Policymakers Priors and Inflation Density Forecasting*

Marco Vega
Banco Central de Reserva del Perú

Resumen

Este trabajo modela un esquema de proyección de densidad de inflación que se aproxima a la conducta de las autoridades de política monetaria en referencia a la determinación de parámetros como el punto modal, la incertidumbre y la asimetría de la densidad de la proyección de inflación.

El esquema combina la información a priori que manejan las autoridades de política sobre los parámetros en cuestión con aquella obtenida de técnicas paramétricas de estimación de densidad usuales utilizando teoría bayesiana. La combinación se basa en las ganancias informativas de las autoridades de política que obtienen a partir de los ejercicios de proyección. La evaluación se realiza por medio de la teoría de información.

Palabras clave: Política monetaria, Metas de inflación, Métodos bayesianos.
Códigos JEL: C53, E37, E58

Abstract

This paper models an inflation density forecast framework that closely resembles policymakers' actual behavior regarding the determination of the modal point, the uncertainty and asymmetry in inflation forecasts.

The framework combines the prior information about these parameters available to policymakers with a standard parametric density estimation technique using Bayesian theory. The combination crucially hinges on an information-theoretic utility function gain for the policymaker from performing the forecast exercise.

Keywords: Monetary Policy, Inflation Targeting, Bayesian Methods.
JEL Codes: C53, E37, E58

INTRODUCTION

The purpose of this paper is to build a methodology to obtain marginal inflation density forecasts. The approach used involves estimating a parametric inflation density forecast where uncertainty, asymmetry and central tendency profiles are derived mainly from the exogenous variables through the use of a forecasting model.

The estimated parameters are combined with policymakers' prior views through an explicit Bayesian approach. The prior views encompass all other factors of risk and uncertainty that may strike at 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 environments where macroeconometric formulation of models is hindered by measurement errors and poor data availability[1]. Nevertheless, even in stable and developed countries with quality data rich environments, prior inputs are essential.

The chapter proceeds as follows: Section 2 outlines the density forecast framework, Section 3 illustrates the methodology with a simple example for forecasting Peruvian inflation and finally Section 4 draws conclusions. The appendix contains technical derivations.

1. DENSITY FORECAST FRAMEWORK

The forecasting literature has recently turned the focus of its attention away from point forecasts towards density forecasts[2]. The reasons for the need to provide complete representations of probability distribution lie in the shortcomings of the certainty equivalence principle in a world characterized by asymmetric risks. This is particularly relevant in the fields of financial risk management and modern monetary policy where decision theory plays a substantial role.

Some central banks. like the US Federal Reserve or the Bank of England, have a long tradition in producing macroeconomic point forecasts. Only recently, the Bank of England pioneered the presentation of density forecasts by means of fan charts. Since then, a

---

1.   This is the case in most emerging-market economies.
2.   See Diebold et.al (1999) and Tay and Wallis (2000)

number of inflation targeters (ITers) have published density forecasts with varying degrees of detail. Twelve of 21 ITers regularly publish such fan charts[3].

Leading density forecast central banks[4] have favored the use of specific parametric methods to construct their density forecasts. The parameters governing forecast densities directly control for uncertainty and the asymmetry of the distribution. This is the approach taken in the next subsection.

The role played by models in forecasting has been recognized by academics and practitioners alike. In a recent survey of central banks practicing IT (Schmidt-Hebbel and Tapia (2002)), all 20 of the banks surveyed refer to the use of some kind of model. The key point here is that most central banks, especially ITers, endorse the use of a core forecasting model in helping to center policy discussions within the bank.

However, the use of models in forecasting does not mean that subjectivity is filtered out in the forecasting process. A point mentioned in the Schmidt-Hebbel and Tapia (2002) survey is that in most central banks published forecasts are a "balanced combination" of technical forecasts and the views of decision makers. The inclusion of subjective approaches to macroeconomic forecasting within central banks is also recognized in Sims (2002) and Goodhart (2001).

Papers like those by Hall and Mitchell (2004a, 2004b, 2005) suggest a powerful method for forecast combinations using in ways that incorporate subjective forecasts. This combination hinges on forecast error minimization. Instead, this paper proposes a methodology based on the interaction between the policymaker and those producing the forecasts. This involves central bank staff undertaking simulations using a forecasting model, with policymakers inputting priors with regard to parameters that reflect uncertainty, the risk balance, and baseline forecast values.

1.1 The parametric density forecast

The economists at a central bank implement a forecasting process at time t with respect to an inflation sequence up to horizon H. This sequence is generated by a forecasting model and is denoted by[5] $\{ \hat{\pi}_s \}_{s=t+1}^{H}$

---

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 (2003), Israel appears as not publishing a fan chart because, exceptionally, the inflation report under assessment lacked one. Colombia is not considered in their sample due to the «limited information» available.

4. For the Bank of England, the references are Britton et al. (1998) and Wallis (2003). For the Riksbank, the reference is Blix and Sellin (1998).

5. Hatted variables are forecasts of either exogenous or endogenous variables. In the case of the instrument setting, it refers to the stance assumed by the policy maker.

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

In equation [1], Y_t denotes the known history of endogenous macroeconomic variables y_t in the model (including inflation π_t). Formally, $Y_t = \{y_t, ..., y_{t-n}\}$

This model-based forecast is conditional upon various factors that can be assumed or induced in the process. These factors are X_t, θ, and I_t. The first denotes the history and likely future course of the exogenous variables: $X_t = \{x_{t-n}, ..., x_t, \hat{\pi}_{t+1} ... \hat{\pi}_{t+H} ...\}$, θ denotes the set of parameters that describes 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: $I_t = \{i_{t-n}, ..., i_t, i_{t+1}, ...\}$

Model M is sufficiently general and need not be explicit as it may correspond to a rational expectations equilibrium solution. I make the following definition:

Definition 1 A central forecast[6] is an inflation sequence $\{\hat{\pi}_{c,s}\}^{H}_{s=t+1}$ obtained by conditioning the model to: (a) the most likely sequence of exogenous variables within the forecast horizon $\{x_{c,s}\}^{H}_{s=t+1}$, (b) parameter values θ_c and (c) the monetary policy instrument setting $I_{c,t}$

Also, the central bank economists involved provide a technical assessment of risk and uncertainty about the inflation forecast. This 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 historical estimates made by technical staff as well as the subjective but informed view of sectoral experts.

Among the various probability density functions that are suitable to perform random draws are the Beta and the Split Normal. The latter is used intensively by Blix and Sellin (1998), Britton et al. (1998) and Vega (2003). These two types of distributions are useful because their parameters illustrate the distributional characteristics that matter most in a density forecast, a central point, a measure of dispersion and skewness.

By performing simulated histories of exogenous variables within the forecast horizon one can determine alternative trajectories of inflation. Evaluated at each point in time within

---

6. In this definition, the subscript c denotes both central forecasts or assumed central values.

the forecast horizon, the different inflation points originated in the simulations can be hypothesized as coming from a generic probability function. Determining the explicit form of the inflation forecast probability distribution function (pdf) that results from this exercise is hindered by two obstacles: (a) the mapping from the exogenous variables to inflation implies a solution like equation [1], which can be highly non-linear; and (b) even if we can establish the exact form of the distribution, communicating it to the policymakers would not be easy. One way to circumvent the problem is to assume a parametric form for the distribution function that serves two purposes by being at once a good approximation to the true *pdf* while providing a communication strategy easily grasped by the policymaker. The Split Normal distribution is a good candidate for the assumed *pdf*, since its parameters are easily communicated in terms of a straightforward balance of risks.

Definition 2 A model-based parametric inflation density forecast is a sequence of parameters $\left\{ \widehat{\Lambda}_{c,s} \right\}_{s=t+1}^{H}$ that describes a probability density function of the inflation forecast at every point in time s. These can be obtained by a likelihood estimation procedure assuming the Split Normal distribution and using the simulated data.

Henceforth, I concentrate on a relevant horizon H dropping time subscripts. After S number of stochastic simulations on the exogenous variables are performed, I obtain a mapping from data conditional on the model parameters and the instrument setting to object ω.

$$(\{X_t\}_{j=1}^{S}, Y_t, \Theta, I_t) \rightarrow \omega \qquad [2]$$

The variable ω contains the elements which both the econometrist and the policymaker care about[7]: the inflation forecast at horizon H, and the three parameters that underlie the policy discussions. I group these three parameters in the vector $\Lambda = (m, \sigma^2, \gamma)$ with γ as the modal point[8], σ^2 the uncertainty measure and γ the skewness of the distribution of the inflation forecast. These three parameters precisely define the Split Normal $SN(m, \sigma^2, \gamma)$. This distribution collapses into a Normal $N(m, \sigma^2)$ whenever the skewness parameter equals zero.

The parameter γ varies on the range $\langle -1,1 \rangle$ and is closely linked to the balance of risks made at central banks (see Appendix B). Specifying ω in a compact way

$$\omega = (\{\pi\}_{j=1}^{S}, \Lambda) \qquad [3]$$

---

7. Observe that the parameter Θ as well as the instrument may remain constant or vary exogenously along the simulations.
8. When risks are asymmetric, there are three measures of tendency that central banks can look at. In practice, central banks tend to pay more attention to modal points. See Goodhart (2001) and Vega (2003).

We treat ω parameters in a Bayesian context[9] and characterize its posterior probability density conditional on all the information acquired after performing S simulations of the model conditional on all the given factors Ω (note that S itself is a conditioning factor)

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

where

Ω is the given information set $\Omega = \left\{ \{X_t\}_{j=1}^{S}, Y_t, \Theta, I_t \right\}$

$p(\Lambda \mid \Omega)$ is the prior density elicited by the policymaker, and $p(\{\pi\}_{j=1}^{S} \mid \Lambda, \Omega)$ is the probability of the simulated inflation forecast data given the information 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: $L(\Lambda \mid \{\pi\}_{j=1}^{S}, \Omega)$.

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

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

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

1.2 Elicitation of the priors as the outcome of policymakers views

Through the outcome of the model-based density forecast, policymakers form their views. These take into account other forms of uncertainty not included in the forecast, including model-uncertainty, measurement errors and others. It is therefore an internal operational task how best to extract these views and to translate them into tractable distribution functions.

For our purposes, we assume that the first subjective view is that the three parameters are independent random variables, so that the joint prior is

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

---

9. Namely, it is itself a random variable.

a.   Prior for uncertainty parameter σ^2

Following the literature (Bauwens et.al (1999)), we assume that σ^2 is driven by the Inverted Gamma-2 distribution $iG_2(a,b)$. The parameters (a,b) are extracted from the policymaker. This distribution has support $<0,\infty>$ and its parameters can be specified using the two moments and the mode of the distribution as guidelines:

$$E\left(\sigma^2|.\right) = \frac{b}{a-2} \quad \text{for } a > 2$$

and:

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

while the mode is:

$$\text{mode}\left(\sigma^2|.\right) = \frac{b}{a+2}$$

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

b.   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, calling this $\tilde{B}(c,d)$. 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 $\langle 0,1\rangle$ with a Beta distribution $B(c,d)$ (the transformation applied is $\gamma = 2z-1$). The first two moments are defined as:

$$E\left(\gamma|.\right) = \frac{c-d}{c+d}$$

and

$$V\left(\gamma|.\right) = \frac{4cd}{(c+d-1)}\left(\frac{1}{c+d}\right)^2$$

with mode

$$\text{mode}\left(\gamma|.\right) = \frac{c-d}{c+d-2}$$

c. Prior for mode parameter m

We impose a non-informative uniform distribution for the mode

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

1.3 The posterior distribution

Given the Split Normal likelihood assumption[10], the kernel of the joint posterior distribution of the three parameters of interest is[11]

$$p\left(\Lambda \mid \{\pi_i\}, \Omega\right) \propto \left(\tfrac{\gamma+1}{2}\right)^{c-1} \left(\tfrac{1-\gamma}{2}\right)^{d-1} \left(\sigma^2\right)^{\frac{-(a+2)}{2}} \exp\left(\tfrac{-b}{2\sigma^2}\right) \cdots$$

$$\left(\frac{\left(\sigma^2\right)^{\frac{-1}{2}}}{\sqrt{1-\gamma} + \sqrt{1+\gamma}}\right)^N \exp\left(\frac{-1}{2}\left\{\sum_{i=1}^{S_1}\left(\frac{\pi_i - m}{\sigma\sqrt{1-\gamma}}\right)^2 + \sum_{i=S_1+1}^{S}\left(\frac{\pi_i - m}{\sigma\sqrt{1+\gamma}}\right)^2\right\}\right) \tag{8}$$

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

$$p\left(\sigma^2 \mid \gamma, m, \{\pi\}, \Omega = \right) \left(\sigma^2\right)^{\frac{-(a+N+2)}{2}} \exp\left(\frac{-(\vartheta(m,\gamma)+b)}{2\sigma^2}\right) \tag{9}$$

where

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

The other two relevant conditional distributions are given by:

$$p\left(m \mid \gamma, \sigma^2, \{\pi\}, \Omega\right) = \exp\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{\pi_i-m}{1-\gamma}\right)^2 + \sum_{i=S_1+1}^{S}\left(\frac{\pi_i-m}{1+\gamma}\right)^2\right\}\right) \tag{10}$$

and

$$p\left(\gamma \mid m, \sigma^2, \{\pi\}, \Omega\right) = \left(\tfrac{\gamma+1}{2}\right)^{c-1} \left(\tfrac{1-\gamma}{2}\right)^{d-1} \left(\frac{2}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right)^S \cdots \tag{11}$$

$$\exp\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{\pi_i-m}{1-\gamma}\right)^2 + \sum_{i=S_1+1}^{S}\left(\frac{\pi_i-m}{1+\gamma}\right)^2\right\}\right)$$

---

10. See appendix [B] for details about this distribution.
11. Inflation data refers to simulated forecasts at a fixed horizon $t+H$.

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

$$E\left(\sigma^2 \mid .\right) = \frac{\frac{a+S}{2}}{\frac{2}{\vartheta(m,\gamma)+b} - 2} \quad \text{for} \quad \frac{2}{\vartheta(m,\gamma)+b} > 2$$

and

$$V\left(\sigma^2 \mid .\right) = \frac{2}{\frac{2}{\vartheta(m,\gamma)+b} - 4}\left(\frac{\frac{a+s}{2}}{\frac{2}{\vartheta(m,\gamma)+b} - 2}\right)^2 \quad \text{for} \quad \frac{2}{\vartheta(m,\gamma)+b} > 4$$

while the mode is:

$$\text{mode}\left(\sigma^2 \mid .\right) = \frac{\frac{a+S}{2}}{\frac{2}{\vartheta(m,\gamma)+b} + 2}$$

From this explicit representation, we observe that as the sample size increases, the posterior mean and mode would collapse to the model-based estimates. In that case, the prior view has a small effect on the posterior outcome. In econometric estimations a larger sample size is always good because it improves the model-based information. The context here is rather different; it is based on the willingness of a Bayesian policymaker to learn about the properties of the inflation forecast from a general perspective as opposed to a non-Bayesian econometrist who wants to learn the properties of the model-based forecast.

1.4 The choice of sample size as an information theoretic design problem
In the proposed methodology, the sample size S is a choice variable as well. If a large enough sample size is considered, the prior view of the policymakers becomes useless. On the other hand, if the sample size is small, then the model-based estimation becomes less accurate and the simulation experiment therefore suffers from being informationally poor.

Policymakers need to weigh the information provided by the model against the prior beliefs they may hold. In practice, this may seem complex as it is tied to the subjective beliefs of the policymakers coupled to the out-of-model information that they may possess.

In such circumstances, the information-theoretic approach[12] common in the field of «experimental design» seems plausible. What does the experiment the policymaker performs consist of? In our view, it is in updating the policymaker's prior beliefs about the

---

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 Müller and Parmigiani (1996). Most of these applications pertain to the design of clinical experiments.

inflation forecast modal point, uncertainty and risks using a forecasting model provided by econometricians. The outcome of this updating process depends crucially on the simulation sample size under evaluation.

The choice of sample size S is made so that policymakers maximize their expected utility resulting from the experiment. In other words:

$$S^{*} = \arg_{S} \max \left\{ KL\left(S\right) - \lambda S \right\} \qquad [12]$$

This expected utility of experimentation with sample size S depends on two factors: a) the Kullback-Leibler (KL) divergence between the posterior and prior distribution of the parameters $KL(S)$; and b) the linear loss function λ. The KL number gives the value of the information provided by the forecasting model used[13]. The loss term is rationalized by the unwillingness to disregard their priors[14]. So, as the sample size increases, the prior of the policymaker is downweighted and has reduced utility for the policymaker who considers his/her priors are indeed important. In this case, the utility parameter λ is the degree of importance of the prior in the overall utility function[15]. The KL divergence number is defined as:

$$KL\left(S\right) = \iint\limits_{\Lambda\ \Pi} log\left(\frac{p\left(\Lambda\mid\Pi, S\right)}{p\left(\Lambda\right)}\right) p\left(\Pi, \Lambda\mid S\right) d\Pi d\Lambda \qquad [13]$$

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

2. AN EXAMPLE

In order to provide an example, I use a simple ad-hoc univariate model[16] for quarterly inflation, estimated using ordinary least squares[17]. 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-month Libor rate at lag 3 (i_{t-3}^{*}) and the terms-of-trade growth at lag 4 (Δtot_{t-4}).

13. KL(S) is increasing in S and concave. See Lindley (1956).
14. These priors might indeed not be correct ex post and as studied by Bigio and Vega (2006); they are influenced by their fears and uncertainties about the driving forces in the economy.
15. λ can also be interpreted as the inverse of policy makers' confidence in the model.
16. The univariate model is used only to facilitate exposition. In practice, structural models, such as those developed in Luque and Vega (2003) and by Llosa et.al. (2005) for Peru, should be used.
17. We use data from Peru. The Central Bank of Peru has recently adopted the Inflation Targeting framework (January 2002).

$$\pi_t = 0.69\pi_{t-1} + 0.24\Delta e_{t-1} + 0.23g_{t-2} - 0.30i_{t-1} + 0.55\dot{i}_{t-3} + 0.66\Delta tot_{t-4} + \varepsilon_t \qquad [14]$$

$$(9.23) \qquad (3.58) \qquad (3.06) \qquad (-1.95) \quad (1.72) \qquad (1.70)$$

The estimation[18] is conducted on the basis of data between the first quarter of 1994 and the second quarter of 2003. Except for lagged inflation, all the variables on the right-hand side are considered as exogenous. Hence, to start producing the density forecast we need to construct a baseline scenario and uncertainty and risk profiles for the set of exogenous variables: $(\Delta e_t, g_t, i_t, \dot{i}_t, \Delta tot_t)$. In particular, I assume the following distributions:

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

Exogenous variable	Balance of risk	Distribution	Mode	σ^2
Libor rate	upside 70%	Split normal	3.57	1.2
Nominal exchange rate percentage change	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

In Figure [1] we show the historical, central scenario and the 90 per cent central prediction interval for the exogenous variables over the forecast periods. The asymmetry as well as the uncertainty increases linearly until it reaches the values specified in Table [1]. In each forecast period, I also consider random realizations of the unforecastable shock ε_t, drawn from a 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 about 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 held at 2.75 per cent over the forecast period.

The inflation density forecast is then achieved by estimating the parameters of an assumed split normal distribution $SN(m,s^2,\gamma)$ for the simulated sample of size[19] S_T for each forecast period.

---

18. In equation [14] the lag structure minimizes the sum of squared residuals. As usual, the t-values are in parenthesis.
19. In this step, the sample size S_T should be as large as possible. The objective here is to obtain the most accurate distributional representation originating from the forecasting model alone.

An important conclusion emerges from this exercise: notwithstanding that exchange rate depreciation, GDP growth and the Libor rate all show considerable asymmetry[20] (especially at the end of the forecast horizon), there is no build-up of asymmetry in either the quarterly or year-on-year inflation measures. In Figure [2] we show the estimated densities at each of the eight forecast periods, along with the estimated parameters; m, σ^2, and γ. The gamma parameter is close to zero in all periods.

There are two main reasons why the increasingly asymmetric nature of exogenous variables does not pass on to inflation, namely the lag structure and the interplay between the variability versus asymmetric forces. With respect to 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 regards the relation between 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 (1998)). Inflation variability grows because the exogenous variability increases linearly and because the persistence of inflation (since it depends strongly on its own lags) tends to exacerbate all the sources of inflation uncertainty, even those which come from the inflation shock itself.

The estimated mode from the simulations is quite different from the one calculated using only the central scenario values of exogenous variables (the modes). There is an upward bias (See Figure [3]) in both the quarterly inflation and the year-on-year inflation rates. The reason for this 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 density forecast.

Once the results of the simulation are known, I proceed to introduce the information provided by the policymaker. To do this, I concentrate in forecast horizon $H = 8$. We need to assume a prior distribution for the set of parameters $\Lambda = (m, \sigma^2, \gamma)$. We take the distributional assumptions outlined in subsection 1.2, namely 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 a year-on-year inflation rate of 2.78 percent.

The uncertainty parameter follows an inverted gamma-2 distribution: $\sigma^2 \sim iG_2(a, b)$. In order to find the parameters, we may consider that the estimated $\hat{\sigma}^2$ from the simulation step is too low. Policymakers may consider that there are other factors that necessarily drive

---

20. In Figure [3] in the appendix the estimated mean 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.

forecast uncertainty to a higher level. For example they may assume that $E_{prior}(\sigma^2) = 1.95$ and $mode_{prior}(\sigma^2) = 1.8$. This implies the corresponding parameters $(a,b) = (38,72)$.

The asymmetry parameter follows a beta type of distribution considered in the previous section: $\gamma \sim B(c,d)$. In this case, policymakers 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 us 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 strongly about this asymmetry $V_{prior}(\gamma) = 0.006$. This implies parameter values $(c,d) = (92.857,50)$.

Before combining the prior information given by the policymaker, it is necessary to establish the size of the sample to be used in the Bayesian procedure. This sample size is obtained from solving the problem in equation [12]. Calculating the utility measure requires us to obtain the KL divergence number via some numerical integration procedure. In Appendix D, I 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 degree of confidence in the prior. In this example, we assume $\lambda = 0.01$. Therefore the optimal sample size is $S^* = 120$ (see Figure [7]).

Next, I sample from the Bayesian conditional posterior distributions. The corresponding mean values are shown in Table [2] where the posterior and prior distributions are shown graphically in Figure [8]).

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

		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

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 policymakers 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[21].

CONCLUSION

This paper contributes to the understanding of how central banks conduct forecasts as part of monetary policy making. It focuses attention on Bayesian policymakers who hold or develop prior views on key features of the inflation density forecast. Policymakers interact with technical staff responsible for running the macroeconomic model-based density forecast.

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

In the approach adopted here, policymakers 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 policymakers' priors and the «faith» in the core forecasting model. If the model commands full «faith» then priors are irrelevant and vice-versa.

A further application of the approach developed in this paper would be to reverse engineer this density forecasting process to extract and thus to find a metric on the degree of importance of judgment relative to pure objective model-based forecasts.

---

21. 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.

BIBLIOGRAPHY

BAUWENS, L.; M. LUMBRANO and J-F. RICHARD
1999   «Bayesian Inference in Dynamic Econometric Models». Oxford University Press.

BIGIO, Saki and Marco VEGA
2006   «Monetary Policy under Balance Sheet Uncertainty». Computing in Economics and Finance 2006, 157. Society for Computational Economics.

BLIX, M. and P. SELLIN
1998   «Uncertainty Bands for Inflation Forecasts». Working Paper No. 65, Sveriges Riksbank, Stockholm.

BRITTON, E.; P. FISHER and J. WHITLEY
1998   «The Inflation Report Projections: Understanding the Fan Chart». In: *Bank of England Quarterly Bulletin*, 38, 30-37.

CHALONER, K. and I. VERDINELLI
1995   «Bayesian Experimental Design: A Review». In: *Statistical Science* 10, 273-304.

CLYDE, M.
«Experimental Design: A Bayesian Perspective». In: *International Encyclopedia of Social and Behavioral Sciences*, Vol. 2.1.

DIEBOLD, F.; J. HAHN and A. TAY
1999   «Real-Time Multivariate Density Forecast Evaluation and Calibration: Monitoring the Risk of High-frequency Returns on Foreign Exchange». The Wharton Financial Institutions Center Working Paper Series 99-05.

FRACASSO, A.; H. GENBERG and C. WYPLOSZ
2003   «How do Central Banks Write? An Evaluation of Inflation Targeting Central Banks». Geneva Reports on the World Economy, Special Report 2.

GOODHART, C.A.E.
2001   «The Inflation Forecast». In: *National Institute Economic Review* No. 175.
2004a Hall, S. and J. Mitchell, «Density Forecast Combination», NIESR Discussion Paper 249.
2004b Hall, S. and J. Mitchell, «Optimal Combination of Density Forecasts», NIESR Discussion Paper 248.

HALL, S. and J. MITCHELL
2004   «Evaluating, Comparing and Combining Density Forecasts Using the KLIC with an Application to the Bank of England and NIESR fan charts of inflation». In: *Oxford Bulletin of Economics and Statistics*, Vol. 67, pp. 995-1033.

LINDLEY, D.
1956   «On a Measure of the Information Provided by an Experiment». In: *The Annals of Mathematical Statistics*, Vol. 27, No. 4, 986-1005.
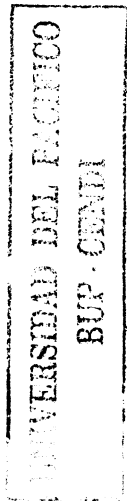
LUQUE, J. and M. VEGA

2003 Usando un modelo semi-estructural de pequeña escala para hacer proyecciones: algunas consideraciones. In: *Revista Estudios Económicos* No. 10, Central Bank of Peru.

LLOSA G.; V. TUESTA and M. VEGA

2004 «A BVAR Forecasting Model for Peruvian Inflation». Central Bank of Peru Working Paper DT No. 2005-007.

MÜLLER, P. And G. PARMIGIANI

1996 «Numerical Evaluation of Information Theoretic Measures». In: D. A. BERRY; K. M. CHALONER, and J. K. GEWEKE (eds.). *Bayesian Analysis in Statistics and Econometrics.* Wiley NY, 397-406.

PARMIGIANI, G. and D. BERRY

1994 «Applications of Lindley Information Measure to the Design of Clinical Experiments». In: Smith, A.F.M and P. Freeman (eds.). *Aspects of Uncertainty: A Tribute to D.V. Lindley.* Wiley NY, 333-352.

RYAN, K.

2003 «Estimating Expected Information Gains for Experimental Designs With Application to the Random Fatigue-Limit Model». In: *Journal of Computational and Graphical Statistics.* Vol. 12, No. 3, 1-19.

SCHMIDT-HEBBEL, K. and M. TAPIA

2002 «Monetary Policy Implementation and Results in Twenty Inflation-Targeting Countries». Central Bank of Chile Working Paper 166.

SIMS, C.

2002 «The Role of Models and Probabilities in the Monetary Policy Process», Mimeo.

TAY, A.S. and K. WALLIS

2000 «Density Forecasting: A Survey». In: *Journal of Forecasting,* 19, 235-254.

VEGA, M.

2003 «Reportando la distribución de la proyección de inflación». In: *Revista Estudios Económicos* No. 10, Central Bank of Peru.

WALLIS, K.

2003a «Chi-squared Tests of Interval and Density Forecasts, and the Bank of England's Fan Charts». In: *International Journal of Forecasting* 19, 165-175.

Inflation Reports published on web pages of those central banks with an inflation targeting regime, namely: 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.

Appendix A: Inflation forecast prior distribution

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)

$$p\left(\sigma^2 \mid .\right) = \left(\Gamma\left(\tfrac{a}{b}\right)\left(\tfrac{2}{b}\right)^{\tfrac{a}{2}}\right)^{-1} \left(\sigma^2\right)^{\tfrac{-(a+2)}{2}} \exp\left(\tfrac{-b}{2\sigma^2}\right) \qquad [\text{A1}]$$

where

$$E\left(\sigma^2 \mid .\right) = \frac{b}{a-2} \quad \text{for } a > 2$$

and

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

while the mode is

$$\text{mode}\left(\sigma^2 \mid .\right) = \frac{b}{a+2}$$

A.2 Prior for γ

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

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

$$E\left(z \mid \Omega\right) = \frac{c}{c+d}$$

and

$$V\left(\gamma \mid \Omega\right) = \frac{4cd}{(c+d-1)\,(c+d)}$$

while mode

$$\text{mode}\left(\gamma \mid .\right) = \frac{c-d}{c+d-2}$$

Then we define γ in terms of the following transformation

$$\gamma = 2z - 1$$

Hence, the prior distribution of γ can be expressed as

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

As a result, the prior distribution of γ is

$$p(\gamma \mid c, d) = \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} \left(\frac{1+\gamma}{2} \right)^{c-1} \left(\frac{1-\gamma}{2} \right)^{d-1} \quad \text{for} \ -1 < \gamma < 1 \qquad [A2]$$

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 since I require sampling from the priors. Hence, I assume $m \sim Uniform\left(m_{low}, m_{high} \right)$

$$p(m \mid .) = \frac{1}{m_{high} - m_{low}} \quad \text{for} \ m_{low} < m < m_{high} \qquad [A3]$$

Appendix B: Model-based density simulation and estimation

B.1   Fitting the simulated data

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

$$f\left(x \mid m, \sigma^2 \gamma \right) = \begin{cases} \frac{2}{\sqrt{\sigma^2}\left(\sqrt{1-\gamma} - \sqrt{1+\gamma}\right)} \phi\left(\frac{x-m}{\sqrt{\sigma}(1-\gamma)} \right) & \text{if } x < m \\[4mm] \frac{2}{\sqrt{\sigma^2}\left(\sqrt{1-\gamma} - \sqrt{1+\gamma}\right)} \phi\left(\frac{x-m}{\sqrt{\sigma}(1+\gamma)} \right) & \text{otherwise} \end{cases}$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-z^2\right)$

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

$$\mathbb{S}_1 = \left\{ \tilde{x}_i \mid \tilde{x}_i < m \right\} \text{ and } \mathbb{S}_2 = \left\{ \tilde{x}_i \mid \tilde{x}_i \geq m \right\}$$

Let S_1 and $S_T - S_2$ the total number of elements of sets $\mathbb{S}_1$ and $\mathbb{S}_2$ respectively, then the likelihood of the sample is given by

$$L\left(x \mid m, \sigma^2, \gamma\right) = \left(\frac{2\sqrt{2\pi\sigma^2}}{\sqrt{1-\gamma} - \sqrt{1+\gamma}}\right)^{S_T} \exp\left(\frac{-1}{2}\left[\sum_{i=1}^{S_1}\left(\frac{x-m}{\sqrt{\sigma^2(1-\gamma)}}\right)^2 + \sum_{i=S_1+1}^{S_T}\left(\frac{x-m}{\sqrt{\sigma^2(1+\gamma)}}\right)^2\right]\right) \qquad [A4]$$

While the log-likelihood is

$$L\left(x \mid m, \sigma^2, \gamma\right) = S_T \log\left(\frac{2}{\sqrt{2\pi}}\right) - \frac{S_T}{2}\log\left(\sigma^2\right) - S_T \log\left(\sqrt{1-\gamma} + \sqrt{1+\gamma}\right) + \ldots$$

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

Estimation of the parameters requires the computation of the first order conditions of the Ikelihood problem:

For the uncertainty parameter we have

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

$$\hat{\sigma}^2 = \frac{1}{S_T(1-\hat{\gamma})}\sum_{i=1}^{S_1}\left(x - \hat{m}\right)^2 + \frac{1}{S_T(1+\hat{\gamma})}\sum_{i=S_1+1}^{S_T}\left(x - \hat{m}\right)^2 \qquad [A5]$$

For the risk parameter we find

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

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

which collapses to the following equation in the estimators:

$$\frac{\sum\limits_{i=S_1+1}^{S_T} (x - \widehat{m})^2}{(1 + \widehat{\gamma})^2} - \frac{\sum\limits_{i=1}^{S_1} (x - \widehat{m})^2}{(1 - \widehat{\gamma})^2} = \widehat{\sigma}^2 S_T \sqrt{1 + \widehat{\gamma}} \sqrt{1 - \widehat{\gamma}} \left(\frac{\sqrt{1-\widehat{\gamma}} - \sqrt{1+\widehat{\gamma}}}{\sqrt{1-\widehat{\gamma}} + \sqrt{1+\widehat{\gamma}}}\right) \qquad [A6]$$

For the mode parameter we have the expression:

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

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

which is simplified as:

$$\frac{\sum\limits_{i=1}^{S_1} x}{(1 - \widehat{\gamma})^2} + \frac{\sum\limits_{i=S_1+1}^{S_T} x}{(1 + \widehat{\gamma})^2} = \left[\frac{S_1}{(1 - \widehat{\gamma})^2} + \frac{S_T - S_1}{(1 + \widehat{\gamma})^2}\right] \widehat{m} \qquad [A7]$$

Equations [B.5], [B.6] and [B.7] are solved to find the triple of MLE parameters

$$\widehat{\Lambda} = (\widehat{m}, \widehat{\sigma}^2, \widehat{\gamma})$$

Appendix C: The posterior distribution

C.1 The joint posterior. The joint posterior distribution is given by:

$$p\left(\Lambda \mid \{\pi\}, \Omega\right) \propto \left(\frac{\gamma + 1}{2}\right)^{c-1} \left(\frac{1 - \gamma}{2}\right)^{d-1} (\sigma^2)^{\frac{-(a+2)}{2}} \exp\left(\frac{-b}{2\sigma^2}\right)$$

$$\left(\frac{(\sigma^2)^{-\frac{1}{2}}}{\sqrt{1 - \gamma} + \sqrt{1 + \gamma}}\right)^{S_T} \exp\left(\frac{-1}{2}\left\{\sum_{i=1}^{S_1}\left(\frac{\pi_i - m}{\sqrt{\sigma^2(1-\gamma)}}\right)^2 + \sum_{i=S_1+1}^{S^*}\left(\frac{\pi_i - m}{\sqrt{\sigma^2(1+\gamma)}}\right)^2\right\}\right)$$

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

$$p\left(\sigma^2 \mid \gamma, m, \{\pi_H\}, \Omega\right) \propto \left(\sigma^2\right)^{\frac{-(a+S^*+2)}{2}} \exp\left(-\frac{\vartheta(m,\gamma;S^*)+b}{2\sigma^2}\right) \qquad [A8]$$

Where

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

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.

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

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

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

While the fitted estimation with simulated data according to equation [A5] gives:

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

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

Starting with the conditional: $\dfrac{b}{a-2} > \dfrac{\vartheta(m,\gamma;S)}{S}$

(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:

$$bS + \vartheta\,(m,\gamma;S)S > (a-2)\,\vartheta\,(m,\gamma;S) + \vartheta\,(m,g;S)S$$

$$S\,\big(b + \vartheta\,(m,\gamma;S)\big) > \vartheta\,(m,\gamma;S)\,(a-2+S)$$

$$\frac{b + \vartheta\,(m,\gamma;S)}{a-2+S} > \frac{\vartheta\,(m,\gamma;S)}{S}$$

The basic result when $E\,(\sigma^2\mid.)_{prior} > \widehat{\sigma}^2_{fit}$ 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; thus the simulated variance does not differ from the posterior.

The other two relevant conditional distributions are given by:

$$p\,\big(m\mid\gamma,\sigma^2,\pi_{t+H},\,\Omega\,\big) \propto \exp\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_{t+H}-m)^2}{(1-\gamma)}\right) + \sum_{i=S_1+1}^{S}\left(\frac{(\pi_{t+H}-m)^2}{(1+\gamma)}\right)\right\}\right) \qquad [A9]$$

and

$$p\,\big(\gamma\mid m,\sigma^2,\pi_{t+H},\,\Omega\,\big) \propto \left(\frac{\gamma+1}{2}\right)^{c-1}\left(\frac{1-\gamma}{2}\right)^{d-1}\left(\frac{2}{\sqrt{1-\gamma}+\sqrt{1+\gamma}}\right)^{S}$$

$$[A10]$$

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

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 [A8]) and thus poses no problem. However, the other two kernels (equations [A9] and [A10] 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 $\sigma_k^2 \sim p\left(\sigma_{k-1}^2 \mid \gamma_k, m_{k\cdot\cdot}\right)$

3. Metropolis step to get m update:

Consider the function from equation [A9]:

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

(a) Calculate a function value: $M_{k-1} = c_m\left(m_{k-1}; \sigma_k^2, \gamma_{k-1}\right)$

(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: $\rho = \min\left(\frac{M_k}{M_{k-1}}, 1\right)$

(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 [A9]:

$$c_\gamma(\gamma; \sigma^2, m) \propto \left(\frac{\gamma+1}{2}\right)^{c-1}\left(\frac{1-\gamma}{2}\right)^{d-1}\left(\frac{2}{\sqrt{1-\gamma} + \sqrt{1+\gamma}}\right)^{S}$$

$$\exp\left(\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{S_1}\left(\frac{(\pi_i-m)^2}{1-\gamma}\right) + \sum_{i=S_1+1}^{S}\left(\frac{(\pi_i-m)^2}{1+\gamma}\right)\right\}\right)$$

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.

Appendix D: The optimal design of the sample size

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

$$S^* = \underset{S \in D}{\operatorname{argmax}} \left\{ KL(S) - \lambda S \right\} \qquad \text{[A11]}$$

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

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

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

$$\widehat{KL}(S) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \log[p(\Pi_i | \Lambda_i, S)] - \log[\hat{p}(\Pi_i | S)] \right\} \qquad \text{[A12]}$$

Where (Π_i, Λ_i) for $i = 1,..., N$ is a sample from $p(\Pi, \Lambda | S)$ and $\hat{p}(\Pi_i | S)$ is an estimator of the marginal density of the data $p(\Pi_i | S)$. The dependent pair (Π_i, Λ_i) drawn from $p(\Pi, \Lambda | S) = p(\Pi | \Lambda, S) p(\Lambda)$, is obtained by first drawing Λ from the prior distribution $p(\Lambda)$ and then Π from the conditional distribution $p(\Pi_i, | S)$.

The estimation of the marginal density of the data is obtained by an importance sampling based estimator as in Ryan (2003):

$$\hat{p}(\Pi_i | S) = \frac{1}{M} \sum_{j=1}^{M} p(\Pi_i | \Lambda_{ij}^*, S) \qquad \text{[A13]}$$

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 [A12] 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 [A12] 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 [A13] 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 $p(\Pi \mid \Lambda_{out_i}, S_k)$ for each $k = 1,...,$ n_d and each $i = 1,...,N$.
5) For $k = 1,...,$ n_d, compute

$$\widehat{KL}^M(S_k) = \frac{1}{N} \sum_{j=1}^N \widehat{KL}_i^M(S_k) \qquad [A14]$$

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

To implement the estimation, we considered the following values: $N_\Lambda = 5000$, $N = 1000$, M = 100, and $n_d = 200$. Also, we considered 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 [A11] to get the utility function shown in figure [7]

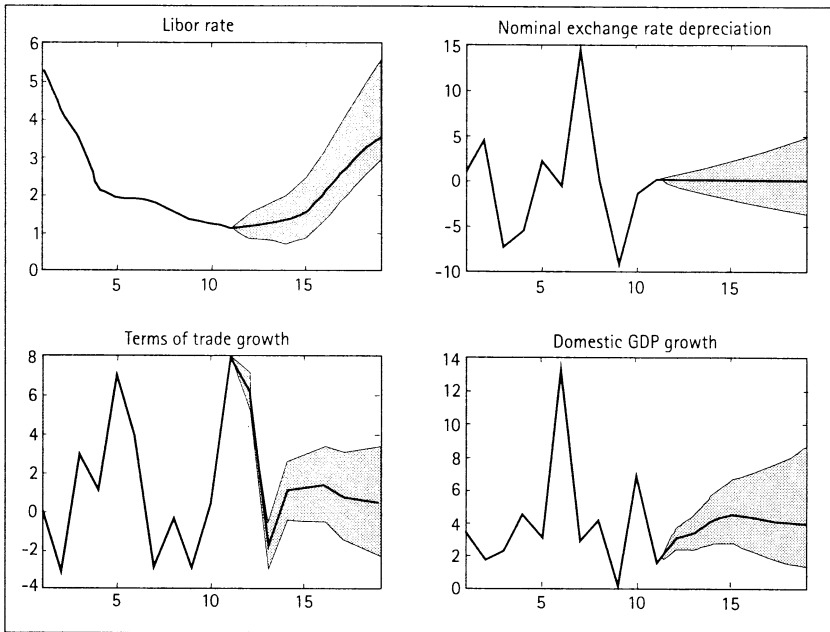Figure 1: Forecast interval and modal forecast

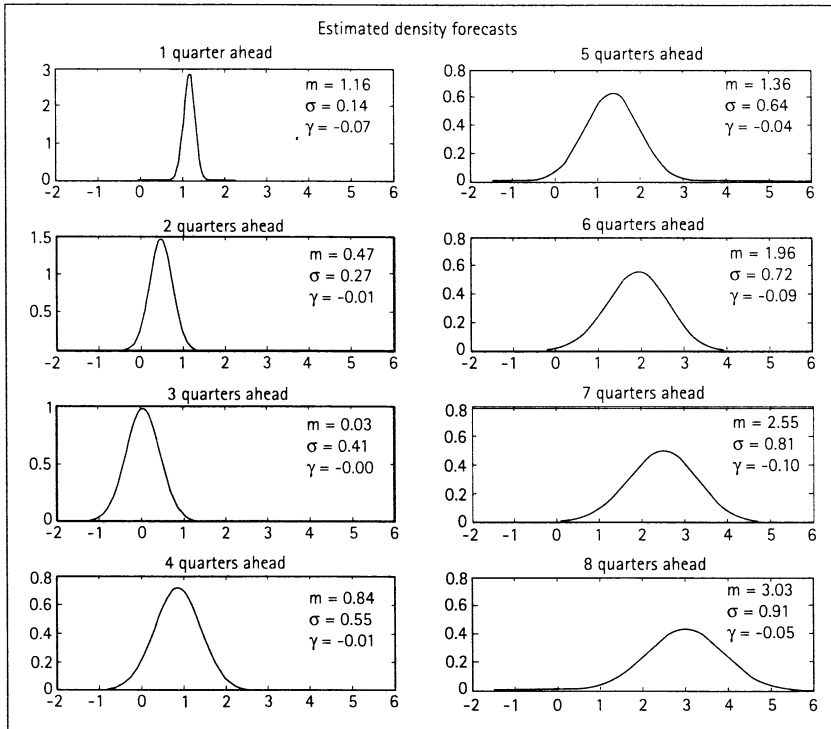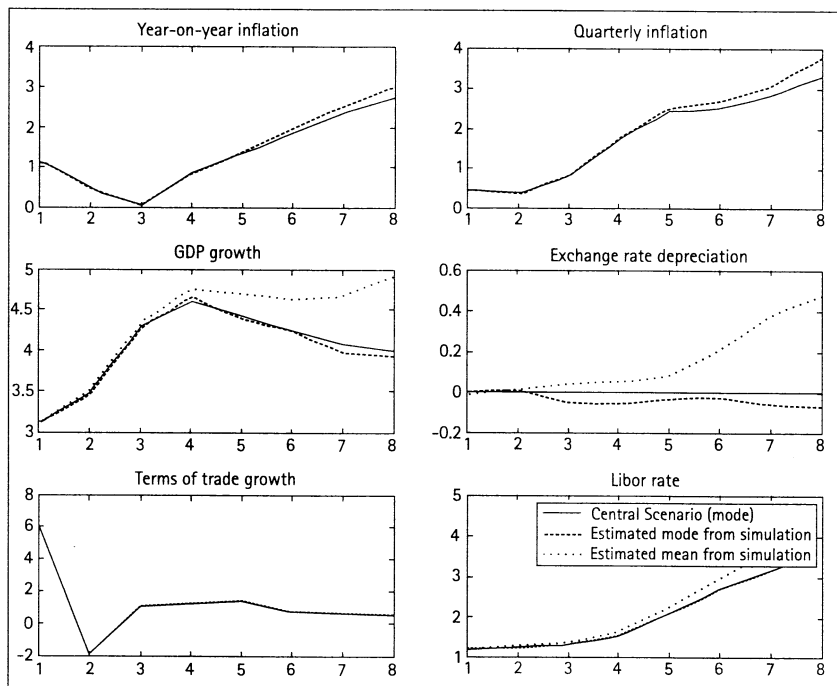## Figure 2: Estimated SN pdf's for the year-on-year inflation forecast.

Estimated density forecasts

Figure 3: Central measures of tendency

Figure 4: Evolution of the skewness paramater

Year-on-year inflation

Quartery inflation
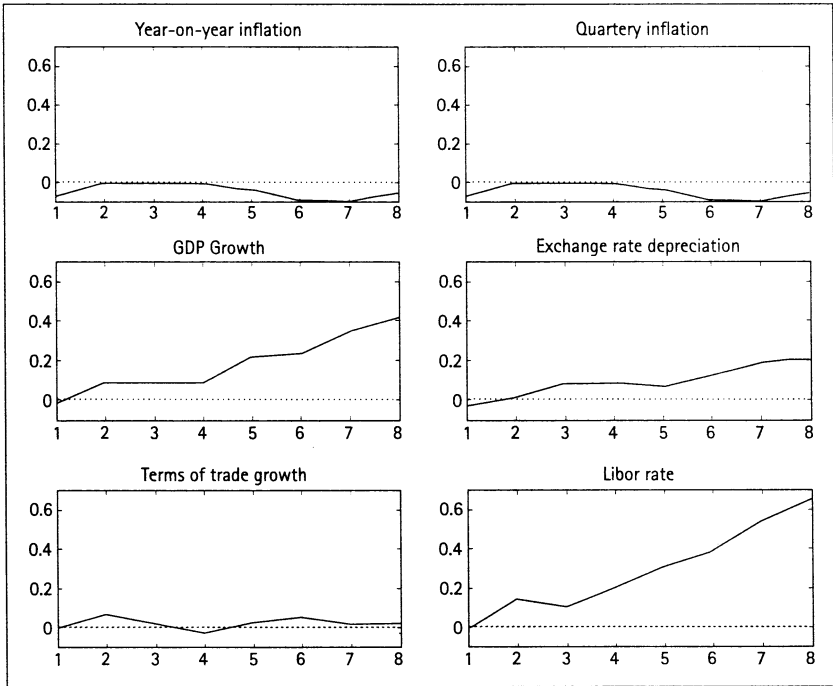
GDP Growth

Exchange rate depreciation

Terms of trade growth

Libor rate

Figure 5: Evolution of the uncertainty parameter
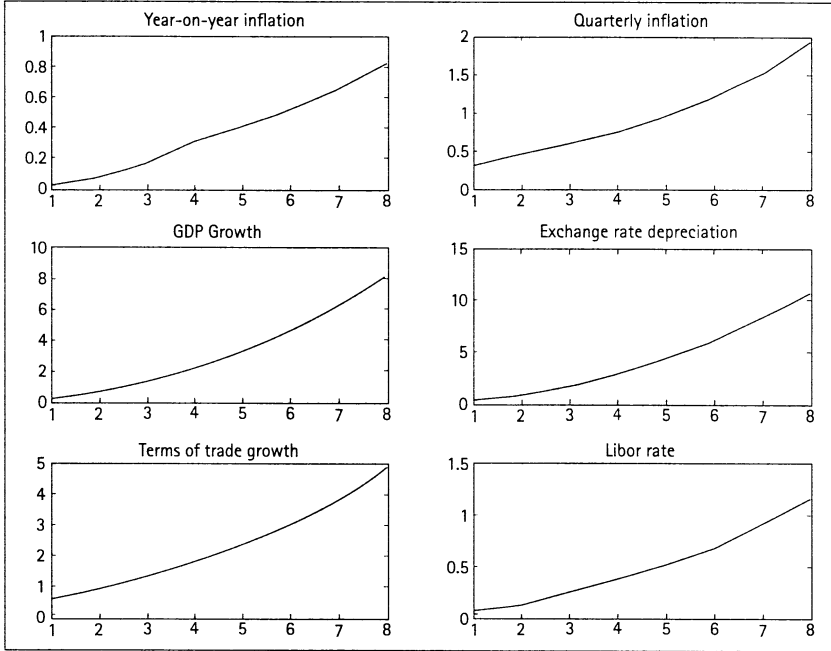
Figure 6: The KL divergence number (a.k.a entropy). The scatter plot is the estimation with monte carlo variation. The line is a smoothed version
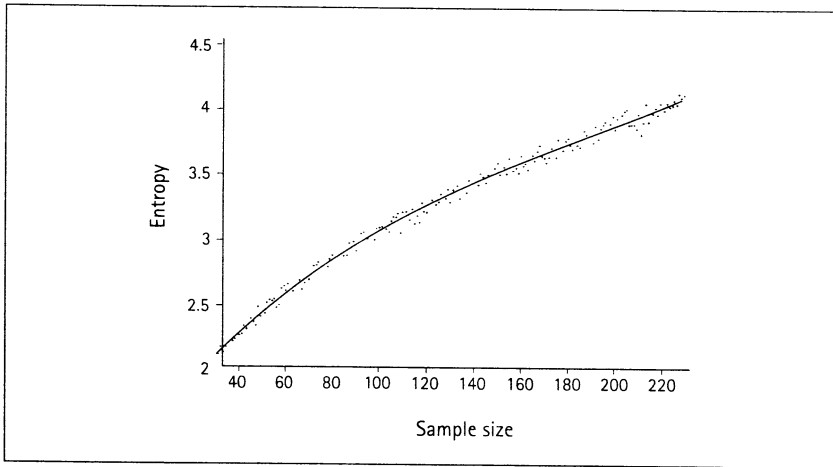

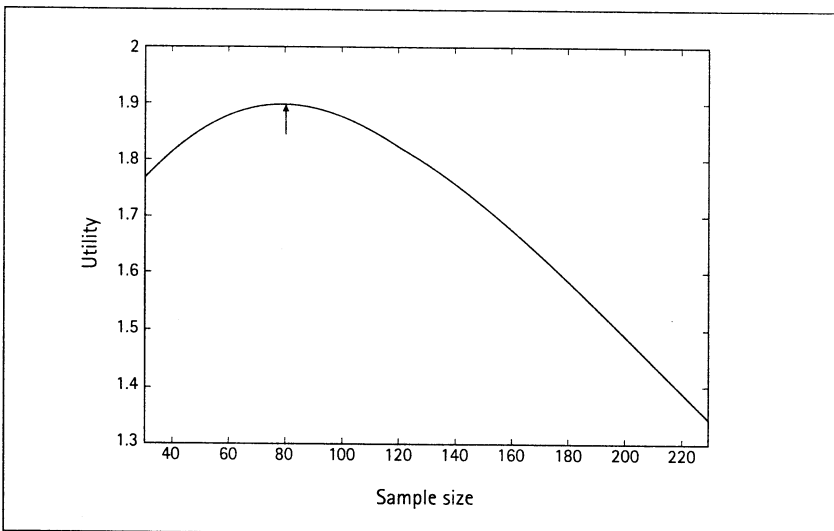Figure 7: The utility function of the policy maker as a function of the simulation sample size

Figure 8: Prior and posterior parameter distributions