Policy Makers Priors and Inflation Density Forecasts

Marco Vega [*]
LSE and Central Bank of Peru.

This version, October 2003

Abstract

This paper models an inflation forecast density framework that closely resembles actual policy makers behaviour regarding the determination of the modal point, the uncertainty and asymmetry in the inflation forecasts.
The framework combines policy makers prior information about these parameters with a standard parametric density estimation technique using Bayesian theory. The combination crucially hinges on an information-theoretic utility function gains of the policy maker from performing the forecast exercise.

JEL Classification: C53, E37, E58

Keywords: Monetary Policy, Inflation Targeting, Bayesian Methods.

1 Introduction

The aim for price stability has lead many central banks to be key inflation forecasters. This fact has been even more noticeable with the advent of the inflation-targeting framework for monetary policy. Inflation forecasts are important in this regime because they can be regarded as intermediate targets in the implementation of the regime as proposed in Svensson (1997).
On the other hand, inflation forecasts made by a central bank together with the formal explanation of the factors behind them serve as a signaling device for central banks to communicate how their actions are taken in relation to the likely outcomes for inflation forecasts at a specified horizon. An important challenging factor is that the forecasts are in practice subject to a myriad of asymmetric risks that unavoidably affect the asymmetry of the inflation forecast itself. This sheer fact has prompted central banks to turn attention to density instead of point forecasts - see Goodhart (2001).
The forecasting practice in central banks generally relies on conditioning the forecast not only to a specific policy stance but also to the outlook of exogenous factors that drive inflation. Our approach lies in estimating a parametric inflation density forecast where uncertainty, asymmetry and central tendency profiles are brought about mainly from the exogenous variables through the use of a core macroeconomic forecasting model. The estimated parameters are

[*]I thank Charles Goodhart for his comments and support. The opinions expressed herein are those of the author and not necessarily represent those of the Central Bank of Peru.

combined with the prior views through an explicit Bayesian approach. The prior views encompass all other factors of risk and uncertainty that may affect the inflation forecast. The formulation postulates that policymakers weigh their confidence in both; their prior beliefs and their model via a utility function of the sorts used in information-theoretic design as proposed by Lindley (1956).

This is a more realistic way of combining prior beliefs with model-based density forecasts. The approach is particularly important, in conditions where macroeconometric formulation of models is hindered by measurement errors and poor data availability. Nevertheless, even in stable and rich countries with quality data reach environments, prior inputs are essential.

The paper proceeds as follows, in Section 2 we outline the density forecast framework, in Section 3 we first describe the central bank forecasting process under investigation and then put forward the basic definitions regarding the density forecasts as well as the prior assumptions. In Section 4 we illustrate the methodology with a simple example for forecasting Peruvian inflation. Finally, in section 5 we draw our concluding remarks.

2 Density forecast framework

The forecasting literature has recently turned attention from point forecasts towards density forecasts[1]. The reasons to provide complete representations of probability distributions lie on the failure of the certainty equivalence principle in a world overwhelmingly characterised by asymmetric risks. This failure is particularly relevant in the fields of financial risk management and modern monetary policy where decision theory plays a substantial role nowadys.

The certainty equivalence principle in a gaussian stochastic linear-quadratic (LQ) environment[2] allows to compute the expected loss of an action taken at time t to achieve a particular objective value at time $t+n$ in terms of computing that loss as a function of the expected value of the objective. However, even if we can entertain the idea of a quadratic loss function for a central bank; it is hard to believe that all users of central bank inflation forecasts have quadratic loss functions. Moreover, some risks are asymmetric in nature, which makes point forecasts or expected-value calculations insufficient to assess expected losses.

Some central banks like the Federal Reserve in the USA or the Bank of England have a long tradition in macroeconomic point forecasts. Only recently, the Bank of England has pioneered the presentation of density forecasts by means of fan charts. Since then, a number of inflation targeting central banks publish density forecasts with varying degrees of detail. On the second half of 2003, twelve out of twenty inflation-targeting central banks published a fan chart[3].

The value of publishing this fan charts has been stressed in Blix and Sellin (1998), Goodhart (2001), Tarkka and Mayes (2000) and Wallis (2003b). The density forecast approach is appropriate because:

[1] See Diebold, Hahn and Tay (1998) and Tay and Wallis (2000).

[2] Linear in constraints and quadratic in loss function.

[3] In alphabetical order: Brazil, Chile, Colombia, Hungary, Iceland, Israel, Norway, Peru, South Africa, South Korea, Sweden, Thailand, and United Kingdom. In Fracasso et al (2002), Israel appears as not publishing a Fan chat because the inflation report under assessment exceptionally did not have one. Colombia is not considered in their sample due to "limited information".

- Central banks can better communicate their assessment of uncertainty and risk through adequate presentations of their density forecasts. This type of information provides a much richer insight of how a central bank views the future, allowing the users of the inflation report to produce probabilistic assessments of future outcomes. As Sims (2002) puts it, this makes it clear that the policymaker does not make mistakes when outcomes deviate from their most likely values by about the expected absolute amount.

- The internal discussions about the forecasts to be published are focused on a number of dimensions representing underlying sources and degrees of uncertainty and risk.

- The density forecasts can be quantitatively evaluated as opposed to the scenarios (or variants) approach advocated as an alternative device to show forecast uncertainty[4].

Leading density forecast central banks[5] have favoured the use of specific parametric methods to construct their densities. The parameters governing the forecast densities directly control for uncertainty and risk.

Uncertainty:
One direct measure of uncertainty is given by the variance of historical forecast errors; which is general enough to aggregate all the sources of uncertainty[6] for forecasts made at previous periods. Nevertheless, determining the uncertainty of forecasts made at a current period is by all means a daunting task given that the policymaker has to be able to somehow aggregate all the sources of uncertainty that will likely prevail now and in the coming future.
The classical econometric literature treats model-based forecast uncertainty coming from parameters and exogenous variables in a fairly standard way: it assumes that all historical conditions will prevail and extrapolates historical records into the future. Central bankers however might worry more about future prospective sources of uncertainty rather than the historical accounts of them. For example, inflation reports - published before the Iraq war in early 2003 - stated concerns about the evolution of oil prices in the future.

Risk:
In the inflation targeting context, the concept of risk denotes the probabilities that the inflation forecasts will lie above and/or below three likely benchmark measures: the central forecast, the inflation target or the inflation-target bounds.
Here, we concentrate on the probability of the inflation realizations being below (above) the central forecast of inflation[7]. The inflation reports analysed in this study seem to convey to the readers the connection between the risk balance of exogenous factors and the resulting risk balance of inflation.

[4]See Don (2000) for an analysis of this approach. In practice, the Bank of Canada or the Czech National Bank follow this scheme.

[5]For the Bank of England the references are Britton, Fisher and Whitley (1998) and Wallis (2003), For the Riksbank the references is Blix and Sellin (1998), Blix and Sellin (1999).

[6]Such as changes in the underlying structure of the economy, model uncertainty, measurement errors, exogenous variables, subjective adjustments to model-based forecasts, etc.

[7]In Kilian and Manganelli (2003) the risk assessment is done in terms of probabilities of being away from the target.

Given that future uncertainty and risk factors affecting the forecast have an essential subjective input, a natural framework to study this behaviour is Bayesian theory. The subjective approach to risk and uncertainty necessary in the monetary policy process can not be conceptualized by classical econometric practice. Instead a Bayesian approach seems to provide a reasonable framework. To this we turn now.

3 Modelling a central bank forecasting process

A central bank forecasting process is basically an institutional arrangement contingent on the historical perspective and organizational structure of each central bank. Regardless of how elaborate and particular this process may be, there are striking common features. In general, both; central banks policy and forecast publications stress explicitly that the output of the forecast is influenced (or conditioned) by the outlook of exogenous variables and the particular conditioning policy. Other conditioning factors such us parameters or model types are mentioned scantily if at all.

To this end, models seem to be key elements that link both exogenous variables and policy decisions with the forecasted endogenous variables. The role of models within this process has been recognized by academics and practitioners alike. In a recent survey of central banks practicing inflation targeting (Schmidt-Hebbel and Tapia (2002)), basically all 20 surveyed banks refer the use of some kind of model. The key evidence is that most central banks, specially inflation targeters endorse the use of one core forecasting model that helps center policy discussions within the bank.

But the use of models in forecasting does not mean that subjective views are filtered out in the forecasting process. In fact, a factor also mentioned in the Schmidt-Hebbel and Tapia (2002) survey is that in most central banks; the published forecasts are a *"balanced combination"* of technical forecasts and decision makers' views. The practice of including subjective approaches to macroeconomic forecasting within central banks is also recognized in Sims (2002) and Goodhart (2001).

From the standpoint of decision makers, the subjective approach is justified in a context of monetary policy making under uncertainty. Central banks use a core forecasting model but they are aware that their model can not capture the richness of reality, not to mention how this reality may evolve in the future. According to Sims (2001), the forecasting process within central banks can be understood using Bayesian decision theory;

> "Policy discussion at central banks uses the language of Bayesian decision theory putting postsample probabilities on models, generating probability distributions for future values of variables that reflect uncertainty about parameter values and subjective judgment, weighing expected losses of alternative courses of action".

We attempt to take this view of central banks forecasting process. We model a forecasting process that somewhat resembles this approach. In our framework; the technical staff interacts with the policy decision makers. The staff implements simulations using a model and policymakers input priors about parameters that reflect uncertainty, risk balance and central forecast values. In Table 1,we depict the tasks involved in the forecasting exercise, the arrows indicate who executes each task and the tasks are numbered in some sequential order. The technical staff starts gathering information relevant to the model and interaction with policy makers produces the assumptions of the model. As mentioned before, the outlook of exogenous factors serves to shed light on the uncertainty surrounding the final forecast. This is task (3) where policy makers provide a general view which is subsequently translated by the technical staff. Tasks (3) and (4) are repeatedly done until a sensible central scenario is achieved. Task (5) produces the density forecast and estimates its parameters. Task (6) corresponds to the elicitation of priors by policy makers. This task is done after receiving a first notion about the density forecast (in form of an initial fan chart). Finally, by performing task (7), the technical staff factors policy makers priors into the density forecast by means of Bayes theorem.

Table 1: A Central Bank forecasting process:

Technical Staff		**Policy Maker**
$\longrightarrow$	(1) Gather information	
$\longrightarrow$	(2) Make assumptions about initial conditions, trends and cycles	$\longleftarrow$
	(3) Produce outlook of exogenous factors	
	Provide a general view	$\longleftarrow$
$\longrightarrow$	Parameterize variance, skewness and central scenario	
$\longrightarrow$	(4) Run forecast rounds and Monte Carlo simulations	
$\longrightarrow$	(5) Parameterize density forecast	
	(6) Input "subjective" view of density forecast parameters	$\longleftarrow$
$\longrightarrow$	(7) Combine subjective views from (6) into (5)	

In what follows attention is put to tasks (5), (6) and (7). We provide formal detail.

3.1 The density forecast

Suppose that the forecasting process at time t about future realizations of an inflation sequence up to horizon H, is denoted by $\{\widehat{\pi}_s\}_{s=t+1}^{H}$ which is generated through a model[8]:

$$\pi_s = M_s(Y_t, X_t; \theta, I_t) \text{ for } s = t+1, t+2...H \tag{1}$$

[8]Hatted variables are forecast of either exogenous or endogenous variables. In the case of the instrument setting, it refers to the stance assumed by the policymaker.

In equation [1], Y_t denotes the known history of endogenous macroeconomic variables y_t in the model including inflation π_t. Formally:

$$Y_t = \{y_t, ..., y_{t-n}\}$$

A macroeconomic model-based forecast is conditional upon various factors that can be controlled in the process. These factors are X_t, θ, and I_t. The first one denotes the history and likely future realizations of the exogenous variables: $X_t = \{x_{t-n}, .., x_t, \widehat{x}_{t+1}...\widehat{x}_{t+H}...\}$, θ denotes the set of parameters that defines the particular economic model in use. This set of parameters is included in the broader set of parameters Θ that defines model uncertainty. The last factor, I_t denotes the history as well as the particular stance of the central bank instrument assumed at time t: $I_t = \left\{i_{t-n,}..., i_t, \widehat{i}_{t+1}, ...\right\}$.
Model M is general enough and need not be explicit as it may correspond to a rational expectations equilibrium solution. We make the following definition:

Definition 1. *A central forecast*[9] *is an inflation sequence* $\{\widehat{\pi}_{c,s}\}_{s=t+1}^{H}$ *obtained by conditioning the model to: (a) the most likely sequence of exogenous variables within the forecast horizon* $\{\widehat{x}_{c,s}\}_{s=t+1}^{H}$*, (b) parameter values* θ_c *and (c) the monetary policy instrument setting* $I_{c,t}$

The central forecast is the result of tasks (1), (2) and (3) as outlined above and achieved in possibly multiple rounds.

On the other hand, the technical assessment of risk and uncertainty relies on random realizations of exogenous variables from suitably calibrated probability distribution functions. The random draws take into account a chosen parameterized standard deviation, skewness and the "most-likely" sequence of exogenous variables. The parameters of these probability density functions reflect the technical staff historical estimates as well as subjective and the informed view of sectorial experts.

Among the distinct probability density functions that are suitable to perform random draws are the Beta and the Split Normal. The latter is used intensively in Blix and Sellin (1998) and Briton, Fisher and Whitley (1998). These two types of distribution are useful because their parameters illustrate the distributional characteristics that matter most in a density forecast; a central point; a measure of dispersion and a measure of skewness.

Performing simulated histories of exogenous variables within the forecast horizon allows to determine alternative trajectories of inflation. Evaluated at each point in time within the forecast horizon, the distinct inflation points originated in the simulation can be hypothesized as coming from a generic probability function conditional on the probability distribution of the exogenous variables.
The determination of the exact probability distribution function of inflation resulting from this exercise is hindered by two facts (a) the mapping from the exogenous variables to inflation imply a solution like [1] which can be highly non-linear and (b) even if we manage to find the

[9] In this definition, the subscript c denotes both central forecasts and assumed central values.

exact form of the distribution; its communication to policy makers would not be easy. A way to circumvent the problem is to assume a parametric form for the distribution function that can serve two purposes; be a good approximation to the true pdf and allow a communication strategy that can easily be grasped by the policy maker. A good candidate for the assumed pdf is the Split Normal, given that its parameters can be easily communicated in terms of straightforward balance of risks.

Definition 2. *A model-based parametric density forecast of inflation is a sequence of parameters* $\left\{\widehat{\Lambda}_{c,s}\right\}_{s=t+1}^{H}$ *describing a probability density function of the inflation forecast at every point in time s.*

The parameters are obtained by a likelihood estimation procedure assuming the Split Normal distribution for inflation.

Henceforth, we are going to concentrate on the relevant horizon H and drop time subscripts. After S number of stochastic simulations on the exogenous variables, we can define the following mapping from data conditional on the model parameters and the instrument setting to object ω:

$$\left(\{X_t\}_{j=1}^{S}, Y_t; \Theta, I_t\right) \rightarrow \omega \tag{2}$$

Object ω contains the elements upon which both the econometrist and the policy maker care about[10]: the inflation forecast at horizon H and the three parameters that underlie policy discussions. We will group these three parameters in the vector $\Lambda = \left(m, \sigma^2, \gamma\right)$, with m being the modal point, σ^2 the uncertainty measure and γ the skewness of the inflation forecast distribution. These three parameters uniquely define the Split Normal $SN\left(m, \sigma^2, \gamma\right)$. This distribution collapses into a Normal $N\left(m, \sigma^2\right)$ whenever the skewness parameter γ equals zero. The γ parameter varies on the range $\langle -1, 1\rangle$ and is closely linked to the balance or risks made at central banks (see Appendix B). We specify ω in a compact way:

$$\omega = \left(\{\pi\}_{j=1}^{S}, \Lambda\right) \tag{3}$$

We treat ω in a Bayesian context. We characterise its posterior probability density conditional on all the information acquired after performing S number of simulations of the model conditional on all the given factors Ω (observe that S itself is a variable to be determined)

$$p(\omega|\ \Omega) = p(\Lambda|\ \Omega)p(\{\pi\}_{j=1}^{S} \mid \Lambda, \Omega) \tag{4}$$

where:
Ω is the given information set: $\Omega = \{\{X_t\}_{j=1}^{S}, Y_t; \Theta, I_t\}$
$p(\Lambda|\ \Omega)$ is the prior density elicited by the policy maker, and

[10]Observe that the parameter as well as the instrument remain constant along the simulations.

$p\left(\{\pi\}_{j=1}^{S} \mid \Lambda, \Omega\right)$ is the probability of the simulated inflation forecast data given the information Ω and the parameters of interest. The likelihood principle implies that this probability is equivalent to the likelihood of the parameters given the simulated data and the information set: $L(\Lambda| \ \{\pi\}_{j=1}^{S}, \Omega)$.

Our interest is to draw probabilistic judgments of the inflation forecast distribution, thus we need to find the posterior conditional distribution of the parameters. This is achieved by making use of Bayes theorem:

$$p(\Lambda| \ \{\pi\}_{j=1}^{S}, \Omega) = \frac{p(\Lambda| \ \Omega)L(\Lambda \mid \ \{\pi\}_{j=1}^{S}, \Omega)}{p(\{\pi\}_{j=1}^{S} \mid \Omega)} \tag{5}$$

Given that the prior distribution as well as the likelihood are known parameterized functions, the posterior distribution can be explicitly determined. Furthermore, by holding constant a pair of parameters we can determine the conditional distribution of the remaining parameter.

3.2 Elicitation of the priors as the outcome of policy makers views:

Upon learning the outcome of the model-based density forecast, policy-maker's views are formed. These views take into account other forms of uncertainties not included in the forecast: model-uncertainty; measurement errors, etc. The way to optimally extract these views and to translate them into tractable distribution functions is an internal operational task.
For our purpose we assume that the first subjective view is that the three parameters are independent random variables, so that the joint prior is:

$$p(\Lambda| \ \Omega) = p(\sigma^2| \ \Omega)p(\gamma| \ \Omega)p(m| \ \Omega) \tag{6}$$

Prior for uncertainty parameter σ^2

As customarily done in the literature, we assume the family of the Inverted Gamma-2 distributions $iG_2(b, a)$ in terms of the parameters (a, b) to be chosen by the policy maker. This distribution has support $\langle 0, \infty\rangle$ and its parameters can be specified using the two moments and the mode of the distribution as guidelines:

$$E(\sigma^2|.) \equiv \frac{b}{a-2} \text{ for } a > 2$$

and

$$V(\sigma^2|.) \equiv \frac{2}{a-4}\left(\frac{b}{a-2}\right)^2 \text{ for } a > 4$$

while the mode is:

$$mode(\sigma^2|.) \equiv \frac{b}{a+2}$$

It can be observed that the mean is always higher than the mode, by taking the estimated $\widehat{\sigma}_c^2$ in Definition 2 as a reference point, possible values of b and a can be evaluated by weighing the resulting mode, mean and variance.

Prior for skewness parameter γ

For the skewness parameter we need a distribution with bounded support. We assume a slight transformation of a Beta distribution; we name it $\widetilde{B}(c, d)$. This allows γ to vary in the interval $\langle -1, 1\rangle$. To do this we make a transformation of a random variable z lying on the interval $\langle 0, 1\rangle$ and following a Beta distribution $B(c, d)$ (the transformation applied is $\gamma = 2z - 1$). The first two moments are defined as:

$$E(\gamma|\Omega) \equiv \frac{c-d}{c+d}$$

and

$$V(\gamma|\Omega) \equiv \frac{4cd}{(c+d+1)\,(c+d)^2}$$

with mode:

$$mode(\gamma|\Omega) \equiv \frac{c-d}{c+d-2}$$

Prior for mode parameter m

We impose a non-informative uniform distribution for the mode m:

$$p(m|a_m, b_m) \propto \text{constant} \tag{7}$$

3.3 The posterior distribution

Given the Split Normal likelihood assumption[11], the kernel of the joint posterior distribution of the three parameters of interest are:

[11]See Appendix [B] for details about this distribution.

$$p\left(\Lambda|\ \pi_{t+H},\Omega\right) \;\propto\; \left(\frac{\gamma+1}{2}\right)^{c-1}\left(\frac{1-\gamma}{2}\right)^{d-1}\left(\sigma^2\right)^{\frac{-(a+2)}{2}} e^{\left(\frac{-b}{2\sigma^2}\right)}$$
$$\left(\frac{\left(\sigma^2\right)^{-\frac{1}{2}}}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right)^N e^{\left(\frac{-1}{2}\left\{\sum_{i=1}^{S_1}\left(\frac{\pi_{t+H}-m}{\sigma\sqrt{1-\gamma}}\right)^2+\sum_{i=S_1+1}^{S}\left(\frac{\pi_{t+H}-m}{\sigma\sqrt{1+\gamma}}\right)^2\right\}\right)} \tag{8}$$

From this joint pdf, we can obtain the posterior conditional distribution of σ^2. As expected, this distribution is also an Inverted Gamma-2.

$$p\left(\sigma^2|\ \gamma,m,\pi_{t+H},\Omega\right)\propto\left(\sigma^2\right)^{\frac{-(a+N+2)}{2}} e^{\left(\frac{-(\vartheta(m,\gamma)+b)}{2\sigma^2}\right)} \tag{9}$$

where $\vartheta(m,\gamma)=\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{1-\gamma}\right)+\sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{1+\gamma}\right)\right\}$

The other two relevant conditional distributions are given by:

$$p\left(m|\ \gamma,\sigma^2,\pi_{t+H},\Omega\right)\propto e^{\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{(1-\gamma)}\right)+\sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{(1+\gamma)}\right)\right\}\right)} \tag{10}$$

and

$$p\left(\gamma|m\ ,\sigma^2,\pi_{t+H},\Omega\right) \;\propto\; \left(\frac{\gamma+1}{2}\right)^{c-1}\left(\frac{1-\gamma}{2}\right)^{d-1}\left(\frac{2}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right)^S$$
$$e^{\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{1-\gamma}\right)+\sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{1+\gamma}\right)\right\}\right)} \tag{11}$$

The conjugacy of the prior distribution of σ^2 allows to express the conditional moments from the posterior from an inverted gamma distribution $iG_2(\frac{a+S}{2},\frac{2}{\vartheta(m,\gamma)+b})$: The moments are:

$$E(\sigma^2|.)\equiv\frac{\frac{a+S}{2}}{\frac{2}{\vartheta(m,\gamma)+b}-2}\text{ for }\frac{2}{\vartheta(m,\gamma)+b}>2$$

and

$$V(\sigma^2|.)\equiv\frac{2}{\frac{2}{\vartheta(m,\gamma)+b}-4}\left(\frac{\frac{a+S}{2}}{\frac{2}{\vartheta(m,\gamma)+b}-2}\right)^2\text{ for }\frac{2}{\vartheta(m,\gamma)+b}>4$$

while the mode is:

$$mode(\sigma^2|.) \equiv \frac{\frac{a+S}{2}}{\frac{2}{\vartheta(m,\gamma)+b}+2}$$

From this explicit representation, we can see that as the sample size increases, the posterior mean and mode collapse to the model-based estimates. In that case, the prior view has a small effect on the posterior outcome. In an econometric estimation environment, a larger sample size is always good because it improves the model-based information. Our context is rather different. It is based on the willingness of a Bayesian policy maker to learn about the properties of the inflation forecast from a general perspective. It is not about a non-Bayesian econometrist who wants to learn the properties of its model-based forecast.

3.4 The choice of sample size as an information theoretic design problem

In the framework we propose, the sample size S is a choice variable as well. If a high enough sample size is considered, the prior view of the policy makers becomes useless. On the other hand, if the sample size is small, then the model-based estimation turns less accurate so that the simulation experiment becomes informationally poor.
Policy makers need to weigh the information provided by the model and the prior beliefs they may hold. In practice, this process appears complex as it is bound to the subjective beliefs of the policy makers coupled with out-of-model information they have.
Under these circumstance, the information-theoretic approach[12] for "experiment design" seems plausible. In our view, the experiment performed by policy makers consists in updating their prior beliefs about the inflation forecast modal point, uncertainty and risks by means of a econometric forecasting model. The outcome of these updating depends crucially on the simulation sample size under evaluation. They choose sample size S so that policy makers maximize their expected utility resulting from the experiment.

$$S^* = \arg\max_S \{KL(S) - \lambda S\} \tag{12}$$

The expected utility of the experiment with sample size S depends on two factors; a) the Kullback-Leibler (KL hereon) divergence between the posterior and prior distribution of the parameters $KL(S)$ and b) the linear loss function λS. The KL number provides the value of the information provided by the forecasting model under use[13]. The loss of the policy maker is rationalised by the unwillingness to disregard their priors. So, as the sample size increases, the prior of the policy maker is downweighted and thus reduces the utility of a policy maker who considers her priors are indeed somewhat important. In our case, the utility parameter λ is the degree of importance of the prior in the overall utility function[14].

[12] This view was proposed by Lindley (1956). Applications of Lindley's approach are found for example in Ryan (2003), Clyde(2001), Parmigiani and Berry(1994?), Chaloner and Verdinelli (1995) and Muller and Parmigiani(1996). Most of these applications are in the design of clinical experiments.

[13] KL(S) is increasing in S and concave. See Lindley (1956)

[14] λ can also be interpreted as the inverse of policymakers credibility on the model.

The KL divergence number is defined as:

$$KL(S) = \int_{\Lambda}\int_{\Pi} \log\left[\frac{p(\Lambda|\Pi, S)}{p(\Lambda)}\right] p(\Pi, \Lambda|S) d\Pi d\Lambda \tag{13}$$

Where $\Pi = \{\pi\}_{j=1}^{S}$ is the simulated inflation data of size S, $p(\Lambda)$ is the prior distribution of the parameters and $p(\Lambda|\Pi, S)$ is the posterior distribution.

4 An example

In order to provide an example, we use a simple ad-hoc univariate model for quarterly inflation estimated using ordinary least squares[15]. We run the inflation rate at quarter t against the following regressors: the exchange rate depreciation at lag 3 (Δe_{t-3}), GDP growth at lag 2 (g_{t-2}), the mean interbank interest rate at lag 1 (i_{t-1}), the mean three months Libor rate at lag 3 (i^*_{t-3}) and the terms-of-trade growth at lag 4 (Δtot_{t-4}).

$$\begin{array}{llllllll} \pi_t = & 0.69\pi_{t-1} & +0.24\Delta e_{t-1} & +0.23 g_{t-2} & -0.30 i_{t-1} & +0.55 i^*_{t-3} & +0.06\Delta tot_{t-4} & +\varepsilon_t \\ & (9.23) & (3.58) & (3.06) & (-1.95) & (1.72) & (1.70) & \end{array} \tag{14}$$

The estimation[16] is carried out using data from the first quarter of 1994 to the second quarter of 2003. Except for lagged inflation, all the variables on the right-hand side are considered as exogenous. Hence, to start the density forecast we need to construct a baseline scenario and uncertainty and risk profiles for the set of exogenous variables: $(g_t, i_t, \Delta e_t, i^*_t, \Delta tot_t)$. In particular, we assume the following distributions:

Exogenous variable	Balance of risk	Distribution	Mode	σ^2
Libor rate	upside 70%	Split normal	3.57	1.2
Nominal exchange rate depreciation	upside 55%	Split normal	0.00	10.6
GDP growth	upside 60%	Split normal	3.90	8.3
Terms of trade growth	neutral	Normal	0.5	4.9

Table 1: Distributional assumptions for exogenous variables at the end of the forecast horizon

In Figure [1] we show the historical, central scenario and the 90 per cent central prediction interval for the exogenous variables along the forecast periods. The asymmetry as well as the uncertainty increases linearly until it reaches the values specified in Table [3]. In each forecast period, we also consider random realizations of the unforecastable shock ε_t, drawn from a

[15]We use data from Peru. The Central Bank of Peru has recently adopted the Inflation Targeting framework (January 2002).

[16]In equation [14] the lag structure minimises the sum of squared residuals. As usual, the t-values are in parenthesis.

normal distribution $N(0, 0.3)$. This last feature is important for two reasons; first it makes the first-period-ahead inflation forecast random given that all the exogenous determinants are predetermined for this horizon. Second, it allows the inflation uncertainty to increase even in the absence of uncertainty in the exogenous variables.
To complete the conditioning factors, we also need to assume a particular monetary policy setting within the forecast horizon. In this case, we consider a constant-interest-rate forecast with the rate kept at 2.75 per cent during the forecast period.

The inflation density forecast is then achieved by estimating the parameters of an assumed split normal distribution $SN(m, \sigma^2, \gamma)$ for the simulated sample of size[17] S_T for each forecast period.

An important conclusion emerges from this exercise: Notwithstanding that the exchange rate depreciation, GDP growth and the Libor rate all show considerable asymmetry[18] (especially at the end of the forecast horizon). There is no build up of asymmetry in both inflation measures; the quarterly and the year-on-year rate. In Figure [2] we show the estimated densities at each of the eight forecast periods along with the estimated parameters; mode m, σ^2 and γ. The gamma parameter is close to zero in all periods.
The reasons why the increasingly asymmetric nature of exogenous variables does not pass on to inflation are twofold; the lag structure and the interplay between the variability versus asymmetric forces. Regarding the lag structure, as the asymmetric exogenous variables affect quarterly inflation with some lags, then full asymmetry is not transferred to inflation at the end of the forecast horizon. As of the relation variability/asymmetry, it is known that when the variability of inflation increases the asymmetric forces that affect inflation are dampened (see for example Blix and Sellin 2000). Inflation variability does grow because the exogenous variability increases linearly and because the persistent nature of inflation (as it depends strongly on its own lags) exacerbates all the sources of uncertainty in inflation, even the one that corresponds to the inflation shock itself.

The estimated mode from the simulations is quite different from the one computed using only the central scenario values of exogenous variables. There is an upward bias (See Figure [3]) in both the quarterly inflation and the year-on-year inflation. The reason is that at the end of the forecast horizon, the simulated distribution is quite symmetric around the mean. The mean is the central tendency that is preserved in both the point and the density forecast.

Once the results of the simulation are known, we proceed to introduce the information provided by the policy maker. To do this, we concentrate in forecast horizon $H = 8$. We need to assume a prior distribution for the set of parameters $\Lambda = \left(m, \sigma^2, \gamma\right)$. We take the distributional assumptions outlined in Section 3. Namely:

[17]In this step, the sample size S_T can be as large as possible. The objective here is to get the most accurate distributional representation originated from the forecasting model alone.

[18]In Figure [3] in the appendix the estimated means differs from the modes of the asymmetric exogenous variables. In Figure [4] the asymmetry parameter γ for the exogenous variables becomes larger towards the end of the forecast horizon.

The mode follows a uniform distribution; $m \sim U(m_{low}, m_{high})$ with parameters $m_{low} = -0.22$ and $m_{high} = 5.78$ such that the distribution is centered in an year-on-year inflation rate of 2.78 percent.

The uncertainty parameter follows an inverted gamma-2 distribution; $\sigma^2 \sim iG_2(b, a)$. In order to find the parameters, we can consider that the estimated $\widehat{\sigma}^2$ from the simulation step is too low. Policy makers may consider that there are other factors that necessarily drive forecast uncertainty to a higher level. For example they can assume that $E_{prior}(\sigma^2) = 1.95$ and the $mode_{prior}(\sigma^2) = 1.8$. This implies the corresponding parameters $(a, b) = (38, 72)$

The asymmetry parameter follows a beta type of distribution considered in Section 3; $\gamma \sim \widetilde{B}(c, d)$. In this case, we assume policy makers believe that the inflation forecast at horizon H will have an upside risk, as opposed to the model-based case which considers a slight downside risk. Let's suppose that the mean prior gamma is $E_{prior}(\gamma) = 0.3$ (which is close to a 60 percent upside risk) and that they believe about this asymmetry quite strongly $V_{prior}(\gamma) \approx 0.006$. This implies parameter values $(c, d) = (92.857, 50)$.

Before combining the prior information given by the policy maker, it is necessary to establish the sample size to use in the Bayesian procedure. The sample size is obtained from solving the problem in equation [12)]. The calculation of the utility measure requires to get the KL divergence number via some numerical integration procedure. In Appendix D, we follow Ryan (2003) by using a MCMC estimation. The optimal value S^* depends on the parameter λ. A small λ about 0.007 is related to a large sample size (about 164), a "large" λ, around 0.017, generates a sample size of about 33. Hence, we interpret the sample size as the weight of confidence in the prior. In our example, we assume $\lambda = 0.01$. Therefore the optimal sample size is $S^* = 120$ (see Figure [7])

Next, we sample from the Bayesian conditional posterior distributions (See Sub-section 3.3 and Appendix C). The corresponding mean values are shown in Table (2) and a graphical representation of conditional posterior against prior distributions is shown in Figure [8]).

		Prior Mean	Model-based Estimation	Posterior Mean
Mode	m	2.78	3.03	2.75
Uncertainty	σ^2	1.95	0.83	0.78
Risk	γ	0.30	-0.05	0.34

Table 2: Mean values of the parameters under the prior distribution, the ML estimation and the posterior distributions.

The distributional means of the prior and posterior turn out to be very close to each other except for the uncertainty parameter σ^2. The model-based estimate of uncertainty is low while the prior belief about this parameter is too high relative to the model. Also, the model-based estimate of the asymmetry is slightly negative (-0.05) as opposed to the prior belief which posits a strong upside risk ($\gamma = 0.3$). It seems that the model strongly rejects the combination of high levels of uncertainty and sizeable upside risks as defined by the prior. Thus, in terms of the posterior, the prior view of the policy makers is taken into account for the modal and the risk forecasts, yet it is not the case for the uncertainty parameter estimation. In fact, the posterior calculation hints that a lower uncertainty seems necessary in order to "make room" for a high value of asymmetry provided in the likelihood[19].

5 Concluding remarks

This paper contributes to the understanding of how central banks do forecasts in the context of monetary policy making. It posits attention to Bayesian policy makers who hold or develop prior views on key features of the inflation density forecast. The decision makers interact with the technical staff in charged of running the macroeconomic model-based density forecast.

In reality, neither the prior views nor the model-based forecast are per se true. Prior views are subject to human imperfection while models are always false. However, policy makers in fact use both types of inputs to make quantitative inference about their forecasts.

In our approach, policy makers weigh both the prior view and the information provided by the model via a utility function advocated in Information Theory. The utility function considers the trade-off between the importance of policy makers priors and the "faith" on the core forecasting model. If the model is given full "faith" then priors are irrelevant and viceversa.

[19]This particular result does not always hold. It depends on the relative prior variances of the parameters. If policy makers are highly confident about their prior view of uncertainty, then the distributional variance is in fact very low. Therefore, the resulting posterior might be closer to this posterior.

References

[1] Bauwens, L., Lumbrano M. and Richard J-F. (1999), "Bayesian Inference in Dynamic Econometric Models", Oxford University Press

[2] Blix, M. and Sellin, P. (1998), "Uncertainty bands for inflation forecasts". Working Paper No.65, Sveriges Riksbank, Stockholm.

[3] Britton, E., Fisher, P. and Whitley, J. (1998), "The Inflation Report Projections: Understanding the Fan Chart", Bank of England Quarterly Bulletin, 38, 30-37.

[4] Chaloner, K. and Verdinelli, I. (1995), "Bayesian Experimental Design: A Review", Statistical Science 10, 273-304.

[5] Clyde, M. (2001), "Experimental Design: A Bayesian Perspective", International Encyclopedia of Social and Behavioral Sciences, Vol 2.1.

[6] Don, F.J.H. (2000), "Forecasting in macroeconomics: a practitioners view", CPB Netherlands Bureau for Economic Policy Analysis, Research Memorandum No. 162.

[7] Ericsson, N. (2001), "Forecast Uncertainty in Economic Modeling", in: Understanding Economic Forecasts (Hendry, D. and Ericsson, N., eds.) The MIT Press, 68-92.

[8] Fracasso, A., Genberg, H. and Wyplosz, C. (2003), "How do central banks write? An evaluation of Inflation Targeting Central Banks", Geneva Reports on the World Economy, Special Report 2.

[9] Geweke, J. (1998), "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication", Mimeo.

[10] Goodhart, C.A.E. (2001), "The Inflation Forecast", National Institute Economic Review No. 175.

[11] Goodhart, C.A.E. (2001), Monetary transmission lags and the formulation of the policy decision on interest rates, FRB of St. Louis Review, July/August, 165-186.

[12] Kilian, L. and Manganelli, S. (2003), "The Central Bank as a Risk Manager: Quantifying and Forecasting Inflation Risks", ECB Working Paper 226..

[13] Lindley, D. (1956), "On a Measure of the Information Provided by an Experiment", The Annals of Mathematical Statistics, Vol. 27, No. 4, 986-1005.

[14] Müller, P. and Parmigiani, G. (1996), "Numerical Evaluation of Information Theoretic Measures", in: Bayesian Analysis in Statistics and Econometrics (D. A. Berry, K. M. Chaloner, and J. K. Geweke eds), Wiley NY, 397-406.

[15] Parmigiani, G. and Berry, D. (1994), "Applications of Lindley Information Measure to the Design of Clinical Experiments", in: Aspects of Uncertainty: A tribute to D.V. Lindley (A.F.M Smith and P. Freeman eds.), Wiley NY, 333-352.

[16] Ryan, K. (2003), "Estimating Expected Information Gains for Experimental Designs With Application to the Random Fatigue-Limit Model", Journal of Computational and Graphical Statistics, Volume 12, No. 3, 1-19.

[17] Schmidt-Hebbel, K. and Tapia, M. (2002), "Monetary Policy Implementation and Results in Twenty Inflation-Targeting countries", Central Bank of Chile Working Paper 166.

[18] Sims, C. (2002), "The Role of Models and Probabilities in the Monetary Policy Process", Mimeo.

[19] Svensson, L. (1997), "Inflation forecast targeting: Implementing and monitoring inflation targets", European Economic Review, Vol. 41, 1111-1146.

[20] Svensson, L. (1999), "Inflation targeting as a monetary policy rule", Journal of Monetary Economics, 43, 607-654.

[21] Tarkka J. and Mayes D. (2000), "The Value of Publishing Official Central Bank forecasts", Mimeo, Bank of Finland.

[22] Tay, A.S. and Wallis, K. (2000), "Density forecasting: a survey", Journal of Forecasting, 19, 235-254.

[23] Wallis, K. (1999), "Asymmetric density forecasts of inflation and the Bank of England's Fan Chart", National Institute Economic Review No. 167.

[24] Wallis, K. (2003a), "Chi-squared tests of interval and density forecasts, and the Bank of Englands fan charts", International Journal of Forecasting 19, 165-175.

[25] Wallis, K. (2003b), Forecast uncertainty, its representation and evaluation. Boletin Inflacion y Analisis Macroeconomico, Universidad Carlos III de Madrid, Special Issue No 100, pp.89-98. Updated 23 January 2003.

[26] Inflation Reports published on web pages of the central banks with an inflation targeting regime: Australia, Brazil, Canada, Chile, Colombia, Hungary, Island, Israel, Mexico, Norway, New Zealand, Peru, Poland, Czech Republic, United Kingdom, South Africa, South Korea, Sweden, Switzerland and Thailand.

A Prior distributions

A.1 Prior for σ^2

In the main text we assume that σ^2 follows an Inverted Gamma 2 distribution with parameters (b, a):

$$p(\sigma^2|.) = \left(\Gamma(\frac{a}{2})\left(\frac{2}{b}\right)^{\frac{a}{2}}\right)^{-1}(\sigma^2)^{\frac{-(a+2)}{2}}e^{\left(-\frac{b}{2\sigma^2}\right)} \tag{A1}$$

Where:

$$E(\sigma^2|.) \equiv \frac{b}{a-2} \text{ for } a > 2$$

and

$$V(\sigma^2|.) \equiv \frac{2}{a-4}\left(\frac{b}{a-2}\right)^2 \text{ for } a > 4$$

while the mode is:

$$mode(\sigma^2|.) \equiv \frac{b}{a+2}$$

A.2 Prior for γ

We start assuming that a random variable z follows a Beta distribution with parameters (c, d):

$$g(z|c,d) = \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} z^{c-1} (1-z)^{d-1} \text{ for } 0 < z < 1$$

with:

$$E(z|\Omega) \equiv \frac{c}{c+d}$$

and

$$V(z|\Omega) \equiv \frac{cd}{(c+d+1)(c+d)^2}$$

with mode:

$$mode(z|\Omega) \equiv \frac{c-1}{c+d-2}$$

Then we define γ in terms of the following transformation:

$$\gamma = 2z - 1$$

Hence, the prior distribution of γ can be expressed as:

$$p(\gamma|.) = g(z(\gamma)|c,d)\left|\frac{d}{d\gamma}z\right|$$

As a result, the prior distribution for γ is:

$$p(\gamma|.) = \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)}\left[\frac{1+\gamma}{2}\right]^{c-1}\left[\frac{1-\gamma}{2}\right]^{d-1} \text{ for } -1<\gamma<1 \tag{A2}$$

A.3 Prior for m

As for m we assume a uniform, non-informative prior. The exact determination for this prior is inconsequential for the Bayesian posterior sampling. However, it is used in the sample size determination given that we require sampling from the priors. Hence, we assume: $m \sim Uniform(m_{low}, m_{high})$

$$p(m|.) = \frac{1}{m_{high} - m_{low}} \text{ for } m_{low} < m < m_{high} \tag{A3}$$

B Model-based density simulation and estimation

B.1 Fitting the simulated data

We define a Split Normal pdf for the data with parameters (m, σ^2, γ) in the following way:

$$f(x; m, \sigma^2, \gamma) = \begin{cases} \frac{2}{\sqrt{\sigma^2}\left(\sqrt{1-\gamma}+\sqrt{1+\gamma}\right)}\phi\left(\frac{x-m}{\sqrt{\sigma^2(1-\gamma)}}\right) & \text{if } x < m \\ \frac{2}{\sqrt{\sigma^2}\left(\sqrt{1-\gamma}+\sqrt{1+\gamma}\right)}\phi\left(\frac{x-m}{\sqrt{\sigma^2(1+\gamma)}}\right) & \text{otherwise} \end{cases}$$

Where $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2}$

Given a simulated sample $\{x\}_{s=1}^{S_T}$; we can sort the data in ascending order and split the ordered data $\{\widetilde{x}\}_{s=1}^{S_T}$ in two sub-samples:

$\mathcal{S}_1 = \{\widetilde{x}_i \mid \widetilde{x}_i < m\}$

$\mathcal{S}_2 = \{\widetilde{x}_i \mid \widetilde{x}_i \geq m\}$

Let S_1 and $S_T - S_1$ be the number of elements of $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. Then the likelihood of the sample is given by:

$$L(x; m, \sigma^2, \gamma) = \left(\frac{2/\sqrt{2\pi\sigma^2}}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right)^{S_T} e^{\left(\frac{-1}{2}\left\{\sum_{i=1}^{S_1}\left(\frac{x-m}{\sqrt{\sigma^2(1-\gamma)}}\right)^2 + \sum_{i=S_1+1}^{S_T}\left(\frac{x-m}{\sqrt{\sigma^2(1+\gamma)}}\right)^2\right\}\right)} \tag{B1}$$

while the log-likelihood is:

$$\mathcal{L}(x;m,\sigma^2,\gamma) = S_T \log\left(\frac{2/\left(2\pi\sigma^2\right)^{\frac{1}{2}}}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right) - \frac{1}{2}\sum_{i=1}^{S_1}\left(\frac{x-m}{\sqrt{\sigma^2(1-\gamma)}}\right)^2 - \frac{1}{2}\sum_{i=S_1+1}^{S_T}\left(\frac{x-m}{\sqrt{\sigma^2(1+\gamma)}}\right)^2$$

and further expressed as:

$$\begin{aligned}\mathcal{L}(x;m,\sigma^2,\gamma) &= S_T\log\left(2/\sqrt{2\pi}\right) - \frac{S_T}{2}\log\left(\sigma^2\right) - S_T\log\left(\sqrt{1-\gamma}+\sqrt{1+\gamma}\right) \\ &\quad -\frac{1}{2\sigma^2}\sum_{i=1}^{S_1}\left(\frac{x-m}{\sqrt{1-\gamma}}\right)^2 - \frac{1}{2\sigma^2}\sum_{i=S_1+1}^{S_T}\left(\frac{x-m}{\sqrt{1+\gamma}}\right)^2\end{aligned}$$

Estimation of the parameters requires the computation of the firs order conditions of the likelihood problem:

For the uncertainty parameter we have:

$$\frac{\partial}{\partial\sigma^2}\mathcal{L}(x;\sigma^2,\gamma,m) = -\frac{S_T}{2\sigma^2} + \frac{1}{2\left(\sigma^2\right)^2}\sum_{i=1}^{S_1}\left(\frac{x-m}{\sqrt{1-\gamma}}\right)^2 + \frac{1}{2\left(\sigma^2\right)^2}\sum_{i=S_1+1}^{S_T}\left(\frac{x-m}{\sqrt{1+\gamma}}\right)^2 = 0$$

$$\widehat{\sigma}^2 = \frac{1}{S_T\left(1-\widehat{\gamma}\right)}\sum_{i=1}^{S_1}\left(x-\widehat{m}\right)^2 + \frac{1}{S_T\left(1+\widehat{\gamma}\right)}\sum_{i=S_1+1}^{S_T}\left(x-\widehat{m}\right)^2 \tag{B2}$$

For the risk parameter we find:

$$\begin{aligned}\frac{\partial}{\partial\gamma}\mathcal{L}(x;\sigma^2,\gamma,m) &= -\frac{S_T/2}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\left(\frac{\sqrt{1-\gamma}-\sqrt{1+\gamma}}{\sqrt{1+\gamma}\sqrt{1-\gamma}}\right) \\ &\quad -\frac{1}{2\sigma^2\left(1-\gamma\right)^2}\sum_{i=1}^{S_1}\left(x-m\right)^2 + \frac{1}{2\sigma^2\left(1+\gamma\right)^2}\sum_{i=S_1+1}^{S_T}\left(x-m\right)^2\end{aligned}$$

which collapses to the following equation in the estimators:

$$\frac{\sum_{i=S_1+1}^{S_T}\left(x-\widehat{m}\right)^2}{\left(1+\widehat{\gamma}\right)^2} - \frac{\sum_{i=1}^{S_1}\left(x-\widehat{m}\right)^2}{\left(1-\widehat{\gamma}\right)^2} = \frac{\widehat{\sigma}^2 S_T\sqrt{1+\widehat{\gamma}}\sqrt{1-\widehat{\gamma}}}{1}\left(\frac{\sqrt{1-\widehat{\gamma}}-\sqrt{1+\widehat{\gamma}}}{\sqrt{1-\widehat{\gamma}}+\sqrt{1+\widehat{\gamma}}}\right) \tag{B3}$$

For the mode parameter we have the expression:

$$\frac{\partial}{\partial m}\mathcal{L}(x;\sigma^2,\gamma,m) = \frac{\sum_{i=1}^{S_1}\left(x-m\right)}{\sigma^2\left(1-\gamma\right)^2} + \frac{\sum_{i=S_1+1}^{S_T}\left(x-m\right)}{\sigma^2\left(1+\gamma\right)^2} = 0$$

$$\frac{\sum_{i=1}^{S_1} x - \sum_{i=1}^{S_1} m}{(1-\gamma)^2} + \frac{\sum_{i=S_1+1}^{S_T} x - \sum_{i=S_1+1}^{S_T} m}{(1+\gamma)^2} = 0$$

which is simplified as:

$$\frac{\sum_{i=1}^{S_1} x}{(1-\widehat{\gamma})^2} + \frac{\sum_{i=S_1+1}^{S_T} x}{(1+\widehat{\gamma})^2} = \left[\frac{S_1}{(1-\widehat{\gamma})^2} + \frac{S_T - S_1}{(1+\widehat{\gamma})^2}\right] \widehat{m} \tag{B4}$$

Equations [B2], [B3] and [B4] are solved to find the triple of MLE parameters $\widehat{\Lambda} = \left(\widehat{m}, \widehat{\sigma}^2, \widehat{\gamma}\right)$.

C The Posterior distribution

C.1 The joint posterior

The joint posterior distribution is given by:

$$\begin{aligned} p\left(\Lambda \vert \ \{\pi\}, \Omega\right) \ \propto \ & \left(\frac{\gamma+1}{2}\right)^{c-1} \left(\frac{1-\gamma}{2}\right)^{d-1} \left(\sigma^2\right)^{\frac{-(a+2)}{2}} e^{\left(\frac{-b}{2\sigma^2}\right)} \\ & \left(\frac{\left(\sigma^2\right)^{-\frac{1}{2}}}{\sqrt{1-\gamma} + \sqrt{1+\gamma}}\right)^{S_T} e^{\left(\frac{-1}{2}\left\{\sum_{i=1}^{S_1}\left(\frac{\pi_i - m}{\sqrt{\sigma^2(1-\gamma)}}\right)^2 + \sum_{i=S_1+1}^{S^*}\left(\frac{\pi_i - m}{\sqrt{\sigma^2(1+\gamma)}}\right)^2\right\}\right)} \end{aligned}$$

In the main text we have determined the conditional posterior distribution kernel of σ^2 by fixing the other two parameters:

$$p\left(\sigma^2 \vert \ \gamma, m, \{\pi_H\}, \Omega\right) \propto \left(\sigma^2\right)^{\frac{-\left(a+S^*+2\right)}{2}} e^{\left(-\frac{\vartheta(m,\gamma;S^*)+b}{2\sigma^2}\right)} \tag{C1}$$

where $\vartheta(m, \gamma; S^*) = \left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_i - m)^2}{1-\gamma}\right) + \sum_{i=1+S_1}^{S^*}\left(\frac{(\pi_i - m)^2}{1+\gamma}\right)\right\}$

The implied posterior distribution of σ^2 is also a iG2 distribution with parameters: $(\vartheta(m, \gamma; S^*) + b, a + S^*)$. From here, it is straightforward to determine the mean of σ^2 under the conditional posterior:

$$E\left(\sigma^2 \vert \ .\right)_{post} = \frac{\vartheta(m, \gamma; S^*) + b}{a + S^* - 2}$$

On the other hand, the prior mean was given by:

$$E\left(\sigma^2 \middle|\, .\right)_{prior} = \frac{b}{a-2}$$

While the fitted estimation with simulated data according to equation [B2] gives:

$$\widehat{\sigma}^2\Big|_{\ fit} = \frac{\vartheta(m,\gamma;S^*)}{S^*}$$

Proposition 3. *If* $E\left(\sigma^2\middle|\, .\right)_{prior} > \widehat{\sigma}^2\Big|_{\ fit}$ *, then* $E\left(\sigma^2\middle|\, .\right)_{prior} > E\left(\sigma^2\middle|\, .\right)_{post} > \widehat{\sigma}^2\Big|_{\ fit}$

Proof. Starting with the conditional: $\frac{b}{a-2} > \frac{\vartheta(m,\gamma;S)}{S}$:
(a) we post multiply and add the term $b(a-2)$ in both sides:
$bS + b(a-2) > (a-2)\,\vartheta(m,\gamma;S) + b(a-2)$
$b\,(a+S-2) > (a-2)\,(\vartheta(m,\gamma;S)+b)$
$\frac{b}{a-2} > \frac{\vartheta(m,\gamma;S)+b}{a+S-2}$
(b) we post multiply and add the term $\vartheta(m,\gamma;S)S$ in both sides:
$bS + \vartheta(m,\gamma;S)S > (a-2)\,\vartheta(m,\gamma;S) + \vartheta(m,\gamma;S)S$
$S\,(b+\vartheta(m,\gamma;S)) > \vartheta(m,\gamma;S)\,(a-2+S)$
$\frac{b+\vartheta(m,\gamma;S)}{a-2+S} > \frac{\vartheta(m,\gamma;S)}{S}$ ∎

The basic result when $E\left(\sigma^2\middle|\, .\right)_{prior} > \widehat{\sigma}^2\Big|_{\ fit}$ is:
$\frac{b}{a-2} > \frac{b+\vartheta(m,\gamma;S^*)}{a-2+S^*} > \frac{\vartheta(m,\gamma;S^*)}{S^*}$
As the simulated sample becomes large, the procedure implemented here downweights the prior; and thus the simulated variance does not differ from the posterior.

The other two relevant conditional distributions are given by:

$$p\left(m\middle|\, \gamma,\sigma^2,\pi_{t+H},\Omega\right) \propto e^{\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{(1-\gamma)}\right)+\sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{(1+\gamma)}\right)\right\}\right)} \tag{C2}$$

and

$$p\left(\gamma|m\ ,\sigma^2,\pi_{t+H},\Omega\right) \ \propto\ \left(\frac{\gamma+1}{2}\right)^{c-1}\left(\frac{1-\gamma}{2}\right)^{d-1}\left(\frac{2}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right)^{S} e^{\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{1-\gamma}\right)+\sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{1+\gamma}\right)\right\}\right)} \tag{C3}$$

C.2 Sampling from the posterior

In order to make inferences about the posterior distribution of the parameters, it is necessary to obtain samples from the three posterior distributions. The posterior distribution of σ^2 is an inverted gamma-2 (equation [C1]) and thus, poses no problem. However, the other two kernels (equations [C2] and [15)] are of unknown form. This calls for a sampling procedure commonly known as Metropolis-Hastings within Gibbs sampling:

The sampling algorithm takes the following steps:

1. Initialize the parameters at an arbitrary value $\left(m_0, \sigma_0^2, \gamma_0\right)$.
2. Generate a k_{th}-draw $\sigma_k^2 \sim p\left(\sigma_{k-1}^2 | \; \gamma_k, m_k, .\right)$
3. Metropolis step to get m update:

   Consider the function from equation [C2]:

$$c_m(m; \sigma^2, \gamma) = e^{\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{(1-\gamma)}\right)+\sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{(1+\gamma)}\right)\right\}\right)}$$

   (a) Calculate a function value: $M_{k-1} = c_m(m_{k-1}; \sigma_k^2, \gamma_{k-1})$
   (b) Generate a candidate draw from: $m_k^* \sim m_{k-1} + cN(0, 1)$; where c is an appropriate constant.
   (c) Calculate the corresponding function value: $M_k = c_m(m_k^*; \sigma_k^2, \gamma_{k-1})$
   (d) Calculate the ratio: $\rho = \min(\frac{M_k}{M_{k-1}}, 1)$
   (e) Draw a uniform random variable between zero and one $\rho_u = Uniform(0, 1)$
   (f) if $\rho_u < \rho$, make the candidate m_k^* draw be the selected draw m_k. Otherwise go back to [a.] and repeat the procedure.

4. Metropolis step to get γ update: Considering the function from equation [C2]:

$$c_\gamma(\gamma; \sigma^2, m) \;\; \propto \;\; \left(\frac{\gamma+1}{2}\right)^{c-1}\left(\frac{1-\gamma}{2}\right)^{d-1}\left(\frac{2}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right)^{S} e^{\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{1-\gamma}\right)+\sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{1+\gamma}\right)\right\}\right)}$$

   And repeat [a.] to [f.] as in Step 3.

After a number of draws, the sampling scheme is equivalent to sampling from the true posterior distributions outlined above. In the example developed in the paper, the number of total draws amounts to 50,000 from which, the first 5,000 were excluded.

D The optimal design of the sample size

As stated in the main text, the optimal sample size design maximizes the expected utility:

$$S^* = \arg\max_{S \in \mathcal{D}} \{KL(S) - \lambda S\} \tag{D1}$$

Where the KL divergence number is defined as:

$$KL(S) = \int_{\Lambda} \int_{\Pi} \log\left[\frac{p(\Lambda|\Pi, S)}{p(\Lambda)}\right] p(\Pi, \Lambda|S) d\Pi d\Lambda$$

Where $\Pi = \{\pi\}_{j=1}^{S}$ is the simulated inflation data of size S, $p(\Lambda)$ is the prior distribution of the parameters and $p(\Lambda|\Pi, S)$ is the posterior distribution.
Following Ryan (2003), it is straightforward to show that the KL information number is

$$KL(S) = \int \int \log\left[p(\Pi|\Lambda, S)\right] p(\Pi, \Lambda|S) d\Pi d\Lambda - \int \log\left[p(\Pi|S)\right] p(\Pi|S) d\Pi$$

Hence, this number can be estimated by a MCMC procedure that does not rely in sampling from the posterior distribution of the parameters. The estimator is:

$$\widehat{KL}(S) = \frac{1}{N} \sum_{i=1}^{N} \{\log[p(\Pi_i|\Lambda_i, S)] - \log[\widehat{p}(\Pi_i|S)]\} \tag{D2}$$

Where (Π_i, Λ_i) for $i = 1, ..., N$ is a sample from $p(\Pi, \Lambda|S)$ and $\widehat{p}(\Pi_i|S)$ is an estimator of the marginal density of the data $p(\Pi_i|S)$. The dependent pair $(\Pi_i,\ \Lambda_i)$ drawn from $p(\Pi, \Lambda|S) = p(\Pi|\Lambda, S)p(\Lambda)$, is obtained by first drawing Λ_i from the prior distribution $p(\Lambda)$ and then Π_i from the conditional distribution $p(\Pi|\Lambda_i, S)$.
The estimation of the marginal density of the data is obtained by an importance sampling based estimator as in Ryan (2003):

$$\widehat{p}(\Pi_i|S) = \frac{1}{M} \sum_{j=1}^{M} p(\Pi_i|\Lambda^*_{ij}, S) \tag{D3}$$

Where $\{\Lambda^*_{ij}\}$ for $i = 1, ..., N$ and $j = 1, ..., M$ are N samples of size M drawn from the prior $p(\Lambda)$ obtained independently of the N pairs (Π_i, Λ_i) drawn before.

The sampling algorithm to get the estimator [D2] follows exactly that of Ryan (2003)

1. Generate a large sample of size N_Λ from $p(\Lambda)$, $\{\Lambda, ..., \Lambda_{N_\Lambda}\}$.

2. Generate an index set for MCMC estimator [D2] as a size $N \leq N_\Lambda$ random sample without repetition of the integers 1 to N_Λ. Call this sample $\{out_i\}_{i=1}^N$

3. Generate index sets for importance sampling estimator [D3] as N independent size $N \leq N_\Lambda$ random samples without repetition of the integers 1 to N_Λ. Call these samples $\{in_{ij}\}_{j=1}^M$ for $i = 1, ..., N$.

4. For $k = 1, ..., n_d$, let S_k represent n_d designs to be compared. Generate one dataset Π_{ki} from $p(\Pi|\Lambda_{out_i}, S_k)$ for each $k = 1, ..., n_d$ and each $i = 1, ..., N$.

5. For $k = 1, ..., nd$, compute

$$\widehat{KL}^M(S_k) = \frac{1}{N}\sum_{j=1}^{N}\widehat{KL}_i^M(S_k) \tag{D4}$$

where

$$\widehat{KL}_i^M(S_k) = \log[p(\Pi_i|\Lambda_{out_i}, S_k)] - \log\left[\frac{1}{M}\sum_{j=1}^{M}p(\Pi_i|\Lambda_{ij}^*, S)\right]$$

To implement the estimation, we considered the following values: $N_\Lambda = 5000$, $N = 1000$, $M = 100$, and $n_d = 200$. Also,we consider sample size higher than 30 via: $S_k = (k-1) + 30$.

In figure [6], we depict the MCMC draws of KL together with a smoothed version of it. The smoothed version is combined with the loss term in [D1] to get the utility function shown in figure [7].

E Figures

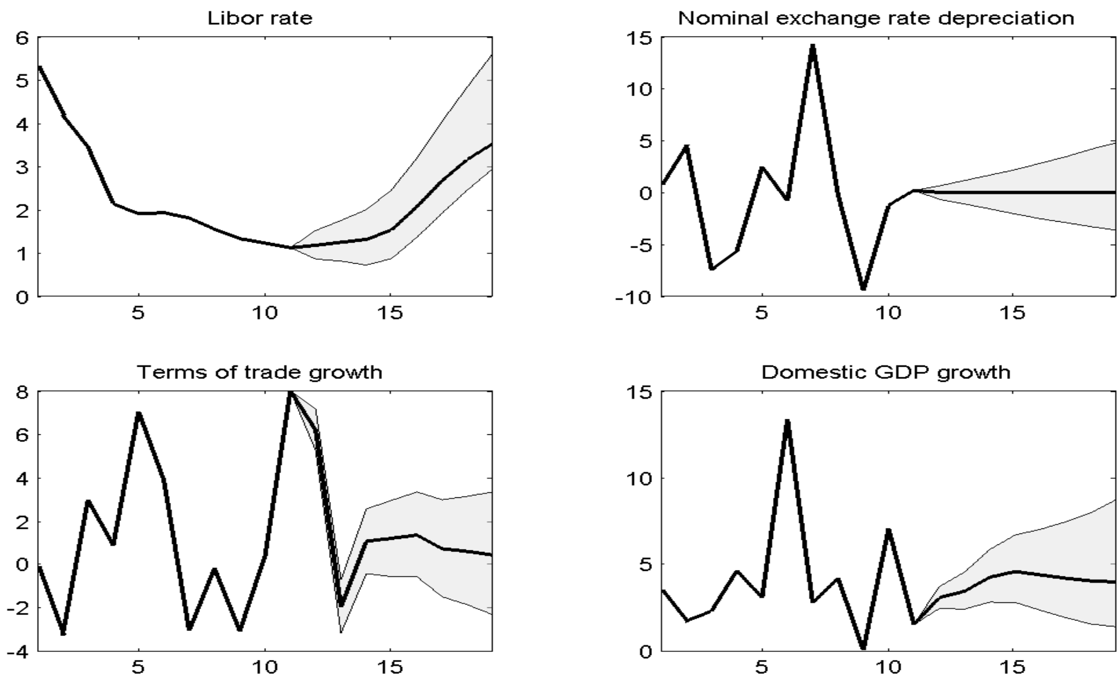

Figure 1: 90 % forecast interval and modal forecast

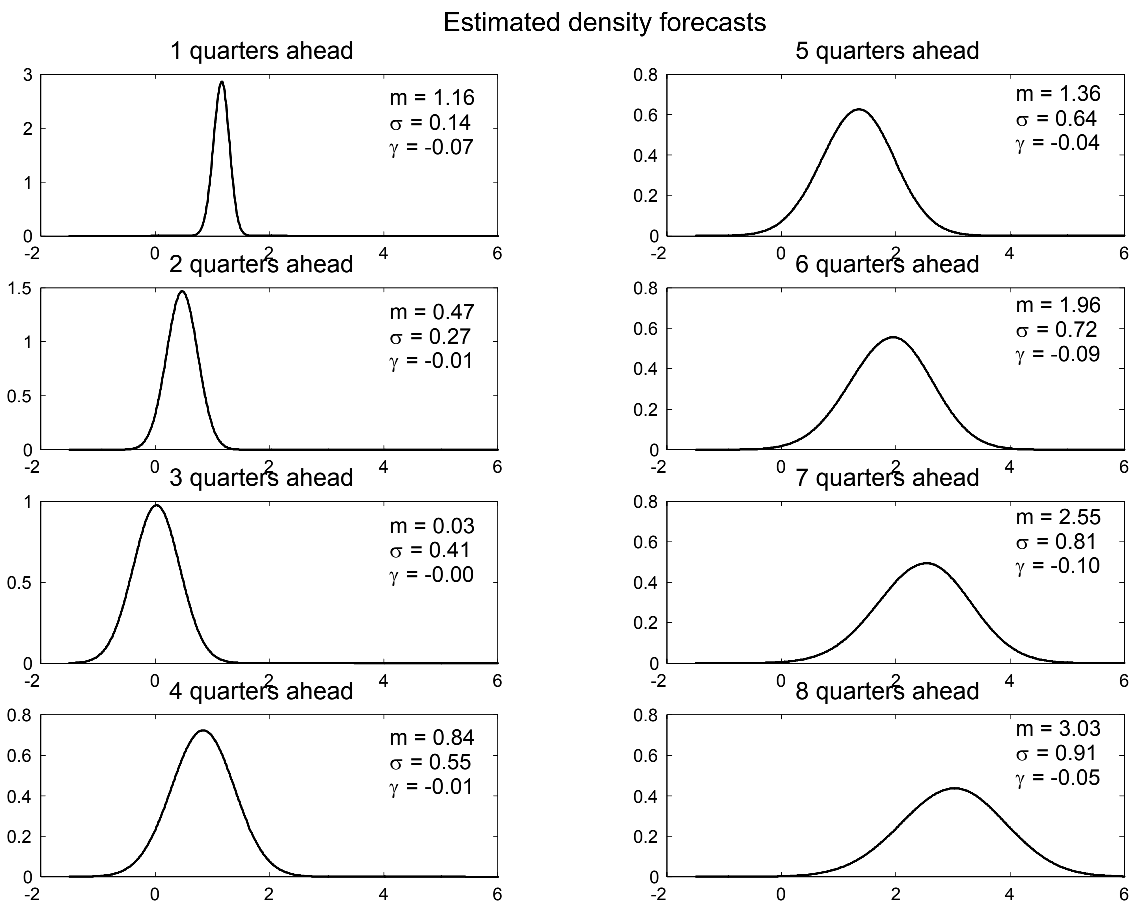

Figure 2: Estimated SN pdf's for the year-on-year inflation forecast

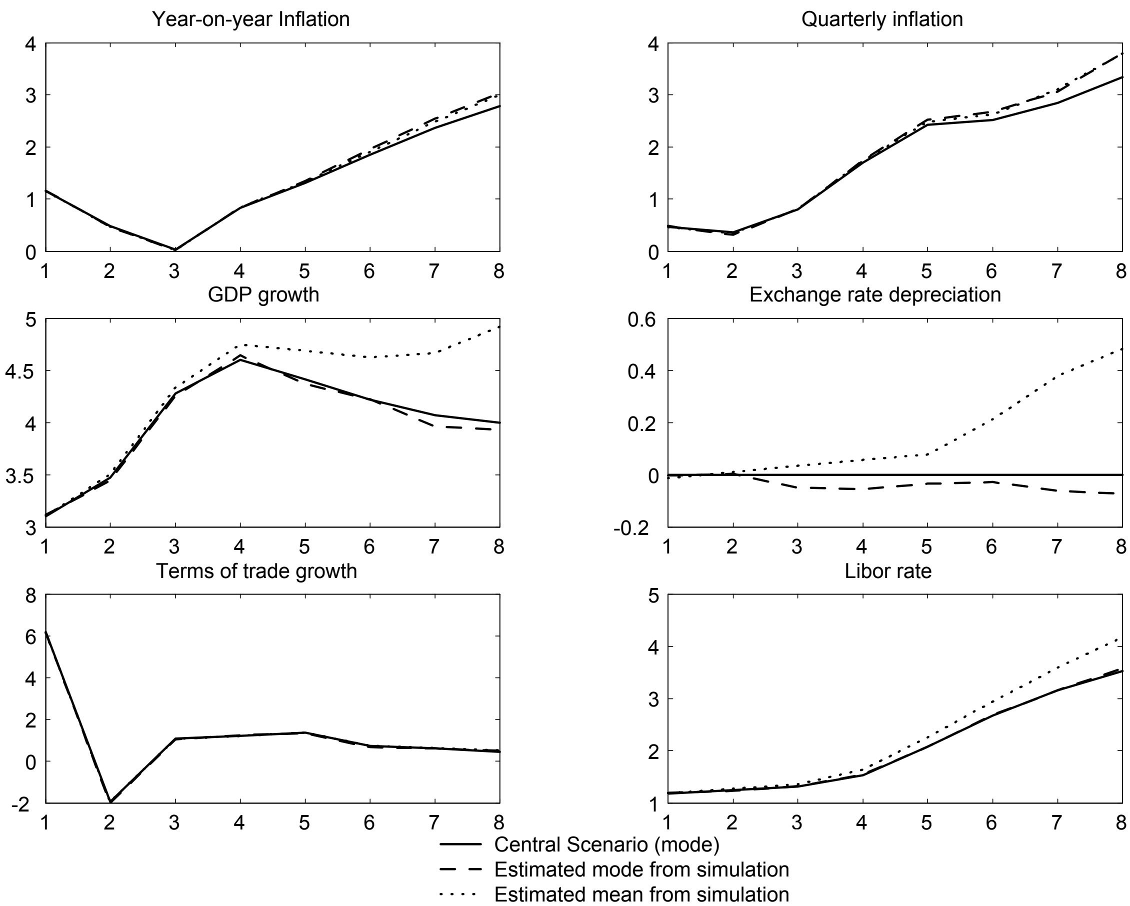

Figure 3: Central measures of tendency

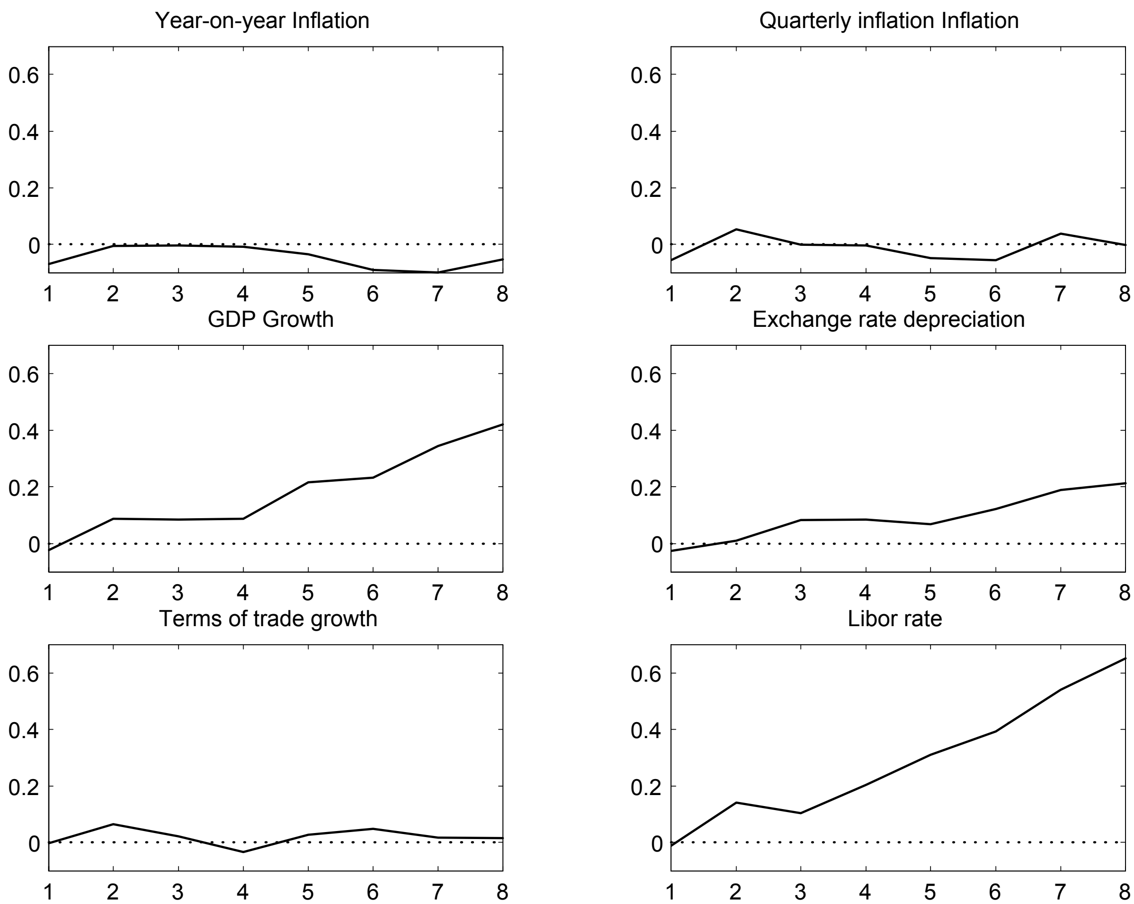

Figure 4: Evolution of the gamma parameter

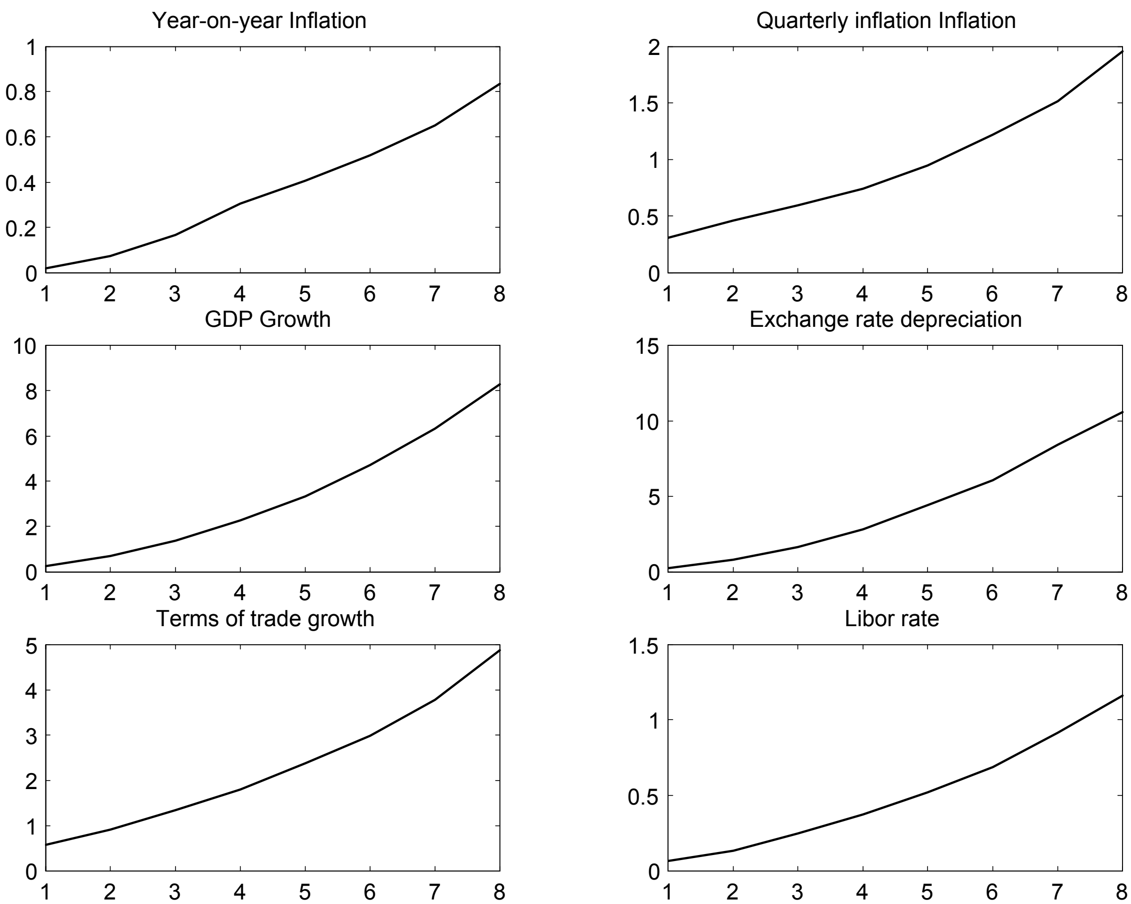

Figure 5: Evolution of the uncertainty parameter

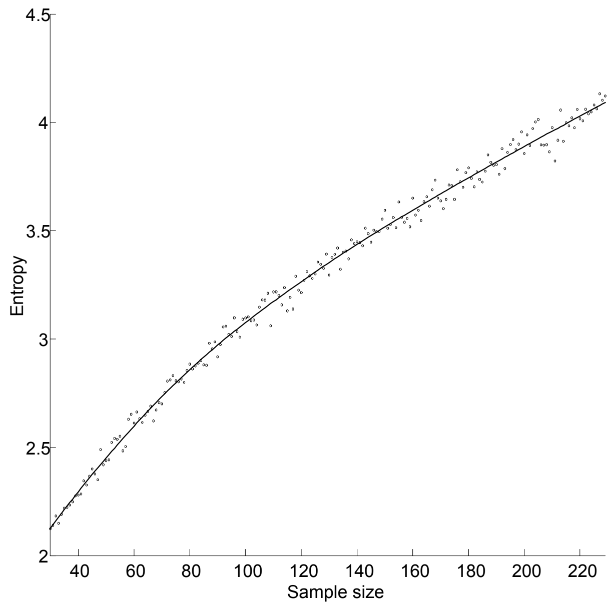

Figure 6: The KL divergence number (a.k.a entropy). The scatter plot is the estimation with monte carlo variation, the line is the smoothed version

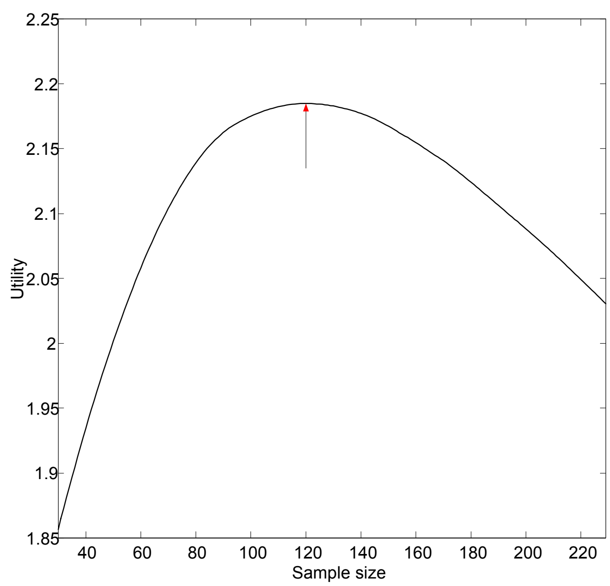

Figure 7: The utility function of the policymaker as a function of the simulation sample size

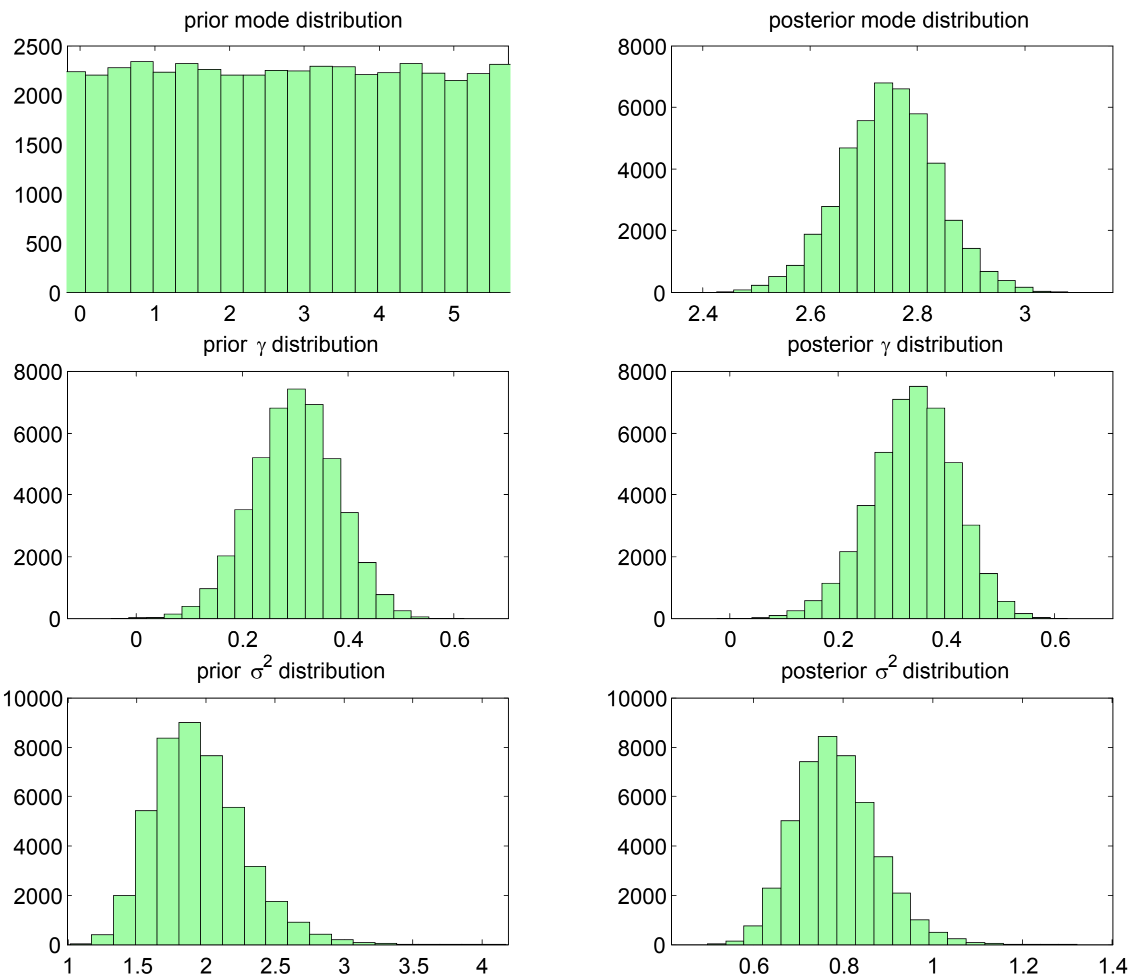

Figure 8: The utility function of the policymaker as a function of the simulation sample size