July 2001

Non-standard central bank loss functions, skewed risks, and the certainty equivalence principle

By Ali al-Nowaihi and Livio Stracca*†**

Abstract

This paper sets out to investigate the role of additive uncertainty under plausible *non-standard* central bank loss functions over future inflation. Building on a substantial body of evidence in the economic psychology literature, this paper postulates (i) period-by-period loss functions that are non-convex, i.e. displaying diminishing or non-increasing sensitivity to losses, and (ii) non-linear weighing of probabilities, hence departing from the expected utility paradigm. In addition, a simple and plausible form of non-time separability of the central bank's intertemporal loss function is also considered in the analysis. The main conclusion of the study is that if the additive uncertainty is caused by a non-Normal distributed additive shock, for instance if the probability distribution of the shock is skewed, then with these departures from the quadratic function *the principle of certainty equivalence does not hold anymore*. Thus, it appears that with additive uncertainty of the non-Normal type the assumption of a quadratic loss function for the central banker may not be as innocuous as it is commonly regarded. Conversely, non-time separability of the central bank intertemporal loss function as studied in this paper does not determine *per se* any departure from certainty equivalence. Moreover, no evidence is found of an optimal policy gradualism as a response to increased additive uncertainty even under the non-standard loss functions considered in this paper.

Keywords: Monetary policy, non-quadratic loss functions, economic psychology, certainty equivalence

JEL codes: E52, E58

---

*University of Leicester, Department of Economics.

†University of Leicester, Department of Economics and European Central Bank (ECB). Corresponding author; e-mail livio.stracca@ecb.int. We thank participants in a seminar at the ECB, especially Frank Smets, an anonymous referee and Charles Goodhart for helpful comments. The opinions expressed in the paper are only those of the authors and are not necessarily shared by the ECB.

1 Introduction

The role of uncertainty in monetary policy-making has attracted considerable interest in the literature in recent years (see, e.g., the review by Batini, Martin and Salmon, 1999). It has long been known that *multiplicative* uncertainty (namely, over the true "value" of the interest rate elasticity of output and inflation) is not neutral for the policy-maker and generally leads to caution and policy gradualism (Brainard, 1967).

There is, however, much less consensus on the role of *additive* uncertainty, which denotes the uncertainty over the true "state of the economy", despite the obvious importance of this matter for policy-makers, who are confronted with this type of uncertainty practically every day.[1] On the one hand, the view is prevailing in the academia that additive uncertainty should not matter, a principle known as "certainty equivalence" (Theil, 1958).[2] Recently, Svensson and Woodford (2000) have provided a general proof that with a quadratic objective function for the central banker, the optimal policy is unaffected by uncertainty about the state of the economy (Svensson and Woodford characterise this situation as the orthogonality of estimation and policy). Moreover, the principle of certainty equivalence does not seem to be a mere artefact of the use of a quadratic loss function. Chadha and Schellekens (1999) – henceforth CS – have shown that the certainty equivalence principle holds for a general class of convex loss functions, and the coefficients of the optimal policy rule are not affected

[1] Data measurement problems, uncertainty over the economy's natural rate of employment (or the natural rate of interest) at any point in time, shocks to the inflation rate and/or to the output gap which occur *after* a certain monetary policy decision but *before* its impulse has fully worked through the economy, are all prominent examples of additive uncertainty.

[2] Formally, let x be a control variable and y a state variable, with $y = f(x) + e$, e being a zero mean additive disturbance. If certainty equivalence holds, the optimal value of x (for instance, the value which minimizes y) is independent of the probability distribution of e.

by additive uncertainty even if the preferences of the central banker are asymmetric.[3] On the other hand, policy-makers generally do not seem to think that additive uncertainty is irrelevant (see for example Blinder, 1998). A casual look at central banks' external communication tends to lend support to this assessment. For instance, in the Sveriges Riiksbank's Inflation Report (quoted in Blix and Sellin, 2000), it was reported that:

> "The element of *uncertainty in the inflation assessment* can accordingly influence monetary policy's construction. A high degree of uncertainty can be a reason for giving policy a more cautious turn" [emphasis ours]

and in the Bank of England's Inflation Report (again quoted from Blix and Sellin):

> "in the light of the central projection and *the risks surrounding it*, the Bank continues to see the need for a moderate tightening of policy". [emphasis ours]

Finally, this statement can be retrieved from the European Central Bank's website:

> The European Central Bank (ECB) confirms its position of 'wait and see' with regard to its monetary policy stance. In an environment of *increased uncertainty over the global economy and its impact on the euro area*, the Governing Council is carefully assessing whether and to what extent upward risks to price stability will continue to decline." [emphasis ours]

In all cases, central banks seem to refer to additive uncertainty (i.e., uncertainty over the state of the economy, not over the effect of the monetary policy levers) as an important

[3] Orphanides and Wieland (2000) also analysed the properties of inflation zone targeting using a nonlinear central bank loss function. Orphanides and Wieland, however, did not deal with the issue of additive uncertainty, at least not directly.

element in the determination of policy. Thus, there seems to be a discrepancy of views between policy-makers and the academia over this key aspect of monetary policy-making. This divergence, in turn, should lead one to wonder whether simple and plausible monetary policy models alternative to those traditionally used in the academic literature on optimal monetary policy can be worked out to give account of the seemingly important role of additive uncertainty in actual policy-making.

Against this background, this paper sets out to analyze the interaction, if any, between non-standard and yet analytically tractable and behaviorally plausible central bank loss functions and *additive* uncertainty modelled as a non-Normal distributed additive shock to the inflation process. Throughout the paper, the assumption will be maintained that the policy-maker has no multiplicative uncertainty, namely a perfect knowledge of the effect of the monetary policy instrument on the target variable(s). This paper relaxes three commonly maintained assumptions on the central bank loss function. First, a curvature different from, and more general than, the quadratic is considered. Second, the effect of a non-linear weighing of probabilities by the central bank is analyzed. Finally, the common assumption of time separability of the intertemporal loss function is relaxed. Ultimately, the objective of the analysis is to establish whether the principle of certainty equivalence carries through to the non-standard loss functions examined here, and hence whether the assumption of a time separable quadratic loss function evaluated according to the expected utility criterion – which permeates the bulk of the literature on optimal monetary policy – is indeed innocuous or not. That no assumption on the probability distribution of the additive shock is maintained is worth stressing, as it distinguishes the analysis in this paper from that in CS[4] . Indeed, the type of uncertainty central bankers are routinely confronted with is certainly not always Normal distributed. Goodhart (2000), for instance, stressed that skewed risks represent a challenge for monetary policy-makers and that the profession should come up with an analytical framework to study how such skewed risks may be meaningfully incorporated in the policy assessment. Overall, this paper seems to be a first step in that direction.

In devising plausible and tractable non-quadratic central bank loss functions, this paper builds on the substantial body of evidence made available by the literature on economic psychology under uncertainty, especially by the strand linked to the names of Daniel Kahneman and Amos Tversky (see Kahneman and Tversky, 2000, for a review and assessment of this literature). With no evidence available thus far on the "typical central banker"'s psychology, the working assumption of this paper is that the patterns and tendencies that the Kahneman-Tversky literature has identified in a number of experimental studies are also valid for those agents in charge of monetary policy.[5] Among the key elements identified in this literature, diminishing sensitivity to losses (i.e., non-convex loss functions) and non-linear weighing of probabilities (i.e., departures from the expected utility paradigm) seem to be plausible and interesting also as a characterization of central bank preferences. Moreover, the paper also investigates a further departure from the "standard" setting, by assuming that the central

[4] CS maintain the assumption of a Normal distributed additive shock to inflation throughout their paper.

[5] It should be stressed that throughout the paper the emphasis will always be on the central bank's *positive* (i.e., descriptive) preferences. The analysis abstains from the determination of the *normative* preferences, for example those that would maximise society's welfare. Of course, if it were possible for society to write down explicitly the central banker's loss function, this would lead us a long way towards the identification of the "true" central bank preferences. However, even in this (rather unrealistic) case the "ex post" preferences will not correspond to the "ex ante" preferences (those which matter for monetary policy-making), for the latter also involve the central banker's attitude towards risk.

bank intertemporal loss function may be non-time separable in its argument (inflation). This appears to be justified by the observation – which only stems from basic psychological insight – that a central banker is likely to be more worried by repeated misses in the same direction, rather than by random misses in different directions, for in the former case his credibility might be harmed, or at least the public may derive an impression of incompetence, which the central banker should want to avoid at all costs.

In sum, the paper finds that the interaction of a non-Normal additive disturbance to inflation and the set of non-quadratic preferences postulated here leads to situations where the principle of certainty equivalence does *not* hold. Thus, additive uncertainty seems to matter. Instead, if the disturbance is Normal distributed (as assumed in CS), the usual result of certainty equivalence continues to be valid. If one observes that non-Normal distributed shocks are an essential element which monetary policy-makers have to deal with, the overall policy message of this paper is that additive uncertainty matters, and that the assumption of a quadratic loss function may not be as innocuous as it is often considered to be. Conversely, introducing non-time separability – at least of the type considered in this paper – does not determine *per se* any departure from certainty equivalence.

The paper is organized as follows. In Section 2 the usual setting is outlined of a monetary policy-maker aiming at minimizing a loss function defined in terms of the inflation rate, with the structure of the economy acting as a constraint on behaviour. The effect of considering non-quadratic central bank loss functions on the role of additive uncertainty is analyzed in Section 3. Finally, Section 4 concludes.

2 A simple optimal control model for discretionary monetary policy

This section lays down a standard optimal control problem for discretionary monetary policy.[6]

The structure of the economy includes an IS curve, whereby the monetary authority can influence the output gap by steering the nominal interest rate, and a backward-looking Phillips curve, linking current inflation to past inflation and to past output gap (see, e.g., Clarida, Gali and Gertler, 1999, and Mankiw, 2001).[7]

The IS curve is specified as follows:

$$x_t = a_1 x_{t-1} - a_2(i_t - \pi_t) + u_t, \tag{1}$$

where the output gap x is affected by the real interest rate $(i - \pi)$, and u is a disturbance term, unknown to the central bank at time t, with $E_t u_t = 0$. Parameters in this equation are $0 < a_1 < 1$ and $a_2 > 0$, known to the central bank.

The Phillips curve is:

$$\pi_{t+1} = b_1 \pi_t + b_2 x_t + v_{t+1}, \tag{2}$$

where v is an additive disturbance, for instance capturing a cost-push shock unknown to the policy-maker at time t, with $E_t v_{t+1} = 0$, and $0 < b_1 < 1$ and $b_2 > 0$ are parameters,

[6] Throughout the paper, the assumption is always maintained that the central bank cannot credibly commit to follow a policy rule; monetary policy is thus carried out in a discretionary manner.

[7] The choice of a backward-looking specification of the Phillips curve is motivated by the fact that it squares better with the available empirical evidence (see in particular Mankiw, 2001). The line of argumentation in the paper, however, would not be substantially changed with a forward-looking Phillips curve, as long as inflationary expectations at time t are exogenous for the central bank (to avoid the simultaneity problems dealt with by Svensson and Woodford, 2000).

of which the central bank has again full knowledge.[8] [9] No further assumption on the probability distribution of v is added at this stage. For simplicity, a zero drift is assumed; this, together with the assumption that $b_1 < 1$, implies that the steady state level of inflation is zero.

For notational simplicity, it is convenient to consolidate the IS and the Phillips curves to obtain a reduced form for inflation as follows:

$$\pi_{t+1} = c_1\pi_t + c_2 x_{t-1} - c_3 i_t + \varepsilon_{t+1}, \tag{3}$$

where $c_1 = b_1 + b_2 a_2$, $c_2 = b_2 a_1$, $c_3 = b_2 a_2$, and $\varepsilon_{t+1} = b_2 u_t + v_{t+1}$ is a zero mean (but possibly non-Normal and non-symmetrically distributed) random disturbance (comprising output gap and cost-push shocks).

Simplifying it further:

$$\pi_{t+1} = z_t - c_3 i_t + \varepsilon_{t+1}, \tag{4}$$

with $z_t = c_1\pi_t - c_2 x_{t-1}$, known to the central bank at time t and exogenous to i_t, representing the "state of the economy" or equivalently the "inflationary pressures" at time t. In general, policy rules are specified as feedback rules $i_t = f(z_t)$, whereby the central bank sets its monetary policy instrument in reaction to changes in the state of the economy.

For simplicity of notation, let us consider the variable $\widetilde{i_t}$ defined as follows:

$$\widetilde{i_t} = c_3 i_t - z_t \tag{5}$$

[8] That the monetary policy instrument i affects inflation with a longer lag than it affects output is consistent with the bulk of the available empirical evidence.

[9] More complex Phillips curve equations may be conceived (see, e.g, the non-linear specification in Clark et al, 2001), but such complications should *prima facie* be immaterial for the purpose of the present analysis.

If $\widetilde{i_t} = 0$, $i_t = \frac{z_t}{c_3}$ and $E_t\pi_{t+1} = 0$. Therefore, $\widetilde{i_t}$ is the deviation of the monetary policy instrument from the value which offsets at time $t+1$ the expected impact of the inflationary pressures observed at time t (i.e., z_t). There follows that the inflation process may also be expressed as:

$$\pi_{t+1} = \varepsilon_{t+1} - \widetilde{i_t}, \tag{6}$$

with ε_{t+1} independent of $\widetilde{i_t}$. Intuitively, the inflation process is equal to the difference between the realization of the additive shock at time $t+1$ and the deviation of the monetary policy instrument from the value for which $E_t\pi_{t+1} = 0$. Thus, the value $-\widetilde{i_t}$ may also be interpreted as the inflation target of the central bank at time t. In the continuation of this paper, we will often refer to $\widetilde{i_t}$ when speaking about monetary policy, without any loss of generality as the monetary policy instrument may be derived straightforwardly as $i_t = \frac{\widetilde{i_t}+z_t}{c_3}$.

The task of monetary policy is to select a value of $\widetilde{i_t}$ which minimizes an intertemporal loss function defined in terms of inflation levels (assuming for simplicity – and without loss of generality – that the inflation objective of the central bank is zero).[10] The principle of certainty equivalence – the main focus of the present analysis – stipulates that the probability distribution of ε should *not* matter in the determination of $\widetilde{i_t}$ (or, i_t).

The standard approach to deal with monetary policy problems is to specify the objective function of the central banker, with the structure of the economy that acts as a constraint on behaviour. Given a period-by-period loss function $L(\pi)$, the central bank is normally

[10] As in CS, we do not consider the possibility that the monetary authority may also care about the output gap. While this is certainly a restrictive and rather unrealistic assumption, it greatly simplifies the algebra. Moreover, the inclusion of the output gap may under certain conditions determine a *departure* from certainty equivalence (see in particular Smets, 1998). Thus, given that the focus of the present analysis is to ascertain whether additive uncertainty *on inflation* matters *per se* under non-quadratic central bank loss functions, we leave this complication aside in order to balance the odds in favour of certainty equivalence as much as possible.

assumed to minimize a time separable intertemporal loss function expressed as the expected value of a discounted sum L (where $0 < \gamma < 1$ is the discount factor):

$$E_t L = E_t \sum_{j=1}^{\infty} \gamma^{j-1} L(\pi_{t+j}), \tag{7}$$

under the constraint $\pi_{t+1} = \varepsilon_{t+1} - \widetilde{i}_t$. It is clear from the way the problem is formulated that the central bank's task may be reduced straightforwardly to the minimization of $E_t L(\pi_{t+1})$. In fact, the problem is specified recursively and the monetary policy instrument at time t only affects inflation one period ahead.

Under quadratic preferences, we have:

$$E_t L(\pi_{t+1}) = E_t \pi_{t+1}^2 = E_t (\varepsilon_{t+1} - \widetilde{i}_t)^2, \tag{8}$$

and thus, after solving the first order condition:

$$\widetilde{i}_t = 0, \tag{9}$$

given that $E_t \varepsilon_{t+1} = 0$. Of course, this result may be derived also by force of the statistical law that the arithmetic mean (zero, in this case) is the measure of central tendency minimizing the average of the squared deviations from it. Hence, $i_t = \frac{z_t}{c_3} = \frac{c_1 \pi_t - c_2 x_{t-1}}{c_3}$, *irrespective* of the probability distribution of ε; this is the standard result of certainty equivalence. In general, a value of $\widetilde{i}_t$ different from zero depending on (some moments of) the probability distribution of ε_{t+1} would signal a departure from certainty equivalence.

Having laid down the framework and the notation for the central bank's optimal control problem, in the ensuing section we move to analyze the effect of imposing non-quadratic loss functions onto the optimal choice of the monetary policy instrument $\widetilde{i}_t$.

3 Non-standard loss functions

In this section three assumptions, underlying the standard central bank intertemporal loss function normally considered in the literature, are relaxed with a view to studying the effect of these departures on the role of additive uncertainty. First, we consider the possibility that the *period-by-period* loss function may have a non-quadratic functional form, possibly being non-convex over its dominion. Second, we analyze the effect of a non-linear weighing of probabilities by the central banker, thus departing from the expected utility paradigm. Third, we relax the assumption that the central bank *intertemporal* loss function is time separable, by introducing a simple and relatively plausible form of non-time separability, whereby the central banker cares about missing the target repeatedly in the same direction comparatively more than missing the target in random directions. For each departure from the standard setting, some behavioral justification will be provided.

As the first two departures are closely linked to, and firmly grounded in, a branch of the economic psychology literature, we consider them together in Section 3.1. A non-time separable intertemporal loss function will be then examined separately in Section 3.2.

3.1 Non-convex central bank preferences with non-linear weighing of probabilities

In order to devise *reasonable* alternatives to the quadratic as a central bank loss function, it should be first pondered about the behavioral assumptions underlying the quadratic function. A key element of the quadratic function is that "large" shocks are penalized proportionally more heavily than "small" shocks.[11] Thus, it appears *prima facie* useful to analyze func-

[11] A standard argument in favour of the quadratic loss function is that for "small" shocks it approximates relatively well any differentiable loss function as a second order Taylor expansion. However, in our analysis

tional forms that do not penalize large shocks as much as the quadratic function does (or, equivalently, do not attribute so little weight to small shocks). For example, Goodhart (2000) criticizes the assumption that large shocks should be penalized proportionally more than small shocks and reports that "I could never see why a 2% deviation from desired outcome was 4x as bad as a 1% deviation, rather than just twice as bad". Moreover, Rabin (2000) and Rabin and Thaler (2001) have shown (albeit in a context unrelated to monetary policy) that with a convex (e.g., quadratic) loss function and even a small aversion to small shocks, the aversion to moderate or large shocks may easily reach astronomical (behaviourally absurd) levels. Indeed, an important strand of the economic psychology literature, mainly popularized by Daniel Kahneman and Amos Tversky, building on basic psychological intuition and on a substantial body of experimental evidence, points to the fact that agents generally show a tendency to a mildly *decreasing* (instead of increasing as postulated in the quadratic function) sensitivity to shocks (computed against a "reference point") as the size of the shock rises (see Kahneman and Tversky, 2000, Rabin, 1998, and Thaler, 2000, for a review of this literature). If the results of this literature are to be taken seriously, this would suggest loss functions with a *curvature* different from the quadratic, and in particular functions that are *non-convex*.[12]

Another key element of this literature is the observed widespread tendency for agents to weigh probabilities in a non-linear manner, hence departing from the expected utility paradigm (see, e.g., Starmer, 2000). A non-linear weighing of probabilies by central banks is very realistic. Cecchetti (2000), for instance, reports that "[...] we would expect policy-

we assume that "large" shocks do exist, albeit maybe with low probability.

[12] The Kahneman-Tversky literature favours loss functions that have a "kink" on the reference value, and a diffferent curvature above and below this value.

makers to take action when the mean and variance of forecast distributions are likely to stay the same, while the probability of some extreme bad event increases. [...] even if the variance is unchanged, an increase in the possibility of a severe economic downturn is likely to prompt action." While this is somewhat at odds with Goodhart (2000), who claims that "[...] main characteristic of risks which policy should *not* try to pre-empt is that they are low-probability events with a high pay-off", a non-linear weighing of probabilities is suggested in both cases. Non-linear weighing of probabilities – such as a disproportionate weight attached to changes in probabilities around zero, a phenomenon known as the "certainty effect" – is sometimes associated with anticipatory feelings and anxiety (see, for instance, Caplin and Leahy, 2001). While central bankers may be cooler than normal human beings, it is a fair assumption that at times they can certainly become anxious and mis-calculate probabilities according to their emotional state.

In sum, central bankers are, above all, humans; the possibility should be at least considered that they display the same attitudes towards risk that have been documented for other types of agents in the economic psychology literature.[13] Hence the main focus of this section of the paper is evaluating the effect of incorporating these attitudes (decreasing sensitivity and non-linear weighing of probabilities) in the central bank's loss function on the optimal setting of the monetary policy instrument and, in particular, on whether certainty equivalence continues to hold.

Turning first to the issue of the curvature of the period-by-period loss function, a relatively general and simple functional form to be considered in this analysis can be the following (see

[13] Another key trait identified by the Kahneman-Tversky literature is the asymmetric treatment of gains and losses. This feature, however, should not matter for our central banker, who can only lose (and not gain) from a shock to inflation.

Kahneman and Tversky, 1992):

$$L(\pi_{t+1}) = |\pi_{t+1}|^{\beta} \tag{10}$$

If $\beta = 2$, this is the standard quadratic loss function. In the Kahneman-Tversky literature (see in particular Kahneman and Tversky, 1992) $0 < \beta < 1$ (although β is normally estimated to be very close to one[14]); thus, the loss function is not convex and has a "kink" at zero (clearly the "reference value" in our setting). A loss function specified as in (10) is able to encompass a large number of functional forms and behavioral assumptions. The parameter β, in particular, drives the curvature of the loss function. Values $\beta < 1$ identify non-convex loss functions, namely functions where the value of the loss grows *less than proportionally* with the size of the shock (as in the function postulated in Kahneman and Tversky, 1992).

According to Kahneman and Tversky, diminishing sensitivity is a general trait of human behaviour. In the case of monetary policy-making, some special circumstances have to be taken into consideration. If our central banker behaves in a Kahneman-Tversky manner (i.e., $\beta < 1$), he will be more upset by a 1% *marginal* change in inflation from 2% to 3% than by a change from, say, 15% to 16% (a quadratic central banker will be more upset by the latter). However, the presence of thresholds and "target zones" for inflation (see Orphanides and Wieland, 2000), as it is common in some way or another in many countries, is likely to impart severe non-linearities to the central bank loss function, because overcoming the threshold(s) may be particularly costly. Thus, a central banker may be *more* concerned by a move of inflation from 2% to 3% than from 1% to 2% if the target zone has an upper bound at, say, 2.5%. Nevertheless, *within* and *outside* the target zone diminishing sensitivity should

[14] For example, Kahneman and Tversky (1992) provided the estimate $\beta = 0.88$.

prevail.

In any case, it is important to stress that the analysis in this paper is not limited to the case $0 < \beta < 1$, and a full range of possibilies is considered. At one extreme, $\beta = 0$ indicates that the central bank cares about *any* shock, independent of the size of the shock. Alternatively, $\beta > 1$ indicates that losses rise *more than proportionally* with the size of the shock, i.e. the loss function is convex (and it is also differentiable in the whole domain, having no "kink" at zero). As $\beta \to \infty$, the central bank becomes much more concerned with large shocks; in the limit, it only cares about the largest possible shock. This type of central banker can be labelled as "minimax".

In this setting, the central bank's problem may be expressed as the minimization of:

$$E_t L(\pi_{t+1}) = E_t \left| \varepsilon_{t+1} - \widetilde{i_t} \right|^{\beta} \tag{11}$$

It is immediate to realize that, in general, the measure of central tendency $\widetilde{i_t}$ that minimizes expression (11) will *differ* from the expected value of ε_{t+1} (i.e., zero), as it is the case under quadratic preferences (for it is well known in probability theory that $\widetilde{i_t} = E_t\varepsilon_{t+1} = 0$ minimizes (11) only if $\beta = 2$).

Let us consider, for instance, the case $\beta = 1$.[15] The central bank's loss function collapses to the absolute value of the deviations of inflation from its target (zero):

$$E_t L(\pi_{t+1}) = E_t \left| \varepsilon_{t+1} - \widetilde{i_t} \right| \tag{12}$$

Under this specification, the optimal value for $\widetilde{i_t}$ is given by $\widetilde{i_t} = M_t\varepsilon_{t+1}$, where M represents

[15] Goodhart (2000) describing his own experience at the Bank of England's MPC reports that "I believe that I could, more or less, interpret my loss function when I was at the MPC (*symmetrically linear* in the deviation from target at the six to eight quarter horizon)" [emphasis ours]. This would imply that for Goodhart $\beta = 1$.

the *median* value of the probability distribution of ε. Hence:

$$i_t = \frac{z_t}{c_3} + \frac{1}{c_3} M_t \varepsilon_{t+1} \tag{13}$$

In general, $M_t \varepsilon_{t+1}$ will be different from zero, and the optimal policy will deviate from that identified under the quadratic loss function ($i_t = \frac{z_t}{c_3}$, or $\widetilde{i_t} = 0$). In particular, if the probability distribution of ε_{t+1} is positively skewed, the median will be negative (i.e., smaller than the mean) and the optimal policy will imply a *smaller* value of i_t than it would have been the case under the quadratic loss function (the opposite, i.e. a *higher* value of i_t, will hold true if the probability distribution of ε_{t+1} is negatively skewed). Only if the probability distribution of ε_{t+1} is symmetric is the standard result recovered ($i_t = \frac{z_t}{c_3}$). More generally, the value $\widetilde{i_t}$ – interpreted as a measure of central tendency minimizing the loss function in (11) for an arbitrary value of β – will certainly depend on the probability distribution of ε_{t+1}, i.e. *the principle of certainty equivalence will not hold (unless $\beta = 2$)*.[16] If, however, ε_{t+1} is Normally distributed, then all measures of central tendency collapse to the mean (at least for any finite β), thus to zero. In this case, the principle of certainty equivalence holds and nothing is lost by using the quadratic function.[17]

A second extension which appears to be both interesting and plausible (as discussed above) is a non-linear weighing of probabilities. Now, instead of minimizing $E_t L(\pi_{t+1})$, the central bank will aim at minimizing $E_t^\delta L(\pi_{t+1})$, with E_t^δ defined as follows:

$$E_t^\delta L(\pi_{t+1}) = \int L(\pi_{t+1}) \delta(P(\pi_{t+1})) d\pi_{t+1}, \tag{14}$$

[16] If $\beta = 0$, the interest rate is set to be equal to the *mode* of ε_{t+1}. The opposite case is $\beta \to \infty$ (i.e., only the largest shock matters), which leads to $i_t = \varepsilon^*_{t+1}$, where ε^*_{t+1} is the value of ε for which $|\varepsilon_{t+1}|$ is largest. Clearly, this corresponds to a "robust control" or "minimax" solution to optimal monetary policy.

[17] This appears to reconcile our results with those of CS, who assumed a Normal distributed additive shock to inflation at time $t + 1$.

where P is either the cumulative probability distribution of π_{t+1} or the probability density, and δ is a weighing function which satisfies $\int \delta(P(\pi_{t+1}))d\pi_{t+1} = 1$, $\delta(P(\pi_{t+1})) \geq 0$. It should be noted that $\delta(P)$ is a function of P and not of π_{t+1}, with the consequence that if $P(\pi_{t+1}) = P(-\pi_{t+1})$, then $\delta(P(\pi_{t+1})) = \delta(P(-\pi_{t+1}))$. In plain words, the δ transformation preserves symmetry.[18]

If the δ function is the identity function and P is the probability density, the standard expected utility formulation is recovered. Otherwise, the δ function may be any non-linear transformation of either the cumulative probability distribution or of the probability density. It is important to stress that, in either case, $\delta(P)$ can be itself interpreted as a probability distribution, and the expression in (14) can be thought of as the mathematical expectation of $L(\pi_{t+1})$ computed according to the *transformed* probability law $\delta(P)$. In the Kahneman-Tversky literature, P is normally the cumulative probability distribution (this property is often referred to as "rank dependence") and the function $\delta(\cdot)$ is normally found to give more weight to "large" probabilities and less weight to "small" probabilities compared to the linear case (see Kahmeman and Tversky, 2000). For instance, in Kahneman and Tversky (1992) the following function is postulated:

$$\delta(P) = \frac{P^{\omega}}{[P^{\omega} + (1-P)^{\omega}]^{\frac{1}{\omega}}}, \tag{15}$$

$\omega > 0$, which encompasses the linear weighing of expected utility models as a special case when $\omega = 1$.[19] If $0 < \omega < 1$, this weighing function is first concave and then convex, crossing the linear weighing in a point which is also determined by the value of ω. This function is

[18] This property is often referred to as "reflection".

[19] In Kahneman and Tversky (1992) the parameter ω is estimated to be close to .6 for gains and to .7 for losses.

only an example, as many other weighing functions have been proposed in the literature.[20]

In order to deal with one complication at a time and to isolate the effect of a non-linear weighing of probabilities *in itself*, we assume that the period loss function is quadratic, i.e. $L(\pi_{t+1}) = \pi_{t+1}^2$. Our central banker is thus called to minimize:

$$E_t^\delta \pi_{t+1}^2 = E_t^\delta (\varepsilon_{t+1} - \widetilde{i}_t)^2, \tag{16}$$

which developing the quadratic term in parentheses (and dropping the strictly positive scalar c_3^2) leads to:

$$E_t^\delta \pi_{t+1}^2 = E_t^\delta (\varepsilon_{t+1}^2 - 2\widetilde{i}_t \varepsilon_{t+1} + i_t^2) \tag{17}$$

The first order condition is thus:

$$\frac{\partial E_t^\delta \pi_{t+1}^2}{\partial \widetilde{i}_t} = 2\widetilde{i}_t - 2E_t^\delta \varepsilon_{t+1} = 0, \tag{18}$$

hence:

$$\widetilde{i}_t = E_t^\delta \varepsilon_{t+1} \tag{19}$$

Recalling that $i_t = \frac{\widetilde{i}_t + z_t}{c_3}$, it follows:

$$i_t = \frac{z_t}{c_3} + \frac{1}{c_3} E_t^\delta \varepsilon_{t+1}, \tag{20}$$

which corresponds to the canonical solution $i_t = \frac{z_t}{c_3}$ only if $E_t^\delta \varepsilon_{t+1} = 0$. If the probability distribution of ε_{t+1} is *symmetric*, then $E_t^\delta \varepsilon_{t+1} = E_t \varepsilon_{t+1} = 0$, due to the property of symmetry preservation of the $\delta(P)$ function, and the usual solution $i_t = \frac{z_t}{c_3}$ (thereby the irrelevance of additive uncertainty) is recovered.[21] In general, however, the non-linear weighing will

[20] Another popular weighing function is the one proposed by Prelec (1998), $\delta(P) = \exp[-(-\ln P)]^\alpha$, $0 < \alpha < 1$.

[21] Of course, a Normal distribution is symmetric and certainty equivalence, as in CS, is also recovered.

not be neutral, i.e. $E_t^{\delta}\varepsilon_{t+1} \neq E_t\varepsilon_{t+1} = 0$, hence the probability distribution of ε_{t+1} will not be irrelevant, in other words, the principle of certainty equivalence will *not* hold. In particular, the interplay between a non-linear weighing of probabilities and a skewed probability distribution of ε_{t+1} determines a departure from the certainty equivalence principle.

As a simple numerical example to illustrate the kind of situation that the analysis in this section refers to, consider a central banker who observes $z_t = 2$, $c_3 = \frac{1}{2}$ (expressed in percentage points). Assume further that the probability distribution of ε_{t+1} is the following:

ε_{t+1}	Prob
-2	.4
-1	.3
1	.1
5	.2

Of course, the assumption is satisfied that $E_t\varepsilon_{t+1} = 0$. Whilst this probability distribution is purely hypothetical, it is nevertheless representative at least of the *type* of uncertainty that central banks often have to deal with in practice. This probability distribution features high-probability and small-size downside risks (-2 and -1, respectively with probabilities .4 and .3), a low-probability, small-size upside risk (+1, with probability .1) and a low-probability, large-size upside risk (+5, with probability .2). For instance, the event "+5" might be associated with "extreme" circumstances such as, say, the collapse of an exchange rate peg. If our central bank has a quadratic loss function and weighs probabilities linearly, it will pick $\widetilde{i_t} = E_t\varepsilon_{t+1} = 0\%$ (and $i_t = \frac{z_t + c_3\widetilde{i_t}}{c_3} = 4\%$). However, central bankers with a different β will pick other values of i_t. For instance, Goodhart (2000) – for who $\beta = 1$ – will select

$i_t = \frac{z_t}{c_3} + \frac{1}{c_3} M_t \varepsilon_{t+1} = 4\% - 2\% = 2\%$ (the median value of ε_{t+1} is, in fact, -1). A "minimax" type of central banker ($\beta \to \infty$) will select $i_t = \frac{z_t}{c_3} + \frac{1}{c_3} \varepsilon^*_{t+1}$ (where ε^*_{t+1} is the value of ε for which $|\varepsilon_{t+1}|$ is largest); hence $i_t = 4\% + 10\% = 14\%$. Finally, a central banker for who $\beta = 0$ will pick $i_t = \frac{z_t}{c_3} + \frac{1}{c_3} \text{mod}(\varepsilon_{t+1})$ – where $\text{mod}(\varepsilon_{t+1})$ is the mode of ε_{t+1}, in this case -2 – and then $i_t = 4\% - 4\% = 0\%$. To sum up:

Central banker type	**Optimal policy**
Quadratic ($\beta = 2$)	$i_t = 4\%$
Linear ($\beta = 1$)	$i_t = 2\%$
Minimax ($\beta \to \infty$)	$i_t = 14\%$
$\beta \to 0$	$i_t = 0\%$

A similar reasoning (and a similar discrepancy of policy outcomes) would follow by imposing non-linear methods of weighing probabilities. Clearly, the large variation in the policy outcomes is the result of the particular set up of the example, but it is certainly not far-fetched – as already argued above – to think of a real-world probability distribution of future inflation resembling, at least *qualitatively*, to our hypothetical distribution.

In synthesis, in this section it is found that (i) a non-quadratic curvature of the loss function can determine a departure from the principle of certainty equivalence unless the probability distribution of ε_{t+1} is *Normal*, and (ii) a non-linear weighing of probabilities (alone) will bring about the same result if the probability distribution of ε_{t+1} is not *symmetric*. Therefore, additive uncertainty matters. It should be added that none of these departures from the principle of certainty equivalence suggests the optimality of policy gradualism: the optimal response to z_t (the "state of the economy" at time t) is independent of the probability distribution of ε_{t+1}, for $\frac{\partial i_t}{\partial z_t} = \frac{1}{c_3}$ in all the considered cases.

3.2 A non-time separable central bank loss function

The standard assumption in the literature is that the central bank intertemporal loss function is time separable, i.e. inflation and in some models also the output gap at time t matter on their own and not depending on their past (or expected future) realizations. However, simply out of basic psychological insight, it may be argued that time separability is an overly restrictive assumption and that the central banker is likely to care about *repeated* target misses *in the same direction* comparatively more than misses in different (e.g., random) directions. For example, making repeated mistakes always in the same direction (e.g., always on the upside) may negatively affect the central bank's credibility and thereby fuel inflationary expectations or give an impression of incompetence to the public, which the central bank arguably wants to eschew.[22] Against this background, this section sets out to study whether under a simple and plausible form of non-time separability of the central banker's loss function the principle of certainty equivalence is maintained or not.

The specific form of time non-separability which we introduce in this section can be labelled as "additive". The period-by-period loss function of the central bank is as follows:

$$L(\pi_t, \pi_{t+1}) = \pi_{t+1}^2 + \theta(\pi_{t+1} + \eta\pi_t)^2, \tag{21}$$

where $\theta \geq 0$ and η are real scalars. Besides caring about future inflation, the central banker also cares about the cumulated (or average) inflation rates in periods t and $t+1$

[22] This reasoning might be formally underpinned, for instance, in a model where the private sector learns (possibly in a Bayesian manner) about the central bank's inflation objective, which may be evolving over time (see, e.g., Bomfin and Rudebusch, 2000). Developing this model in full, however, goes beyond the scope of the present analysis.

(where the weight attributed to current inflation is determined by the parameter η). For simplicity, and to single out the effect of non-time separability *per se*, the functional form of the loss function is now quadratic again and a linear weighing of probabilities is assumed. If $\theta = 0$, the traditional time separable specification is recovered (”baseline” case). In general, the parameter θ might depend on the central bank's credibility (for instance, central banks with low credibility – and willing to improve it through reputation – may have a higher θ in their loss function, on the basis of the consideration that repeated misses in a certain direction would more likely be interpreted by the public as a lack of commitment to the achievement of the target). The parameter η specifies the exact form of the non-separability. If $\eta > 0$, the central bank wants to avoid missing the target repeatedly in the same direction, following the arguments presented above. Otherwise, a $\eta < 0$ indicates a willingness to err repeatedly on the same side, although this latter tendency seems to be rather difficult to rationalize.

In this setting, the central bank wishes to minimize the expectation of the quantity L:

$$E_t L = E_t L(\pi_t, \pi_{t+1}) + \gamma E_t L(\pi_{t+1}, \pi_{t+2}), \tag{22}$$

which may be rewritten as:

$$L = E_t[\pi_{t+1}^2 + \theta(\pi_{t+1} + \eta\pi_t)^2] + \gamma E_t[\pi_{t+2}^2 + \theta(\pi_{t+2} + \eta\pi_{t+1})^2] \tag{23}$$

Solving the first order condition around the steady state (see the Annex for details) it is found:

$$\widetilde{i}_t = \frac{\theta\eta\pi_t}{1 + \theta + \gamma\theta\eta^2 + \gamma\theta\eta}, \tag{24}$$

or:

$$i_t = \frac{z_t}{c_3} + \frac{\theta\eta\pi_t}{c_3(1+\theta+\gamma\theta\eta^2+\gamma\theta\eta)} \tag{25}$$

If $\theta = 0$ or $\eta = 0$ – i.e., the loss function is characterized by full time separability – the standard result $i_t = \frac{z_t}{c_3}$ is recovered.[23] As it is evident from equation (24), the deviation of $\widetilde{i_t}$ from zero depends on the current realization of inflation (π_t), and the sign of this deviation hinges on the sign of the parameter η. Given that a positive sign of η seems warranted on behavioral grounds, in general $\widetilde{i_t}$ will tend to co-move with π_t, implying a more aggressive response of monetary policy to current inflation than in the baseline case. Indeed, the optimal setting of monetary policy identified by (25) can be interpreted as a Taylor-type feedback rule where the central bank reacts to inflation and the output gap according to:

$$i_t = \alpha_1 x_{t-1} + [\alpha_2 + \alpha_3]\pi_t, \tag{26}$$

with $\alpha_1 = -\frac{c_2}{c_3}$, $\alpha_2 = \frac{c_1}{c_3}$, and $\alpha_3 = \frac{\theta\eta}{c_3(1+\theta+\gamma\theta\eta^2+\gamma\theta\eta)}$, the latter parameter reflecting the non-time separability of the loss function. If $\theta\eta > 0$, this implies that $\alpha_3 > 0$, $\alpha_2 + \alpha_3 > \alpha_2$, thus a more aggressive policy reaction than in the baseline case.

Coming back to the our central focus on the role of additive uncertainty, although the optimal monetary policy rule under non-time separability differs from the baseline case, the principle of certainty equivalence nevertheless holds exactly as under a time separable loss function. In fact, the optimal value of i_t in (25) does not depend in any way on the probability distribution of the additive shock. Thus, it seems that time-non separability

[23] As pointed out by a referee, the loss function in (21) may be simplified to $L(\pi_t, \pi_{t+1}) = (\pi_{t+1} + \eta\pi_t)^2$, whereby $\widetilde{i}_t = \frac{\eta}{1+\gamma\eta(\eta+1)}$. The slightly more complex specification in (21) might be justified, however, in order to distinguish the inter-temporal concern of the central banker *per se* (captured by the parameter θ) from the precise form of the non-time separability, captured by the parameter η.

– at least of the additive type introduced in this paper – does not explain the apparently important role of additive uncertainty in monetary policy-making. Overall, the evidence in this section provides an extension (to the non-time separable specification) of the general result by Svensson and Woodford (2000) that certainty equivalence always holds under a quadratic loss function, although this conclusion is clearly only valid for the specification of the non-time separability proposed in this paper.

4 Conclusions

This paper has investigated the robustness of the principle of certainty equivalence, i.e. the irrelevance of the uncertainty over additive shocks, to simple and behaviorally plausible departures of the central bank's loss function from the standard time separable expected quadratic loss. The analysis essentially follows up a paper by Chadha and Schellekens (1999). Compared to that study, this paper relaxes a key assumption, namely that the additive shock to inflation is Normal distributed. This seems to be quite an important innovation because non-Normal probability distributions of future inflation (e.g., skewed risks as stressed by Goodhart, 2000) are arguably the "daily bread" of central bankers.

The analysis in this paper has considered two classes of departures from the expected time separable quadratic loss, which is a main analytical tool of the literature on optimal monetary policy. *First*, building on a substantial body of literature on economic psychology mainly linked to the names of Daniel Kahneman and Amos Tversky, this paper has considered the possibility that the central bank loss function displays diminishing sensitivity to losses (namely, be non-convex over its argument) and that the central bank weighs probabilities in a non-linear manner (thus departing from the expected utility paradigm). This analysis appears to be interesting and to make sense, because these tendencies are firmly grounded in the economic psychology literature and have been tested and confirmed in a large number of experimental studies (Kahneman and Tversky, 2000). Therefore, it appears *prima facie* reasonable to assume that central bankers – who are humans after all – may also display such attitudes. From this analysis it is found that this first form of non-quadratic loss functions generally brings about a departure from the principle of certainty equivalence, unless the probability distribution of the additive shocks to inflation is *Normal* (as far as the property of diminishing sensitivity is concerned, and in general if the curvature of the function is not of the quadratic type) or *symmetric* (as far as a non-linear weighing of probabilities is concerned). Overall, thus, the assumption of a quadratic loss function does not seem to be innocuous when considering the effect of non-Normal distributed additive shocks to inflation onto the determination of the optimal policy.

The second departure from the standard setting analyzed in this paper is the possibility that the central bank loss function is non-time separable, and in particular it is argued that repeated errors in the same direction may matter (cost) more than errors in different directions, for instance due to the credibility loss that the central bank may incur ("additive" non-time separability). It is found that under this non-time separable specification of central bank preferences the principle of certainty equivalence continues to hold, although the optimal setting of the monetary policy instrument deviates from the standard solution, and in particular implies a more aggressive response to current inflation than under time separable preferences.

In sum, this paper has shown that with additive uncertainty of a non-Normal type, the assumption of quadratic loss functions may not be completely innocuous. Thus, this paper tends to limit the generality of the conclusions of Chadha and Schellekens, who favored the idea that the assumption of a quadratic loss function is not so restrictive as far as the role of additive uncertainty is concerned.[24] This result should caution against an a-critical reliance on simple policy rules, such as the Taylor rule, in which certainty equivalence holds and the uncertainty on the state of the economy does not play any role in the optimal setting of monetary policy instruments.

In addition, this paper provides an analytical framework to study the role of the higher moments of (non-Normal) probability distribution of the additive shock to inflation, for example its skewness, for optimal policy-making. This seems to be a very interesting line of research because policy-makers are continuously confronted with this type of uncertainty.[25] At the same time, this paper does not shed any light on explaining the alleged positive correlation between policy gradualism and uncertainty over the state of the economy in actual policy-making. It seems that imposing some form of multiplicative uncertainty is necessary to explain central banks' tendency to act gradually and to react cautiously to new information.

[24] The difference in results between this paper and CS is entirely explained by the fact that here the assumption of Normality of the probability distribution of the additive shock to inflation is relaxed.

[25] Goodhart (2000) again reports that "unlike uncertainty and variance, skew and risk mapped directly into the interest rate decision". We could not agree more with this statement.

References

Batini, N., Martin, B. and C. Salmon (1999): "Monetary policy and uncertainty", Bank of England Quarterly Bulletin, pp. 183-189.

Blinder, A. (1997): "What central bankers can learn from academia – and vice versa", Journal of Economic Perspectives, 11, 2, pp. 3-19.

Blinder, A. (1998): *Central Banking in Theory and Practice*, The Lionel Robbins Lectures.

Blix, M. and P. Sellin (1998): "Uncertainty bands for inflation forecasts", Sveriges Riksbank Working Paper n. 65.

Blix, M. and P. Sellin (2000): "A Bivariate Distribution for Inflation and Output Forecasts", Sveriges Riksbank Working Paper n. 102.

Bomfin, A. N. and G. D. Rudebusch (2000): "Opportunistic and Deliberate Disinflation under Imperfect Credibility", Journal of Money, Credit and Banking, 32, 4, pp. 707-721.

Brainard, W. (1967): "Uncertainty and the Effectiveness of Policy", American Economic Review, 57, pp. 411-425.

Caplin, A. and J. Leahy (2001): "Psychological Expected Utility Theory and Anticipatory Feelings", Quarterly Journal of Economics, February, pp. 55-79.

Cecchetti, S. (2000): "Making Monetary Policy: Objectives and Rules", Oxford Review of Economic Policy, 16, 4, pp. 43-59.

Chadha, J. S. and P. Schellekens (1999): "Utility Functions for Central Bankers – The Not So Drastic Quadratic", LSE Financial Markets Group Discussion Paper No. 308 (also available as "Monetary Policy Loss Functions: Two Cheers for the Quadratic", Bank of England Working Paper No. 101).

Clarida, R., Gali, J. and M. Gertler (1999): "The Science of Monetary Policy", Journal of Economic Literature, 37, 4, pp. 1661-1707.

Clark, P., Laxton, D. and D. Rose (2001): "An Evaluation of Alternative Monetary Policy Rules in a Model with Capacity Constraints", Journal of Money, Credit and Banking 33, 1, pp. 42-64.

Dennis, R. (2000): "Steps Toward Identifying Central Bank Policy Preferences", Federal Reserve Bank of San Francisco, mimeo.

Goodhart, C. A. E. (1999): "Central bankers and uncertainty", Bank of England Quarterly Bulletin, pp. 102-114.

Goodhart, C. A. E. (2000): "The Inflation Forecast", mimeo, London School of Economics.

Granger, C.W.J. (1999): "Outline of Forecast Theory Using Generalised Cost Functions", Spanish Economic Review, 1, pp. 161-173.

Kahneman, D. and A. Tversky (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty", Journal of Risk and Uncertainty, 5, pp. 297-323.

Kahneman, D. and A. Tversky (eds.) (2000): *Choices, Values, and Frames,* Cambridge University Press.

Horowitz, A. R. (1987): "Loss Functions and Public Policy", Journal of Macroeconomics, 9, 4, pp. 489-504.

Laxton, D., Rose, D. and R. Tetlow (1993): "Monetary policy, uncertainty, and the presumption of linearity", Technical Report No. 63, Bank of Canada.

Mankiw, N. G. (2001): "The Inexorable Trade-off between Inflation and Unemployment", Economic Journal, 111 (May), pp. 45-62.

Orphanides, A. and V. Wieland (2000): "Inflation Zone Targeting", European Economic Review, 44, 7, pp. 1351-1387.

Pearlman, J. R. (1992): "Reputational and non-reputational policies under partial information", Journal of Economic Dynamics and Control, 16, pp. 339-357.

Prelec, D. (1998): "The Probability Weighting Function", Econometrica, 66, 3, pp. 497-527.

Rabin, M. (1998): "Psychology and Economics", Journal of Economic Literature, 36, 1, pp. 11-46.

Rabin, M. (2000): "Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion" in *Choice, Values and Frames*, by D. Kahneman and A. Tversky (eds.), Cambridge University Press.

Rabin, M. and R. Thaler (2001): "Anomalies: Risk Aversion", Journal of Economic Perspectives, 15, 1, pp. 219-232.

Rudebusch, G. D. (2001): "Is the Fed Too Timid? Monetary Policy in an Uncertain World", The Review of Economics and Statistics, 83, 2, pp. 203-217.

Ruge-Murcia, F. J. (2001): "Inflation targeting under asymmetric preferences", Banco de Espana, Servicio de Estudios, Documento de Trabajo n. 0106.

Smets, F. (1998): "Output gap uncertainty: does it matter for the Taylor rule?", BIS working paper No. 60.

Starmer, C. (2000): "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk", Journal of Economic Literature, 37, pp. 332-382.

Svensson, L. (1997): "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets", European Economic Review, 41, 6, pp. 1111-1146.

Svensson, L. (1999): "Inflation Targeting: Some Extensions", Scandinavian Journal of Economics, 101, 3, pp. 337-361.

Svensson, L. and M. Woodford (2000): "Indicator Variables for Optimal Policy", NBER Working Paper No. 7953.

Tambakis, D. N. (1999): "Monetary Policy with a Nonlinear Phillips Curve and Asymmetric Loss", Studies in Nonlinear Dynamics and Econometrics, 3, 4, pp. 223-237.

Thaler, R. H. (2000): "From Homo Economicus to Homo Sapiens", Journal of Economic Perspectives, 14, 1, pp. 133-141.

Theil, H. (1958): *Economic forecasts and policy*, Amsterdam: North Holland.

Annex: derivation of the optimal policy rule under non-time separability in text.

The central bank intertemporal loss function postulated in Section 3.2 of the text may be written down as:

$$E_t L = E_t[\pi_{t+1}^2 + \theta(\pi_{t+1} + \eta\pi_t)^2 + \gamma\pi_{t+2}^2 + \gamma\theta(\pi_{t+2} + \eta\pi_{t+1})^2] \tag{27}$$

Developing the L term it is obtained:

$$L = \pi_{t+1}^2(1 + \theta + \gamma\theta\eta^2) + \pi_{t+1}(2\theta\eta\pi_t + 2\gamma\theta\eta\pi_{t+2}) + \theta\eta\pi_t^2 + \gamma\pi_{t+2}^2(1 + \theta) \tag{28}$$

Differentiating L by $\widetilde{i_t}$ (recalling that $\pi_{t+1} = \varepsilon_{t+1} - \widetilde{i_t}$) and setting $a = 1 + \theta + \gamma\theta\eta^2$, $b_t = \theta\eta\pi_t + \gamma\theta\pi_{t+2}$:

$$\frac{\partial L}{\partial \widetilde{i_t}} = -2a(\varepsilon_{t+1} - \widetilde{i_t}) - 2b_t, \tag{29}$$

hence the first order condition (recalling $E_t\varepsilon_{t+1} = 0$) is:

$$E_t \frac{\partial L}{\partial \widetilde{i_t}} = 0 \Rightarrow \widetilde{i_t} = \frac{E_t b_t}{a} \tag{30}$$

Now recall that $\pi_{t+2} = \varepsilon_{t+2} - \widetilde{i_{t+1}}$ and $E_t\varepsilon_{t+2} = 0$; therefore, $E_t\pi_{t+2} = -\widetilde{i_{t+1}}$. There follows:

$$\widetilde{i_t} = \frac{\theta\eta\pi_t - \gamma\theta\eta\widetilde{i_{t+1}}}{a}, \tag{31}$$

whereby after some simple algebra the steady state value of $\widetilde{i_t}$ can be obtained:

$$\widetilde{i_t} = \frac{\theta\eta\pi_t}{1 + \theta + \gamma\theta\eta^2 + \gamma\theta\eta} \tag{32}$$

Recalling that $i_t = \frac{z_t + \widetilde{t_t}}{c_3}$, the monetary policy instrument is therefore set as follows:

$$i_t = \frac{z_t}{c_3} + \frac{\theta \eta \pi_t}{c_3(1 + \theta + \gamma\theta\eta^2 + \gamma\theta\eta)}, \tag{33}$$

which is exactly expression (25) in text.