The Barro-Gordon Model

M. McMahon
University of Warwick
July 29, 2014

This note outlines the Barro-Gordon model of time-consistent monetary policy, discussing the meaning of the equations and how to solve the model. I also present a game-theoretic outline of what is going on in the model which may help some of you to understand the material more easily. This note is not a substitute for understanding the lecture material but will hopefully act as a complement. As always, I welcome comments and suggestions on ways of improving this note.

1 The Idea of Time Inconsistency

The key idea to take from this model is that when the government is given an opportunity to cheat workers, they will - but knowing that they will be cheated, rational agents will lead to a higher level of inflation but with no gain in terms of lower unemployment. In this model, the incentive to cheat derives from the fact that inflation expectations are set in advance of the government setting current policy and so there is a chance for the government to exploit a trade-off between inflation and unemployment (the Phillips curve) - the government will choose higher inflation (above expected inflation) to drive down the unemployment rate. All this, despite the fact that everyone knows that in advance, the government can do no better than achieve the inflation target and the natural rate of unemployment - in technical terms, there are differences between ex-ante and ex-post incentives to set good policy.

This idea is known as the time inconsistency problem and is actually not due to Barro or Gordon! It is originally due to the work of Kydland and Prescott for which, along with real business cycle theory, they were awarded The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel in 2004. The Kydland-Prescott work was not related only to monetary policy but to all types of government policy - they

\[\text{Footnote: This is what we called the Economics Nobel prize but it is actually not a proper Nobel prize... but still worth winning!}\]
showed that when the economic agents were rational, and where expectations influence the
decision of government policy, discretionary policymaking made everyone worse off compared with a situation in which the government could commit to a particular policy path - i.e. there was a credibility problem, or a time inconsistency problem.  

The 1983 works of Barro and Gordon were focused on the issue of monetary policy and in particular highlighted the role for monetary rules as a potential means to overcome the time inconsistency problem in monetary policy. In the next few pages I will cover the solution to the basic model and hopefully drive home the intuition to the results obtained.

2  The Simple Model

2.1 Set-up

The model comprises of the following key elements:

**Loss Function**  \[ L = (U - U^*)^2 + a(\pi - \pi^*)^2 \]

This tells us that society (the economy) is better off in the sense that it has smaller losses when inflation (\(\pi\)) is close to target (\(\pi^*\)), and unemployment (\(U\)) is also close to target (\(U^*\)). In fact, the best that society can achieve according to this loss function is \(L = 0\) by setting \(\pi = \pi^*\) and \(U = U^*\). There are a number of important features of this loss function:

- the parameter \(a\) determines the relative importance of inflation relative to unemployment - for example when \(a\) is very high, inflation deviations are very bad for social welfare.
- because of the square terms, the loss can never go negative;
- also because of the square terms, the loss is quadratic and so increases very quickly as deviations from target increase - e.g. if \(\pi^* = 0\) but \(\pi = 1\), then the loss is 1, but when inflation doubles (\(\pi = 2\)) the loss quadruples!

**Phillips Curve**  \[ U = U^N - b(\pi - \pi^e) \]

This is the economy that we have to deal with in this model - it is a very simply relationship between the level of unemployment and inflation. It tells us that unemployment

\[^2\]There is a great article written about the work of Kydland and Prescott which is available at: http://nobelprize.org/economics/laureates/2004/ecoadv.pdf

is the natural rate of unemployment \( (U^N) \) plus or minus something determined by the difference between inflation \( (\pi) \) and inflation expectations \( (\pi^e) \). This relationship can be justified if we think about the real wage facing firms - inflation \( (\pi) \) measures the amount that their output prices have increased, while workers’ inflation expectations \( (\pi^e) \) will be reflected in wage demands and so reflect the change in firms wage costs. So if \( \pi > \pi^e \) then the firm is increasing output prices more quickly than costs and so real wages are falling - this leads firms to increase demand for labour and therefore unemployment falls. On the other hand, if \( \pi < \pi^e \) then the real wage is increasing and firms reduce demand for labour and therefore unemployment rises.

A few things are worth discussing. \( U^N \), the natural or structural rate of unemployment, is the level of unemployment that prevails when the economy is in steady-state (people often look at average rates over the business cycle to give an indication of what the natural rate is). There are many things, all of them structural, that determine the size of the natural rate of unemployment. The coefficient \( b \) tells us the slope of the Phillips curve which determines the extent to which

**Policy instrument** Government/Central Bank chooses current inflation \( (\pi) \)

In this model, we assume for simplicity that the government/central bank can choose the current rate of inflation at the start of the period precisely - that is, we don’t assume any problems in achieving a particular rate of inflation. In reality, we know that central banks have to choose an interest rate today, which will then act through a number of channels in the economy, to influence the rate of inflation in 18 months or 2 years time. This would introduce more complex dynamics into the model but the end result should be about the same - even if we take into account the fact the monetary policy acts with long and variable lags!

Therefore, in solving the model we will assume that the government chooses the level of current inflation \( (\pi) \) to minimise the loss \( (L) \).

**Assumed Political Preferences** \( U^* < U^N \)

We assume that the government wants to have lower unemployment than is the natural or structural rate. There are many easily believed stories as to why this may be the case.

---

4. This is called the monetary transmission mechanism and is a key concept in macroeconomic policymaking. It is also a concept that we discuss when talking about how the same interest rate set by the ECB might affect different countries in different ways - an important consideration when discussing the costs and benefits of joining a single currency. For more on the transmission mechanism, see the Bank of England article from May (1999) at: http://www.bankofengland.co.uk/publications/other/monetary/montrans.pdf

5. Of course, a very genuine complaint about this model is that in a world where all workers behave rationally, how is it rational for the government to try to achieve an unemployment rate that is below the natural rate? They should instead focus their efforts on reducing the \( U^N \) through supply-side policies such as increasing the flexibility of the labour market.
- the most simple one being that unemployed workers might vote a government out of office and so governments will try to keep unemployment low to signal that they are doing a good job. Another very nice story surrounds unions who may put political pressure on governments to try to create higher employment in the economy and so drive $U^*$ down - this is a particularly neat story because as we mentioned above, unions can drive $U^N$ up and as we will see below the gap between $U^N$ and $U^*$ will be the key to determine the extent of the problem of the inflation bias.

Notice that because of this assumption, the optimal solution of $L = 0$ is not attainable as setting $\pi = \pi^*$ will result in $U = U^N > U^*$. Therefore, there will be some losses in this economy in steady-state because the government is trying to achieve something that is inherently unachievable.

**Timing of the Model**

The timing in the model is key to the models predictions - in particular, it is key that inflation expectations are formed before the government sets the current rate of inflation. Otherwise, there would be no opportunity for the government/central bank to cheat since worker expectations would be formed after current inflation is set and so they would never be fooled. The precise timing of the model is as follows:

**Pre-Time** Before the economy begins, the parameters $a, b, c$, as well as the targets $U^*$ and $\pi^*$ are determined.

**Start of Period** At the beginning of each period, the workers must form there expectations of inflation in the forthcoming period ($\pi_e$)

**Middle of Period** The government/central bank selects the policy for the current period ($\pi$) taking $\pi_e$ as given.

**End of Period** Given the choice of $\pi$ and $\pi_e$. The Phillips curve determines $U$ and social welfare ($L$) is therefore determined.

**Worker Expectations**

In solving this model, the way in which workers form their expectations of inflation (and therefore the wage increases that they seek) is key. In solving the model, we consider 3 different approaches - namely, naive expectations, rational expectations (RE) and also briefly adaptive expectations, though the RE solution is the most important one.
2.2 Naive Solution

Let us first do the naive solution to highlight the incentive to cheat. In this case, the workers believe that since the economy can never do better than setting $\pi = \pi^*$ and letting $U = U^N$. The problem for the government is now:

$$\min_{\{\pi\}} L = (U - U^*)^2 + a(\pi - \pi^*)^2 \quad (1)$$

$$s.t. U = U^N - b(\pi - \pi^e) \quad (2)$$

and $\pi^e = \pi^*$ given \( (3) \)

So we substitute in the Phillips curve:

$$\min_{\{\pi\}} L = (U^N - b(\pi - \pi^e) - U^*)^2 + a(\pi - \pi^*)^2$$

and take the derivative wrt $\pi$ to minimise the function with $\pi^e = \pi^*$ taken as given (we should also really check the second order condition):

$$\frac{dL}{d\pi} = 0 \iff 2(U^N - b(\pi - \pi^e) - U^*)(-b) + 2a(\pi - \pi^*) = 0$$

$$\Rightarrow (b^2(\pi - \pi^e)) + a(\pi - \pi^*) = b(U^N - U^*) \quad (4)$$

And then we can use the fact that $\pi^e = \pi^*$:

$$\Rightarrow (b^2(\pi - \pi^*)) + a(\pi - \pi^*) = b(U^N - U^*) \quad (5)$$

$$\Rightarrow \pi = \pi^* + \frac{b}{a + b^2}(U^N - U^*) > \pi^e = \pi^* \quad (6)$$

Equation 6 shows us how when the workers set $\pi^e = \pi^*$, there is ex-post (after the event) incentive for the government to run higher inflation than expected and try to drive down unemployment below the natural rate (since $\pi > \pi^e$, the Phillips curve tells us that unemployment falls as real wages have declined).

2.3 Rational Expectations Solution

Obviously the naive result is somewhat foolish as it relies on workers being consistently fooled in thinking that the government will not cheat them despite the fact that they will have the incentive and the capability to do so. The rational expectations (RE) solution is the major contribution of Kydland and Prescott and formed part of the rational
expectations revolution in macroeconomics. What it means is that agents in the economy
will not consistently allow themselves to be fooled by policymakers.

In order to solve this version of the model, we will assume that agents are correct in
their expectations - i.e. that \( \pi^e = \pi \) in the solution to the model. What is critical is
that we only apply this condition to the first order condition (equation 4) and not before.
Therefore, in the rational expectations solution becomes:

\[
\Rightarrow (b^2(\pi - \pi^e)) + a(\pi - \pi^*) = b(U^N - U^*) \tag{7}
\]

And then we can use the fact that \( \pi^e = \pi \):

\[
\Rightarrow (b^2(\pi - \pi)) + a(\pi - \pi^*) = b(U^N - U^*) \tag{8}
\]

\[
\Rightarrow \pi^e = \pi = \pi^* + \frac{b}{a}(U^N - U^*) > \pi^* \tag{9}
\]

Equation 9 shows us how in a world of rational agents, the level of inflation is above
the target (\( \pi > \pi^* \)) and so we generate an extra loss (\( L \)). Also notice, that since \( \pi = \pi^e \)
the level of unemployment is \( U = U^N \) so we get higher inflation but no "reward" in terms
of lower unemployment! This is the positive inflation bias from the time inconsistency
problem - inflation is higher than it would optimally be even though \( U = U^N \).

2.4 Adaptive Expectations Solution

I will not solve the adaptive expectations model, but it is easy to see that it is a slow
progression between the Naive solution (give reference equation) and the RE solution
(give reference equation). For example with fully adaptive expectations, after being
fooled by the policymaker setting \( \pi = \pi^* + \frac{b}{a+b^2}(U^N - U^*) \) in period 1, the agents now
set \( \pi^e = \pi^* + \frac{b}{a+b^2}(U^N - U^*) \) in period 2. You can now solve the model again taking
\( \pi^e = \pi^* + \frac{b}{a+b^2}(U^N - U^*) \) as given and you will find that the incentive is still for the
CB to set \( \pi > \pi^* \). You can continue doing this until the answer converges on the RE
solution but it takes quite a long time and is full of really annoying algebra.\[^{7}\]

\[^{6}\]One way to think about this is that the workers figure out that if they set their inflation expec-
tations as \( \pi^e = \pi^* \) then the government will cheat them and they can work out what the government will do (it
is assumed the workers all know the model and the necessary mathematics). They will keep trying to
find a level of \( \pi^e \) such that the government does not have an incentive to cheat - i.e. the level such that
\( \pi = \pi^e \). Alternatively, we can use this assumption to short-cut to the correct RE solution.

\[^{7}\]In fact, during my MSc year at LSE I tried this once for a couple of hours and can tell you now that
I will never be trying it again!
2.5 Insights and policy implications

So to summarise the model, we can use the following table:

<table>
<thead>
<tr>
<th>Inflation ($\pi$)</th>
<th>Unemployment ($U$)</th>
<th>Loss ($L$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Solution</td>
<td>$\pi &gt; \pi^e = \pi^*$</td>
<td>$U &lt; U^N$</td>
</tr>
<tr>
<td>RE solution</td>
<td>$\pi^e = \pi &gt; \pi^*$</td>
<td>$U = U^N$</td>
</tr>
</tbody>
</table>

Therefore the key insight is that simply because the government cannot commit to a policy path, the economy ends up in a less optimal equilibrium. The key policy insight is therefore that we should look at ways of allowing the government/central bank to commit not cheating to lower inflation. Numerous suggestions have been put forward and used in the real world; for example, delegating monetary policy to an independent central bank who only cares about an inflation target (inflation targeting ala the Bank of England).

Of course, the other possible solutions to this model highlight what are the key assumptions for solving the problem:

- Timing assumption - if the world was more flexible, and workers set the wages more regularly (or after the Central Bank has set policy), then there is no possibility to cheat and the problem goes away.

- Assumption that $U < U^N$ - as discussed above, the size of the inflation bias depends on the difference between $U$ and $U^N$, and therefore reducing this gap to zero eliminates the problem. See also footnote 5 above.

2.6 Alternative approaches

Of course, the simple model outlined here is not the only specification of the model that is possible. There are many similar specifications of the model that may work (and can be solved in the same way) or even applications of this idea to other situations such as fiscal policy and optimal taxation.
3 Game-theoretic outline of the model

Here is the basic "game" played by the authorities and the wage setters (workers)\(^8\). In the beginning of each period, it is optimal for the government or CB to set \( \pi = \pi^* \) - this is the ex-ante incentive. However, in the next stage of the game workers must set \( \pi^e \) and then only after \( \pi^e \) are set, will the government actually set \( \pi \). Once \( \pi^e \) is fixed, the government now has an incentive to lower \( L \) by choosing slightly higher inflation but getting \( U \) below \( U^N \) and so closer to \( U^* \). This is the ex-post incentive and depends on the parameters of the model as we have seen above.

Notice that although the most desirable outcome for social welfare is \((1,1)\)^9, this outcome will not be achieved. That is because rational workers know that if they set \( \pi^e = \pi^* \), the government will then choose \( \pi > \pi^* \) which yields a higher outcome for the government (2) but lower for workers who lose out. Therefore, the workers will always choose to set \( \pi^e > \pi^* \), then the government is forced to choose \( \pi = \pi^e > 0 \) which yields an outcome \((0,0)\). This is lower than the best outcome, but is time consistent!

Of course, those of you who are good with game theory will realise that the problem of time inconsistency is reduced if we consider the policy game as a repeated game rather than a one-shot game. In a repeated game, we can build up a reputation which leads to the optimal outcome - in fact, building monetary policy credibility is one of the main aims of central banks around the world, a trend due largely to the work of Barro and Gordon and others.

---


\(^9\)The payoffs are defined as:

( payoffs to government, payoffs to workers )
Figure 1: Game Theoretic Outline of the Barro-Gordon Model

[Central Bank announces that it would ideally like to set $\pi = \pi^*$]

- **Workers believe that $\pi = \pi^*$,**
  - $\therefore$ set $\pi^e = \pi^*$
  - CB sets $\pi = \pi^*$
    - $U = U^N$
  - CB sets $\pi > \pi^*$
    - $U < U^N$

- **Workers not believe that $\pi = \pi^*$,**
  - $\therefore$ set $\pi^e > \pi^*$
  - CB sets $\pi = \pi^*$
    - $U > U^N$
  - CB sets $\pi = \pi^e > 0$
    - $U = U^N$

<table>
<thead>
<tr>
<th>Workers believe</th>
<th>Workers not believe</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1,1]</td>
<td>[2,-1]</td>
</tr>
<tr>
<td>[-1,2]</td>
<td>[0,0]</td>
</tr>
</tbody>
</table>