsilon_t \\ \pi_t^e &= \tilde{\varphi}\alpha\kappa \\ y_t &= y_n + \frac{1}{1+\tilde{\varphi}\alpha^2}\varepsilon_t \end{aligned} \tag{3.2.2}$$

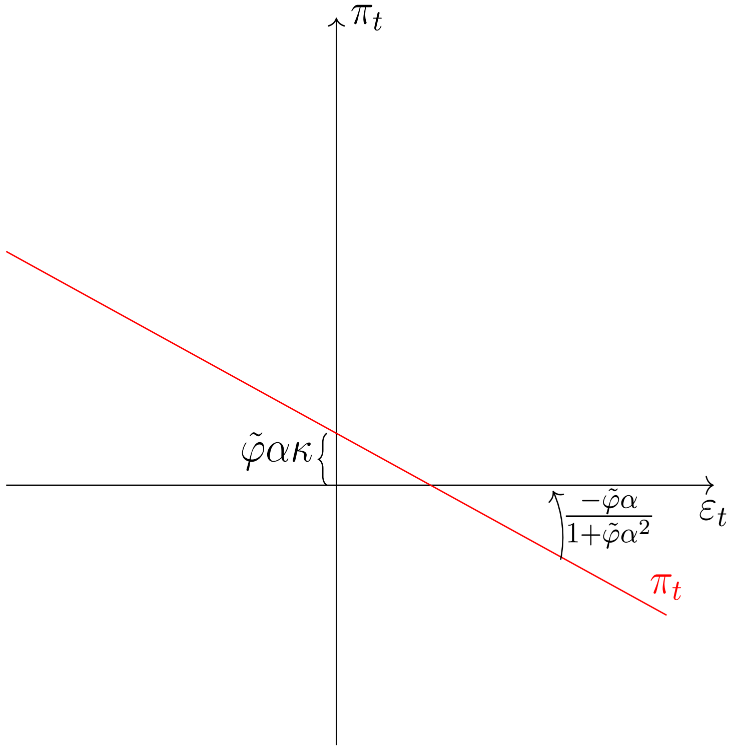
What happens if the policymaker decides to set $\tilde{\varphi} = 0$? Then we are back to the simple rule case (3.1.1). And with $\tilde{\varphi} = \varphi$ we are in the discretionary case (2.2.8). But the policymaker can choose an intermediary scenario with $\tilde{\varphi} \in [0, \varphi]$. The policymaker assigns the optimal $\tilde{\varphi}$ that minimizes the expected social welfare loss. That is, he solves the following optimization problem:

$$\min_{\tilde{\varphi}} \mathbb{E}\left[\left(\tilde{\varphi}\alpha\kappa - \frac{\tilde{\varphi}\alpha}{1+\tilde{\varphi}\alpha^2}\varepsilon_t\right)^2 + \varphi\left(-\kappa + \frac{1}{1+\tilde{\varphi}\alpha^2}\varepsilon_t\right)^2\right] \tag{3.2.3}$$

Note that this is the social welfare loss function, not the central bank's. Remark the φ instead of $\tilde{\varphi}$ at the beginning of the second term.

Unfortunately, it's impossible to find a simple tractable equation with the optimal parameter $\tilde{\varphi}$ as a function of the other variables and parameters. But in the

FIGURE 3.2. Inflation under the independent central bank



appendix A.3.1 I show that the optimal level is in the interval $(0, \varphi)$. This means that society is better off with a central bank that is more conservative, in the sense that it puts a higher weight on inflation-fighting.

Although it's not possible to find a nice expression for the optimal $\tilde{\varphi}$, it's possible to do some comparative statics. Let us denote $\tilde{\varphi}^*$ as the $\tilde{\varphi}$ such that (3.2.3) is minimized. Then:

$$\begin{aligned} \frac{\partial \tilde{\varphi}^*}{\partial \varphi} > 0 \, , \quad \frac{\partial \tilde{\varphi}^*}{\partial \sigma^2} > 0 \\ \frac{\partial \tilde{\varphi}^*}{\partial \kappa} < 0 \, , \quad \frac{\partial \tilde{\varphi}^*}{\partial \alpha} < 0 \end{aligned} \tag{3.2.4}$$

Rogoff's solution - the policymaker nominating a more conservative (inflation-focused) central banker - was very influential. One could argue that it influenced the support for independent central banks. The model makes an empirical prediction: countries with independent central banks should have lower inflation and higher output variance. Alesina (1988), Grilli et al. (1991), and many others find that monetary authority independence is associated with low inflation. Indeed, since the '90s, the institutional design of central banks has shifted towards greater independence[6]. But,

[6]See Cukierman (1996).

contrary to what Rogoff's model suggests, the link between independence and real indicators doesn't seem to be very strong. Alesina and Summers (1993) find little evidence that central bank independence benefits or harms economic performance.

A related branch of literature is the political economy of monetary policy. The models I develop here all assume that policymakers are benevolent, a strong assumption. What happens if political actors with short-term goals control policymaking? Most of the political economy models predict that the policymakers will distort policy to reap short-horizon gains. Nordhaus (1975) famously showed that politicians have incentives to attempt expansionary policies when seeking reelection. See Persson and Tabellini (1999) for a survey of the literature. I mention this here because it's another justification to isolate monetary policymaking from policymakers.

I want to end this section with a provocation. We assume that the policymaker delegates monetary policy to an independent agent. Couldn't the policymaker discretionarily take this independence back? Is the delegation decision itself time consistent? Imagine, for instance, that the policymaker has decided to grant monetary policy to an independent central bank that highly values inflation-fighting. But then an extremely negative offer shock hits the economy. The policymaker has incentives to overrule the independence of the monetary authority and accommodate the shock. This possibility is absent in this model because we assumed that the policymaker is not able to do that. This assumption is not very realistic given that, in practice, policymakers have the discretion to make these changes. I relax this assumption in the next section.

3.3 STATE-CONTINGENT RULE

In the previous section, I explored one famous solution for the inflation bias problem: the independent conservative central bank. The optimal choice of conservatism trades off the lower inflation with the distorted stabilization policy. The result is a higher expected social welfare. Take a careful look at this previous sentence. Note the word "expected." If the offer shock is large, then society can be worse off with the conservative central bank. We can see this somewhat reflected in the model by observing that the higher the volatility of the offer shocks σ^2, the more aligned the monetary authority will be with society ($\tilde{\varphi}$ near φ).

Flood and Isard (1989) and also Lohmann (1992) address this problem explicitly with a state-contingent rule. The policymaker nominates a central banker with a relative preference coefficient $\tilde{\varphi}$, but the monetary authority isn't entirely independent. The policymaker reserves itself the right to overrule the monetary authority decision and set inflation directly. However, when the policymaker does that, society loses $\tau > 0$ units of welfare.

We can think of τ as influenced by exogenous factors, such as a reputation cost of overruling the monetary authority. But τ could arguably be endogenous: the policymaker can influence how costly it is to override the central bank's independence through institutional bindings. Granting the central bank de jure independence through legislation can be seen as setting a high τ. It would take a high political cost to change the legislation. Nonetheless, it's hard to justify τ as being infinite as it's implicit in the model of the previous section. In times of crisis, such as war, institutions are often overridden by policymakers, even in strongly democratic countries.

Why would the policymaker overrule the monetary authority? The lower inflation bias conquered through a conservative central bank has a price: a dampened response to offer shocks. So in scenarios where the offer shocks are large (be them positive or negative), it may be optimal to overrule the central bank's decision. That's a state-contingent rule: it's conditional on the magnitude of the offer shock.

We saw in the previous section that we could think of the discretionary model and the simple zero-inflation rule as being specific cases of the independent central bank model. We can think of the independent central bank model as a case of the state-contingent rule model: When the cost of overruling the monetary authority τ is infinite, we're back to the traditional independent central bank. Basically, we're adding a new margin of adjustment.

The timing of events:

1. Policymaker determines the monetary authority's preference parameter $\tilde{\varphi}$
2. Private agents form π_t^e
3. ε_t is realized
4. Policymaker decides to overrule or not
5. π_t is realized

The superscript $\mathcal{I}$ refers to the independent central bank and $\mathcal{O}$ for the overruling scenario. For example, $\pi_t^{\mathcal{I}}$ refers to the inflation the monetary authority would set if it kept its independence.

As in the previous section, the problem that the central bank solves is:

$$\min_{\pi_t} \pi_t^2 + \tilde{\varphi}(y_t - y^*)^2 \tag{3.3.1}$$

The reaction function of the independent central bank, conditional on the inflation expectations and offer shock, is as we previously found:

$$\pi_t^{\mathcal{I}} = \frac{\tilde{\varphi}\alpha}{1 + \tilde{\varphi}\alpha^2} \left[\kappa + \alpha\pi_t^e - \varepsilon_t\right] \tag{3.3.2}$$

That's the reaction for the independent conservative monetary authority. What would be the inflation that the policymaker would set if he decided to overrule? Because the policymaker is benevolent, he would produce inflation to minimize social loss:

$$\min_{\pi_t} \pi_t^2 + \varphi(y_t - y^*)^2 \tag{3.3.3}$$

Which the first-order condition yields the same result of the familiar discretionary monetary policy:

$$\pi_t^{\mathcal{O}} = \frac{\varphi\alpha}{1 + \varphi\alpha^2} \left[\kappa + \alpha\pi_t^e - \varepsilon_t\right] \tag{3.3.4}$$

Let's compute the welfare loss associated with each scenario: the central bank setting inflation $\pi_t^{\mathcal{I}}$, and with the policymaker overruling the decision and producing $\pi_t^{\mathcal{O}}$. Because we're talking about a state-contingent rule, we must not take the expected value.

$$\begin{aligned} \mathcal{L}^{\mathcal{I}} &= \frac{\varphi + \tilde{\varphi}^2\alpha^2}{(1 + \tilde{\varphi}\alpha^2)^2} \left[\kappa + \alpha\pi_t^e - \varepsilon_t\right]^2 \\ \mathcal{L}^{\mathcal{O}} &= \frac{\varphi}{1 + \varphi\alpha^2} \left[\kappa + \alpha\pi_t^e - \varepsilon_t\right]^2 + \tau \end{aligned} \tag{3.3.5}$$

Appendix A.4.1

The central bank keeps its independence and is allowed to set inflation if $\mathcal{L}^{\mathcal{O}} > \mathcal{L}^{\mathcal{I}}$, and this happens if and only if:

$$\tau > \frac{\alpha^2(\varphi - \tilde{\varphi})^2}{(1 + \varphi\alpha^2)(1 + \tilde{\varphi}\alpha^2)^2} \left[\kappa + \alpha\pi_t^e - \varepsilon_t\right]^2 \tag{3.3.6}$$

Appendix A.4.2

Let us formally define the set "$\mathcal{I}$" for the situations in which (3.3.6) is satisfied, such that the offer shock doesn't trigger intervention. Similarly, denote the set "$\mathcal{O}$" for the circumstances in which the ε_t is of such magnitude that the policymaker intervenes.

$$\begin{aligned}\mathcal{I} &= \left\{ \varepsilon_t \in \mathbb{R} : \tau > \frac{\alpha^2(\varphi - \tilde{\varphi})^2}{(1+\varphi\alpha^2)(1+\tilde{\varphi}\alpha^2)^2} \left[\kappa + \alpha\pi_t^e - \varepsilon_t\right]^2 \right\} \\ \mathcal{O} &= \mathbb{R} \backslash \mathcal{I}\end{aligned} \tag{3.3.7}$$

The right-hand side of the equation (3.3.6) can be seen as the welfare benefits of overruling and the left-hand side as the costs. Observe that, as $\tilde{\varphi}$ gets nearer φ, the gains of overruling get smaller, given that the central bank is already very aligned with society. A central bank that is much more conservative than society (large $\varphi - \tilde{\varphi}$) will get overruled more often. To put it more formally:

$$\begin{aligned}\lim_{\tilde{\varphi}\to\varphi} \mathcal{I} &= \mathbb{R} \\ \lim_{\tilde{\varphi}\to\varphi} \mathcal{O} &= \varnothing\end{aligned} \tag{3.3.8}$$

That is, as the central bank gets aligned with society, the set of values of ε_t that trigger overriding becomes smaller. Another exposition for the expression (3.3.6) is the following:

$$\kappa + \alpha\pi_t^e - \Omega \leq \varepsilon_t \leq \kappa + \alpha\pi_t^e + \Omega \tag{3.3.9}$$

Where $\Omega \equiv \frac{1+\tilde{\varphi}\alpha^2}{\alpha(\varphi-\tilde{\varphi})}\sqrt{\tau(1+\varphi\alpha^2)}$. Appendix A.4.2

Figure 3.3 makes it easier to visualize this expression. It's the plot of a generic probability density function for ε_t (remember that $\varepsilon_t \sim N(0,\sigma^2)$). If the shock is in the blue area of the graph, then the condition (3.3.9) is satisfied, and the central bank sets inflation independently. For shocks in the orange area, the policymaker intervenes. Remark the lack of symmetry of the areas. A negative shock triggers intervention more easily than a positive one: a negative shock may trigger intervention while a positive shock of the same magnitude may not.

Now we can specify a rule for inflation:

$$\pi_t = \begin{cases} \dfrac{\tilde{\varphi}\alpha}{1+\tilde{\varphi}\alpha^2}\left[\kappa + \alpha\pi_t^e - \varepsilon_t\right] & \text{if } \epsilon_t \in \mathcal{I} \\ \dfrac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa + \alpha\pi_t^e - \varepsilon_t\right] & \text{if } \epsilon_t \in \mathcal{O} \end{cases} \tag{3.3.10}$$

FIGURE 3.3. Probability density function of ε_t and overruling situations

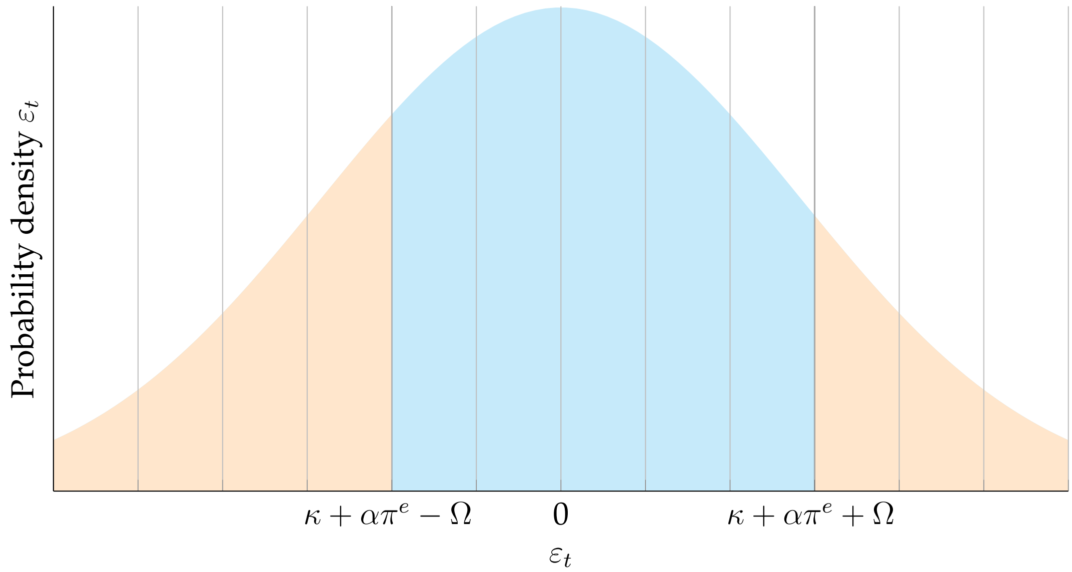
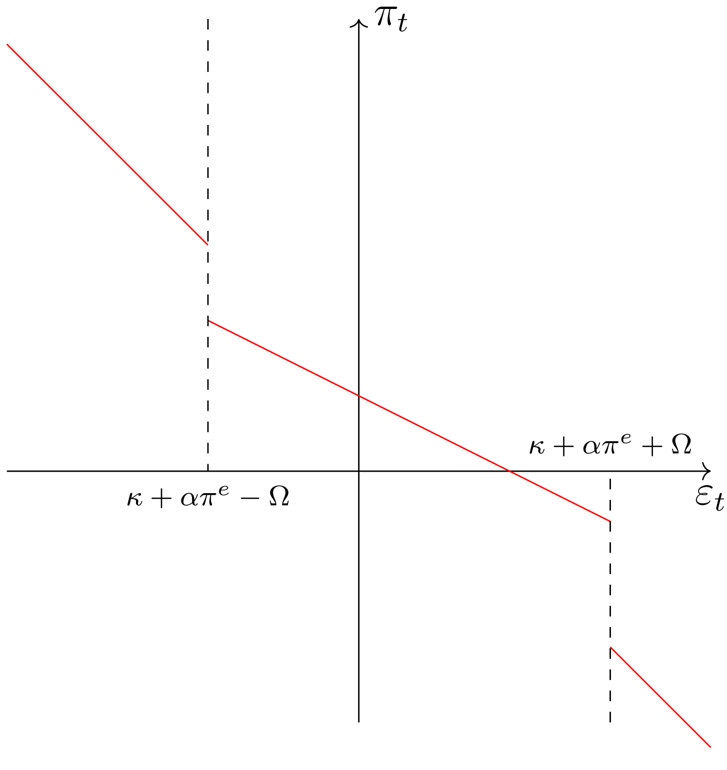


FIGURE 3.4. Inflation under the state-contingent rule



One thing that I haven't addressed so far is the inflation expectation formation. The expectations are a bit more complex in this model because they need to account for the possibility of an overruling (given that it will affect inflation). Let us denote $p \equiv \Pr\{\varepsilon_t \in \mathcal{I}\}$, then we have:

$$\pi_t^e = p\mathbb{E}[\pi_t^{\mathcal{I}}|\varepsilon_t \in \mathcal{I}] + (1-p)\mathbb{E}[\pi_t^{\mathcal{O}}|\varepsilon_t \in \mathcal{O}] \tag{3.3.11}$$

Expected inflation is trickier because the expected value of the shock ε_t conditional on the state of intervention isn't zero. To grasp this, I recommend that the reader take another look at the probability density graph (3.3). The area in which the central bank remains independent skews right (it's centered around $\frac{\kappa+\alpha\pi^e}{2} > 0$). The consequence is that $\mathbb{E}[\varepsilon_t|\varepsilon_t \in \mathcal{I}] > 0$. By the symmetry of the normal distribution, it follows that $\mathbb{E}[\varepsilon_t|\varepsilon_t \in \mathcal{O}] < 0$.

How should we interpret this? The intuitive answer is that the policymaker will usually intervene when negative shocks hit. This explanation is correct, but the reason is that the optimal output is higher than the natural output. Because positive shocks push the output closer to the optimal level (to a certain degree), the policymaker's intervention is asymmetrical.

Accommodation for large shocks and a conservative approach in "normal" times is, arguably, a desirable feature[7]. Having this escape clause allows the society to choose an intermediate option between a fully independent central banker and discretionary policymaking.

I also argue that this model is more realistic than the fully independent central bank model. Just as policymakers have the power to delegate monetary policy, they have the power to take it back. τ can be very high if independence is granted firmly by institutional means, such as through a constitutional amendment, but it's not infinite. And probably shouldn't.

This model can shed light on why countries with low institutional strength (to be interpreted as a low τ) have difficulty maintaining an independent central bank that is much more conservative than society (a high $\varphi - \tilde{\varphi}$). If the $\mathcal{I}$ set is small, such that even moderate shocks trigger intervention, then π_t^e will be very near the discretionary inflation $\pi_t^{\mathcal{O}}$. The benefits of a de jure independent monetary authority only exist if this independence is de facto credible.

[7]See Cukierman and Gerlach (2003) for another model with this feature, but through a very different mechanism.

4. Exploring reputation as a solution

This chapter explores how the inflation bias problem is alleviated if the policymaker has a mechanism to build a reputation. The first section builds a model of an infinitely repeated game that deviates from the assumption of rational agents. The second section shows that even in a finite period game and with rational agents the reputational equilibrium is possible.

4.1 Reputation in an infinite game

Besides institutional arrangements such as an independent central bank, there might be other solutions for the inflation bias problem. Barro and Gordon (1983b) famously explore one of these solutions. In their setup, the policymaker might want to build a reputation of setting low inflation, even being a discretionary agent.

We need to understand how the reputational model differs from the traditional discretionary model to achieve these results. First, the agents play an infinite-period game: the monetary authority [8] setting inflation and the private agents forming inflation expectation. Before agents set their expectations, the central bank announces that inflation will be $\bar{\pi}$ for the next period. The private agents play a tit-for-tat strategy: they start believing the central bank, but if the central bank does not set inflation as promised, then the agents will always expect the discretionary level of inflation. The assumption of a tit-for-tat strategy is essential for this model. It's a departure from rational expectations, but it's a justifiable one given that there are several real situations in which people do seem to play tit-for-tat.

A savvy reader might already grasp the mechanisms that allow the monetary authority to build a reputation. If the central bank honors his promise, then he can keep inflation at the lower announced level. But the monetary authority also has an opportunity in the first period[9]: because inflation expectations are lower than the

[8]Here there's no distinction between the central bank and the policymaker. I will use the terms interchangeably.

[9]Actually this opportunity also exists in the other periods, but because if the central bank would

discretionary level, the one-period best response would be higher than announced inflation. The intertemporal cost of this deception is that they will return to the discretionary equilibrium afterward. The central bank weights the short-term benefits of one-period deception against the long-term benefits of cooperation. Naturally, the decision will depend on how impatient the central bank is. More formally, it means the monetary authority minimizes the intertemporal welfare loss function:

$$\sum_{i=0}^{\infty} \delta^i \mathbb{E}\left[\pi_t^2 + \varphi(y_t - y^*)^2\right] \tag{4.1.1}$$

Where δ is the central bank's intertemporal discount factor.

I need to make a strong simplification to develop the model more easily: there isn't offer shocks. I will comment at the end of this section on why we need this assumption and how things may change by relaxing it. The simplified offer curve for this chapter will be:

$$y_t = y_n + \alpha(\pi_t - \pi_t^e) \tag{4.1.2}$$

Another common simplification, but one I won't be making here, is to use an alternative society loss function linear in output, instead of a quadratic one. The reason for this simplification is that it makes it much easier to find tractable equations. I will stick with our usual loss function for consistency's sake, even though the mathematical expression won't end up so neat. The reader can see Walsh (2017) for the development of this model with the simpler loss function.

The timing of events:

1. Policymaker announces that inflation for next period will be $\bar{\pi}$
2. Private agents form $\pi_t^e = \bar{\pi}$
3. Policymaker sets π_t
4. Private agents observe if policymaker kept his promise and form π_{t+1}^e playing a tit-for-tat strategy
5. The game continues ad infinitum ...

The game starts with the monetary authority announcing that inflation in the next period will be $\bar{\pi}$, lower than the discretionary equilibrium inflation. The private

deceive people, it would certainly do that in the first period because in the future periods the gains would be the intertemporally smaller given the discount factor.

agents believe the announcement, forming expectations such that $\pi_t^e = \bar{\pi}$. Let's first assume that the central bank decided to keep its promise. The expected social loss associated with the cooperation equilibrium is:

$$\mathbb{E}[\mathcal{L}] = \bar{\pi}^2 + \varphi\kappa^2 \tag{4.1.3}$$

Appendix A.5.1

If the central bank maintains the inflation at the announced level in all periods, then the expected intertemporal social loss function is:

$$\sum_{i=0}^{\infty} \delta^i \mathbb{E}\left[\mathcal{L}\right] = \frac{1}{1-\delta}\left[\bar{\pi}^2 + \varphi\kappa^2\right] \tag{4.1.4}$$

Appendix A.5.1

Now let's assume that the monetary authority chooses to break the promise in the first period. The optimal inflation response, given the expected inflation of $\pi^e = \bar{\pi}$ is:

$$\pi_t = \frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa + \alpha\bar{\pi}\right] \tag{4.1.5}$$

We can compute the associated welfare loss of this deception:

$$\mathbb{E}[\mathcal{L}] = \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa + \alpha\bar{\pi}\right]^2 \tag{4.1.6}$$

Appendix A.5.2

It's not hard to show that the one-period deception is better for society (lower welfare loss) than the cooperation equilibrium (demonstrated in the appendix A.5.3). But the central bank needs to consider the consequences of the betrayal. After deception, they return to the discretionary equilibrium:

$$\begin{aligned} \pi_t &= \varphi\alpha\kappa \\ \pi_t^e &= \varphi\alpha\kappa \\ y_t &= y_n \end{aligned} \tag{4.1.7}$$

The expected one-period welfare loss associated with the discretionary equilibrium is the same we already calculated in equation (2.2.11) (without the σ^2):

$$\mathbb{E}[\mathcal{L}] = \varphi\kappa^2(1+\varphi\alpha^2) \tag{4.1.8}$$

The expected intertemporal social loss function associated with an initial deception from the central bank and then going back to discretionary equilibrium is:

$$\sum_{i=0}^{\infty}\delta^i\mathbb{E}\left[\mathcal{L}\right] = \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa+\alpha\bar{\pi}\right]^2 + \frac{\delta}{1-\delta}\left[\varphi\kappa^2(1+\varphi\alpha^2)\right] \tag{4.1.9}$$

Appendix A.5.4

We can represent the one-period payoffs (actually negative payoffs, since we're using a loss function) of this repeated game in a payoff-matrix 4.1. "Trust" means that the private agents expect the announced inflation. If they "Distrust", then they expect the discretionary inflation, that is $\pi_t^e = \varphi\alpha\kappa$. Remember that the implicit welfare function of rational private agents is given by the expression (2.2.6).

FIGURE 4.1. Payoff-matrix: Reputation Game

		Private Agents	
		Trust	Distrust
Policymaker	Set $\bar{\pi}$	$\bar{\pi}^2+\varphi\kappa^2$, 0	$\bar{\pi}^2+\varphi\left[\alpha\bar{\pi}-\kappa(1+\varphi\alpha^2)\right]^2$, $(\bar{\pi}-\varphi\alpha\kappa)^2$
	Deceive	$\frac{\varphi\left[\kappa+\alpha\bar{\pi}\right]^2}{1+\varphi\alpha^2}$, $(\varphi\alpha\kappa-\bar{\pi})^2$	$\varphi\kappa^2(1+\varphi\alpha^2)$, 0

I've colored the optimal one-period response of the private agents and the policymaker given the action of the other agent, blue and red respectively. Deceiving is a dominant one-period strategy for the policymaker. As it's expected the Nash Equilibrium of the one-period game is (Deceive, Distrust). That's the same as the discretionary equilibrium.

The outcomes (Set $\bar{\pi}$, Trust) are Pareto superior to the Nash Equilibrium (Deceive, Distrust), meaning that both agents would be better off under that equilibrium. If the game were finite, then the only sequential equilibrium possible would be the Nash Equilibrium[10]. But this is an infinitely repeated game, and these

[10] The monetary authority would have incentives to deceive in the last period. The private agents would know that and would anticipate that movement. By backward induction, the equilibrium would be the discretionary one.

games have a nice feature. If the central bank is sufficiently future-sighted (δ near 1), then by the Folk Theorem[11] there's an equilibrium in which both agents cooperate and reach the Pareto superior equilibrium (Set $\bar{\pi}$, Trust).

Another way to demonstrate that the reputational equilibrium is viable is by showing that the associated welfare intertemporal loss is lower than the deception welfare loss:

$$\frac{\varphi}{1+\varphi\alpha^2}\left[\kappa+\alpha\bar{\pi}\right]^2+\frac{\delta}{1-\delta}\left[\varphi\kappa^2(1+\varphi\alpha^2)\right]>\frac{1}{1-\delta}\left[\bar{\pi}^2+\varphi\kappa^2\right] \tag{4.1.10}$$

It's not easy to isolate δ in this expression. I will reserve for the appendix A.5.5 to show that, under certain conditions, a high discount factor makes it worthwhile for the policymaker to pursue lower inflation.

The takeaway is that if the policymaker is sufficiently future-sighted, then it's worth it to build a reputation. Reputation building models are attractive because they are intuitive and have real-world resemblance. Some caveats are not present in this model that are worthy of mention. The reputational equilibrium might not be possible if the public cannot easily monitor the policy instrument. Here the private agents could observe if monetary authority was following the non-discretionary path. If inflation was not the announced one, they knew they were being deceived. But if the monetary instrument or its effects is not clearly observable, then accountability to the announcement becomes much trickier. This possibility is absent by assumption in the model, but it should be a problem in real-life scenarios. Stokey (2002) discusses this issue, focusing on the trade-off between observability and precision of instruments.

Also, remember the simplification $\varepsilon_t = 0\ \forall t$. The reason for this assumption may be clear after reading the state-contingent rule section. If a large shock hits the economy, then even a very future-sighted policymaker may decide to abandon the reputational equilibrium to stabilize output. This possibility makes it very hard to maintain the reputational equilibrium. That said, if the penalty structure of the private agents were different, then this isn't such a big deal. For example, imagine that instead of always expecting the discretionary after deception, the private agents only expect it for one period. That is, the private agents continue to give "second chances" to the policymaker. Under this penalty scheme, it's possible to have offer shocks and still build a reputation.

[11] This theorem is called that way because it was well known among game theorists in the 50s', but no one had published it. It refers to a class of theorems with some resemblance and marginal differences. See, for example, Friedman (1971) .

4.2 REPUTATION IN A FINITE GAME

Two of the main criticisms over Barro and Gordon (1983b) are the departure from rational expectations and the necessity of an infinite period[12]. The following model, inspired by Backus and Driffill (1985), addresses these issues. It shows that, under certain circumstances, a discretionary central bank might want to build a reputation even with finite time and rational agents.

The setup of this model is a game of incomplete information[13]. The central bank may be of two types: conservative or discretionary. The conservative[14] always sets inflation equal to zero, independent of expectations. The discretionary wants to minimize the social loss function. The private agents don't know which type the central bank is, only the probability of each scenario. The agents infer the likelihood of each type by looking at the monetary authority's actions.

The timing of events:

1. Nature drafts the type of the central bank with known probability
2. Private agents form π_1^e
3. Central bank sets π_1
4. Private agents observe π_1 and update their priors, forming π_2^e
5. Central bank sets π_2

The intuition is that the discretionary central bank would like to be perceived as conservative. He can portray this impression by setting zero inflation, as a conservative monetary authority would do. The private agents, knowing about this possibility, update their beliefs in a Bayesian way. The discretionary central bank can accrue short-term gains by setting positive inflation in the first period or imitate a conservative central bank to obtain a larger benefit deceiving in the second period. That's the trade-offs he must weigh.

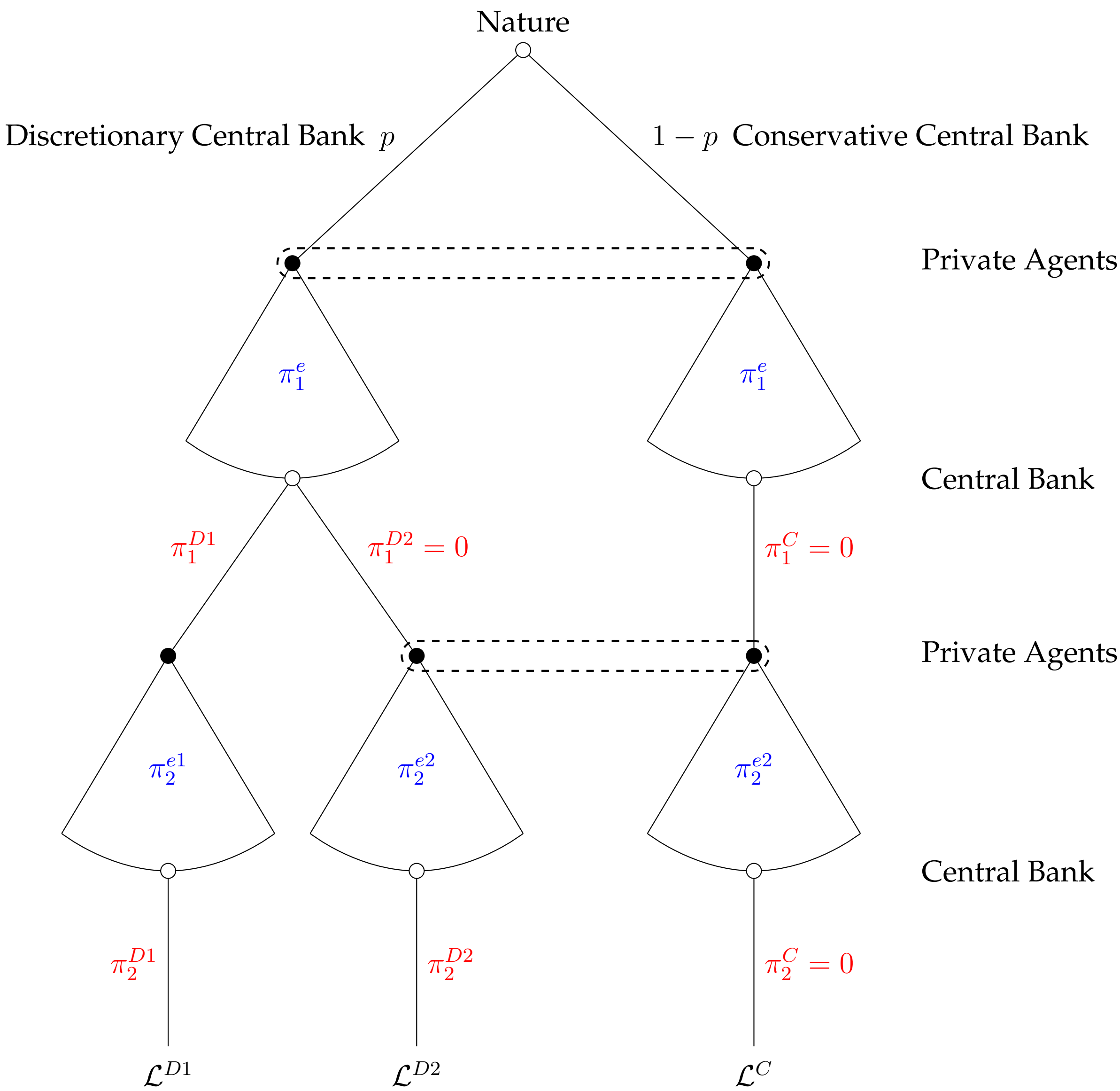
We can start building the formal model to show these results. For simplicity's sake, lets us illustrate this in a two-period game, but it also works for $T \geq 2$ periods. It helps to build the game in the extensive form to make things easier to grasp, as can be

[12]Actually, the game can be finite if it doesn't have a deterministic end. That is: if it ends with a probability each period.

[13]The solution concept is the sequential equilibria formulated by Kreps and Wilson (1982).

[14]In the previous models I used "conservative" to indicate a central bank that had a lower preference parameter than society. Here conservative has a stronger meaning: the central banks only care about inflation fighting. You could think of it as the case in which $\tilde{\varphi} = 0$.

FIGURE 4.2. Extensive Form: Reputation Game



seen in figure 4.2.

The private agents form the expected inflation in the first period by taking an average of the discretionary inflation and conservative inflation (zero), weighted by their priors. Let us denote the prior probability that the central bank is discretionary as $p \equiv \Pr(D)$. Similarly, the probability that the central bank is conservative is $1-p \equiv \Pr(C)$. From the previous models we know that the discretionary type reaction function is:

$$\pi_t = \frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa + \alpha\pi_t^e\right] \tag{4.2.1}$$

So the inflation the private agents will expect is:

$$\pi_1^e = \frac{p\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa \tag{4.2.2}$$

Appendix A.6.1

In the first period, the discretionary central bank can either set the optimal discretionary inflation or mimic a conservative central bank setting inflation at zero. Let us denote the former as π_1^{D1} and the latter as π_1^{D2}. Given π_1^e (4.2.2), the optimal discretionary response would be:

$$\pi_1^{D1} = \frac{\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa \tag{4.2.3}$$

Appendix A.6.2

If the private agents observe π_1^{D1}, they know that the central bank is discretionary. Then the expected inflation for the second period will be the familiar discretionary one. Let us denote π_2^{e1} as this scenario:

$$\pi_2^{e1} = \varphi\alpha\kappa \tag{4.2.4}$$

If the private agents observe $\pi_1 = 0$, they don't know whether the central bank is really conservative or if he's just pretending to be. They need to update their priors about the probability of each scenario, given the new information (zero inflation in the first period). This is a Bayesian updating, and the expected inflation in this scenario (π_2^{e2}) will have the form:

$$\begin{aligned}\pi_2^{e2} &= \Pr(D \mid \pi_1 = 0)\mathbb{E}[\pi_2^{D1}] + \Pr(C \mid \pi_1 = 0)[0] \\ &= \Pr(D \mid \pi_1 = 0)\mathbb{E}[\pi_2^{D1}]\end{aligned} \tag{4.2.5}$$

The term $\Pr(D \mid \pi_1 = 0)$ is the probability of the central bank being discretionary given that the private agents have observed zero inflation in the first period. We can use the Bayes Theorem to find this probability:

$$P(D \mid \pi_1 = 0) = \frac{P(\pi_1 = 0 \mid D)p}{P(\pi_1 = 0 \mid D)p + 1 - p} \tag{4.2.6}$$

If we define $\omega \equiv P(\pi_1 = 0 \mid D)$, then the expected inflation in this scenario will be:

$$\pi_2^{e2} = \frac{p\omega\varphi\alpha}{(1-p)(1+\varphi\alpha^2)+p\omega}\kappa \tag{4.2.7}$$

Appendix A.6.3

In the second period, the discretionary monetary authority will always play the optimal discretionary policy. If the central bank has set π_1^{D1}, then, given π_2^{e1} the response will be:

$$\pi_2^{D1} = \varphi\alpha\kappa \tag{4.2.8}$$

If he instead set π_1^{D2}, such that private agents expect π_2^{e2} (4.2.7), then the optimal response is:

$$\pi_2^{D2} = \frac{(1+p(\omega-1))\varphi\alpha}{(1-p)(1+\varphi\alpha^2)+p\omega}\kappa \tag{4.2.9}$$

Appendix A.6.4

Now that we have the game well set up, we can compare the outcomes to see which decision the discretionary central bank would take. Let us denote $\mathcal{L}^{D1}$ as the intertemporal welfare loss associated with the central bank deceiving people in the first round. Analogously, $\mathcal{L}^{D2}$ is the intertemporal welfare loss of the central bank pretending to be conservative.

$$\begin{aligned}
\mathcal{L}^{D1} &= \varphi\kappa^2[1+\varphi\alpha^2]\left[\left(\frac{1}{1+\varphi\alpha^2(1-p)}\right)^2+\delta\right] \\
\mathcal{L}^{D2} &= \varphi\kappa^2[1+\varphi\alpha^2]\left[\frac{1+\varphi\alpha^2}{(1+\varphi\alpha^2(1-p))^2}+\delta\left(\frac{1+p(\omega-1)}{(1-p)(1+\varphi\alpha^2)+p\omega}\right)^2\right]
\end{aligned} \tag{4.2.10}$$

Appendix A.6.5

It isn't straightforward to compare the equations. I will reserve for the appendix A.6.6 the demonstration that there are situations in which the discretionary central bank will behave conservatively.

This model shows that reputation building can be the equilibrium even in a finite game with rational agents. But it doesn't address the issues mentioned in the previous section: the importance of instrument observability. The difficulty to precisely monitor the policy instruments makes it harder to build a reputation.

5. Does it still matter?

After almost four decades of low inflation in the advanced economies, a skeptical reader might question: "Does this matter at all now?" I argue that yes, it still matters. In this chapter, I explain why I think that this literature is relevant even today.

First, we must not forget that one of the reasons (although not the only one) most advanced countries have low inflation today is their independent central bank. Even though it's hard to establish causal inference, most evidence suggests this was a relevant factor. This is especially clear for the regime change from the '60s and '70s highly discretionary monetary policy to the more independent central banks of the '80s and '90s. That said, I concede that inflation, particularly after the Financial Crisis, is a complex phenomenon, having other reasons to be that low.

Second, and most importantly, there are crucial lessons to be learned from the rules versus discretion debate. The relevance of time consistency isn't restrained to the inflation bias, it matters for other challenges we face today. Woodford (2013) investigates how time inconsistency can affect the effectiveness of Forward Guidance. With the Zero-Lower Bound binding, the central bank would like to signal that monetary policy will be expansionary for a long time. But if the private agents believe this and adjust their expectations, then the normalization of conditions would happen sooner, such that the central bank would have incentives to deviate from the announced path. This policy is time-inconsistent. Filardo and Hofmann (2014) provides a similar analysis, and Nakata and Sunakawa (2019) build a formal model of the credibility of Forward Guidance.

Finally, I argue that even outside the context of monetary policy, the rules versus discretion debate is useful. Consider patent policy. Imagine that the policymaker promises generous life-long patent protection, incentivizing entrepreneurs to invent new products and ideas. But once this stock of good ideas exists, the policymaker could deviate from the promise and break patent protection. The patent policy of a country must be credible to be time consistent.

In any situation in which people's expectations interact with policy actions, we might have time consistency problems. We can build institutional solutions that might alleviate that, but it isn't without its trade-offs. Also, institutions are not immutable. Policymaking in democracies is, by a consequence of the democratic process, discretionary. Policymakers can unbind their hands just as they can bind them. It's not just a matter of institutional solutions, but credible institutional solutions.

6. Conclusion

As we saw, discretionary policymaking in a framework with rational agents can lead to suboptimal results. In the context of monetary policy, the inefficiency emerges as an inflation bias. The problem of the discretionary policy is that a better equilibrium isn't time consistent. The ability of the policymaker to ad-hoc decide a policy is what drives this result. The outcomes can be improved by using rules, rather than discretion.

The first explored solution was to bind the central bank to a strict rule, such as the zero-inflation rule. It manages to defeat the inflation bias but at the cost of eliminating the stabilization of the offer shocks. The policymaker can seek an intermediate solution by delegating monetary policy to an independent monetary authority. A conservative central bank can achieve a lower inflation bias without distorting too much the stabilization policy. The powerful result is that it's always socially optimal to have a central bank values fighting inflation more than society itself.

The last institutional solution model exposed another margin of adjustment: how much independence to grant. The state-contingent rule allows the central bank to set inflation in normal periods, but the policymaker holds the option of overruling when a crisis hits. I also discuss the time consistency of institutional solutions themselves. Policymakers have discretionary power to build institutions, so they have the power to modify them. We shouldn't consider them as immutable. This last model deals with this problem more explicitly by assuming that policy is distorted during exceptional times.

Finally, reputation building by a discretionary central bank can diminish the inflation bias problem. This equilibrium doesn't necessarily rely on infinite time or non-rational expectations, although instrument observability is essential for the reputational equilibrium.

Bibliography

ALESINA, A. Macroeconomics and politics. In: *NBER Macroeconomics Annual 1988, Volume 3*. MIT Press, 1988. p. 13–62. Available from Internet: <http://www.nber.org/chapters/c10951>.

ALESINA, A.; GRILLI, V. The European Central Bank: Reshaping Monetary Politics in Europe. n. 3860, out. 1991. Available from Internet: <https://ideas.repec.org/p/nbr/nberwo/3860.html>.

ALESINA, A.; SUMMERS, L. H. Central bank independence and macroeconomic performance: Some comparative evidence. *Journal of Money, Credit and Banking*, [Wiley, Ohio State University Press], v. 25, n. 2, p. 151–162, 1993. ISSN 00222879, 15384616. Available from Internet: <http://www.jstor.org/stable/2077833>.

BACKUS, D.; DRIFFILL, J. Inflation and Reputation. *American Economic Review*, v. 75, n. 3, p. 530–538, June 1985. Available from Internet: <https://ideas.repec.org/a/aea/aecrev/v75y1985i3p530-38.html>.

BARRO, R. J.; GORDON, D. B. A positive theory of monetary policy in a natural rate model. *Journal of Political Economy*, University of Chicago Press, v. 91, n. 4, p. 589–610, oct 1983. ISSN 1537534X. Available from Internet: <https://www.journals.uchicago.edu/doi/abs/10.1086/261167>.

BARRO, R. J.; GORDON, D. B. Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, v. 12, n. 1, p. 101–121, 1983. ISSN 03043932.

BLINDER, A. *Central Banking in Theory and Practice*. 1. ed. The MIT Press, 1999. v. 1. Available from Internet: <https://EconPapers.repec.org/RePEc:mtp:titles:0262522608>.

CALVO, G. A. On the time consistency of optimal policy in a monetary economy. *Econometrica*, [Wiley, Econometric Society], v. 46, n. 6, p. 1411–1428, 1978. ISSN 00129682, 14680262. Available from Internet: <http://www.jstor.org/stable/1913836>.

CLARIDA, R.; GALÍ, J.; GERTLER, M. The science of monetary policy: A new Keynesian perspective. *Journal of Economic Literature*, v. 37, n. 4, p. 1661–1707, 1999. ISSN 00220515.

CUKIERMAN, A. *Central Bank Strategy, Credibility, and Independence: Theory and Evidence*. The MIT Press, 1992. v. 1. (MIT Press Books, 0262031981). ISBN ARRAY(0x5041e0c0). Available from Internet: <https://ideas.repec.org/b/mtp/titles/0262031981.html>.

CUKIERMAN, A. *The Economics of Central Banking*. [s.n.], 1996. ISBN 978-1-349-26074-4. Available from Internet: <https://ideas.repec.org/p/tiu/tiucen/e8f7a23c-1058-4c68-b6b5-dcbe0499d1bd.html>.

CUKIERMAN, A.; GERLACH, S. The inflation bias revisited: theory and some international evidence. *Manchester School*, v. 71, n. 5, p. 541–565, September 2003. Available from Internet: <https://ideas.repec.org/a/bla/manchs/v71y2003i5p541-565.html>.

FILARDO, A.; HOFMANN, B. Forward guidance at the zero lower bound. *BIS Quarterly Review*, March 2014. Available from Internet: <https://ideas.repec.org/a/bis/bisqtr/1403f.html>.

FISHER, I. Stabilizing the dollar. *The American Economic Review*, American Economic Association, v. 9, n. 1, p. 156–160, 1919. ISSN 00028282. Available from Internet: <http://www.jstor.org/stable/1813996>.

FLOOD, R. P.; ISARD, P. Monetary policy strategies. *IMF Staff Papers*, v. 36, n. 3, p. 612–632, 1989. Available from Internet: <https://EconPapers.repec.org/RePEc:pal:imfstp:v:36:y:1989:i:3:p:612-632>.

FRIEDMAN, J. W. A Non-cooperative Equilibrium for Supergames12. *The Review of Economic Studies*, v. 38, n. 1, p. 1–12, 01 1971. ISSN 0034-6527. Available from Internet: <https://doi.org/10.2307/2296617>.

FRIEDMAN, M. A monetary and fiscal framework for economic stability. *The American Economic Review*, American Economic Association, v. 38, n. 3, p. 245–264, 1948. ISSN 00028282. Available from Internet: <http://www.jstor.org/stable/1810624>.

GRILLI, V.; MASCIANDARO, D.; TABELLINI, G.; MALINVAUD, E.; PAGANO, M. Political and monetary institutions and public financial policies in the industrial countries. *Economic Policy*, [Center for Economic Studies, Maison des Sciences de l'Homme, Centre for Economic Policy Research, Wiley], v. 6, n. 13, p. 342–392, 1991. ISSN 02664658, 14680327. Available from Internet: <http://www.jstor.org/stable/1344630>.

KREPS, D. M.; WILSON, R. Reputation and imperfect information. *Journal of Economic Theory*, v. 27, n. 2, p. 253–279, 1982. ISSN 0022-0531. Available from Internet: <https://www.sciencedirect.com/science/article/pii/0022053182900308>.

KYDLAND, F. E.; PRESCOTT, E. C. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, University of Chicago Press, v. 85, n. 3, p. 473–491, 1977. ISSN 00223808, 1537534X. Available from Internet: <http://www.jstor.org/stable/1830193>.

LOHMANN, S. Optimal commitment in monetary policy: Credibility versus flexibility. *The American Economic Review*, American Economic Association, v. 82, n. 1, p. 273–286, 1992. ISSN 00028282. Available from Internet: <http://www.jstor.org/stable/2117615>.

MCCALLUM, B. T. Chapter 23 issues in the design of monetary policy rules. In: . Elsevier, 1999, (Handbook of Macroeconomics, v. 1). p. 1483–1530.

Available from Internet: <https://www.sciencedirect.com/science/article/pii/S1574004899100363>.

NAKATA, T.; SUNAKAWA, T. Credible Forward Guidance. n. 2019-037, maio 2019. Available from Internet: <https://ideas.repec.org/p/fip/fedgfe/2019-37.html>.

NORDHAUS, W. D. The political business cycle. *The Review of Economic Studies*, [Oxford University Press, Review of Economic Studies, Ltd.], v. 42, n. 2, p. 169–190, 1975. ISSN 00346527, 1467937X. Available from Internet: <http://www.jstor.org/stable/2296528>.

PERSSON, T.; TABELLINI, G. Chapter 22 Political economics and macroeconomic policy. *Handbook of Macroeconomics*, v. 1, n. PART C, p. 1397–1482, 1999. ISSN 15740048.

ROGOFF, K. The Optimal Degree of Commitment to an Intermediate Monetary Target. *The Quarterly Journal of Economics*, v. 100, n. 4, p. 1169–1189, 1985. Available from Internet: <https://ideas.repec.org/a/oup/qjecon/v100y1985i4p1169-1189.html>.

SCHMITT-GROHé, S.; URIBE, M. Chapter 13 - the optimal rate of inflation. In: FRIEDMAN, B. M.; WOODFORD, M. (Ed.). Elsevier, 2010, (Handbook of Monetary Economics, v. 3). p. 653–722. Available from Internet: <https://www.sciencedirect.com/science/article/pii/B9780444534545000013>.

STOKEY, N. L. "rules vs. discretion" after twenty-five years. *NBER Macroeconomics Annual*, University of Chicago Press, v. 17, p. 9–45, 2002. ISSN 08893365, 15372642. Available from Internet: <http://www.jstor.org/stable/3585272>.

TAYLOR, J. B. *Rules Versus Discretion: Assessing the Debate Over the Conduct of Monetary Policy*. [S.l.], 2017. (Working Paper Series, 24149). Available from Internet: <http://www.nber.org/papers/w24149>.

TAYLOR, J. B.; WILLIAMS, J. C. Chapter 15 - simple and robust rules for monetary polic. In: FRIEDMAN, B. M.; WOODFORD, M. (Ed.). Elsevier, 2010, (Handbook of Monetary Economics, v. 3). p. 829–859. Available from Internet: <https://www.sciencedirect.com/science/article/pii/B9780444534545000037>.

WALSH, C. E. *Monetary Theory and Policy, Fourth Edition*. 4. ed. The MIT Press, 2017. v. 1. (MIT Press Books, 9780262035811). ISBN 9780262035811. Available from Internet: <https://mitpress.mit.edu/books/monetary-theory-and-policy-fourth-edition>.

WICKSELL, K. The Influence of the Rate of Interest on Prices. *History of Economic Thought Articles*, v. 17, p. 213–220, 1907. Available from Internet: <https://ideas.repec.org/a/hay/hetart/wicksell1907.html>.

WOODFORD, M. *Forward Guidance by Inflation-Targeting Central Banks*. [S.l.], 2013. Available from Internet: <https://ideas.repec.org/p/cpr/ceprdp/9722.html>.

A. MATHEMATICAL APPENDIX

A.1 MODEL 1: DISCRETIONARY POLICYMAKER

A.1.1 OPTIMAL INFLATION FOR THE DISCRETIONARY CENTRAL BANK

Here I solve the optimization problem depicted in (2.2.4). First, we must substitute the output y_t with the supply function (2.2.1). Then the problem becomes:

$$\min_{\pi_t} \ \pi_t^2 + \varphi(y_n + \alpha(\pi_t - \pi_t^e) + \varepsilon - y^*)^2 \tag{A.1.1}$$

Then take the derivative of this function with respect to π_t, considering all the other variables as constants, and equal to zero to find the critical point:

$$2\pi_t + (2)(\alpha)\varphi(y_n + \alpha(\pi_t - \pi_t^e) + \varepsilon - y^*) = 0 \tag{A.1.2}$$

By rearranging π_t we get:

$$\pi_t = \frac{\varphi\alpha}{1 + \varphi\alpha^2} [y^* - y_n + \alpha\pi_t^e - \varepsilon_t] \tag{A.1.3}$$

The same expression of (2.2.5), only that there we denoted $\kappa \equiv y^* - y_n$

A.1.2 DISCRETIONARY INFLATION EXPECTATION

Here I develop the steps required to obtain the result in (2.2.7). The expected inflation is obtained by taking the expected level of the reaction function of the policymaker:

$$\pi_t^e = \mathbb{E}[\pi_t] = \mathbb{E}\left[\frac{\varphi\alpha}{1 + \varphi\alpha^2} [\kappa + \alpha\pi_t^e - \varepsilon_t]\right] \tag{A.1.4}$$

Remember that the expected value is a linear operator, so we can break the expression and take the constants "out":

$$\pi_t^e = \frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\mathbb{E}[\kappa] + \alpha\mathbb{E}[\pi_t^e] - \mathbb{E}[\varepsilon_t]\right] \tag{A.1.5}$$

The output gap κ is a constant, so $\mathbb{E}[\kappa] = \kappa$. The expected value of the expected inflation is, naturally, the expected inflation, therefore $\mathbb{E}[\pi_t^e] = \pi_t^e$. And the expected value of the offer shocks is zero (remember that $\varepsilon_t \sim N(0, \sigma^2)$): $\mathbb{E}[\varepsilon_t] = 0$. So our expression simplifies to:

$$\pi_t^e = \frac{\varphi\alpha}{1+\varphi\alpha^2}\kappa + \frac{\varphi\alpha^2}{1+\varphi\alpha^2}\pi_t^e \tag{A.1.6}$$

By isolating π_t^e we find the expression in (2.2.7):

$$\pi_t^e = \varphi\alpha\kappa \tag{A.1.7}$$

A.1.3 EXPECTED WELFARE IN THE DISCRETIONARY EQUILIBRIUM

Here we explain how to obtain the expression (2.2.11). First, we need to plug the equilibrium outcomes (2.2.8) into the (standard) welfare loss function (2.2.3):

$$\mathcal{L} = \left(\varphi\alpha\kappa - \frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t\right)^2 + \varphi\left(y_n - y^* + \alpha\left(\varphi\alpha\kappa - \frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t - \varphi\alpha\kappa\right) + \varepsilon_t\right)^2 \tag{A.1.8}$$

We can easily simplify the second term because the inflation bias will cancel and we can find a nicer expression for the offer shocks:

$$\mathcal{L} = \left(\varphi\alpha\kappa - \frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t\right)^2 + \varphi\left(-\kappa + \frac{1}{1+\varphi\alpha^2}\varepsilon_t\right)^2 \tag{A.1.9}$$

Now we need to open the quadratic terms and take the expected value. It's important to remember that the expected value of the square of a random variable with zero mean is its variance: $\mathrm{Var}[\varepsilon_t] = \mathbb{E}[\varepsilon_t^2] - \mathbb{E}[\varepsilon_t]^2 = \sigma^2 - 0 = \sigma^2$. Also, because the expected value of ε_t is zero, all the terms with it (not the ones squared as mentioned before) will have zero expected value. So we have:

$$
\begin{aligned}
\mathbb{E}[\mathcal{L}] &= \varphi^2\alpha^2\kappa^2 + \left(\frac{\varphi\alpha}{1+\varphi\alpha^2}\right)^2\sigma^2 + \varphi\left[\kappa^2 + \left(\frac{1}{1+\varphi\alpha^2}\right)^2\sigma^2\right] \\
&= (\varphi + \varphi^2\alpha^2)\kappa^2 + \left(\frac{\varphi+\varphi^2\alpha^2}{(1+\varphi\alpha^2)^2}\right)\sigma^2 \\
&= \varphi(1+\varphi\alpha^2)\kappa^2 + \frac{\varphi}{1+\varphi\alpha^2}\sigma^2 \\
&= \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa^2\left(1+\varphi\alpha^2\right)^2 + \sigma^2\right]
\end{aligned}
\tag{A.1.10}
$$

This is the same expression as (2.2.11).

A.1.4 EXPECTED WELFARE IN THE SOCIALLY OPTIMAL EQUILIBRIUM

Here I obtain the expression (2.2.12). As in the other cases, we need to plug the equilibrium values (2.2.10) into the welfare social loss function:

$$
\mathcal{L} = \left(-\frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t\right)^2 + \varphi\left(y_n - y^* + \alpha\left(-\frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t\right) + \varepsilon_t\right)^2 \tag{A.1.11}
$$

Note the similarity between this expression and the one of the discretionary equilibrium. As before, we can simplify the expression:

$$
\mathcal{L} = \left(-\frac{\varphi\alpha}{1+\varphi\alpha^2}\varepsilon_t\right)^2 + \varphi\left(-\kappa + \frac{1}{1+\varphi\alpha^2}\varepsilon_t\right)^2 \tag{A.1.12}
$$

As previously done, we can open the quadratic terms and take the expected value:

$$
\begin{aligned}
\mathbb{E}[\mathcal{L}] &= \left(\frac{\varphi\alpha}{1+\varphi\alpha^2}\right)^2\sigma^2 + \varphi\left[\kappa^2 + \left(\frac{1}{1+\varphi\alpha^2}\right)^2\sigma^2\right] \\
&= \varphi\kappa^2 + \left(\frac{\varphi+\varphi^2\alpha^2}{(1+\varphi\alpha^2)^2}\right)\sigma^2 \\
&= \varphi\kappa^2 + \frac{\varphi}{1+\varphi\alpha^2}\sigma^2
\end{aligned}
\tag{A.1.13}
$$

And this is the expression found in (2.2.12)

A.2 MODEL 2: ZERO-INFLATION RULE

A.2.1 EXPECTED WELFARE IN A SIMPLE RULE EQUILIBRIUM

Here I obtain the expression (3.1.2). As before, I plug the equilibrium outcomes (3.1.1) into the social loss function:

$$\begin{aligned}\mathcal{L} &= (0)^2 + \varphi\left(y_n - y^* + \alpha\left(0 - 0\right) + \varepsilon_t\right)^2 \\ &= \varphi(-\kappa + \varepsilon_t)^2\end{aligned} \tag{A.2.1}$$

Given the simplicity of this simple rule, it's very easy to obtain (3.1.2):

$$\mathbb{E}[\mathcal{L}] = \varphi(\kappa^2 + \sigma^2) \tag{A.2.2}$$

A.2.2 SIMPLE RULE VERSUS DISCRETION

Here we find the expression (3.1.3), comparing the expected welfare of the simple rule against the discretionary equilibrium.

Remember that we are working with a social **loss** function, so the lower the value of the loss function, the better. The zero-inflation rule is better than the discretionary equilibrium if the expected welfare loss of the rule is lower than the discretionary one. This happens when:

$$\varphi\left[\kappa^2 + \sigma^2\right] < \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa^2\left(1+\varphi\alpha^2\right)^2 + \sigma^2\right] \tag{A.2.3}$$

It's not very hard to simplify this expression, we can throw the κ to the right-hand side and the σ to the left:

$$\varphi\sigma^2 - \frac{\varphi}{1+\varphi\alpha^2}\sigma^2 < \varphi\kappa^2\left(1+\varphi\alpha^2\right) - \varphi\kappa^2 \tag{A.2.4}$$

Now we can divide the expression by φ (which is always positive so the inequality doesn't changes signs) and simplify:

$$\left[\frac{1+\varphi\alpha^2 - 1}{1+\varphi\alpha^2}\right]\sigma^2 < \left[1+\varphi\alpha^2 - 1\right]\kappa^2 \tag{A.2.5}$$

Finally, by simplifying and rearranging the terms we obtain (3.1.3):

$$\sigma^2 < \kappa^2(1+\varphi\alpha^2) \tag{A.2.6}$$

A.3 MODEL 3: INDEPENDENT CENTRAL BANK

A.3.1 OPTIMIZATION FOR THE CONSERVATIVE CENTRAL BANK

Here I develop the optimization problem (3.2.3). First, we can open the quadratic terms and take the expected value (similarly as we have done in the previous problems), such that the new optimization problem yields:

$$\begin{aligned} &\min_{\tilde{\varphi}} \mathbb{E}\left[\left(\tilde{\varphi}\alpha\kappa - \frac{\tilde{\varphi}\alpha}{1+\tilde{\varphi}\alpha^2}\varepsilon_t\right)^2 + \varphi\left(-\kappa + \frac{1}{1+\tilde{\varphi}\alpha^2}\varepsilon_t\right)^2\right] \\ \iff &\min_{\tilde{\varphi}} \kappa^2\left(\varphi + \tilde{\varphi}^2\alpha^2\right) + \sigma^2\frac{\varphi + \tilde{\varphi}^2\alpha^2}{(1+\tilde{\varphi}\alpha^2)^2} \end{aligned} \tag{A.3.1}$$

By taking the derivative with respect to $\tilde{\varphi}$, the first order conditions yields:

$$2\tilde{\varphi}\alpha^2\kappa^2 + \sigma^2\left[\frac{(2\tilde{\varphi}\alpha^2)(1+\tilde{\varphi}\alpha^2)^2 - 2(\alpha^2)(1+\tilde{\varphi}\alpha^2)(\varphi+\tilde{\varphi}^2\alpha^2)}{(1+\tilde{\varphi}\alpha^2)^4}\right] = 0 \tag{A.3.2}$$

We can simplify the expression by dividing everything by $2\alpha^2$ and throwing the expression with σ^2 to the right-hand side :

$$\tilde{\varphi}\kappa^2 = \sigma^2\left[\frac{(1+\tilde{\varphi}\alpha^2)(\varphi+\tilde{\varphi}^2\alpha^2) - \tilde{\varphi}(1+\tilde{\varphi}\alpha^2)^2}{(1+\tilde{\varphi}\alpha^2)^4}\right] \tag{A.3.3}$$

Then we can divide the denominator and numerator of the right-hand side by $(1+\tilde{\varphi}\alpha^2)$:

$$\tilde{\varphi}\kappa^2 = \sigma^2\left[\frac{\varphi+\tilde{\varphi}^2\alpha^2 - \tilde{\varphi}(1+\tilde{\varphi}\alpha^2)}{(1+\tilde{\varphi}\alpha^2)^3}\right] \tag{A.3.4}$$

This expression simplifies to:

$$\tilde{\varphi}\kappa^2 = \sigma^2\frac{\varphi - \tilde{\varphi}}{(1+\tilde{\varphi}\alpha^2)^3} \tag{A.3.5}$$

As mentioned, we can't isolate $\tilde{\varphi}$ and find a nice expression as in the previous cases, but we can prove that there exists $\tilde{\varphi} \in (0, \varphi)$ such that this expression is respected. Lets throw σ^2 to the left-hand side and define two functions, $g(\tilde{\varphi})$ and $h(\tilde{\varphi})$:

$$g(\tilde{\varphi}) \equiv \tilde{\varphi}\frac{\kappa^2}{\sigma^2} = \frac{\varphi - \tilde{\varphi}}{(1+\tilde{\varphi}\alpha^2)^3} \equiv h(\tilde{\varphi}) \tag{A.3.6}$$

We must show that $\exists\tilde{\varphi}$ that respect (A.3.6). First, note that both functions are continuous and differentiable. Second, $g(0) = 0$ and it's strictly increasing, that is:

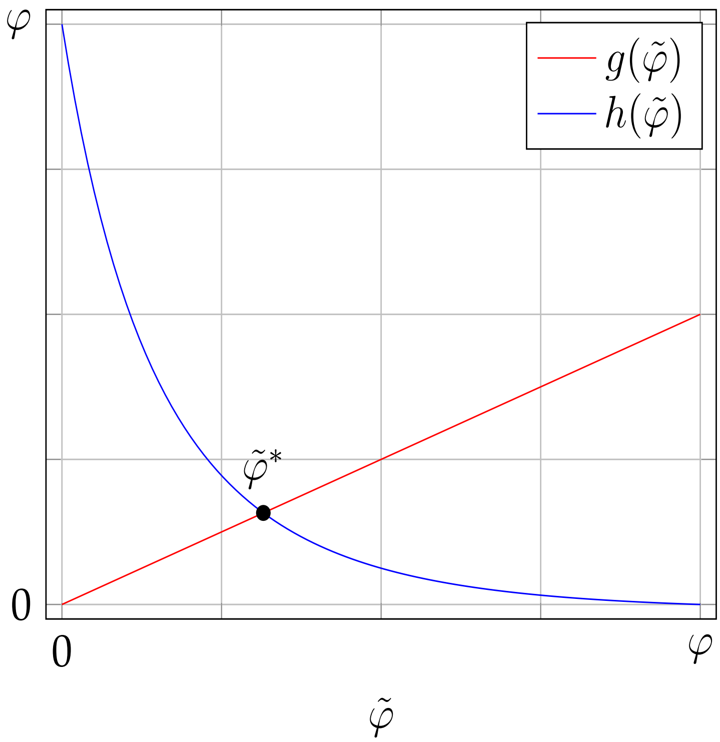
$$\begin{aligned} g(0) &= 0 \\ \frac{\partial g(\tilde{\varphi})}{\partial \tilde{\varphi}} &= \frac{\kappa^2}{\sigma^2} > 0 \; \forall \; \tilde{\varphi} \end{aligned} \tag{A.3.7}$$

Also note that:

$$\begin{aligned} h(0) &= \varphi \\ h(\varphi) &= 0 \end{aligned} \tag{A.3.8}$$

Because both functions are continuous, then there must be at least one point in $(0, \varphi)$ such that (A.3.6) is satisfied. To visualize why this is the case, I plot in figure A.1 a generic example of these functions. To make this plot I used $\varphi = 2, \sigma^2 = 4, \kappa = 1, \alpha = 1$, but it will work with any values that respect our assumptions (such as positive variance and output gap).

FIGURE A.1. Representation of $g(\tilde{\varphi})$ and $h(\tilde{\varphi})$



A.4 MODEL 4: STATE-CONTINGENT RULE

A.4.1 SOCIAL WELFARE IN STATE-CONTINGENT RULE

Here I show how to find the social welfare function in (3.3.5). The process is very similar to the previous sections. The only difference is that I won't take the expected value. Also, the expected inflation is consider as given. Plugging the reaction function of a conservative central bank into the social welfare loss we obtain:

$$\mathcal{L}^{\mathcal{I}} = \left[\frac{\tilde{\varphi}\alpha}{1+\tilde{\varphi}\alpha^2}(\kappa + \alpha\pi_t^e - \varepsilon_t)\right]^2 + \varphi\left[-\kappa + \frac{\tilde{\varphi}\alpha^2}{1+\tilde{\varphi}\alpha^2}(\kappa + \alpha\pi_t^e - \varepsilon_t) - \alpha\pi_t^e + \varepsilon_t\right]^2 \quad \text{(A.4.1)}$$

By developing the second squared term we can find a more tractable equation:

$$\mathcal{L}^{\mathcal{I}} = \left[\frac{\tilde{\varphi}\alpha}{1+\tilde{\varphi}\alpha^2}\right]^2 (\kappa + \alpha\pi_t^e - \varepsilon_t)^2 + \frac{\varphi}{(1+\tilde{\varphi}\alpha^2)^2}(-\kappa - \alpha\pi_t^e + \varepsilon_t)^2 \quad \text{(A.4.2)}$$

We can simplify the expression even further by remembering that, because the expressions are squared, it holds that $(\kappa + \alpha\pi_t^e - \varepsilon_t)^2 = (-\kappa - \alpha\pi_t^e + \varepsilon_t)^2$. Such that we can arrive at the desired expression:

$$\mathcal{L}^{\mathcal{I}} = \left[\frac{\varphi + \tilde{\varphi}^2\alpha^2}{(1+\tilde{\varphi}\alpha^2)^2}\right] (\kappa + \alpha\pi_t^e - \varepsilon_t)^2 \quad \text{(A.4.3)}$$

For the social welfare function if the policymaker overrules it makes it easier to note that it will be the same of the conservative welfare function, only changing $\tilde{\varphi}$ for φ and adding the cost term τ:

$$\begin{aligned} \mathcal{L}^{\mathcal{O}} &= \left[\frac{\varphi + \varphi^2\alpha^2}{(1+\varphi\alpha^2)^2}\right] (\kappa + \alpha\pi_t^e - \varepsilon_t)^2 + \tau \\ &= \left[\frac{\varphi}{1+\varphi\alpha^2}\right] [\kappa + \alpha\pi_t^e - \varepsilon_t]^2 + \tau \end{aligned} \quad \text{(A.4.4)}$$

A.4.2 CONDITIONS FOR OVERRULING

Here I show how to find (3.3.6) and also (3.3.9).

The policymaker won't overrule if and only if $\mathcal{L}^o > \mathcal{L}^c$. That is:

$$\left[\frac{\varphi}{1+\varphi\alpha^2}\right] [\kappa + \alpha\pi_t^e - \varepsilon_t]^2 + \tau > \left[\frac{\varphi + \tilde{\varphi}^2\alpha^2}{(1+\tilde{\varphi}\alpha^2)^2}\right] (\kappa + \alpha\pi_t^e - \varepsilon_t)^2 \quad \text{(A.4.5)}$$

We can isolate τ in the left-hand side and develop:

$$\begin{aligned}\tau &> \left(\frac{\varphi + \tilde{\varphi}^2\alpha^2}{(1+\tilde{\varphi}\alpha^2)^2} - \frac{\varphi}{1+\varphi\alpha^2}\right)(\kappa + \alpha\pi_t^e - \varepsilon_t)^2 \\ \tau &> \left(\frac{(1+\varphi\alpha^2)(\varphi + \tilde{\varphi}^2\alpha^2) - \varphi(1+\tilde{\varphi}\alpha^2)^2}{(1+\tilde{\varphi}\alpha^2)^2(1+\varphi\alpha^2)}\right)(\kappa + \alpha\pi_t^e - \varepsilon_t)^2\end{aligned} \tag{A.4.6}$$

If you open the terms you will see that the expression simplifies nicely to the expression in (3.3.6):

$$\tau > \frac{\alpha^2(\varphi - \tilde{\varphi})^2}{(1+\varphi\alpha^2)(1+\tilde{\varphi}\alpha^2)^2}(\kappa + \alpha\pi_t^e - \varepsilon_t)^2 \tag{A.4.7}$$

To find the expression in (3.3.9) we must first isolate the squared term. Also, instead of working with $(\kappa + \alpha\pi_t^e - \varepsilon_t)^2$, lets work with the equivalent $(-\kappa - \alpha\pi_t^e + \varepsilon_t)^2$:

$$(-\kappa - \alpha\pi_t^e + \varepsilon_t)^2 < \tau\frac{(1+\varphi\alpha^2)(1+\tilde{\varphi}\alpha^2)^2}{\alpha^2(\varphi - \tilde{\varphi})^2} \tag{A.4.8}$$

Note that this is an inequality with an squared term, so when we take the root the inequality transforms to:

$$-\sqrt{\tau\frac{(1+\varphi\alpha^2)(1+\tilde{\varphi}\alpha^2)^2}{\alpha^2(\varphi - \tilde{\varphi})^2}} < -\kappa - \alpha\pi_t^e + \varepsilon_t < \sqrt{\tau\frac{(1+\varphi\alpha^2)(1+\tilde{\varphi}\alpha^2)^2}{\alpha^2(\varphi - \tilde{\varphi})^2}} \tag{A.4.9}$$

By adding $\kappa + \alpha\pi_t^e$ to both sides and simplifying the squared roots we find:

$$\kappa + \alpha\pi_t^e - \frac{1+\tilde{\varphi}\alpha^2}{\alpha(\varphi - \tilde{\varphi})}\sqrt{\tau(1+\varphi\alpha^2)} \leq \varepsilon_t \leq \kappa + \alpha\pi_t^e + \frac{1+\tilde{\varphi}\alpha^2}{\alpha(\varphi - \tilde{\varphi})}\sqrt{\tau(1+\varphi\alpha^2)} \tag{A.4.10}$$

This is the same expression as in (3.3.9), I only denoted $\Omega \equiv \frac{1+\tilde{\varphi}\alpha^2}{\alpha(\varphi-\tilde{\varphi})}\sqrt{\tau(1+\varphi\alpha^2)}$ to make the expression simpler.

A.5 MODEL 5: REPUTATION IN AN INFINITE GAME

A.5.1 EXPECTED WELFARE OF COOPERATION

Here we develop the expression (4.1.3) and (4.1.4). It's quite easy to find the expressions given that in our simpler output function (4.1.2) we don't have output shocks. As in the previous models, we simply take the expected value of the welfare loss function considering the equilibrium valuers:

$$\begin{aligned}\mathbb{E}[\mathcal{L}] &= \mathbb{E}\left[(\bar{\pi})^2 + \varphi\left[-\kappa + \alpha(\bar{\pi} - \bar{\pi})\right]^2\right] \\ &= \bar{\pi}^2 + \varphi\kappa^2\end{aligned} \tag{A.5.1}$$

This is the expression exposed in (4.1.3). For the intertemporal expected social welfare function we have a convergent geometric series:

$$\begin{aligned}\sum_{i=0}^{\infty} \delta^i \mathbb{E}\left[\mathcal{L}\right] &= \bar{\pi}^2 + \varphi\kappa^2 + \delta\left[\bar{\pi}^2 + \varphi\kappa^2\right] + \delta^2\left[\bar{\pi}^2 + \varphi\kappa^2\right] + \dots \\ &= \frac{1}{1-\delta}\left[\bar{\pi}^2 + \varphi\kappa^2\right]\end{aligned} \tag{A.5.2}$$

If you don't remember, the general rule for a geometric series is the following:

$$\sum_{n=0}^{\infty} ax^n = a + ax + ax^2 + ax^3 + \ldots = \frac{a}{1-x} \quad \text{if } |x| < 1 \tag{A.5.3}$$

In our example we have $a = \bar{\pi}^2 + \varphi\kappa^2$ and $x = \delta$. By assumption we have $\delta \in [0, 1)$, which guarantees that the series converges.

A.5.2 EXPECTED WELFARE OF DECEPTION: ONE-TIME GAIN

Here we find the expression (4.1.6). We must plug the inflation (4.1.5) into the social loss and assume that the expected inflation is $\bar{\pi}$. Arriving at this expression:

$$\mathcal{L} = \left(\frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa + \alpha\bar{\pi}\right]\right)^2 + \varphi\left(-\kappa + \alpha\left(\frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa + \alpha\bar{\pi}\right] - \bar{\pi}\right)\right)^2 \tag{A.5.4}$$

By taking the expected value and developing we arrive at (4.1.6).

$$\mathbb{E}[\mathcal{L}] = \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa + \alpha\bar{\pi}\right]^2 \tag{A.5.5}$$

A.5.3 COMPARING DECEPTION WITH COOPERATION (FIRST PERIOD)

I want to show here that the expected welfare loss of deception is always at least as low as the expected welfare loss of cooperation in the first period. That is, show that (4.1.6) is lower than (4.1.3):

$$\frac{\varphi}{1+\varphi\alpha^2}\left[\kappa + \alpha\bar{\pi}\right]^2 \leq \bar{\pi}^2 + \varphi\kappa^2 \tag{A.5.6}$$

We can multiply the right-hand side with $1+\varphi\alpha^2$ and pass all the terms to the left-hand side, opening the squared term:

$$0 \leq \bar{\pi}^2 + \varphi\alpha^2\bar{\pi}^2 + \varphi\kappa^2 + \varphi^2\alpha^2\kappa^2 - \varphi(\kappa^2 + 2\alpha\bar{\pi}\kappa + \alpha^2\kappa^2) \tag{A.5.7}$$

This expressions simplifies to

$$0 \leq \bar{\pi}^2 + \varphi^2\alpha^2\kappa^2 - 2\varphi\alpha\bar{\pi}\kappa \tag{A.5.8}$$

A savvy reader can see that this expression can be exposed as:

$$0 \leq (\bar{\pi} - \varphi\alpha\kappa)^2 \tag{A.5.9}$$

This inequality will always be true, given that a squared term cannot be negative. Also, we can infer that the expected welfare loss of deception is always lower than cooperation (in the first period), only being equal if $\bar{\pi} = \varphi\alpha\kappa$.

A.5.4 INTERTEMPORAL WELFARE FOR THE DECEPTION EQUILIBRIUM

Here we find the expression (4.1.9). In the first period we will have the expected welfare associated with deception and then we return to the discretionary equilibrium, so the intertemporal expected social welfare is:

$$\begin{aligned}\sum_{i=0}^{\infty}\delta^i\mathbb{E}\left[\mathcal{L}\right] &= \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa+\alpha\bar{\pi}\right]^2 + \delta\varphi\kappa^2(1+\varphi\alpha^2) + \delta^2\varphi\kappa^2(1+\varphi\alpha^2) + \dots \\ &= \frac{\varphi}{1+\varphi\alpha^2}\left[\kappa+\alpha\bar{\pi}\right]^2 + \frac{\delta}{1-\delta}\left[\varphi\kappa^2(1+\varphi\alpha^2)\right]\end{aligned} \tag{A.5.10}$$

A.5.5 SHOWING THAT THE DISCRETIONARY POLICYMAKER MIGHT SET ZERO INFLATION

Here I show that, under certain conditions, (4.1.10) is satisfied. Instead of proving for the general case, I will demonstrate that with a specific example. Let's postulate that $\varphi = 1,\ \alpha = 1,\ \bar{\pi} = 0$, then (4.1.10) simplifies to

$$\frac{1}{1+1}\kappa^2 + \frac{\delta}{1-\delta}[\kappa^2(1+1)] > \frac{1}{1-\delta}\kappa^2 \tag{A.5.11}$$

This expression is satisfied if and only if

$$\delta > \frac{1}{3} \approx 33\% \tag{A.5.12}$$

A.6 MODEL 6: REPUTATION IN A FINITE GAME

A.6.1 EXPECTED INFLATION IN THE FIRST PERIOD

Here I show how to obtain the expected inflation in the first period (4.2.2). It's actually quite simple. The agents expect an average between what a conservative central bank would set (zero) and a discretionary central bank would set (4.2.1). The weights are the priors probabilities of each type:

$$\pi_1^e = \Pr(D)\mathbb{E}[\pi_1^{D1}] + \Pr(C)(0) \tag{A.6.1}$$

Denoting $p \equiv \Pr(D)$ and taking the expected value we find the expression.

$$\begin{aligned} \pi_1^e &= p\mathbb{E}\left[\frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa + \alpha\pi_t^e\right]\right] \\ &= \frac{p\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa \end{aligned} \tag{A.6.2}$$

A.6.2 OPTIMAL DISCRETIONARY RESPONSE TO π_1^e

What would the best-response for a discretionary central bank given the π_1^e. That's the expression (4.2.3). We find it by plugging the expected inflation (4.2.2) into the reaction function (4.2.1):

$$\begin{aligned} \pi_1^{D1} &= \frac{\varphi\alpha}{1+\varphi\alpha^2}\left[\kappa + \alpha\left(\frac{p\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa\right)\right] \\ &= \frac{\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa \end{aligned} \tag{A.6.3}$$

A.6.3 EXPECTED INFLATION IF $\pi_1 = 0$

Here I show how to find (4.2.7). Again the expected inflation will be an average, but now, instead of working with a prior, the agents will update the probabilities of the central bank being each type. The agents will use the information $\pi_1 = 0$ to update their beliefs. By using the Bayes Theorem the posterior of probability is:

$$\begin{aligned} P(D \mid \pi_1 = 0) &= \frac{P(\pi_1 = 0 \mid D)P(D)}{P(\pi_1 = 0 \mid D)P(D) + P(\pi_1 = 0 \mid C)P(C)} \\ &= \frac{P(\pi_1 = 0 \mid D)p}{P(\pi_1 = 0 \mid D)p + 1 - p} \\ &= \frac{p\omega}{1 + p(\omega - 1)} \end{aligned} \tag{A.6.4}$$

In which we denoted $\omega \equiv P(\pi_1 = 0 \mid D)$.

Now we do the same thing as before, but using this updated probability instead of the pior p:

$$\begin{aligned} \pi_2^{e2} &= \frac{p\omega}{1 + p(\omega - 1)} \mathbb{E}[\pi_2] \\ &= \frac{p\omega}{1 + p(\omega - 1)} \left[\frac{\varphi\alpha}{1 + \varphi\alpha^2} \left[\kappa + \alpha\pi_2^{e2}\right] \right] \\ &= \frac{p\omega\varphi\alpha}{(1 - p)(1 + \varphi\alpha^2) + p\omega} \kappa \end{aligned} \tag{A.6.5}$$

A.6.4 Optimal discretionary response to π_2^{e2}

Here I show how to obtain (4.2.9). The process is the same as before. We plug the expected inflation π_2^{e2} into the reaction function of the central bank:

$$\begin{aligned} \pi_2^{D2} &= \frac{\varphi\alpha}{1 + \varphi\alpha^2} \left[\kappa + \alpha\pi_2^{e1}\right] \\ &= \frac{\varphi\alpha}{1 + \varphi\alpha^2} \left[\kappa + \alpha \frac{p\omega\varphi\alpha}{(1 - p)(1 + \varphi\alpha^2) + p\omega} \kappa \right] \\ &= \frac{(1 + p(\omega - 1))\varphi\alpha}{(1 - p)(1 + \varphi\alpha^2) + p\omega} \kappa \end{aligned} \tag{A.6.6}$$

A.6.5 Social welfare loss

Here I show how to obtain (4.2.10). We need to compute the payoffs of each period with the discount factor. Because this game only has two periods, it's simply:

$$\begin{aligned} \mathcal{L}^{D1} &= \mathcal{L}_1^{D1} + \delta \mathcal{L}_2^{D1} \\ \mathcal{L}^{D2} &= \mathcal{L}_1^{D2} + \delta \mathcal{L}_2^{D2} \end{aligned} \tag{A.6.7}$$

Where, for example, $\mathcal{L}_2^{D2}$ is the second-period social welfare loss associated with the discretionary central bank that has imitated a conservative central bank in the first period. To compute those welfare losses we need to input the respective inflation and expected inflation into the welfare loss function.

To find $\mathcal{L}_1^{D1}$ we plug (4.2.3) and (4.2.2) into the welfare function:

$$\mathcal{L}_1^{D1} = \left[\frac{\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa\right]^2 + \varphi\left[-\kappa + \alpha\left(\frac{\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa - p\frac{\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa\right)\right]^2 \tag{A.6.8}$$

By doing simplifications as we did with the previous examples we arrive at:

$$\mathcal{L}_1^{D1} = \kappa^2\frac{\varphi(1+\varphi\alpha^2)}{[1+\varphi\alpha^2(1-p)]^2} \tag{A.6.9}$$

Finding $\mathcal{L}_2^{D1}$ is easier, given that is analogous to the discretionary equilibrium of the previous chapters. We plug (4.2.8) and (4.2.4) into the welfare function:

$$\begin{aligned}\mathcal{L}_2^{D1} &= (\varphi\alpha\kappa)^2 + \varphi(-\kappa + \alpha(\varphi\alpha\kappa - \varphi\alpha\kappa))^2 \\ &= \varphi\kappa^2(1+\varphi\alpha^2)\end{aligned} \tag{A.6.10}$$

$\mathcal{L}^{D1}$ is the intertemporal sum of $\mathcal{L}_1^{D1}$ and $\mathcal{L}_2^{D1}$:

$$\begin{aligned}\mathcal{L}^{D1} &= \kappa^2\frac{\varphi(1+\varphi\alpha^2)}{[1+\varphi\alpha^2(1-p)]^2} + \delta\varphi\kappa^2(1+\varphi\alpha^2) \\ &= \varphi\kappa^2[1+\varphi\alpha^2]\left[\left(\frac{1}{1+\varphi\alpha^2(1-p)}\right)^2 + \delta\right]\end{aligned} \tag{A.6.11}$$

Similarly, to find $\mathcal{L}_1^{D2}$ we plug $\pi_1^{D2} = 0$ and (4.2.2) into the welfare function:

$$\mathcal{L}_1^{D2} = (0)^2 + \varphi\left[-\kappa + \alpha\left(0 - p\frac{\varphi\alpha}{1+\varphi\alpha^2(1-p)}\kappa\right)\right]^2 \tag{A.6.12}$$

Simplifying we obtain

$$\mathcal{L}_1^{D2} = \varphi\kappa^2\left[\frac{1+\varphi\alpha^2}{(1+\varphi\alpha^2(1-p))}\right]^2 \tag{A.6.13}$$

To obtain $\mathcal{L}_2^{D2}$ we plug (4.2.9) and (4.2.7) into the social welfare loss function:

$$\begin{aligned}\mathcal{L}_2^{D2} = &\left[\frac{(1+p(\omega-1))\varphi\alpha}{(1-p)(1+\varphi\alpha^2)+p\omega}\kappa\right]^2 + \\ &\varphi\left[-\kappa + \alpha\left(\frac{(1+p(\omega-1))\varphi\alpha}{(1-p)(1+\varphi\alpha^2)+p\omega}\kappa - \frac{p\omega\varphi\alpha}{(1-p)(1+\varphi\alpha^2)+p\omega}\kappa\right)\right]^2\end{aligned} \tag{A.6.14}$$

After lots of algebra, the equation simplifies to

$$\mathcal{L}_2^{D2} = \varphi\kappa^2[1+\varphi\alpha^2]\left(\frac{1+p(\omega-1)}{(1-p)(1+\varphi\alpha^2)+p\omega}\right)^2 \tag{A.6.15}$$

The intertemporal welfare loss of the discretionary policymaker imitating the conservative central bank is

$$\mathcal{L}^{D2} = \varphi\kappa^2[1+\varphi\alpha^2]\left[\frac{1+\varphi\alpha^2}{(1+\varphi\alpha^2(1-p))^2}+\delta\left(\frac{1+p(\omega-1)}{(1-p)(1+\varphi\alpha^2)+p\omega}\right)^2\right] \tag{A.6.16}$$

A.6.6 SHOWING THAT THE DISCRETIONARY POLICYMAKER MIGHT BEHAVE CONSERVATIVELY

To show that the discretionary policymaker might want to set inflation to zero, imitating a conservative central bank, we would need to show that $\mathcal{L}^{D2} < \mathcal{L}^{D1}$. That is:

$$\begin{aligned}\varphi\kappa^2[1+\varphi\alpha^2]&\left[\frac{1+\varphi\alpha^2}{(1+\varphi\alpha^2(1-p))^2}+\delta\left(\frac{1+p(\omega-1)}{(1-p)(1+\varphi\alpha^2)+p\omega}\right)^2\right]\\ &<\varphi\kappa^2[1+\varphi\alpha^2]\left[\left(\frac{1}{1+\varphi\alpha^2(1-p)}\right)^2+\delta\right]\end{aligned} \tag{A.6.17}$$

This expression simplifies to:

$$\frac{1+\varphi\alpha^2}{(1+\varphi\alpha^2(1-p))^2}+\delta\left(\frac{1+p(\omega-1)}{(1-p)(1+\varphi\alpha^2)+p\omega}\right)^2<\left(\frac{1}{1+\varphi\alpha^2(1-p)}\right)^2+\delta \tag{A.6.18}$$

Instead of providing full proof that shows the conditions under this expression are satisfied, I will show that for a certain condition this expression is satisfied. I will postulate that $\varphi = 1,\ \alpha = 1,\ p = 1/2,\ \omega = 2/3$. Then the conditions are

$$\frac{1+1}{(1+(1-1/2))^2}+\delta\left(\frac{1+(1/2)(2/3-1)}{(1-1/2)(1+1)+(1/2)(2/3)}\right)^2<\left(\frac{1}{1+(1-1/2)}\right)^2+\delta \tag{A.6.19}$$

This expression is satisfied if and only if

$$\delta > \frac{256}{351} \approx 73\% \tag{A.6.20}$$

So, under the conditions of the example, if the central bank has a discount factor higher than 73%, then it's worth setting the low inflation in the first period.