Why speed doesn't kill: Learning to believe in disinflation

E Schaling and M Hoeberichts

1 Introduction

Central banks throughout the world are moving to adopt long-run price stability as their primary goal. As pointed out by Kahn (1996), whether operating under the mandate of official inflation targets, virtually all central banks have recognised the desirability of achieving price stability over time. Countries with moderate to high inflation are adopting policies to reduce inflation, and countries with low inflation are adopting policies to maintain price stability.

There is agreement among central bankers, academics and financial market representatives that low or zero inflation is the appropriate long-run goal of monetary policy[1].

However, there is less agreement on what strategies should be adopted to achieve price stability. For example, on the one hand we have the view of Ball (1997) that a major cause of rising unemployment in the 1980s in Organisation for Economic Co-operation and Development (OECD) countries was the tight monetary policy that those countries pursued to reduce inflation. On the other hand, we have the view of Sargent (1986), who has argued that a sharp disinflation may be preferable to gradualism because the latter invites speculation about future reversals or “U-turns” in policy.

The preference of central bankers – as expressed by King (1996) at the Kansas City Fed symposium on Achieving Price Stability at Jackson Hole – seems to be for a gradual timetable, with inflation targets consistently set below the public's inflation expectations.

Throughout, King (1996) emphasises the role of learning by central banks and the public. He shows how the optimal speed of disinflation depends crucially on whether the private sector immediately believes in the new low inflation regime or not. If it does, the best strategy is to disinflate quickly, since the output costs are zero. Of course, if expectations are slower to adapt, disinflation should be more gradual as well.

King illustrates both these polar cases, labelling the first a “completely credible regime switch” and the second “exogenous learning”. He also considers the case where private sector expectations are a weighted average of the central bank's long-run inflation target and the lagged inflation rate. This is termed “endogenous learning”.

Obviously, case 3 is a mixture of cases 1 and 2. Expectations do not adjust immediately (they depend on actual inflation experience, and hence on the policy choices made during the transition), but are not completely exogenous either.

But case 3 is problematic, since the learning process implies that the learning parameter does not depend on the monetary regime. Put differently, the updating mechanism does not reflect the actual speed of disinflation, and thus it is not clear whether the private sector expectations mechanism is rational. Our first suggestion is that, alternatively, case 3 could be modelled with the aid of Bayesian learning. Therefore, in the first part of the paper we modify the King model with Bayesian learning.

In his discussion of *endogenous learning,* King (1996, p. 68) says that there are good reasons for the private sector to suppose that in trying to learn about the future inflation rate, many of the relevant factors are exogenous to the path of inflation itself.

[1] For a survey of positions see Kahn (1996).

But a central bank may try to convince the private sector of its commitment to price stability by choosing to reduce its inflation target towards zero too quickly. King calls this "teaching by doing". Here the choice of a particular inflation rate influences the speed at which expectations adjust to price stability.

The problem, however, with the King (1996) model is that – as in the case of exogenous learning – the central bank decides on its optimal disinflation plan given those private sector expectations. Thus, although King calls this case “teaching by doing”, a more accurate description would be "doing without teaching". The result is that – as in the case of *exogenous learning* – price stability is reached gradually and is therefore not surprising.

In the second part of the paper "teaching by doing" is modelled differently. We allow the central bank's "doing" to affect private sector learning. Of course, if the central bank recognises its potential for active "teaching", its incentive structure changes. More specifically, it should realise that by disinflating faster, it can reduce the associated output costs by "teaching" the private sector that it means business. Thus, the dependence of private sector expectations on the actual inflation rate should be part of its optimisation problem.

This is, in fact, what we find: Allowing for "teaching by doing" effects always speeds up the disinflation *vis-à-vis* the case where this effect is absent. This paper clarifies why "speed" in the disinflation process does not necessarily "kill" in the sense of creating large output losses.

Another way to think about this is that here central bank *credibility* – a crucial variable in defining the output loss of the disinflation – is endogenous. In fact, the central bank's credibility – defined by Blinder (2000) as the extent to which words match deeds – can be increased by the central bank by starting off the disinflation process by putting its money where its mouth is.

However, there is one caveat. In the above analysis – given expectations – the output costs of disinflation are constant and given by the slope of the Phillips curve. In the present model this parameter is normalised at unity. It follows that a 1 percentage point of disinflation causes a 1 per cent fall in output. However, if we allow the output costs of disinflation to vary with the inflation rate, the central bank's incentives change substantially. The state-contingent output costs of disinflation are analysed by means of allowing for a convex Phillips curve.

With this modification, the outcome is then a disinflation strategy that is more gradualist than in a standard linear-quadratic model (with or without "teaching by doing" effects).

The remainder of the paper is organised as follows. In Section 2 we modify the King model with Bayesian learning. Section 3 extends the analysis with state-contingent output costs. Section 4 examines "teaching by doing". The plan of the paper is summarised in Table 1.1.

Table 1.1: Classification of states

Learning	Phillips curve	
	Linear	Convex
Bayesian	(Section 2)	II (Section 3)
Teaching by doing	III (Section 4.1)	IV (Section 4.2)

Our conclusions are given in Section 5. The appendices provide the derivations of the optimal degree of gradualism, optimal disinflation and the degree of monetary accommodation under “teaching by doing”.

2 A basic model of disinflation

A benchmark model that is useful to analyse the optimal speed of disinflation is the King framework. King (1996) shows that the optimal speed of disinflation depends crucially on whether the private sector immediately believes in the new low inflation regime or not. If it does, the best strategy is to disinflate quickly, since the output costs are zero. Of course, if expectations are slower to adapt, disinflation should be more gradual as well.

King illustrates both these polar cases, labelling the first a "completely credible regime switch" and the second "exogenous learning". Without loss of generality we may recap those results by using a two-period version of the King model.

Our model analyses the interaction between private sector expectations and the monetary regime, and in particular, the speed at which the inflation target implicit in the latter converges to price stability. The model features nominal rigidity and an optimising central bank that trades off price (inflation) versus output stabilisation. Without loss of generality we have normalised the inflation and output targets at zero.

It follows that monetary policy is a choice of an *ex ante* inflation rate π_t, where the central bank minimises (no discounting)

$$\underset{\{\pi\}}{E}\sum_{t=1}^{2}\frac{1}{2}\left[a\pi_t^2 + y_t^2\right] \tag{2.1}$$

subject to

$$y_t = \pi_t - \pi_t^e \qquad (2.2)^2$$

where $\pi_t^e = E_{t-1}\pi_t$

In the case of a *fully credible regime switch*, we have

$$\pi_t^e = \pi_t \tag{2.3}$$

$$y_t = 0 \tag{2.4}$$

Thus, if expectations adjust straight away to the new inflation target, there will be no output costs associated with the disinflation, and in that sense reducing inflation is costless.

King also analyses the opposite case that is labelled *exogenous learning*. In this case, the output costs of disinflation are non-trivial but depend solely on the mechanics of the inflation expectations which, in turn, do not reflect the monetary regime. Here the loss during the transition to a situation of price stability is:

$$\sum_{t=1}^{2}\frac{1}{2}\left[a(\pi_t)^2 + \left(\pi_t - \pi^e\right)^2\right] \tag{2.5}$$

[2] For analytical convenience, the stochastic productivity shocks, and the slope of the Phillips curve (the parameter b in the King model) is set equal to unity.

Differentiating with respect to π_t gives the optimal monetary policy as

$$\pi_t^* = \frac{1}{1+a}\pi_t^e \tag{2.6}$$

Of course, from Equation 2.6 it is clear that if expectations are slower to adapt, the disinflation should be more gradual as well. The inflation rate should decline as a constant proportion of the exogenous expected inflation rate.

King also considers a third case where private sector expectations are a weighted average of the central bank's long-run inflation target (which we have normalised at zero) and the lagged expected inflation rate. This is expressed as

$$\pi_t^e = \rho\pi_{t-1}^e \tag{2.7}$$

This is termed *endogenous learning*. The smaller ρ is, the faster the learning process. For a positive value of ρ, expected inflation converges asymptotically to the inflation target. Given this expectations mechanism, King derives the central bank's optimal disinflation policy.

It follows that case 3 is a mixture of cases 1 and 2. Expectations do not adjust immediately (they depend on actual inflation experience, and hence on the policy choices made during the transition), but are not completely exogenous either.

But case 3 is problematic, since the learning process implies that the learning parameter ρ does not depend on the monetary regime. Put differently, the updating mechanism does not reflect the actual speed of disinflation, and thus it is not clear whether the private sector expectations mechanism is rational. Our first suggestion is that, alternatively, case 3 could be modelled with the aid of Bayesian learning. Therefore, in the first part of the paper we modified the King model with Bayesian learning.

2.1 Bayesian learning

Consider again a two-period version of the model. The idea is to bring inflation down from its initial level, π_0 say, to a situation of price stability where inflation is zero. Thus, the central bank has to disinflate the economy by π_0 percentage points. We assume that this "inflation stabilisation plan" has full credibility and that the only uncertainty is about its *timing*. Thus, at the end of period 2 inflation has to be 0 under all scenarios[3], i.e.

$$\pi_2 = 0 \tag{2.8}$$

as is believed by the private sector. The question now is how should the disinflation be spread over time?

One strategy is a *cold turkey* approach. In this case, the central bank disinflates the economy in period 1 by π_0 percentage points and does nothing in period 2.

[3] This is also assumed in the King (1996) model.

The other strategy is a *gradualist* approach where the central bank inflates according to

$$\Delta\pi_{1,2} = \begin{cases} -q\pi_0 & in \quad period \quad 1 \\ & \qquad \text{where } 0 < q < 1 \\ -(1-q)\pi_0 & in \quad period \quad 2 \end{cases} \tag{2.9}$$

For instance, if q =0.5 and π_0 =10 in period 1, inflation falls from 10 to 5%, and in period 2 from 5 to zero[4].

At the start of period 1, under Bayesian learning wage setters assign a *prior* probability x_1 to the event that the central bank disinflates everything in one go ($x_1 = \Pr ob\langle \pi_1 = 0 \rangle$), i.e. it follows the cold turkey policy and $(1 - x_1)$ to the complementary event that the central bank follows a gradualist policy ($(1 - x_1) = \Pr ob\langle \pi_1 > 0 \rangle$).

By observing monetary policy in period 1, wage setters learn something about the true nature of the policy maker. Wage setters' beliefs, x_1 , are then revised according to Bayes' rule[5]. Period 2 nominal wages are then set on the basis of the *posterior* beliefs, x_2 .

If wage setters observe either a *positive* ($\pi_1 > 0$) or a *zero* inflation rate ($\pi_1 = 0$) in period 1, *Bayes' rule*[6] suggests how to rationally update these prior beliefs as

$$x_2 = \Pr ob\langle q = 1 | \pi_1 \rangle = \frac{\Pr ob\langle \pi_1 = 0 \rangle . \Pr ob\langle \pi_1 | q = 1 \rangle}{\Pr ob\langle \pi_1 \rangle} = \frac{x_1 \Pr ob\langle \pi_1 | q = 1 \rangle}{\Pr ob\langle \pi_1 \rangle} \tag{2.10}$$

Hence, the posterior probability that the central bank follows a cold turkey policy is given by the prior multiplied by the conditional probability of observing the policy π_1 given that the central bank follows a cold turkey policy, divided by the unconditional (*prior*) probability of observing the policy π_1 .

Clearly, if the gradualist policy is followed in period 1, Equation 2.10 gives $x_2 = 0$ [7], since a central bank that follows a cold turkey strategy would never have accommodated inflation expectations ($\Pr ob\langle \pi_1 > 0 | q = 1 \rangle$ =0).

[4] Of course, if $q \longrightarrow 1$ the above collapses to a cold turkey strategy.

[5] This is somewhat similar to the analysis by Huh et al. (2000) where agents update their prior assessment of the true inflation target in a (quasi) Bayesian way on the basis of the central bank's success or failure in reducing inflation over time.

[6] Bayes' rule is $\Pr ob\langle A | B \rangle = \dfrac{\Pr ob\langle A \rangle . \Pr ob\langle B | A \rangle}{\Pr ob\langle B \rangle}$.

[7] $x_2 = \dfrac{x_1 \Pr ob\langle \pi_1 > 0 | q = 1 \rangle}{\Pr ob\langle \pi_1 > 0 \rangle} = \dfrac{x_1 . 0}{(1 - x_1)} = 0$.

Similarly, if the cold turkey strategy is followed in period 1, Equation 2.10 gives $x_2 = 1$[8], since a central bank that follows a cold turkey strategy disinflates everything in period 1 with probability 1 ($\Pr ob\langle \pi_1 = 0 | q = 1 \rangle$ =1).

Thus we have the following rational private learning process:

$$x_2 = \begin{cases} 1 & if \quad \pi_1 = 0 \\ 0 & otherwise \end{cases} \tag{2.11}$$

Since only a gradualist central bank leaves any inflation in the economy, and if it does at the rate $\pi_0 - q\pi_0 = (1-q)\pi_0$, expected inflation at time 0 for period 1 is

$$E_0\pi_1 = (1-x_1)(1-q)\pi_0 \tag{2.12}$$

Note that this expression can be interpreted also in the context of King's case of "endogenous learning". Using $\tilde{\rho}$ as shorthand for $(1-x_1)(1-q)$, Equation 2.12 can be written as

$$E_0\pi_1 = \tilde{\rho}\pi_0 \text{ where } 0 < \tilde{\rho} < 1 \tag{2.12a}$$

Thus, here rational expectations can display some of the backward-looking characteristics of traditional adaptive expectations.

From Equation 2.12a it follows that – unlike in the King (1996) model – now the "speed of learning" $\tilde{\rho}$ does depend on the monetary regime[9]. For example, if the prior probability that the central bank follows a cold turkey policy increases, the private sector will attach less weight to the past inflation rate, as a basis for forecasting next year's inflation[10]. That is, $\frac{\partial x}{\partial \tilde{\rho}} < 0$. Of course, by observing the central bank's disinflation in period 1, in period 2 expectations completely converge to actual inflation.

The central bank's loss during this (partial) transition in period 1 is

$$\frac{1}{2}\left[a(\pi_1)^2 + (\pi_1 - E_0\pi_1)^2\right] \text{ where } 0 < a < \infty \tag{2.13}$$

Differentiating with respect to π_1 gives the optimal monetary policy as

$$\pi_1^* = \frac{1}{1+a}E_0\pi_1 \tag{2.14}$$

8 $x_2 = \frac{x_1 \Pr ob\langle \pi_1 = 0 | q = 1 \rangle}{\Pr ob\langle \pi_1 = 0 \rangle} = \frac{x_1 .1}{x_1} = 1$.

9 For an application of linear updating rules in an empirical model of the US economy, see Bomfim et al. (1997).

10 Thus, here the speed of learning is defined as the inverse of the weight attached to past inflation as a basis for forecasting future inflation.

Substituting from Equation 2.12 yields

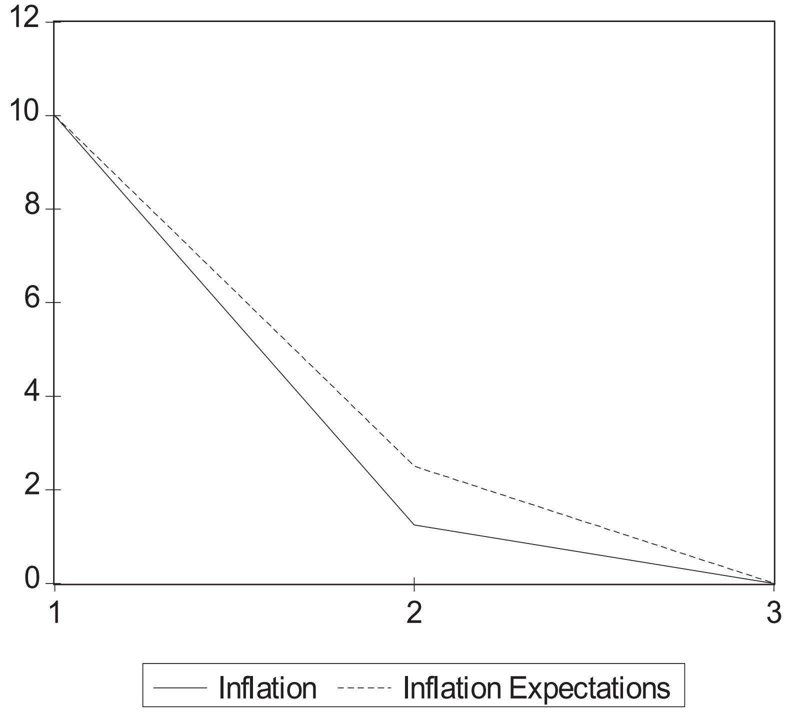
$$\pi_1^{*} = \frac{(1-x_1)}{1+a} * (1-q)\pi_0 \tag{2.15}$$

Equation 2.15 suggests several things. Assuming that the central bank *cares* about output, $0 < a < \infty$, the first is that the central bank will *only* follow a cold turkey strategy ($\pi_1^{*} = 0$) if the private sector is convinced it will ($x_1 = 1$). So, in this case *beliefs* (or rather, priors) are self-fulfilling.

Next, if the private sector thinks the central bank might be *gradualist* ($x_1 < 1$) the central bank will indeed be gradualist. However, it will be *less gradual* than the private sector expects.

Hence, here beliefs are *partly* self-fulfilling. The reason is that the central bank not only cares about output, and hence about appeasing labour market participants, but also about *inflation* itself. Of course, the more it cares about inflation (the bigger a) the greater the incentive to "speed up" the disinflation. For instance if q =0.5 and $x_1 = 0.5$, then the expected ("probability weighted") "degree of gradualism" $(1-x_1)(1-q)$ is 0.25. This means that the private sector rationally expects the inflation rate to fall in period 1 by 7.5 percentage points from 10 to 2.5% and to zero the year after. If the central bank only cared about output – that is, if a were equal to zero – then it would exactly accommodate the above expectations and follow the same timing. If it does care about inflation as well ($a > 0$), it will speed up things and disinflate *faster* than the private sector expects. For instance, if $a = 1$ it will disinflate by *8.75 percentage points* in year 1, and squeeze the last 1.25 % out of the economy in year two. This is illustrated in Graph 2.1.

Graph 2.1: Gradual disinflation



How “gradualist” the central bank will be exactly, can be seen from Equation 2.15a.

$$(1-q)^{*} = \frac{(1-x_1)(1-q)}{(1+a)} \tag{2.15a}$$

So the “ex post” degree of gradualism is always smaller than the “ex ante” degree. Of course the latter is the private sector's prior. Thus, the central bank will always disinflate *faster* than the private sector thinks. This is reminiscent of the King model where the optimal disinflation strategy is to partially accommodate inflation expectations (see Appendix A for details).

Finally, this paragraph suggests that the central bank’s optimal disinflation strategy is to be gradualist if first, it cares about output and second the prior that it *might* follow a gradualist strategy is non-zero ($1-x_1>0$). Moreover, what is also interesting about this set-up is that if central bank statements influence private sector priors, such central bank ”talk” is *not* cheap. This means that you cannot communicate an inflation stabilisation programme without at the same time being constrained by your words.

3 Bayesian learning with a non-linear Phillips curve

In the above paragraphs – given expectations – the output costs of disinflation are constant and given by the slope of the Phillips curve. In the present model this parameter is normalised at unity. It follows that a 1 percentage point of disinflation causes a 1 per cent fall in output. However, if we allow the output costs of disinflation to vary with the inflation rate, the central bank's incentives change substantially.

One way of extending the model with state-contingent output costs of disinflation is by means of a non-linear Phillips curve[11]. For example, following the recent paper by Schaling (2004) aggregate supply in period t, y_t might be specified as the following reduced-form *convex* supply function (or short-run asymmetric Phillips curve):

$$y = \frac{\pi_t - \pi^e}{1+\varphi(\pi_t - \pi^e)} \tag{3.1}$$

With this modification, the previous model can then be combined with a completely credible regime switch, exogenous learning and Bayesian learning as before. We follow the last option. Thus, the rest of the model is the same as before, except that Equation 2.2 has been modified to allow for nonlinearities[12].

This equation can be inverted as

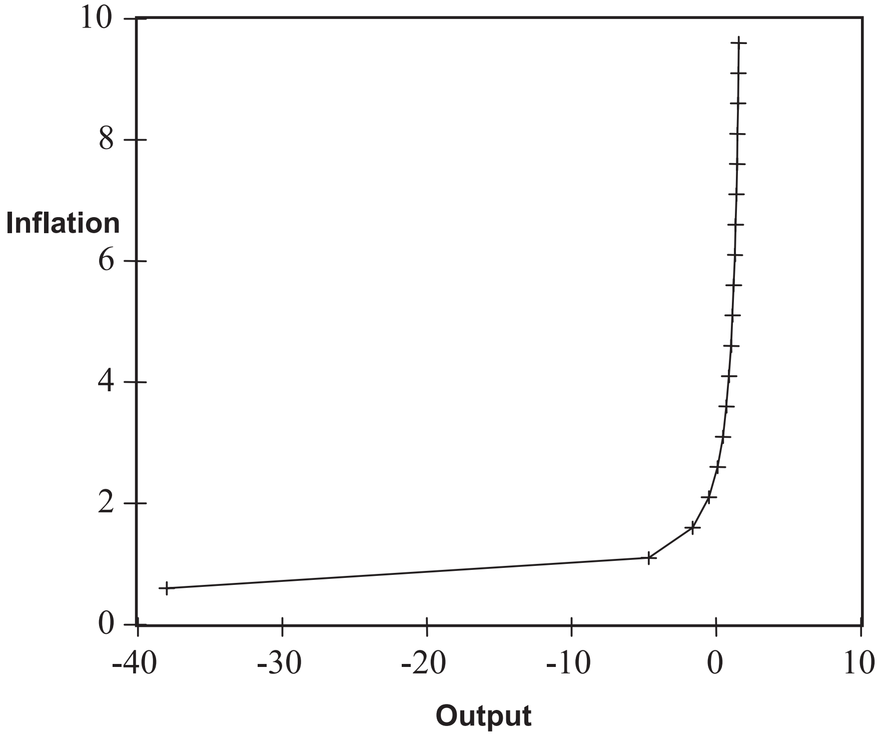
$$\pi_t - \pi_t^e = f(y_t) = \frac{y_t}{1-\varphi y_t} \tag{3.2}$$

[11] For an alternative approach where the welfare costs and benefits of disinflation vary with the inflation rate and the speed of disinflation, see Ireland (1997).

[12] Note, however, that the *qualitative* results of the paper do not depend on the particular specification of the curve. All that matters is that it is convex in output inflation space. For example, using the quadratic $\pi_t - \pi_t^e = f(y_t) = \alpha_1 y_t + \frac{\varphi}{2} y_t^2$ would give rise to similar qualitative results. For an alternative view where the curve is concave when output is below trend, and convex when output is above trend, see Filardo (1998).

We normalise the natural rate of output *in the absence of uncertainty* to zero[13]. This means that y is the (log) of output relative to potential, i.e. the output gap. Equation 3.2 is graphed in Graph 3.1.

Graph 3.1: A non-linear Phillips curve



Its relevant properties can be derived by looking at the first derivative of $f(y_t)$ – that is, the slope of the output inflation trade-off, written as

$$f^{'}(y_t) = \frac{1}{[1-\varphi y_t]^2} \tag{3.3}$$

Following Schaling (2004), it is useful to consider the limiting values of $f(y_t)$ and its derivative for some specific values of φ and y

$$\lim_{\varphi \to 0} f^{'}(y_t) = 1 \tag{3.3.a}$$

$$\lim_{y \to \frac{1}{\varphi}} f'(y_t) = \infty, f(y_t) = \infty \tag{3.3.b}$$

$$\lim_{y \to -\infty} f'(y_t) = 0, f(y_t) = \frac{-1}{\varphi} \tag{3.3.c}$$

$$f'(0) = 1, f(0) = 0 \tag{3.3.d}$$

[13] With uncertainty the natural rate of output in the non-linear model will always be *below* that of the linear model. See for instance Schaling (2004).

Equation 3.3.a shows that, as the parameter φ becomes very small, the Phillips curve approaches the linear relationship as expressed in Equation 2.2, hence the parameter φ indexes the curvature. Equation 3.3.b indicates that the effect on inflation rises without bound as output approaches $1/\varphi$. Hence, $1/\varphi$ represents an *upper bound* beyond which output cannot increase in the short run.

The central bank's first-order condition now becomes

$$\pi_1 = \frac{-\left(\pi_1 - E_0\pi_1\right)}{a\left(1+\varphi\left(\pi_t - E_0\pi_1\right)\right)^3} \tag{3.4}$$

Note that an important limiting case of Equation 3.4 is when φ becomes very small. In the latter case the Phillips curve approaches the standard linear functional form and the Freedom of Choice (FOC) collapses to Equation 2.14.

From Equation 3.4 it can be seen that *both* the left-hand side and the right-hand side of this expression depend on the period 1 inflation rate. Therefore, it is not straightforward to derive an *explicit* function that maps inflation expectations into the appropriate level of the period 1 inflation rate. Instead, one resorts to numerical methods to find the level of the inflation rate that is implicitly determined by Equation 3.4. To find the appropriate level one numerically computes the inflation rate that solves the first order condition. The results can be found in Table 3.1.

Table 3.1: Nonlinearity in the Phillips curve and optimal disinflation

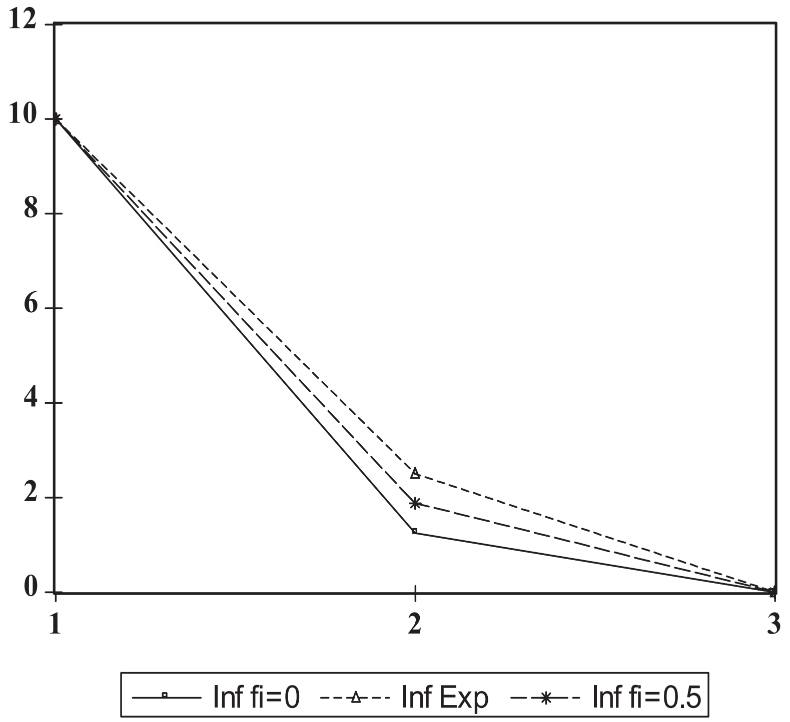
φ	π_1^*	φ	π_1^*
0	1.25	2	2.24
0.1	1.46	3	2.31
0.2	1.61	4	2.35
0.3	1.72	5	2.37
0.4	1.81	7	2.41
0.5	1.88	8	2.42
0.9	2.06	10	2.43
1.1	2.11	12	2.44

Table 3.1 shows the optimal disinflation path for the central bank with loss function (2.13) for different curvatures of the Phillips-curve (φ). We assume that the central bank cares as much about inflation as output (a = 1), as in Section 2 initial inflationary expectations for period 1 are 2.5% ($\pi_1^e = 2.5$) and the learning behaviour is adaptive ($\pi_2^e = \pi_1$).

From the above numerical optimisation exercise it is clear that the more *convex* the Phillips curve (in the output inflation space), the more gradualist the optimal disinflation strategy[14]. This is illustrated in Graph 3.2.

[14] For an analytical proof see Appendix C.

Graph 3.2: Disinflation with a non-linear Phillips curve



The reason is that the higher the output costs, the lower the inflation rate. Since the central bank takes this phenomenon into account in its optimal disinflation decision, it will "avoid" the part of the curve[15] where output costs are relatively high (i.e. when inflation is very low). The outcome is then a disinflation strategy that is more gradualist than in a standard linear-quadratic model.

Thus, it is optimal for the central bank to start off disinflating quickly and then to slacken off as inflation falls.

4 Disinflation with teaching by doing

In his discussion of *endogenous learning*, King (1996, p. 68) says that there are good reasons for the private sector to suppose that in trying to learn about the future inflation rate many of the relevant factors are exogenous to the path of inflation itself.

But a central bank may try to convince the private sector of its commitment to price stability by choosing to reduce its inflation target towards zero too quickly. King calls this "teaching by doing". Then the choice of a particular inflation rate influences the speed at which expectations adjust to price stability.

The problem, however, with the King (1996) model is that – as in the case of exogenous learning – the central bank decides on its optimal disinflation plan *given* those private sector expectations. Thus, although King calls this case "teaching by doing", a more accurate description would be "doing without teaching". The result – as in the case of *exogenous learning* – that price stability is reached gradually is therefore not surprising.

In this section "teaching by doing" is modelled differently. We allow the central bank's "doing" to affect private sector learning. Of course, if the central bank recognises its potential for active "teaching" its incentive structure changes. More specifically, it should realise that by disinflating faster, it can reduce the associated output costs by "teaching" the private sector that it means

[15] The part to "avoid" is the steep segment of the *concave* Phillips curve in inflation output space.

business. Thus, the dependence of private sector expectations on the actual inflation rate – Equation 2.7 in the King (1996) model – should be part of its optimisation problem. We model “teaching by doing” for both a linear and a convex Phillips curve.

4.1 Linear Phillips curve

Solving the model backwards for period $t = 2$ gives the optimal inflation rate in the second period:

$$\pi_2^* = \frac{\pi_2^e}{1+a} \tag{4.1}$$

Following Equation 2.7, now suppose that π_2^e depends on the realised rate of inflation in the first period. This is written as

$$\pi_2^e := \pi_2^e(\pi_1) \text{ with } \frac{\partial \pi_2^e}{\partial \pi_1} > 0 \tag{4.2}$$

Then, the central bank's problem in period $t = 1$ will be

$$\underset{\pi_1}{Min} \frac{1}{2}\left[a(\pi_1)^2 + y_1^2\right] + \frac{1}{2}\left[a(\pi_2(\pi_1))^2 + y_2^2\right] \tag{4.3}$$

The first order condition for minimising Equation 4.3 subject to Equation 2.2 and Equation 4.2 is

$$\pi_1^* = \frac{1}{1+a}\left[\pi_1^e - \frac{a}{1+a}\frac{\partial \pi_2^e}{\partial \pi_1}\pi_2^e\right] \tag{4.4}$$

If we contrast Equation 4.4 with Equation 2.6, it is clear that disinflation with "teaching by doing" is always faster than a "doing without teaching" policy. This can be seen from the presence of the term $-\frac{a}{1+a}\frac{\partial \pi_2^e}{\partial \pi_1}\pi_2^e$, which is unambiguously negative.

Proposition 1: In a model with a linear Phillips curve, disinflation with "teaching by doing" is always faster than a "doing without teaching" policy.

The reason for this proposition is that the choice of a particular inflation rate influences the speed at which expectations adjust to price stability. In fact, a quicker disinflation policy "pays for itself" by speeding up the adjustment of expectations. Of course, the central bank takes this fact into account when deciding on its disinflation programme. Another way to think about this is that central bank credibility – a crucial variable in defining the output loss of the disinflation – in this case is endogenous. In fact, the central bank's credibility – defined by Blinder (2000) as the extent to which words match deeds – can be increased by the central bank by starting off the disinflation process by putting its money where its mouth is.

In order to illustrate this graphically, consider the following algebraic preliminaries:

The central bank’s loss in the case of a linear Phillips curve equals

$$L = \tfrac{1}{2}\left(a\pi_1^2 + (\pi_1 - \pi_1^e)^2 + a\pi_2^2 + (\pi_2 - \pi_2^e)^2\right) \tag{4.3a}$$

Using the central bank's optimal strategy for period 2 as expressed in Equation 4.1 we get

$$L = \tfrac{1}{2}\left(a\pi_1^2 + (\pi_1 - \pi_1^e)^2 + a\left(\frac{1}{1+a}\pi_2^e\right)^2 + \left(\frac{-a}{1+a}\pi_2^e\right)^2 \right) \tag{4.3b}$$

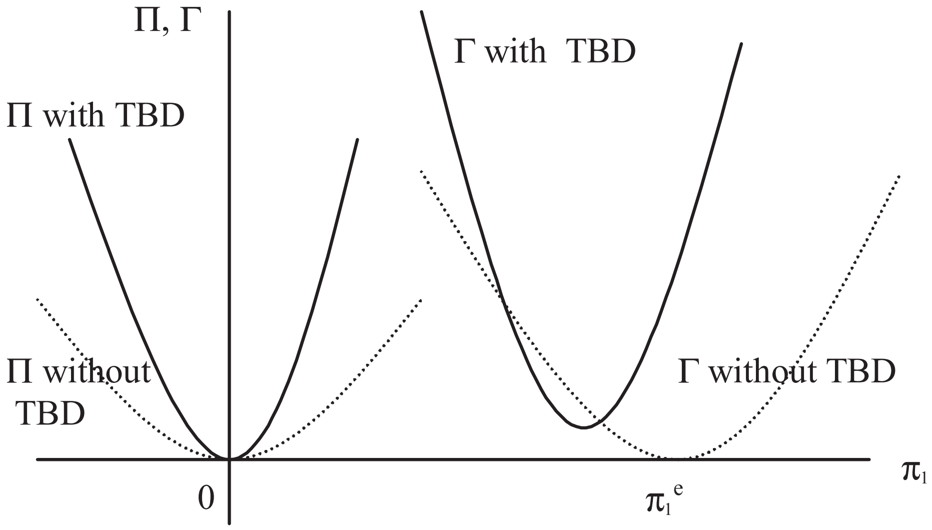
Assuming simple adaptive learning behaviour ($\pi_2^e = \pi_1$) we define

$$\Pi = a\left(1 + \frac{1}{(1+a)^2}\right)\pi_1^2$$
$$\Gamma = (\pi_1 - \pi_1^e)^2 + \left(\frac{a}{1+a}\right)^2 \pi_1^2 \tag{4.3c}$$

where Π represents the central bank's loss due to inflation and Γ represents the central bank's loss caused by a sub-optimal output level. Thus, if Π is positive the central bank is unhappy with the inflation rate, as it falls short of price stability. Similarly, if Γ is positive the central bank is unhappy because it suffers output losses – in other words, a recession – during the transition to price stability. Ideally, it would like both terms to be zero. However, this can only be achieved under a fully credible regime switch. Thus, in general, both terms are non-zero and the trick is to find the solution that steers the economy just between the Scylla of high inflation and the Charybdis of a recession.

These expressions allow one to analyse the effect of internalised learning graphically.

Graph 4.1: Decomposition of the central bank's loss with and without teaching by doing



The solid line indicates the situation under "teaching by doing", the dotted line without. In the case without "teaching by doing", Π is minimised where the rate of inflation equals zero. Obviously, this is also the case with "teaching by doing".

The curves are steeper under "teaching by doing" because the stakes are higher: Policy decisions in period 1 determine expectations in period 2. The difference in the Γ-curve is due to the fact that the central bank knows in period 1 that disinflation will take place (partly) in period 2 (this is implied by Equation 4.1). Reducing inflationary expectations in period 2 by setting a lower inflation rate in period 1 makes disinflation less painful in terms of output loss. This is anticipated in period 1 through having a preference for doing part of the disinflation in period 1 (see Appendix B for details).

4.2 Non-linear Phillips curve

The model is the same as above, except that we now modify Equations 2.2 to 3.1.

Again, solving the model backwards for period $t=2$ gives the optimal inflation rate in the second period. This is written as

$$\pi_2 = \frac{-(\pi_2 - E_1\pi_2)}{a(1+\varphi(\pi_2 - E_1\pi_2))^3} \tag{4.5}$$

The logic is the same as in Section 3. As is shown in Appendix C, the degree of accommodation $\frac{\partial \pi_2}{\partial \pi_2^e}$ depends positively on the convexity of the Phillips curve φ. This implies that a more convex Phillips curve will slow the disinflation process.

As before, suppose that π_2^e depends on the realised rate of inflation in the first period: $\pi_2^e := \pi_2^e(\pi_1)$.

Then, the central bank's problem in period $t=1$ will be:

$$\underset{\pi_1}{Min}\frac{1}{2}\left[a(\pi_1)^2 + y_1^2\right] + \frac{1}{2}\left[a(\pi_2(\pi_1))^2 + y_2^2\right] \tag{4.6}$$

subject to Equations 3.1, 4.2 and 4.5.

The first order condition for minimising Equation 4.6 is

$$a\pi_1 + y_1.\frac{\partial y_1}{\partial \pi_1} + \left[a\pi_2^*(\pi_1)\frac{\partial \pi_2^*(\pi_1)}{\partial \pi_1} + y_2\frac{\partial y_2}{\partial \pi_1}\left(1 - \frac{\partial \pi_2^*(\pi_1)}{\partial \pi_1}\right)\right] = 0 \tag{4.7}$$

which implicitly defines the optimal inflation rate in period 1. Here $\pi_2^*(\pi_1)$ is shorthand notation for Equation 4.5**.** That is, the dependence of the optimal inflation rate in period 2 on π_2^e – which in turn depends on period 1's inflation rate – is emphasised by writing $\pi_2^*(\pi_2^e) = \pi_2^*(\pi_1)$ (see Appendix B for details).

It is clear from Equation 4.7 that analytical solutions are not possible. Thus, numerically solving the first order condition is resorted to. Results can be found in Table 4.1.

Table 4.1: Non-linear Phillips curve and optimal disinflation under teaching by doing

φ	π_1	π_2	φ	π_1	π_2
0	1.00	0.50	2	2.21	1.96
0.1	1.23	0.67	3	2.29	2.11
0.2	1.42	0.84	4	2.33	2.18
0.3	1.56	0.99	5	2.36	2.24
0.4	1.67	1.13	7	2.40	2.31
0.5	1.76	1.25	8	2.41	2.33
0.9	1.98	1.58	10	2.43	2.36
1.1	2.04	1.68	12	2.44	2.38

Table 3 shows the optimal disinflation path for the central bank with loss function as expressed in Equation 4.6 for different curvatures of the Phillips curve (φ). We assume that the central bank cares as much about inflation as output (a =1) and as in Sections 2 and 3 initial inflationary expectations for period 1 are 2.5% ($\pi_1^e = 2.5$) and the anticipated learning behaviour is adaptive ($\pi_2^e = \pi_1$).

Table 3 confirms our finding in Appendix C that a convex Phillips curve reduces the speed of the disinflationary process. Therefore, it counteracts the effect of internalised learning, which makes the disinflation process faster.

Proposition 2: A convex Phillips curve reduces the speed of the disinflation process. It counteracts the effect of "teaching by doing" which makes the disinflation process faster.

5 Summary and concluding remarks

This paper extends the King (1996) disinflation model in several directions.

First, we have modified what King calls "endogenous learning". This refers to a situation where inflation expectations do not adjust immediately (as would be the case under a completely credible regime switch), but are not completely exogenous either. In fact, under endogenous learning, expectations are a weighted average of previous inflation expectations and the previous inflation rate.

The problem with this case is that in the King model the speed of learning does not depend on the monetary regime, and thus it is not clear whether the private sector expectations mechanism is rational. For example, whether the central bank is expected to disinflate fast or slow does not affect the learning parameter.

We address this issue by alternatively modelling the expectations process using Bayesian learning. In this case the expected disinflation profile *does* affect the learning parameter. For example, if the private sector attaches a higher prior probability that the central bank might follow a cold turkey policy, the speed of learning increases.

Furthermore, we find that the central bank's optimal disinflation strategy is to be gradualist if it cares about output and thus follow a gradualist strategy that is non-zero. However, the central bank will always disinflate *faster* than the private sector thinks.

In the second part of the paper we address the issue of "teaching by doing". This refers to the idea that a central bank may try to convince the private sector of its commitment to price stability by choosing to reduce its inflation target towards zero quickly. In this case, the choice of a particular inflation rate influences the speed at which expectations adjust to price stability.

However, the problem with the King (1996) model is that – as in the case of exogenous learning – the central bank decides on its optimal disinflation plan *given* those private sector expectations.

In the second part of the paper "teaching by doing" is modelled differently. We allow the central bank's "doing" to affect private sector learning. Of course, if the central bank recognises its potential for active "teaching" its incentive structure changes. More specifically, it should realise that by disinflating faster, it can reduce the associated output costs by "teaching" the private sector that it means business. This is, in fact, what is found in the paper: Allowing for “teaching by doing” effects always speeds up the disinflation *vis-à-vis* the case where this effect is absent. Thus, it is clarified why "speed" in the disinflation process does not necessarily "kill" in the sense of creating large output losses.

However, there is one caveat. In the above analysis – given expectations – the output costs of disinflation are constant and given by the slope of the Phillips curve. In the present model this parameter is normalised at unity. It follows that a 1 percentage point of disinflation causes a 1% fall in output. However, if we allow the output costs of disinflation to vary with the inflation rate, the central bank's incentives change substantially. We analyse state-contingent output costs of disinflation by means of allowing for a non-linear Phillips curve.

With this modification, the previous model can then be combined with Bayesian learning and teaching by doing as before. We find that all case convexities in the Phillips curve slow down the central bank's optimal disinflation. The reason is the higher the output costs, the lower the inflation rate. Since the central bank takes this phenomenon into account in its optimal disinflation decision, it will “avoid” the part of the curve where output costs are relatively high (i.e. when inflation is very low). The outcome is then a disinflation strategy that is more gradualist than in a standard linear-quadratic model (with or without teaching by doing effects). Table 5.1 summarises the discussion above.

Table 5.1: Main results

Learning	Phillips curve	
	Linear	Convex
Bayesian	**I**: Benchmark case	**II**: $\pi_1(II) > \pi_1(I)$ Slowest disinflation
Teaching by doing	**III**: $\pi_1(III) < \pi_1(I)$ $\therefore \pi_1(III) < \pi_1(I) < \pi_1(II)$ Fastest disinflation	**IV**: $\pi_1(IV) \le \pi_1(II)$ $\pi_1(IV, \varphi < 0.2) < \pi_1(I)$ $\pi_1(IV, \varphi \ge 0.2) > \pi_1(I)$ $\therefore \pi_1(I) < \pi_1(IV, \varphi \ge 0.2) \le \pi_1(I$ Intermediate case: Slower than benchmark, faster than case **II**.

Of course, the model can be made more realistic. Possible extensions are to include more periods, uncertainty about the end point – that is about the total *amount* of disinflation π_0[16]– and the inclusion of stochastic productivity shocks. In the latter case our conjecture is that this will shift the balance in favour of gradualism, the reason being that a cold turkey strategy does not allow for a flexible response to supply shocks. Hence, with such shocks hitting the economy, in general being an “inflation nutter” will become even more unattractive (because of the implied output volatility) than in the present set-up.

[16] For an analysis along these lines see Huh et al. (2000).

Appendix A: The optimal degree of gradualism

$$\pi_1 = \pi_0 + \Delta\pi_1 \tag{A.1}$$

Substituting for $\Delta\pi_1$ from Equation 2.9 is expressed as

$$\pi_1 = (1-q)\pi_0 \tag{A.3}$$

Substituting for π_1^* from Equation 2.15 yields

$$\frac{1}{1+a}E_0\pi_1 = (1-q)^*\pi_0 \tag{A.4}$$

which can be rearranged as

$$(1-q)^* = \frac{E_0\pi_1}{(1+a)\pi_0} \tag{A.5}$$

Finally, substituting for $E_0\pi_1$ from Equation 2.12 yields Equation 2.15a in the main text.

Appendix B: Optimal disinflation under teaching by doing

Linear Phillips curve

The central bank's first order condition in period 1 is

$$a\pi_1 + y_1.\frac{\partial y_1}{\partial \pi_1} + a\pi_2.\frac{\partial \pi_2}{\partial \pi_1} + y_2.\frac{\partial y_2}{\partial \pi_1} = 0 \tag{B.1}$$

Expanding $\frac{\partial \pi_2}{\partial \pi_1}$ as $\frac{\partial \pi_2}{\partial \pi_2^e}.\frac{\partial \pi_2^e}{\partial \pi_1}$ and using Equation 2.2 we get

$$a\pi_1 + \left(\pi_1 - \pi_1^e\right) + a\pi_2.\frac{\partial \pi_2}{\partial \pi_2^e}.\frac{\partial \pi_2^e}{\partial \pi_1} + \left(\pi_2 - \pi_2^e\right).\frac{\partial y_2}{\partial \pi_1} = 0 \tag{B.2}$$

In turn, expanding $\frac{\partial y_2}{\partial \pi_1}$ as $\frac{\partial y_2}{\partial \pi_2^e}.\frac{\partial \pi_2^e}{\partial \pi_1} + \frac{\partial y_2}{\delta \pi_2}\frac{\partial \pi_2}{\partial \pi_2^e}\frac{\partial \pi_2^e}{\partial \pi_1} = -\frac{\partial \pi_2^e}{\partial \pi_1}\left[1 - \frac{\partial \pi_2}{\partial \pi_2^e}\right]$ and using Equation 4.1 we can derive

$$(1+a)\pi_1 - \pi_1^e + \left[\frac{a}{(1+a)^2}\pi_2^e + \frac{a^2}{(1+a)^2}\pi_2^e\right].\frac{\partial \pi_2^e}{\partial \pi_1} = 0 \tag{B.3}$$

where we have used that

$$\frac{\partial y_2}{\partial \pi_2^e} = -\frac{a}{1+a}$$

Rearranging gives Equation 4.4 in the main text.

Non-linear Phillips curve

To find the central bank's disinflation plan for $t = 1,2$, we need to solve the following set of simultaneous non-linear equations:

$$\pi_2 = \frac{-(\pi_2 - E_1\pi_2)}{a(1+\varphi(\pi_2 - E_1\pi_2))^3} \tag{4.5}$$

which is the FOC for the optimal rate of inflation in period 2 and

$$a\pi_1 + y_1.\frac{\partial y_1}{\partial \pi_1} + \left[a\pi_2\frac{\partial \pi_2}{\partial \pi_2^e} + y_2\frac{\partial y_2}{\partial \pi_2^e}\left(1-\frac{\partial \pi_2}{\partial \pi_2^e}\right)\right].\frac{\partial \pi_2^e}{\partial \pi_1} = 0 \tag{B.4}$$

which is the FOC for the optimal rate of inflation in period 1, where (as before) we assume that $E_1(\pi_2) = \pi_1$, so that $\frac{\partial \pi_2^e}{\partial \pi_1} = 1$. Using this in Equations 4.5 and B.4 we get:

$$\pi_2 = \frac{-(\pi_2 - \pi_1)}{a(1+\varphi(\pi_2 - \pi_1))^3} \tag{B.5}$$

$$a\pi_1 + y_1.\frac{\partial y_1}{\partial \pi_1} + \left[a\pi_2\frac{\partial \pi_2}{\partial \pi_1} + y_2\frac{\partial y_2}{\partial \pi_1}\left(1-\frac{\partial \pi_2}{\partial \pi_1}\right)\right] = 0 \tag{B.6}$$

Equation B.5 can be rewritten as $\pi_2^* = \pi_2^*(\pi_1)$. Substituting this expression into Equation B.6 we get the following expression for π_1:

$$a\pi_1 + y_1.\frac{\partial y_1}{\partial \pi_1} + \left[a\pi_2^*(\pi_1)\frac{\partial \pi_2^*(\pi_1)}{\partial \pi_1} + y_2\frac{\partial y_2}{\partial \pi_1}\left(1-\frac{\partial \pi_2^*(\pi_1)}{\partial \pi_1}\right)\right] = 0 \tag{B.7}$$

We know that $y_2 = \frac{\pi_2^*(\pi_1) - \pi_1}{1+\varphi(\pi_2^*(\pi_1) - \pi_1)}$, so that $\frac{\partial y_2}{\partial \pi_1} = \frac{-1}{[1+\varphi(\pi_2^*(\pi_1) - \pi_1)]^2}$.

Using this, Equation B.7 can be written as:

$$a\pi_1 + y_1.\frac{\partial y_1}{\partial \pi_1} + \left[\begin{array}{l} a\pi_2^*(\pi_1)\dfrac{\partial \pi_2^*(\pi_1)}{\partial \pi_1} + \\ \left(\dfrac{\pi_2^*(\pi_1) - \pi_1}{1+\varphi(\pi_2^*(\pi_1) - \pi_1)}\right)\left(\dfrac{-1}{[1+\varphi(\pi_2^*(\pi_1) - \pi_1)]^2}\right)\left(1-\dfrac{\partial \pi_2^*(\pi_1)}{\partial \pi_1}\right)\end{array}\right] = 0 \tag{B.8}$$

Finally, using $y_1 = \frac{\pi_1 - E_0\pi_1}{1+\varphi(\pi_1 - E_0\pi_1)}$, and $\frac{\partial y_1}{\partial \pi_1} = \frac{1}{[1+\varphi(\pi_1 - E_0\pi_1)]^2}$ we find that

$$a\pi_1 + \left(\frac{\pi_1 - E_0\pi_1}{1+\varphi(\pi_1 - E_0\pi_1)}\right)\cdot\left(\frac{1}{[1+\varphi(\pi_1 - E_0\pi_1)]^2}\right)$$
$$+\left[\begin{array}{l} a\pi_2^*(\pi_1)\dfrac{\partial\pi_2^*(\pi_1)}{\partial\pi_1} + \\ \left(\dfrac{\pi_2^*(\pi_1)-\pi_1}{1+\varphi(\pi_2^*(\pi_1)-\pi_1)}\right)\left(\dfrac{-1}{[1+\varphi(\pi_2^*(\pi_1)-\pi_1)]^2}\right)\left(1-\dfrac{\partial\pi_2^*(\pi_1)}{\partial\pi_1}\right)\end{array}\right] = 0 \tag{B.9}$$

Appendix C: Monetary accommodation under teaching by doing

This Appendix investigates the dependence of the degree of monetary accommodation – that is the extent to which the central bank allows inflation *expectations* to show up in the *actual* inflation rate – on the convexity of the Phillips curve. In general we find that more convexity leads to more accommodation. Results in this Appendix apply to both Sections 3 and 4 of the paper.

With respect to the term $\dfrac{\partial\pi_2}{\partial\pi_2^e}$ in period's 2 FOC Equation 4.5, using implicit differentiation we can derive that

$$\frac{\partial\pi_2}{\partial\pi_2^e} = -\frac{\left(\dfrac{\partial FOC}{\partial\pi_2^e}\right)}{\left(\dfrac{\partial FOC}{\partial\pi_2}\right)} = \frac{2\phi(\pi_2 - \pi_2^e) - 1}{2\phi(\pi_2 - \pi_2^e) - 1 - a(1+\phi(\pi_2 - \pi_2^e))^4} \tag{C.1}$$

Note that an important limiting case of Equation C.1 is when φ becomes very small. In the latter case we have $\lim\limits_{\varphi\to 0}\dfrac{\partial\pi_2}{\partial\pi_2^e} = \dfrac{1}{1+a}$, which is in line with Equation 4.1.

Rearranging terms yields the following:

$$\frac{\partial\pi_2}{\partial\pi_2^e} = \frac{1-2\phi(\pi_2 - \pi_2^e)}{1-2\phi(\pi_2 - \pi_2^e) + a(1+\phi(\pi_2 - \pi_2^e))^4} < 1 \text{ if } 1-2\phi(\pi_2 - \pi_2^e) > 0 \tag{C.2}$$

Note that the latter condition is always satisfied if $\pi_2 - \pi_2^e < 0$. In turn, this is the likely outcome under any (optimal) disinflation policy where expectations are only partially accommodated. Whether or not this condition is in fact (ex post) satisfied is checked below.

Equations C.1 or, equivalently, C.2 can be seen as a measure of the degree of monetary *accommodation* of inflationary expectations. We can now investigate the sensitivity of this degree of accommodation with respect to the convexity of the Phillips curve, measured by the parameter ϕ.

First, it is useful to define a variable *B* as follows:

$$B(\phi) := 1 - 2\phi(\pi_2 - \pi_2^e) \tag{C.3}$$

In times of disinflation (where $\pi_2 < \pi_2^e$), we know that:

$$B(\phi) > 0 \text{ and } \frac{\partial B(\phi)}{\partial \phi} > 0 \tag{C.4}$$

Using Equations C.2 and C.3, the degree of accommodation can be written as:

$$ACC := \frac{\partial \pi_2}{\partial \pi_2^e} = \frac{B(\phi)}{B(\phi) + a(1+\phi(\pi_2 - \pi_2^e))^4} \tag{C.5}$$

Therefore:

$$\frac{\partial ACC}{\partial \phi} = \frac{B'(\phi)a(1+\phi(\pi_2 - \pi_2^e))^4 - B(\phi)4a(1+\phi(\pi_2 - \pi_2^e))^3(\pi_2 - \pi_2^e)}{\left(B(\phi) + a(1+\phi(\pi_2 - \pi_2^e))^4\right)^2} > 0 \tag{C.6}$$

Using the properties of the function *B*, the condition that $1+\phi(\pi_2 - \pi_2^e) > 0$ and the fact that we have disinflation, it is easy to prove inequality in Equation C.6. Therefore, more convexity leads to more accommodation.

Finally, note that the above results also apply to Section 3. That is, along similar lines we find that $\frac{\partial \pi_1}{\partial \pi_1^e}$ increases in φ.

References

Ball, L. 1997. Disinflation and the NAIRU In Romer, Christina D. and David H. Romer (eds.), Reducing inflation: Motivation and strategy, *NBER Studies in Business Cycles*, Vol 30. Chicago and London: University of Chicago Press, pp. 167-192.

Blinder, A S. 2000. Central Bank Credibility: Why do We Care? How Do We Build it? *American Economic Review* 90(5): 1421-31. December.

Bomfim, A, Robert, T, von zur Muehlen, P and Williams, J. 1997. Expectations, Learning and the Costs of Disinflation: Experiments Using the FRB/US Model. Board of Governors of the Federal Reserve System, *Finance and Economics Discussion Series*, No 42, August.

Filardo, A J. 1998. New Evidence on the Output Cost of Fighting Inflation *Economic Review, Federal Reserve Bank of Kansas City*. Third Quarter, pp 31-61.

Ireland, P N. 1997. Stopping Inflations, Big and Small. *Journal of Money, Credit, and Banking,Vol* 29(4), Part 2, pp 759-75. November.

Huh, C G. and Lansing, K J. 2000. "Expectations, Credibility, and Disinflation in a Small Macroeconomic Model", *Journal of Economics and Business*, Vol. 51, pp 51-86.

Kahn, G A. 1996. Achieving Price Stability: A Summary of the Bank's 1996 Symposium. *Economic Review_Federal Reserve Bank of Kansas City*, Vol 81, No 4. Fourth Quarter, pp 53-59. Available [online] www.kc.frb.org.

King, M. 1996. How Should Central Banks Reduce Inflation? – Conceptual Issues. *Economic Review Federal Reserve Bank of Kansas City,* Vol. 81, No 4. Fourth Quarter, pp 25-52. Available [online] www.kc.frb.org.

Schaling, E. 2004. The Non-linear Phillips Curve and Inflation Forecast Targeting: Symmetric vs. Asymmetric Policy Rules. *Journal of Money, Credit, and Banking*. Part 1, Vol. 36, No 3: pp 361-86. June.

Sargent, T. 1986. *Rational Expectations and Inflation* New York: Harper and Row.