4. Rules versus discretion: Credibility problems (I)*

Henrik Jensen
Department of Economics
University of Copenhagen

April 11, 2006

Abstract

Notes for the course "Monetary Economics: Macro Aspects," Spring 2006. The relevant literature behind these notes is:

Walsh (2003, Chapter 8, pp. 363-384; pp. 393-425).

1 Introductory remarks

- So far, monetary policy has been treated by some exogenous process, or policy rule. Is this an innocent assumption? No, as one has not specified *how* the policymaker can be guaranteed to adhere to such rules

- What if the policymaker under some circumstances had the incentive to conduct policy that differs from the rule? This will feed back into private sector behavior, as that behavior is based upon expectations about policy. It thus becomes of great importance whether the policy rule is credible

- If it is not credible, private sector expectations will adjust towards what is believed that the policymaker will do instead of following the rule. It therefore becomes important to understand the policymaker's incentives, as this will determine whether it will adhere to the rule, or act *discretionary;* i.e., act as it finds optimal at any given point in time. This, in turn, provides an understanding about the determination of private sector expectations and thus the macroeconomic equilibrium

- The central theme when policy is endogenized builds on a very general insight in frameworks where private sector expectations are important:

  - Policies that are optimal *ex ante*, are often not optimal *ex post*
  - They are *time inconsistent*

- This implies that the policymaker *does* have an incentive to deviate from the optimal rule when private sector expectations are formed

- This is then taken into account in expectations formation, and the economy may end up in an equilibrium which is time consistent, but *suboptimal*:

  - Given the expectations that incorporate the policymaker's incentive to deviate from the optimal rule. . .
  - . . . monetary policy may be "tempted," ending up delivering a poor result; due to lack of credibility
  - This is the positive/descriptive aspect of time inconsistency models

- If this is important for various policy scenarios, it raises the issue:

  - How can the policymaker be endowed with **commitment abilities**, which overcome the time-inconsistency problem of the optimal plan? I.e., securing credibility of the optimal plan?
  - I.e., how can society though legislation, or institutional design, change the policymaker's incentives such that the poor macroeconomic equilibrium is avoided?
  - This is the normative aspect of the time inconsistency models

2 Inflation and discretionary monetary policy

- The importance of the potential problems arising from discretionary monetary policymaking is clearly seen in a simple model of inflation
- The model, the "Barro and Gordon model," exemplifies how optimal monetary policy may be time inconsistent, and which negative consequences this has on the economy

The Barro Gordon "inflation bias model"

- The economic structure is given by a very simple variant of the MIU model with one-period nominal wage contracts. Aggregate supply of output y_t is given by:

$$y_t = y_n + a\left(\pi_t - \pi_t^e\right) + e_t, \quad a > 0 \tag{8.3}$$

  I.e., output is determined by a Lucas-style relationship, by which unanticipated high prices/inflation cause production to increase above the natural rate, y_n. The term e_t denotes some exogenous source of variation in output (say, variation in firm's input costs, variations in firms' mark-up; what is important is that the shock is viewed as economically inefficient). It is assumed that $\mathrm{E}_{t-1}e_t = 0$, where E denotes the mathematical expectations operator (i.e., rational expectations are assumed). Aggregate demand is given by a simplified quantity equation:

$$\pi_t = \Delta m_t + v_t \tag{8.4}$$

  where money growth only creates inflation, and where inflation is also subject to a stochastic disturbance. For the purpose of understanding the basics of the model, one only needs an even simpler AD schedule:

$$\pi_t = \Delta m_t. \tag{8.4'}$$

  This simplification enables one to treat the inflation rate directly as the monetary policy instrument, which is slightly easier.

- There are two variants of the model, which differs in term of how the central bank's preferences are modelled. The first variant is one where utility of the central bank depends linearly on output and quadratic on inflation:
- **Variant 1. Utility is linear in output:**

$$U_t^1 = \lambda\left(y_t - y_n\right) - \frac{1}{2}\pi_t^2, \quad \lambda > 0 \tag{8.1'}$$

  The central bank thus wishes output to exceed the natural rate as much as possible, but it dislikes deviations in inflation from zero (a normalized value for the inflation goal). The second variant is one where preferences are also quadratic in output:

- **Variant 2. Utility is quadratic in output**

$$U_t^2 = -\frac{\lambda}{2}\left(y_t - y_n - k\right)^2 - \frac{1}{2}\pi_t^2, \quad k > 0, \tag{8.2'}$$

  In this case, the central bank:

  - dislikes deviations in output from $y_n + k > y_n$
  - dislikes deviations in inflation from zero

  Hence, the differences between the two variants is that under variant 2, the central bank prefers a particular output level, and one that is higher than the natural rate (since $k > 0$), and dislikes variations in output. Under the first variant, output variability is not of importance, only a high level of output as possible is of importance. To understand this, note that mathematically there is a close relationship between the two functions, as

$$U_t^2 = -\frac{\lambda}{2}\left(y_t - y_n\right)^2 + \lambda k\left(y_t - y_n\right) + \frac{\lambda}{2}k^2 - \frac{1}{2}\pi_t^2.$$

  This is the same functional form as U_t^1 except for the term $-\left(\lambda/2\right)\left(y_t - y_n\right)^2$, which indeed reflects a preference for output stability (the other differences are not important; U_t^2 has $\lambda k y_t$ instead of λy_t, and includes constants $-\lambda k y_n + \left(\lambda/2\right)k^2$ which have no influence on behavior)

- Under both variants, the timing of events — which is very important — within the period is:

  1. π_t^e is formed
  2. e_t is realized
  3. π_t is set
  4. y_t is determined

- Note the following central features of the move structure:

  - The private sector is "committed" to a fixed inflation expectation, e.g., as nominal wage contracts are written in the beginning of the period
  - The private sector does not know e_t at the time they sign contracts; hence, π_t^e cannot conditioned on e_t
  - The central bank performs policy after having observed the shock, and **after** inflation expectations are formed
    - This opens room for stabilization of the economy against the e-shock, and even in absence of shocks, it opens the room for setting inflation differently from expected
    - I.e., there is possibility, for given expectations, for creation of unanticipated inflation leading to output effects when policy is carried out

- However, it is assumed that the private sector knows the incentives of the central bank. Technically, this means that the private sector know that at "point 3." the central bank will set π_t with the aim of maximizing utility. This knowledge is crucial for the determination of π_t^e, as central bank incentives affects actual inflation, and therefore also expectations about inflation. It is assumed that $\pi_t^e = \mathrm{E}_{t-1}\pi_t$, i.e., that expectations are formed rationally based on the information set price setters have when entering period t. We now look at the solution of the model under the two assumptions about preferences.

Solution of model under Variant 1

- The solution concept is a time-consistent Nash equilibrium
  - Nash equilibrium: Agents' choices are best responses to other agents' choices
  - Time consistency: Agent's choices, as prescribed by the equilibrium, will not change as time passes[1]
- To attain time consistency, one solves model by backward induction (by solving the model backwards, one secure that optimal responses as time passes are secured; thus, no deviations will occur). First, determine the optimal monetary policy, for given expectations and supply shock ("point 4." we do not need to consider as such, since it is merely a passive determination of y_t given what has happened at "points 1, 2 and 3"):

$$\begin{aligned}\max_{\pi_t} U_t^1 &= \lambda\left(y_t - y_n\right) - \frac{1}{2}\pi_t^2 \\ &= \lambda\left(y_n + a\left(\pi_t - \pi_t^e\right) + e_t - y_n\right) - \frac{1}{2}\pi_t^2 \\ &= \lambda\left[a\left(\pi_t - \pi_t^e\right) + e_t\right] - \frac{1}{2}\pi_t^2\end{aligned}$$

  where the second line follows by inserting (8.3) into (8.1'). The relevant first-order condition is

$$\lambda a = \pi_t. \qquad (8.5')$$

  The marginal gain of inflation (in terms of higher output) equals the marginal loss (in terms of inflation *per se*). This readily provides the solution for monetary policy,

[1] In this model, time consistency is the equivalent of a sub-game perfect Nash equilibrium. In general, sub-game perfection is a stronger restriction than time consistency. One can, in dynamic settings, have time-consistent equilibrium paths, which are not sub-game perfect. I.e., it does not pay to deviate from an equilibrium path as time passes (Nash, and time consistency are satisfied), but *if* deviation should occur (an out-of-equilibrium action is taken by some), one would not find it optimal to return to the original path (sub-game perfection is violated).

without using inflation expectations or the realization of the shock (both terms that the central bank knows when deciding on monetary policy). From (8,5'), inflation expectations are found easily:

$$\pi_t^e = \mathrm{E}_{t-1}\pi_t = \lambda a$$

Since actual inflation is independent of the shock, inflation is perfectly anticipated. The solution for output then becomes

$$y_t = y_n + e_t.$$

It thus follows that monetary policy has no effect on output, since it is perfectly anticipated by the public. Output thus equals the natural rate, unless an exogenous disturbance hits. The interesting feature of the equilibrium of the model is that even though monetary policy has no real effects, there will be a positive rate of inflation. This is harmful for the central bank according to its utility function (and presumably then also harmful for the economy as such). Therefore, the solution is said to feature an *inflation bias.*

- Why is that? A moment's reflection will convince you that everybody would be better off if output was given by $y_t = y_n + e_t$ and $\pi_t^e = \pi_t = 0$. This would clearly improve over the above mentioned solution, as output is the same, but inflation is absent. But the problem is that this outcome that makes everyone better of is not time consistent. If the private sector started out in "point 1." by expecting zero inflation (economically: signing low wage growth contracts), the central bank would conduct policy according to (8.5') thereby making $\pi_t > \mathrm{E}_{t-1}\pi_t$ in order to raise output.[2] But because the private sector understands this central-bank incentive to raise output by creating an "inflation surprise," the private sector incorporates this into expectations, implying $\mathrm{E}_{t-1}\pi_t = \lambda a$. Policy would still be made according to (8.5'), and there will be no output gain, but too high inflation. Hence, the term an inflation *bias*. So, although $\pi_t^e = \pi_t = 0$ is a better solution for all, it does not involve a credible monetary policy when the central bank's incentives are as they are in model. These incentives refrains the central bank from committing to a zero inflation policy. Instead the central bank is said to conduct policy in a *discretionary* fashion, i.e., by optimizing some objective function at every point in time possible.

- The model thus provides a very strong example of how a well-meaning policymaker without credibility, i.e., without an ability to commit, can bring the economy into a "bad" equilibrium. The finding that optimal plans may be time inconsistent (and thus potentially difficult to obtain) is a feature that appears in a wide range of dynamic policy settings. It was firstly clarified in the literature by Finn Kydland and Edward Prescott (1977, *Journal of Political Economy*), and was one out of two

[2] Output would become $y_t = \lambda a^2 + y_n + e_t$.

fundamental contributions to economics that brought them the Nobel prize in 2004.[3] The fact that Kydland and Prescott indeed as an example of their more general idea used a monetary policy example close to the one here, makes it even more strange that the model is known as the Barro and Gordon model. But Robert J. Barro and David Gordon published two papers in 1983 that considered the monetary policy angle in much more detail than Kydland and Prescott. In any case it made a huge impact onto the debate on whether policymaking should be based on simple rules (to which one could commit), or discretion (a behavior where one needs no commitment as it is time consistent). In the simple example here, the simple rule "$\pi = 0$" clearly dominates discretion: "$\pi = \lambda a$."

Solution of model under Variant 2

- The main message about the time inconsistency of optimal policy also applies when central-bank utility is quadratic in output, but we will see that policy now responds to the supply shock, as well as directly to inflation expectations.[4] Again, when policy is implemented, inflation expectations and the supply shock are taken as given, and the central bank solves

$$\begin{aligned}\max_{\pi_t}\ U_t^2 &= -\frac{\lambda}{2}\left(y_t - y_n - k\right)^2 - \frac{1}{2}\pi_t^2 \\ &= -\frac{\lambda}{2}\left(a\left(\pi_t - \pi_t^e\right) + e_t - k\right)^2 - \frac{1}{2}\pi_t^2\end{aligned}$$

  where the second line follows from insertion of (8.3) into (8.2). the relevant first-order condition is

$$-\lambda a\left(a\left(\pi_t - \pi_t^e\right) + e_t - k\right) = \pi_t.$$

  Again, the marginal benefit of inflation in terms of output is equalled to the marginal loss in terms of inflation per se.

- The central bank's behavior can now be seen as a "reaction function" towards inflation expectations and the shock. One gets

$$\pi_t = \frac{\lambda a}{1 + \lambda a^2}\left(a\pi_t^e - e_t + k\right) \tag{8.6}$$

  This function has the following features:

[3] The other contribution was their approach to modelling business cycles, which initiated the RBC research program in the early 1980s, and which has great impact today on quantitative economic research today.

[4] The quadratic preference specification is the most realistic, as it captures the notion of some optimal output level given by the consumption-leisure trade-off (with linear utility of output, it appears as if more output comes as a free lunch to society). New theories even show how a second-order Taylor approximation of the representative household's utility function can lead to specifications like U_t^2 (see, e.g., Walsh, 2003, chapter 11, *not* curriculum).

– Higher inflation expectations drive up inflation, as the marginal gain in terms of higher output increases (higher inflation expectations all things equal reduces output, causing the central bank to respond expansively)

– A positive supply shock is met by a contractive monetary policy, as stabilization of output now is a concern (a positive e_t drives up y_t inefficiently, and all things equal this contributes to a contractive monetary policy).

– The more the output goal exceeds the natural rate (the higher k) the higher will inflation be

- Since rational expectations are assumed, the private sector can foresee this reaction function. However, it does not know the value of the supply shock. The rational expectations assumption implies, however, that the private sector knows the shock's distributional characteristics, so the private sector knows that it has the value zero on average. Taking expectations on both sides of (8.6) thus yields

$$\begin{aligned} \mathrm{E}_{t-1}\pi_t &= \frac{\lambda a}{1+\lambda a^2}\left(a\mathrm{E}_{t-1}\pi_t^e + \mathrm{E}_{t-1}e_t + k\right) \\ &= \frac{\lambda a}{1+\lambda a^2}\left(a\mathrm{E}_{t-1}\pi_t + k\right), \end{aligned}$$

where the second line follows from $\mathrm{E}_{t-1}e_t = 0$ and the fact that $\mathrm{E}_{t-1}\pi_t^e = \mathrm{E}_{t-1}\mathrm{E}_{t-1}\pi_t = \mathrm{E}_{t-1}\pi_t$.[5] We can then solve for inflation expectations as

$$\mathrm{E}_{t-1}\pi_t = \pi_t^e = \lambda ak > 0.$$

Actual discretionary inflation then follows by inserting this into the central bank's reaction function as

$$\pi = \lambda ak - \frac{\lambda a}{1+\lambda a^2}e_t$$

Actual discretionary output then becomes:

$$y_t = y_n + \frac{1}{1+\lambda a^2}e_t.$$

Again an inflation bias, now λak, prevails with no effect on output. That is, average output, or, expected output, is equal to the natural rate y_n. The supply shock, however, now presents the central bank with a stabilization trade off each period. The shock affects output, and this is to some extent neutralized by monetary policy, but

[5] The last expression follows from the so-called *law of iterated projections*. This law says that if the information sets A and B are such that $A \subseteq B$, then for some random variable X, $\mathrm{E}[\mathrm{E}[X|B]|A] = \mathrm{E}[X|A]$. In words, your "$A$-based" expectations about your "B-based" expectations of X is just your "A-based" expectations of X. Even though B may contain more information, and thus making $\mathrm{E}[X|B]$ potentially more precise than $\mathrm{E}[X|A]$, this is of little to no help when you must find $\mathrm{E}[X|B]$ based on only A-information. In that situation, i.e., when finding $\mathrm{E}[\mathrm{E}[X|B]|A]$, your best estimate is just $\mathrm{E}[X|A]$. You cannot improve the estimate by knowing that better information exists, if you don't posses that information (you only have access to A, you don't have access to B even though you may know it is out there). In the model of the main text, $A = B$.

not fully as the central bank cares about inflation stability. Note indeed that a low value of λ implies high effect of the shock on output, but low effect on inflation (and *vice versa* for a high value of λ). The reason why monetary policy has output stabilization properties is that expectations are fixed before the shock hits. This creates a (short-run) nominal rigidity, which the central bank can use to stabilize output. The model thus features policy irrelevance only in the weak form; policy cannot affect average output, but the actual path of output (i.e., the fluctuations around average) is affected by the central bank in this rational expectations framework.

- Now note that if commitment to a policy rule $\pi_t = b_0 - b_1 e_t$ was possible, it follows that optimal values of coefficients are

$$b_0 = 0, \qquad b_1 = \frac{\lambda a}{1 + \lambda a^2}$$

The optimal stabilization properties of discretionary policy are retained, but the inefficient average properties due to the inflation bias will be removed. Ignore the shock for a moment. If the rule is credible (i.e., the central bank can commit to it), the private sector sets $\pi_t^e = 0$. But absence of credibility will imply that there are output gains of setting $\pi_t > \pi_t^e$ as $k > 0$ (i.e., the central bank wants output above the natural rate), and thus in anticipation of this, expectation is set at $\lambda a k$ such that any potential output gains cannot outweigh the inflation costs for the central bank when it performs policy. I.e., a central bank acting under discretion cannot credibly implement the optimal policy rule!

- Note that earlier debates about rules versus discretion in monetary policy would end up favoring discretion, if there was a need and possibility for stabilization against economic shocks.[6] But with time-consistency problems of optimal policies, this may not be the case. E.g., a rigid rule that does not respond to shocks, $\pi_t = 0$, may even be desirable in the Barro and Gordon model under Variant 2 preferences (it is obviously preferable under Variant 1). This is the case if the variability of e_t is not too large. Then the gain of getting rid of the inflation bias outweigh the excess volatility in output due to the shock (see the precise condition on p. 377 in Walsh, 2003).

2.1 An alternative exposition[7]

- Many practitioners have criticized the Barro Gordon model for giving a distorted picture of monetary policymaking. They claim that monetary policymakers are not

[6] Those favoring rules, like fixed rules for nominal money growth, did that because they questioned the ability of policymakers to respond to economic shocks appropriately, since policy was viewed as influencing the economy by long and variable lags (this idea is usually attributed to Milton Friedman). Hence, trying to stabilize shocks would just lead to policy-induced economic fluctuations.

[7] This subsection builds on Jensen (2003): "Explaining an Inflation Bias without Using the Word "Surprise," which is found from a link at the course website. The article and this subsection is not curriculum, but included for the interested.

attempting to "surprise" the public. Well, note that the model's equilibrium does not predict that either. In equilibrium, inflation expectations are set so high, that the central bank's optimal strategy indeed is not to "surprise" the public!

- In any case, one can explain the excessive inflation result — without using the "surprises" terminology — by reformulating the model. Consider an "inverted" version of the model's aggregate supply curve:

$$\pi_t = \pi_t^e + (1/a)(y_t - y_n) - (1/a)e_t$$

This is a conventional expectations augmented Phillips curve. Then assume that monetary policy affects aggregate demand through the nominal interest rate, but for simplicity, take y_t to be the policy instrument (assuming that the monetary authority can only affect demand imprecisely, does not affect the qualitative features of the results). The central bank then solves (assuming "Variant 1" preferences)

$$\begin{aligned}\max_{y_t} U_t^1 &= \lambda(y_t - y_n) - \frac{1}{2}\pi_t^2 \\ &= \lambda(y_t - y_n) - \frac{1}{2}\left(\pi_t^e + (1/a)(y_t - y_n) - (1/a)e_t\right)^2,\end{aligned}$$

and the relevant first-order condition is

$$\lambda = (1/a)\left(\pi_t^e + (1/a)(y_t - y_n) - (1/a)e_t\right)$$

Marginal gain in terms of output is equal to the marginal loss in terms of inflation

- When forming expectations, the private sector knows that this first-order condition holds, and taking expectations on both sides:

$$\lambda = (1/a)\pi_t^e$$

or,

$$\pi_t^e = \lambda a$$

which is the same as in the standard Barro-Gordon model. Setting this back into the first order condition gives the solution for output:

$$y_t = y_n + e_t$$

which again is the same as before. Finally, actual inflation follows from the Phillips curve as $\pi_t = \pi_t^e = \lambda a$, which again is the same as before. Hence, the equilibrium outcomes are unaffected by the alternative exposition (this is, of course, not terrible surprising as we in either case are searching for an optimal point in the inflation-output space given the aggregate supply constraint; whether that constraint is formulated as inflation as function of output, or output as a function of inflation, will not matter).

- There is will be an excessively high inflation rate, i.e., an inflation bias, but the "surprise" explanation can be avoided
  - The private sector knows that the central bank has a preference for stimulating demand and output
  - They know that this has inflationary effects via the Phillips curve
  - Therefore they push up inflation expectations, which feeds into actual inflation
  - In equilibrium, inflation expectations will be so high that the central bank will not try to bring output above the natural rate, as the additional inflationary consequences will be too severe.
  - No "surprises" needed to explain the intuition of the result.

3 Reputational solutions to credibility problems

- If the Barro and Gordon model has merit for real world monetary policy, a central question is whether it is possible to avoid the excessive inflation result? The short answer is yes. Two ways are examined in the literature. One builds on a "market-based" solution, the other builds on an "institutional based." This section looks at an example of a "market based" solution, namely that of reputation building
- Interactions between the private sector and policymakers are rarely of a one-shot nature. They occur repeatedly, and this opens the possibility for reputation building: If the central bank shows it delivers low inflation, the private sector may gradually start to believe that it will not attempt to create "surprise inflation." This basic idea can be modelled in a very simple manner. Assume that the interaction between the private sector and the central bank, the game, is repeated infinitely. Then, "doing bad" today, *may cause loss of reputation tomorrow* that in effect prevents the central bank from "doing bad" today For this to be the case, the private sector must "punish" bad behavior of the central bank; i.e., punish a deviation from a promised inflation rate.[8]
- Assume there are no supply shocks, and that we are examining Variation 1 utility (linear in output). The central bank promises some $\overline{\pi} < \lambda a$. The potential punishment behavior by the private sector is modelled by a simple "trigger strategy":

$$\begin{aligned} \pi_t^e &= \overline{\pi} < \lambda a \quad \text{if} \quad \pi_{t-1} = \pi_{t-1}^e \\ \pi_t^e &= \lambda a \quad \text{otherwise} \end{aligned}$$

[8] From Game Theory you may recall the Prisoners' Dilemma. In a one-shot version of that game, the Nash equilibrium is inefficient (both prisoners "sing"). In a repeated-game version it is possible that the efficient solution (both prisoners stay silent) is supportable as a Nash equilibrium. This is the case if each prisoner's *temptation* to deviate by betraying the other prisoner, is outweighed by the prospect of enduring a *punishment* phase with, for example, the inefficient equilibrium. The inflation bias model has many similarities with the Prisoners' Dilemma.

I.e., if the central bank did not "surprise" in the previous period the public keeps expecting the promise, $\overline{\pi}$. If not, it reverts to the "bad" equilibrium inflation expectations for one period. This is called a "tit for tat" strategy in game-theoretic language. Note that this is just one of infinitely many punishment strategies, and one can rightfully discuss how the private sector agents can coordinate on a particular one. This issue is ignored here, where the main purpose is to show that it is *possible* to obtain an equilibrium where inflation is lower than the one-period inflation bias. With such a private sector strategy, the central bank now takes decisions knowing that breaking a promise will have implications for future private sector behavior. The central bank's objective is to maximize the discounted sum of per-period utilities:

$$\max \sum_{t=0}^{\infty} \beta^t U_t^1, \quad 0 < \beta < 1,$$

where a crucial parameter is β, the discount factor (if, e.g., $\beta = 0$ the central bank only cares about the present, and doesn't care about future punishments).

- The question is now: which inflation rates, if any, can the central bank get a reputation for delivering?

- First assume $\pi_t^e = \overline{\pi}$. What is the optimal deviation policy in a period? From before, we know that the per-period optimal inflation rate is $\pi_t = \lambda a$. The associated net gain — *temptation* — for the central bank:

$$\lambda a \left(\lambda a - \overline{\pi}\right) - \frac{1}{2}\left(\lambda a\right)^2 + \frac{1}{2}\left(\overline{\pi}\right)^2 = \frac{1}{2}\left(\lambda a - \overline{\pi}\right)^2 = G\left(\overline{\pi}\right)$$

This is always positive. It depicts that the central bank gains by the higher output (even though inflation has increased).

- What is optimal policy under the punishment $\pi_t^e = \lambda a$? From before: $\pi_t = \lambda a$. Associated net loss — *enforcement* — against the central bank

$$-\frac{\beta}{2}\left(\overline{\pi}\right)^2 + \frac{\beta}{2}\left(\lambda a\right)^2 = \frac{\beta}{2}\left[\left(\lambda a\right)^2 - \left(\overline{\pi}\right)^2\right] = C\left(\overline{\pi}\right) \qquad (8.13)$$

This is also positive (as $\overline{\pi} < \lambda a$), and depicts the loss of having high inflation raters than low, and the same output.

- An inflation rate policy $\overline{\pi}$ is then credible if $C\left(\overline{\pi}\right) \geq G\left(\overline{\pi}\right)$; i.e., if the temptation is not stronger than the enforcement. This implies that

$$\frac{\beta}{2}\left[\left(\lambda a\right)^2 - \left(\overline{\pi}\right)^2\right] \geq \frac{1}{2}\left(\lambda a - \overline{\pi}\right)^2$$

must hold. This means that any policy $\overline{\pi}$

$$\lambda a \geq \overline{\pi} \geq \frac{1-\beta}{1+\beta}\lambda a$$

is credible in the repeated game; or, in game-theoretic language, it is a Nash equilibrium of the infinitely repeated game. Note that the minimum sustainable inflation rate is decreasing in β, and for $\beta \to 1$, it goes to zero, i.e., the most preferable inflation rate. This is essentially because the value of the enforcement of a deviation increases the more forward looking is the central bank. While a very stylized example, it models formally the simple idea that a policymaker may act in a better way, if it acknowledges that "bad" behavior today has "bad" implications in the future.

4 Delegation and independent central banks

- To recapitulate, time-inconsistency problems of optimal monetary policy create suboptimal outcomes when the central bank cannot credibly commit, and the Barro and Gordon model provides a simple example of the basic nature of such credibility problems. Importantly, models of credibility problems provide a natural platform for start thinking about how society can overcome credibility problems and their associated costs.[9] In the previous section some form of reputation building was examined. Such a "market based" solution, however, has problems. If a central bank is dominated by a short-sighted government, it will act as if it has a low discount factor, i.e., β is low. Then, only small improvements over the time-consistent inflation rate can be obtained. Also, the issue of coordination among the private sector agents raised the issue about how it can coordinate expectations around a low inflation rate (in the infinitely repeated game model, any inflation rate between the minimal $\overline{\pi}$ and λa can be an equilibrium). Finally, reputation may take time to build. Therefore, research in the past decade has focused on "institutional" approaches to solving credibility problems. These are normative theories of how to overcome time-inconsistency problems by altering the central bank's incentives in a direction that mitigates the incentives to deviate from the optimal plan

- In the context of the Barro and Gordon model, this take the form of theories of *delegation.* I.e., handing monetary policy over to central banks that are independent of the government, and who faces the appropriate incentives. The importance of independence from the government, arises because the incentive to conduct "surprise" inflation in order to increase output (or, equivalently, reduce unemployment)

[9] In the particular model the cost in an inflation bias. A multitude of other examples and costs can be set up. E.g.,:

- Speculative attacks on non-credible exchange rate pegs,
- Hyperinflations if governments cannot credibly refrain from inflationary finance,
- Disinflations typically have significant output costs,
- Suboptimal savings and investments, if governments cannot credible refrain from impose "surprise" taxes on wealth.

is often interpreted as arising from political pressures. Hence, a "solution" to time-inconsistency problem could be achieved by delegating monetary policy conduct to independent central banks. This means that society creates monetary institutions, which secure independence and appropriate policy incentives.

- Several solutions proposed in the literature
- Analyses here are cast in versions of the Barro and Gordon model with "variant 2" utility, I.e.,

$$U = -\frac{\lambda}{2}(y - y_n - k)^2 - \frac{1}{2}\pi^2$$

and with the usual aggregate supply schedule:

$$y = y_n + a(\pi - \pi^e) + e.$$

(We ignore time sub-scripts for now.) Again, it is assumed that inflation is the policy instrument. Recall, that the socially optimal policy (which is not time consistent) is:

$$\pi = -\frac{\lambda a}{1 + \lambda a^2} e$$

- When considering delegation, a new stage is introduced in the move structure; namely the "institutional design stage" where the monetary delegation scheme is implemented:

1. Establishment of monetary delegation regime
2. π^e is formed (rationally)
3. e is realized
4. π is set
5. y is determined

We first consider one of the most famous delegation schemes of the literature (which also has intuitive appeal), namely the one where monetary policymaking is put in the hands of a central banker, which has preferences that are different from those representing society.

Delegation to a "conservative" central banker

- The idea is to appoint a central banker that is "conservative" in the sense that it puts relatively more weight on inflation stabilization than society. Formally, monetary policy conduct is delegated to central banker with preferences

$$U^c = -\frac{\lambda}{2}(y - y_n - k)^2 - \frac{1+\delta}{2}\pi^2, \qquad \delta > 0$$

This represents the conservative central banker as modelled by Kenneth Rogoff (1985, *Quarterly Journal of Economics*). The parameter $\delta > 0$ measures the "degree of conservativeness." When this central banker performs discretionary, time-consistent monetary policy, its behavior is — taking as given π^e and e — characterized by the first-order condition

$$-\lambda a \left(a \left(\pi - \pi^e\right) + e - k\right) = \left(1 + \delta\right) \pi. \tag{*}$$

Importantly, when $\delta > 0$ the marginal cost of inflation is higher. Rational inflation expectations follow by taking expectations on both sides of (*):

$$\lambda a k = \left(1 + \delta\right) \pi^e,$$

leading to

$$\pi^e = \frac{\lambda a k}{1 + \delta} < \lambda a k$$

It is immediately seen that with a Rogoff-conservative central banker, the inflation bias is reduced from $\lambda a k$ to $\lambda a k / \left(1 + \delta\right)$. The reason is that the private sector foresees the central banker's reduced incentive to increase inflation to achieve output gains, when $\delta > 0$.

- This is a benefit of having a conservative central banker. Conservativeness in this model, however, has a cost. To see this, examine the solution for actual inflation (this is found by plugging the solution for π^e back into (*) and isolating π):

$$\pi = \frac{\lambda a k}{1 + \delta} - \frac{\lambda a}{1 + \delta + \lambda a k} e \tag{8.18'}$$

Stabilization of the shock is distorted. Compared to the socially optimal response to a supply shock, a conservative central bank responds less to the shock, and the result is too stable inflation and too unstable output (this is intuitive, and what should be expected, when policy is in the hands of an agent that puts relatively more weight on inflation stability. Hence, appointing a conservative central bank (and determining how conservative is should be) involves a trade-off between lower average inflation (the benefit) and poorer macroeconomic stabilization (the loss). The question is now, whether it can always be a good idea to appoint a conservative central banker? In model terms, will it always be optimal to have $\delta > 0$?

- The answer is unambiguously "yes." The reason is the following:

  - At $\delta = 0$ a marginal increase in δ involves a *first-order social gain* in terms of lower average inflation (at $\delta = 0$ average inflation is suboptimal)
  - At $\delta = 0$ a marginal increase in δ involves a *second-order social loss* of poorer stabilization (at $\delta = 0$ stabilization is already optimal)

– A positive value of δ is therefore *always* optimal

- While finding the exact value of δ within this model can only be done implicitly, the conservative central banker — appointing a governor with particular preferences — partially solves the time-inconsistency problem of monetary policymaking. Nevertheless, the cost of poorer stabilization begs the question of whether other preferences, or incentive structures, may solve the problem completely (i.e., remove the inflation bias while retaining an optimal response to the shock)[10] This is question is asked (and answered) in the "incentive contracts approach" to delegation.

Incentive contracts

- Under this approach, the government appoints a central bank, and offers him/her a *performance contract.* This contract rewards or punishes the central bank depending on its performance. Note that the contract could be "pecuniary" (i.e., when the central bank does well, the central banker will get a high salary) but more generally, it could represent public embarrassment if the central bank doesn't fulfill its "contract." A real-world analogy exists: The Federal Reserve Act of 1989 in New Zealand stipulates that the central bank governor can be fired, if he performs poorly. Formally, the central bank is offered a contract, such that it maximizes

$$U + t$$

where t is the contract transfer

- Now, assume that the contract transfer cannot be made contingent on the supply shock, so only a transfer depending on observed inflation is considered. I.e., $t = t(\pi)$. The task of government is to choose the optimal $t(\pi)$ at the institutional design stage. When the central bank performs policy, it takes expectations, the supply shock and the contract function as given, and maximizes

$$-\frac{\lambda}{2}\left(a\left(\pi - \pi^e\right) + e - k\right)^2 - \frac{1}{2}\pi^2 + t(\pi).$$

The relevant first-order condition is

$$-\lambda a\left(a\left(\pi - \pi^e\right) + e - k\right) = \pi - t'(\pi). \qquad (**)$$

Note that if $t'(\pi) < 0$, the marginal cost of inflation is higher than without the transfer; i.e., the contract punishes inflation increases. Rational inflation expectations follow by taking expectations on both sides of (**):

$$\lambda ak = \pi^e - \mathrm{E}\left[t'(\pi)\right],$$

[10] Note that refinements of the Rogoff solution exists, where the government overrides the central banker (and conducts policy by itself) in face of particular large supply shocks. This will reduce the loss from poorer stabilization

which implies

$$\pi^e = \lambda ak + \mathrm{E}\left[t'(\pi)\right]$$

Insert these expectations back into (**) to get actual inflation:

$$-\lambda a\left(a\left(\pi - \lambda ak - \mathrm{E}\left[t'(\pi)\right]\right) + e - k\right) = \pi - t'(\pi),$$

and thus

$$\pi = \lambda ak + \frac{\lambda a^2}{1+\lambda a^2}\mathrm{E}\left[t'(\pi)\right] + \frac{t'(\pi)}{1+\lambda a^2} - \frac{\lambda a}{1+\lambda a^2}e$$

Now note that optimal policy for society will be implemented (in this time consistent Nash equilibrium), if the transfer function satisfies

$$\lambda ak + \frac{\lambda a^2}{1+\lambda a^2}\mathrm{E}\left[t'(\pi)\right] + \frac{t'(\pi)}{1+\lambda a^2} = 0$$

In that case $\pi^e = 0$, and the policy response to the shock is $-\lambda a/\left(1+\lambda a^2\right)$. Remark that this equation holds if

$$t'(\pi) = -\lambda ak.$$

A contract, or, transfer function, with this property is:

$$t(\pi) = t_0 - \lambda ak\,\pi,$$

where t_0 is an arbitrary constant. Note that this is a *linear inflation contract* (Carl E. Walsh, 1995, *American Economic Review*). It is linear because the incentive to "surprise" the private sector is a constant in equilibrium. Hence, a constant marginal punishment eliminates the inflation bias (also for non-quadratic utility)

- In contrast to a conservative central banker, the linear inflation contract portrays the optimal **incentive structure**, implementing
  - **optimal average inflation** *and*
  - **optimal shock stabilization**
- I.e., under the optimal contract the central bank retains flexibility to respond optimally towards shocks

- Often monetary institutions, however, are set up to limit this flexibility in order to reduce, e.g., the central bank's vulnerability towards political pressures. This is often modelled as *targeting rules* prescribing goals that the central bank should achieve through policy. I.e., the central bank is judged on its ability to attain these goals (note that this is analogous to the contract approach). Examples are exchange rate targeting, inflation targeting, money supply targeting, etc.

Targeting rules

- Targeting rules can be either flexible or strict. Flexible allows some concern for "social welfare" in addition to attaining the target, while under strict targeting rules, attaining the target is the overriding objective of monetary policy. An example of flexible inflation targeting: The central bank's preferences are

$$U - \frac{h}{2}\left(\pi - \pi^T\right)^2, \qquad h > 0$$

where π^T is the inflation target, and where h measures how strict the targeting rule is; higher h means higher emphasis on attaining $\pi = \pi^T$, relative to addressing the preferences given by U (which very loosely could be interpreted as society's).

- The central bank optimal behavior, taking as given expectations and the supply shock, is then to maximize

$$-\frac{\lambda}{2}\left(a\left(\pi - \pi^e\right) + e - k\right)^2 - \frac{1}{2}\pi^2 - \frac{h}{2}\left(\pi - \pi^T\right)^2.$$

The relevant first-order condition is

$$-\lambda a\left(a\left(\pi - \pi^e\right) + e - k\right) = \pi + h\left(\pi - \pi^T\right) \qquad (***)$$

So, inflation targeting implies a higher marginal cost of inflation when $\pi > \pi^T$ (i.e., when the inflation target is overshot). Expected inflation follows from (***) as

$$\begin{aligned} \lambda ak &= \pi^e + h\left(\pi^e - \pi^T\right) \\ &\implies \quad \pi^e = \frac{h}{1+h}\pi^T + \frac{1}{1+h}\lambda\alpha k \end{aligned}$$

Hence, $h > 0$ reduces the inflation bias associated with λak, and a lower than social optimal value of π^T can also reduce average inflation. Actual inflation follows by inserting π^e into (***) and rearranging:

$$\pi = \frac{h}{1+h}\pi^T + \frac{1}{1+h}\lambda\alpha k - \frac{\lambda a}{1+h+\lambda a^2}e$$

Shock stabilization is distorted, $h > 0$, as inflation variability per se is given relatively high priority when $h > 0$ (as opposed to under the linear contract, where inflation is given a constant marginal penalty; inflation variability is not punished). Indeed, the flexible inflation targeting regime "mimics" a conservative central banker for $\pi^T = 0$, as the relevant utility function of the central banker is

$$-\frac{\lambda}{2}\left(y - y_n - k\right)^2 - \frac{1+h}{2}\pi^2.$$

It can be seen that h has just replaced δ. So a flexible inflation targeting rule is equivalent to appointing a Rogoff conservative central banker.

- An analogy between flexible inflation targeting and a linear inflation contract has been provided by Lars Svensson (1997, *American Economic Review*). Consider the utility function:

$$-\frac{\lambda}{2}(y-y_n-k)^2-\frac{1+h}{2}\left(\pi-\pi^T\right)^2.$$

With this, the central bank's first-order condition reads

$$-\lambda a\left(a\left(\pi-\pi^e\right)+e-k\right)=(1+h)\left(\pi-\pi^T\right)$$

and inflation expectations become

$$\begin{aligned}\lambda ak &= (1+h)\left(\pi^e-\pi^T\right)\\ &\implies \quad \pi^e=\frac{\lambda ak}{1+h}+\pi^T\end{aligned}$$

Actual inflation then follows as

$$\pi=\frac{\lambda ak}{1+h}+\pi^T-\frac{\lambda a}{1+h+\lambda a^2}e$$

Optimal policy for society is achieved with $h=0$ and $\pi^T=-\lambda ak$. Hence, no conservativeness ($h=0$) but a constant marginal penalty of inflation ($\pi^T=-\lambda ak<0$); just like under the linear contract.

- Strict targeting rules will typically involve trade-offs. A strict inflation targeting rule corresponds to $h\rightarrow\infty$ and implies $\pi^T=0$
  - Excessively unstable output
  - Undesirable if supply shock variance is large relative to k — the average output distortion causing the credibility problem

- It may, however, be easier to monitor and enforce

Evidence on time-inconsistency problems

- Not many direct tests of the model per se; mostly indirect evidence. However, Peter Ireland (1999, *Journal of Monetary Economics*) finds time-series support for the Barro and Gordon model on US data, when it is extended with a time-varying natural rate
  - The lower US inflation in the 1980s, compared with the 1970s, is consistent with the model as the natural rate of unemployment has fallen (according to the model, this should weaken the incentives to expand thus implying a smaller inflation bias)

- David Romer (1993, *Quarterly Journal of Economics*) finds cross-country evidence, as an open economy version of the Barro and Gordon model predicts that more open economies should have a lower inflation bias

  - The reason is that the cost of a monetary expansion is higher in an open economy:

    - The associated real depreciation creates additional higher CPI inflation

  - The negative relationship between inflation and openness is found for non-OECD countries

- Much indirect evidence shows that higher degrees of central bank independence are negatively correlated with inflation across countries

  - The lower inflation does not, however, come at the cost of more output instability as Rogoff model would suggest
  - Is this indication of institutional design shaping incentives in the right direction?

5 Summary

- The Barro and Gordon model: A simple vehicle for thinking about credibility problems. It provides a framework for thinking about how to overcome them. It has had enormous influence on real-life policy design (the ECB, inflation targeting regimes, etc.). Various institutional mechanisms can be seen as reducing or eliminating the time inconsistency problem. The model and related frameworks have thus had great influence on how policymakers and academics think about "Designing institutions for monetary stability" (the title of Torsten Persson and Guido Tabellini's, 1993, *Carnegie-Rochester Conference Paper Series* article on the contract approach)

Appendix

A Key concepts you should know

Inflation and discretionary monetary policy

- The concept of time inconsistency of optimal policies in macroeconomic policy
- The concept of the suboptimality of time consistent, discretionary policies
- The concept of the optimality of time inconsistent, commitment-based policies
- The Barro Gordon 1983 model
  - Version with utility linear in output (no output stabilization motive)
  - Version with utility quadratic in output (an inflation stabilization motive)
- Real effects of unanticipated inflation
- The resulting inflation bias under discretion in version with utility linear in output
- The time inconsistency of commitment policy
- The resulting inflation bias under discretion in version with utility quadratic in output: An output objective above the natural rate
- Optimal stabilization properties of discretion and commitment policies
- The model turned "upside down" avoids "inflation surprise" explanations

Reputational solutions to credibility problems

- Reputation building when interaction is repeated
- The temptation to deviate versus the enforcement following a deviation
- The role of the discount factor for securing low inflation

Delegation and independent central banks

- The idea of setting up institutions shaping central banks' incentives so as to mitigate the inflation bias/credibility problems
- The "conservative" central banker

  - Trade-off: Lower inflation, but too unstable output
  - Some conservatism, however, always optimal

- Incentive contracts

  - The linear inflation contract
  - No trade-off: Eliminates inflation bias and delivers optimal stabilization

- Targeting rules

  - Flexible versus strict
  - Flexible inflation targeting and relationship with Rogoff-conservativeness
  - Flexible inflation targeting and relationship with linear contract