



DP2016/10

Assessing the economic value of probabilistic forecasts in the presence of an inflation target

**Christopher McDonald, Craig Thamootheram,
Shaun P. Vahey and Elizabeth C. Wakerly**

June 2016

JEL classification: C53, E52

www.rbnz.govt.nz

Discussion Paper Series

ISSN 1177-7567

DP2016/10

**Assessing the economic value of probabilistic forecasts
in the presence of an inflation target***

**Christopher McDonald, Craig Thamotheram, Shaun P.
Vahey and Elizabeth C. Wakerly[†]**

Abstract

We consider the fundamental issue of what makes a 'good' probability forecast for a central bank operating within an inflation targeting framework. We provide two examples in which the candidate forecasts comfortably outperform those from benchmark specifications by conventional statistical metrics such as root mean squared prediction errors and average logarithmic scores. Our assessment of economic significance uses an explicit loss function that relates economic value to a forecast communication problem for an inflation targeting central bank. We analyse the Bank of England's forecasts for inflation during the period in which the central bank operated within a strict inflation targeting framework in our first example. In our second example,

* The Reserve Bank of New Zealand's discussion paper series is externally refereed. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Reserve Bank of New Zealand. Financial support from the ARC (LP0991098), Norges Bank, the Reserve Bank of New Zealand, and the Reserve Bank of Australia is gratefully acknowledged. Computations were performed using the PROFOR toolbox. Download from <http://www2.warwick.ac.uk/fac/soc/wbs/subjects/emf/profor/>. The toolbox was developed with support from Warwick Business School, the Bank of England, and Norges Bank. We thank Christie Smith and seminar participants at the Bank of England, Norges Bank, the Reserve Bank of New Zealand, Narodowy Bank Polski, the Bank of Canada, and the CIRANO Real-time Workshop September 2015 for comments.

[†] Christopher McDonald: Economics Department, Reserve Bank of New Zealand, 2 The Terrace, PO Box 2498, Wellington, New Zealand. Email: Chris.McDonald@rbnz.govt.nz. Craig Thamotheram: Cardiff Business School, Cardiff University, Cardiff CF103EU, UK. Email: ThamotheramC@cardiff.ac.uk. Shaun P. Vahey: Warwick Business School, University of Warwick, Scarman Road, Coventry CV47AL, UK. Email: svahey@warwick.ac.uk. Elizabeth C. Wakerly: 7 Redcote Manor, Walton Park, Milton Keynes MK77HF, UK. Email: lizwakerly@gmail.com.

we consider forecasts for inflation in New Zealand generated from vector autoregressions, when the central bank operated within a flexible inflation targeting framework. In both cases, the economic significance of the performance differential exhibits sensitivity to the parameters of the loss function and, for some values, the differentials are economically negligible.

Non-technical Summary

Inflation targeting central banks devote considerable resources to modelling and assessing the future path of inflation, and to understanding previous forecast errors. Nearly all monetary policymakers also issue periodic ex ante inflation event warnings—telling the public when the path of inflation is expected to move outside “normal” levels. In this paper, we show that strong forecast performance by conventional statistical metrics is not sufficient for a candidate forecasting method to be helpful when the bank wishes to warn the public about inflation events. We use two specific examples to illustrate the importance of this finding for both strict and flexible inflation targeting central banks. The first example considers the UK inflation forecasts published by the Bank of England, operating within a strict inflation targeting regime. The second example evaluates inflation forecasts produced from vector autoregressions, using a sample of New Zealand data, drawn from the period in which the Reserve Bank operated within a flexible inflation targeting framework. In both examples, a conventional statistical analysis would have misled a policymaker about the sensitivity of the forecast performance from an economic perspective. Specifically, for some parameters of the loss function used in our examples, the performance gains over the benchmark models are negligible.

1 Introduction

“Given . . . (the) asymmetric costs or benefits of particular outcomes, a central bank needs to consider not only the most likely future path for the economy, but also the distribution of possible outcomes about that path. The decision-makers then need to reach a judgment about the probabilities, costs, and benefits of the various possible outcomes . . .”

Greenspan (2004, p 37).

Although central banks devote considerable resources to assessing the probabilities of inflation events, many of the evaluation techniques used in practice assess forecast performance from a statistical rather than an economic perspective. One approach is to ignore the ex ante uncertainty about the point forecast and simply use a test of relative forecast accuracy based on Root Mean Squared Prediction Error (RMSPE). An alternative is to use the average logarithmic score of candidate forecast densities. The second approach—less commonly adopted by practitioners but increasingly prevalent in academic work—does take account of the ex ante probabilistic information contained in the forecast density. Examples of recent studies in the macro forecasting literature deploying at least one of these approaches include Bjørnland et al (2011), Garratt et al (2011), Clark (2011), Confitti, De Mol and Giannone (2015), Aastveit et al (2014), Del Negro, Hasegawa and Schorfheide (2014), and Kruger, Clark and Ravazzolo (2015).

The novel contribution of our paper is that we gauge the extent to which statistical measures of forecast performance misrepresent the usefulness of a forecast to an inflation targeting central bank. We provide two examples of inflation forecasting using data where the central banks concerned operate in strict and flexible inflation targeting frameworks, respectively. We gauge forecast performance in both statistical and economic terms in each case. The statistical measures used are the RMSPE and the average logarithmic

score. In contrast, economic significance relates to a specific loss function in which as Greenspan (2004) suggests “. . . the probabilities, costs and benefits” of communication actions matter to the inflation targeting central bank. Specifically, the central bank makes public statements about the risk that inflation will exceed the upper bound of the target interval. The policymaker faces asymmetric costs in the sense that the social cost of an anticipated inflation event is less than the cost of an unanticipated event. In contrast, (among others) Diks et al (2011), Galbraith and van Norden (2011 and 2012) and Gneiting and Ranjan (2011) move beyond RMSPE and the average logarithmic score by examining alternative statistical performance measures, rather than relating the forecasts directly to the economic value of the forecast user as we do in this paper.

In our first example, we analyse the Bank of England’s published forecasts for inflation during the period in which the central bank operated within a strict inflation targeting framework. In our second example, we consider forecasts for inflation in New Zealand generated from vector autoregressions, using a sample drawn from the period in which the central bank operated within a flexible inflation targeting framework. In both examples, the candidate forecasts overwhelmingly outperform statistically those from an autoregressive benchmark using conventional metrics (RMSPE and the average logarithmic score). However, the economic value of the performance differential (relative to the autoregressive benchmark) varies with the degree of asymmetry in economic costs associated with anticipated and unanticipated inflation events. In economic terms, the performance advantage in our examples lies between zero and (just over) 60 percent. For the two central banks considered here, operating under different inflation targeting frameworks, the extreme sensitivity of relative forecast performance would be missed by a policymaker restricting attention to more conventional measures of statistical forecast evaluation.

The remainder of this paper is structured as follows. In Section 2, we set out background information about inflation targeting in the UK and New Zealand and our decision-

theoretic framework for forecast evaluation based on the communication of ex ante inflation risks. In Section 3, we describe the empirical methods, data and results for our Bank of England example. In Section 4, we provide an analysis of model-based forecasting in New Zealand. We draw some conclusions and ideas for extending the analysis in the final section.

2 Inflation Targeting and Decision-theoretic Forecast Evaluation

In this section, we begin by describing the forecast communication strategies followed by the Bank of England and the Reserve Bank of New Zealand. Then, we introduce our decision-theoretic framework for forecast evaluation and relate it to forecast communication in practice.

2.1 Inflation forecast communication

Nearly all central banks routinely produce reports on the state of the macroeconomy, focusing specifically on explicit forecasts for inflation (along with other key macroeconomic variables). However, for central banks operating within an inflation target, the inflation forecasts and the assessments of ex ante inflation risks have particular importance.

The attention afforded to the probability of inflation events in central bank press releases provides some indication of the importance of pre-emptive risk assessment and communication. For example, the Governor’s *Opening Remarks* to the Bank of England’s *Inflation Report* Press Conference in February 2010 stressed that “The January figure for CPI inflation is likely to have exceeded 3% . . . This would be the third episode when inflation has temporarily moved above the target . . .”

The importance of inflation events relative to the UK’s target interval reflects, in

part, an unusual feature of the Bank of England's strict inflation targeting regime. If inflation deviates from the central target (of 2 percent) by more than 1 percentage point, the Governor must send an open letter to the Chancellor.¹ The Governor sent 14 letters to the Chancellor between April 2007 and February 2012, all of which explained the reasons for high inflation in the short term. In each case, the preceding *Inflation Report* forewarned of the event.

Even though a central bank facing a "flexible" medium-term target, such as the Reserve Bank of New Zealand, has greater scope to ignore the bounds of the target interval in the short term, inflation event warnings are still common. The Reserve Bank's Policy Target Agreement "... defines price stability as annual increases in the Consumers Price Index (CPI) of between 1 and 3 percent on average over the medium term, with a focus on keeping future average inflation near the 2 percent target midpoint". Notwithstanding the medium-term perspective, Governor Alan Bollard's Policy Assessment in the *Monetary Policy Statement* in June 2006 forewarned that inflationary pressures were expected "... to keep headline CPI inflation above 3 percent well into 2007". And in September 2011, the Governor stressed that "... headline inflation continues to be above the Bank's 1 to 3 percent target band". The Reserve Bank issued 11 verbal warnings of high inflation events (defined relative to the bounds of the medium-term target interval) via the Monetary Policy Statement between December 2000 and June 2012.²

In common with a number of other central banks, the inflation target has varied in New Zealand. The Reserve Bank faced a target of 0 – 2 from March 1990, 0 – 3 from December 1996, and 1 – 3 percent from September 2002.³ Regardless of these shifts in the target boundaries, and the "flexibility" of the target, ex ante inflation warnings remain a

¹The letters and replies can be downloaded from the Bank of England's website: <http://www.bankofengland.co.uk/monetarypolicy/Pages/inflation.aspx>.

²The Reserve Bank also warned in June 2009 of the danger that inflation could drop below the target band.

³Information on the the Policy Target Agreement can be downloaded from http://www.rbnz.govt.nz/monetary_policy/policy_targets_agreement/.

prevalent feature of New Zealand’s monetary policy communication framework.

2.2 Decision-theoretic framework

Greenspan (2004) describes monetary policymaking in terms of decisions based on both the probabilities of events and the asymmetric costs of the available options. In this section, we describe a one-period framework in which these considerations influence the decision to issue an ex ante inflation warning to the public.

In the examples that follow, the central banks concerned nearly always forewarned of high inflation events—when inflation is forecast to be above the upper bound of the inflation target with high probability. Low inflation events are rare in our data, with just one outturn below the lower bound of the (medium) term inflation target in New Zealand, and none in the United Kingdom. So for expositional ease, we restrict attention to a single threshold that defines the event. (See Granger and Pesaran (2000a and 2000b) on extensions to multiple thresholds.) Also for expositional convenience, we treat the threshold of interest as time invariant. The upper bounds of the targets are constant for both central banks in the examples presented below.

In our decision-theoretic forecast evaluation framework, we assume that the central bank is concerned with the accuracy of forecasts relative to a given upper bound of the inflation target (interval), denoted $\bar{\pi}$.⁴ Based on the forecast, the policymaker takes one of two communication actions. Either the policymaker issues a preemptive inflation event warning to the public, or not. If the central bank warns of the event, there is a one-period economic cost to society of C , reflecting the adjustment of the private sector to the information. (In principle, the cost could be contingent on actual outturn for inflation but we abstract from this for simplicity.) If the central bank fails to issue an inflation

⁴To be consistent with the literature on cost-loss ratios, in what follows we use the standard terminology and notation adopted in other applied statistical fields. See, for example, Berrocal et al (2010).

event warning but the event occurs, society incurs a one-period loss, L , where $0 < C < L$. This cost reflects the economic loss to the private sector of an unanticipated inflation event. The costs are asymmetric in that the cost to society of an anticipated inflation event, C , is less than the cost of the unanticipated event, L .

The optimal rule for publishing the warning can be derived for a given inflation threshold, $\bar{\pi}$. Since society incurs the per period cost C if the warning is issued (regardless of whether the event subsequently occurs), the expected cost of the warning is C . The expected loss from the alternative communication strategy—that is, “no warning”—is $p_{\bar{\pi}}L$, where $p_{\bar{\pi}}$ represents the probability that the inflation event occurs. Hence, the policymaker needs a probabilistic forecast in order to decide whether to issue an inflation event warning, or not. (We assume that this is a one step ahead forecast and discuss the multiple horizon case in the subsequent subsection.) Assuming the policymaker minimises expected cost, an inflation event warning will be issued provided that $R = C/L < p_{\bar{\pi}}$. On the other hand, if $R = C/L \geq p_{\bar{\pi}}$, then the policymaker issues no warning. Note that the framework implies that smaller probabilities trigger inflation event warnings for low R .⁵ In our applied work, we shall treat R as an unknown parameter and examine the robustness of forecast performance across a range of values.

Similar frameworks have been deployed in previous research to study applied forecasting problems in meteorology, finance and economics. See, for examples unrelated to monetary policy, Murphy (1977), Katz and Murphy (1997), Granger and Pesaran (2000a and 2000b), Berrocal et al. (2010) and Garratt, Vahey and Wakerly (2012). Monetary policy applications include Clements (2004), Garratt, Mitchell and Vahey (2014), Coenen and Warne (2014) and Coe and Vahey (2015). Clements (2004) links interest rate setting to forecast evaluation using forecasts published by the Bank of England but does not con-

⁵Although the relative cost R affects the decision to issue the warning, in principle both C and L could be time varying; hence, there could be variations in R . In the application that follows, we consider forecasts in a low and stable inflation environment.

sider forecast communication. Garratt, Mitchell and Vahey (2014) and Coe and Vahey (2015) evaluate model-based forecasts of US inflation. Coenen and Warne (2014) use a dynamic optimisation-based open economy model to assess the balance of risks to inflation in the euro area. These papers do not consider the scope for conventional measures of forecast performance to give misleading information about the usefulness of a forecast from the perspective of an inflation targeting central bank.

There is a considerable literature considering statistical measures of probabilistic forecast performance. Contributions include Garratt et al (2003), Adolfson et al (2007), Lahiri and Wang (2007), Garratt et al (2009), Kryshko, Schorfheide and Sill (2010), Berge and Jorda (2011), Clark (2011), Diks, Panchenko and van Dijk (2011), Galbraith and van Norden (2011 and 2012), Gneiting and Ranjan (2011), and Luciani and Ricci (2015). In these papers, the loss function is not explicitly motivated by an economic decision, as is done in this paper. Arguably, within an inflation targeting framework, policymakers should be concerned with the economic significance of forecast performance, which in our examples depends on (what Greenspan describes as) "... the probabilities, costs and benefits" of their communication actions.

2.3 Discussion

Our framework is motivated by the policymaker's desire to forewarn the public of inflation events relative to the inflation target. The cost asymmetry between anticipated and unanticipated inflation event provides a rationale for the Bank of England and the Reserve Bank of New Zealand to forewarn of inflation risks. Given the concern that the public might misinterpret quantitative signals about inflation, verbal predictions of inflation events (based on probabilistic forecasts) provide a rigorous communication strategy. In practice, nearly all central banks (including the Bank of England and the Reserve Bank of New Zealand) use verbal communication to forewarn the public of inflation events. Some

have followed the Bank of England in publishing the whole forecast density but many, like the Reserve Bank, have not.

Since policy changes affect the real economy with a considerable time lag, and real-time data are released with a publication lag, the short-term path of inflation is in effect predetermined. In the two applications that follow, for expositional ease, we will restrict attention to one step ahead forecasts. That is, given the time lag in releases of macroeconomic data, we consider “nowcasts” in our applied work.

There are similarities between the nowcast inflation event warnings given by inflation targeting central banks and the verbal warnings commonly published by meteorologists. For a meteorology office to warn of a storm requires a probabilistic forecast of the extreme event.⁶ Meteorology offices publish forecast densities but often prefer to communicate with the public via explicit verbal warnings because many forecast users find complete densities difficult to interpret. Similarly, a central bank concerned with the bounds of an inflation target may prefer verbal inflation event warnings (regardless of whether the forecast densities themselves are published) because the public prefers this communication strategy.

In practice, and in contrast to the communication framework deployed in this paper, central bankers tend to forewarn of inflation events over multiple horizons, with policymakers often stressing the date at which inflation will likely return to target.⁷ Beyond the inflation nowcast considered in our analysis, the central bank’s forecasts (and interest policy decisions) influence the future—a complexity not confronted by weather forecasters.⁸

⁶See, for example, the UK’s storm warnings <http://www.metoffice.gov.uk/guide/weather/warnings>.

⁷For example, in the first open letter to the Chancellor (page 4, letter dated April 16, 2007), the Bank of England Governor warned that “CPI inflation is likely to fall back within a matter of few months”.

⁸Short-term projections by central banks tend to be more accurate than longer-term forecasts. In part, this reflects the feedback between monetary policy and the forecasts at longer horizons.

3 Bank of England Forecasts

In this application, we examine the Bank of England’s published forecasts from the perspective of our decision-theoretic framework. We begin by fleshing out the background to the production of the forecasts, then describe the forecasts themselves, and subsequently present the results.

3.1 Background

The forecasts published by the Bank of England represent the views of the Monetary Policy Committee (MPC). The Bank of England has published probabilistic forecasts for inflation since February 1996.⁹ Published in a format known as “fan charts”, the series of predictive densities display forecasts at multiple horizons on a single plot from a given forecast origin. Franta et al (2014) discuss the prevalence of fan charts as a communication tool for inflation targeting central banks, and describe in detail the construction of the Bank of England’s approach. See also the discussion in Galbraith and van Norden (2012).

The Bank of England’s forecasting record has been subject to considerable scrutiny. See, for example, Wallis (2004), Clements (2004), Gneiting and Ranjan (2011) and Galbraith and van Norden (2012). The general finding is that the Bank of England’s one quarter ahead forecasts outperform statistical benchmarks on samples prior to the most recent recession. Galbraith and van Norden (2012) note that the forecasts are well calibrated but typically lack resolution—meaning that the probabilities are close to those of the unconditional probability of the relevant event but often fail to give accurate conditional probabilities. Gneiting and Ranjan (2011) find that at longer horizons, on a pre-recession sample, the Bank’s forecast tend to be over dispersed.

⁹Probabilistic forecasts for GDP growth have been published since November 1997.

3.2 Forecasts

We examine four inflation forecasts per year, corresponding to the release date of the Bank of England’s Inflation Report. The inflation forecasts refer to the annual percentage changes in prices. We examine the sequence of one step ahead probabilistic forecasts with forecast origin dates from 2003.04 to 2013.03. In total, we consider 40 forecasts of CPI inflation.

The Bank’s published forecast densities prior to 2004.01 target an underlying measure of inflation known as RPIX. From the 2004.01 publication (using forecast origin 2003.04), the Bank switched to a CPI measure. Our evaluations compare the Bank of England’s forecasts with those from a first-order autoregression, AR(1), following the benchmark choice of Gneiting and Ranjan (2011).¹⁰ For parameter estimation in the benchmark model, we use recursive OLS estimation, using observations from 1993.01 to the forecast origin for each recursion, with the first forecast origin being 2003.04.

The Bank of England’s forecasts are generated from a two-piece normal; see the discussions in Wallis (2004, 2005). The parameters can be obtained from the Bank of England.¹¹ As Galbraith and van Norden (2012) and Franta et al (2014) note, the two-piece normal methodology permits the members of the MPC to allow for asymmetric risks in their projections. The parameters of each fan chart result from a combination of model-based forecasts with expert-judgement, and represent the views of the MPC. At each point in time, the members assume that interest rates will remain constant thereafter at the current level.¹²

Following, Gneiting and Ranjan (2011), the two-piece normal distribution is written:

¹⁰Results are very similar with an AR(4) benchmark.

¹¹<http://www.bankofengland.co.uk/publications/Pages/inflationreport/irprobab.aspx>.

¹²The Bank of England also publishes inflation forecasts conditional on market expectations of the forward interest rate but the conditioning information matters little for our one step ahead inflation forecasting application.

$$f(y) = (\pi/2)^{-0.5}(\sigma_1 + \sigma_2)^{-1} \exp(-(y - \mu)^2/2\sigma_1^2) \quad \text{if } y \leq \mu$$

and

$$f(y) = (\pi/2)^{-0.5}(\sigma_1 + \sigma_2)^{-1} \exp(-(y - \mu)^2/2\sigma_2^2) \quad \text{if } y > \mu$$

where the parameters μ , σ_1 and σ_2 denote the mode, and standard deviation of the left and right hand side of the distribution, respectively.

3.3 Results

Before undertaking the decision-theoretic evaluation of the Bank of England's one step ahead forecasts for inflation relative to the AR(1) benchmark, we begin by considering more conventional forecast metrics. In terms of point forecasting, using the median of the forecast density, based on the forecast origins 2003.04 through 2013.03 (targets 2004.01 through 2013.04), the Bank's Root Mean Squared Prediction Error (RMSPE) is 0.1838. In contrast, the AR(1) model produces a RMSPE of 0.5709. A gain in forecast accuracy relative to the benchmark of nearly 70%. A Diebold and Mariano (1995) test with the small sample adjustment suggested by Harvey, Leybourne and Newbold (1997) comfortably rejects the null hypothesis of equal forecast accuracy at the 1% significance level (in a two-sided test).

Turning to a forecast evaluation based on the whole predictive density, the results are similar to the point forecasting case. For example, the Bank of England's forecasts produce an average logarithmic score of -0.1278; whereas, the AR(1) benchmark scores -1.037. A test of relative forecasting accuracy based on the differential in the average logarithmic scores (see Bao et al, 2007) comfortably rejects the null of equal forecast

accuracy at the 2% level.¹³

We display the forecasts produced by the Bank of England in Figure 1, which plots the (sequence of) one step ahead forecast densities and the realisations of UK inflation (solid line), together with the quantiles (0.01; 0.05; 0.10; 0.33; 0.66; 0.90; 0.95; 0.99). Note that the Bank’s one step ahead forecast densities exhibit only mild asymmetries, even through the Great Recession (longer horizons periodically have greater skew).

Interestingly, realisations rarely occur in the outer probability percentiles of the Bank’s forecast densities—suggesting that MPC lack confidence in their own projections. This characteristic is particularly striking post-2009. Apparently, the MPC members perceived greater forecast uncertainty for inflation after the Great Recession.

We plot the (end-evaluation) Probability Integral Transforms (PITs) in Figure 2 to provide another perspective on the Bank of England’s forecasting performance. Well-calibrated forecast densities should produce a uniform distribution for the PITs.¹⁴ In particular, the first two forecast bins are empty—the realisations of inflation never fall in the lower tail of the forecast density. That is, the MPC members see too much downside risk to inflation.

We plot the corresponding objects of interest for the AR(1) benchmark in Figure 3 and Figure 4, respectively. Notice that the AR(1) benchmark generates little increase in risk after 2009 (Figure 3) and the realisations are somewhat more evenly distributed in the forecast density (Figure 4) than the Bank of England (Figure 2) but nevertheless the AR(1) benchmark has too many realisations in the upper and lower tails of the forecast densities. In particular, the upper tail misses a number of observations during 2008 (Figure 3). Often the AR(1) projection is a lagging indicator of inflation, in contrast to the Bank (see Figure 1). To some extent, the Bank’s performance advantage over

¹³We abstract from the method used to produce the forecast densities. Amisano and Giacomini (2007) discuss the limiting distribution of related test statistics.

¹⁴The overlapping nature of the Bank of England’s multi-step forecasts almost certainly induces a moving average structure for inflation and complicates inference based on the PITs.

the benchmark reflects the timeliness of the Bank’s forecast. The Bank possess some intra-quarter information that the quarterly AR(1) model lacks.

Overall then, there are two main findings from our study of the Bank of England’s forecasts. First, the Bank’s density forecasts are poorly calibrated, particularly in the lower tail, and are too diffuse, especially post-2008. Second, we find that the Bank still has a considerable advantage in performance in terms of point and density forecasting over the benchmark by the standard evaluation metrics. With these results in mind, we turn now to our decision-theoretic based evaluation to address the issue of economic significance. Specifically, we examine whether the advantage in forecast performance—suggested by standard point and density forecast evaluation techniques—is sufficient for the forecasts to be assessed as “good” when communicating the risk of high inflation events to the public within the inflation targeting regime. In particular, we consider the robustness across a range of parameters reflecting the asymmetric costs in the loss function.

Recall that Total Economic Loss, TEL , depends on both action and realisation. The “action” is that the policymaker issues a preemptive event warning of the inflation event. Namely, that inflation will exceed the 3% threshold of the inflation target. If the realisation is that the inflation realisation exceeds that threshold, then a letter of explanation must be sent by the Bank’s Governor to the Chancellor. As discussed earlier, we assume that the preemptive inflation event warning results is a one-period (time-invariant) cost of C , irrespective of whether the event occurs. On the other hand, in the absence of a warning, an economic loss, L , is incurred only if the inflation event occurs. The relative cost ratio, $R = C/L$, will lie in the region $0 < R < 1$, since the costs are asymmetric.

We denote the number of observations in which action was taken (the warning issued), but the inflation event did not occur, as n_{01} . We denote the number of observations in which action was taken and the event did occur n_{00} . In both of these cases, the cost is

$C > 0$. We denote the number of times that no action was taken and the event did not happen n_{11} . In this case, society incurs no cost. We denote the number of observations in which no action was taken but the event occurs n_{10} , giving a loss of L .

Then, the (end of evaluation) Total Economic Loss, TEL , can be expressed as:

$$TEL = Ln_{10} + C(n_{01} + n_{00}).$$

We plot TEL against R for the Bank of England's forecasts and for the AR(1) benchmark (normalised to 1) in Figure 5. The Bank's forecasts produce a lower TEL than the benchmark for all values of R less than 0.95. But the largest gain is for low values of R , where small probabilities (of the event) trigger the preemptive inflation warning. For R less than 0.2, the Bank's (relative) TEL is minimised at less than 0.4—indicating a performance advantage over the benchmark of just over 60%. For the case with R between 0.20 and 0.75, the performance gain in favour of the Bank lies in the range of 10 to 30%. But, the performance gain drops to zero where R is greater than 0.95. The Bank's (relative) TEL converges on 1 as R rises.

We stress that the performance gain shown in Figure 5 is always smaller, and for high R much smaller than that suggested by, for example, RMSPE, which indicated a 70% improvement over the benchmark. Recall that, given our forecast communication framework, values of R reflect differences between the economic costs of anticipated and unanticipated inflation events. Since the performance differential is negligible for R greater than 0.95, there is no economic value to using the Bank's forecasts if the costs of unanticipated inflation are 95% (or more) of the costs of anticipated inflation. The over-dispersion of the Bank's forecast densities triggers too many inflation warnings in general. And for high R , in particular, this erodes the performance differential. (As noted in discussing Figure 1 and Figure 2, the MPC's perception of the risk of low inflation realisations appears

particularly misplaced but this matters little for “high” inflation events.)

We plot the one step ahead probabilities of the event (that inflation exceeds the 3% threshold) for the Bank’s projections and for the AR(1) forecasts in Figure 6. This plot is consistent with the interpretation that the Bank’s capacity to give accurate one step ahead forecasts stems from the timely nature of its forecasts, based on intra-quarterly information. The probabilities are similar but the Bank typically leads the benchmark by one quarter. Also, the Bank’s probabilities of exceeding the inflation target upper bound are often slightly higher than the AR(1), reflecting the additional uncertainty perceived by the MPC members, particularly since the Great Recession; see Figure 1 and Figure 3.¹⁵

In summary, the density forecasts published by the Bank of England provide helpful information about the probability of inflation being outside the inflation target. But, the extent of the improvement in forecast performance (relative to an AR(1) benchmark) varies with asymmetry in costs, captured by the parameter R . The sensitivity of forecast performance would be missed by a policymaker examining only more conventional statistical measures of forecast performance. Since RMSPE and average logarithmic predictive scores abstract from the communication actions (and their relative costs) available to the policymaker in our framework, they overstate the performance gain that accrues from the Bank’s forecasts for some parameter values.

4 VAR Forecasts for NZ Inflation

In this application, we generate out of sample forecasts for inflation from vector autoregressions, and evaluate the probabilistic predictions from the perspective of the Reserve Bank of New Zealand’s flexible inflation target.

¹⁵The average point forecasts differ little between the two approaches.

4.1 Background

The forecasts published by the Reserve Bank of New Zealand represent the views of the Governor, having taken into account analysis and discussions with the staff.¹⁶ Although the Reserve Bank considers forecast uncertainty and evaluates forecast performance using a number of different measures of point and forecast accuracy, communication with the public focuses on the expected path for key macro variables. For example, there are four inflation forecasts per year published in the Reserve Bank’s *Monetary Policy Statement*. These documents display and analyse the expected path for inflation, along with verbal descriptions of the uncertainties, but do not usually present forecast densities. Periodically, the Reserve Bank examines its own forecast performance in articles published in the *Reserve Bank Bulletin*. Turner (2006) and Labbé and Pepper (2009) provides recent examples of the Reserve Bank’s forecast evaluation. The general finding is that the Reserve Bank’s own forecasts modestly outperform those from statistical benchmarks, and those from professional forecasters, on data prior to 2007.

In the absence of publicly available forecast densities for inflation from the Reserve Bank, we generate inflation forecasts from a bivariate VAR for inflation and the output gap, estimated on a sample from New Zealand’s inflation targeting regime.¹⁷

The Reserve Bank’s current inflation targeting framework specifies that:

“(T)he policy target shall be to keep future CPI inflation outcomes between 1 per cent and 3 per cent on average over the medium term, with a focus on keeping future average inflation near the 2 per cent target midpoint.”

Clause 2b, *Policy Targets Agreement*, September 20, 2012.

¹⁶The Monetary Policy Committee of the Reserve Bank of New Zealand is an internal body which considers projections together with analysis, and provides advice to help inform policy decisions.

¹⁷In practice, other endogenous and exogenous variables may help forecast New Zealand inflation. To simplify our example contrasting the economic and statistical significance of evaluation metrics for probabilistic forecasts, we restrict attention here to a simple bivariate relationship.

A specific clause in the same document describes the responsibility of the Reserve Bank to account for projections that are outside the (medium-term) target range:

“... when such occasions are projected, the Bank shall explain in Policy Statements ... why such outcomes have occurred, or are projected to occur ...”.

Clause 4a, *Policy Targets Agreement*, September 20, 2012.

With these features of the framework in mind, our analysis focuses on a quarterly sample of observations for the annual percentage changes in the Consumers Price Index (CPI) excluding the periodic one-off effects of the Good and Service Tax. Research within the Reserve Bank typically uses this measure of inflation for projections.

4.2 Forecasts

To generate our model-based forecast densities, we utilise a bivariate VAR model space for inflation and the output gap (the deviation of real output from potential). The output gap measure is derived by univariate detrending, utilising Bayesian estimation methods. Similar “off-model” filters have been deployed by Orphanides and van Norden (2002, 2005), and Edge and Rudd (2012). Although Orphanides and van Norden (2002, 2005) document the instability of individual measures of the output gap and their weak predictive content for inflation, Garratt et al (2011) combine forecasts across model specifications to improve forecast performance. They find that adopting this ensemble modelling approach allows VARs to outperform simple autoregressive benchmarks across a variety of statistical measures of forecast performance, including RMSPE and average logarithmic score.

The pragmatic off-model filtering approach is often used by central banks in practice, including the Reserve Bank of New Zealand.¹⁸ The Appendix describes the detrending

¹⁸Although a multivariate approach would be more efficient there is considerable scope for specification error with simultaneous systems.

method, the priors, the Bayesian estimation methods and the forecast combinations in detail. Given the focus aim of this paper is to consider the economic significance of forecast performance differentials from the perspective of the inflation target, we focus on the ensemble VAR forecasts for inflation, rather than the interpretation of the output gap, or the weights on particular forecasts.

Although inflation targeting was established in New Zealand in 1989, the transition years were characterised by volatile inflation expectations. Hence, recursive VAR parameter estimation begins in 1992.01, using an expanding window up to the forecast origin. Evaluation is based on a sample defined by the forecast targets from 2000.01 to 2013.03 (with forecast origins 1999.04 to 2013.02). In total, we consider 55 forecasts. As in the previous example involving the Bank of England’s forecasts, we consider an AR(1) benchmark utilised by Gneiting and Ranjan (2011).¹⁹

4.3 Results

As with the previous application, before undertaking the decision-theoretic evaluation, we begin by considering more conventional forecast metrics. In terms of point forecasting, the VAR’s RMSPE is 0.3868. In contrast, the AR(1) benchmark produces a RMSPE of 0.6137. A gain in forecast accuracy relative to the benchmark of just under 40%. A Diebold and Mariano (1995) test with the small sample adjustment suggested by Harvey, Leybourne and Newbold (1997) comfortably rejects the null hypothesis of equal forecast accuracy at the 1% significance level (in a two-sided test).²⁰

Looking at forecast evaluations based on the whole predictive density, the results are similar to the point forecasting case. The VAR produces an average (end of evaluation)

¹⁹Results are very similar with an AR(4) benchmark.

²⁰This commonly-used test should be interpreted as a rough guide to statistical significance in this context as the data are subject to revision and one component model of the VAR ensemble nests the benchmark.

logarithmic score of -0.7861. The corresponding figure for the benchmark AR(1) is -1.155. As a rough guide to statistical significance, our test of relative forecasting accuracy based on the differential in the average logarithmic scores (see Bao et al, 2007) rejects the null of equal forecast accuracy at the 3% level.²¹

We plot the forecasts produced by the VAR and the AR(1) benchmark specification in Figures 7 and 8. These display the (sequence of) one step ahead forecast densities and the NZ inflation realisations (solid line), together with the quantiles (0.01; 0.05; 0.10; 0.33; 0.66; 0.90; 0.95; 0.99). Like the Bank of England's forecasts, the one step ahead VAR forecast densities for inflation in New Zealand exhibit only mild asymmetries, even through the Great Recession. In general, for New Zealand data, the forecast densities from the VAR are considerably less diffuse than those from the AR(1) benchmark and diffusion alters little through the Great Recession. As a consequence of this, towards the end of the evaluation the benchmark model attributes small probabilities to negative inflation events, whereas the VAR forecasts imply that, for example, below zero inflation is a zero probability event throughout the evaluation.

Figures 9 and 10 display the PITs. In both cases, the forecast densities display departures from the uniform distribution. The sharpness of the VAR forecast densities result in too many realisations in the upper tail. The AR(1) benchmark suffers somewhat from the same characteristic—despite having relatively diffuse forecast densities.

Figure 11 plots the one step ahead probability of inflation events with realisations above the 3% threshold. The VAR gives a slightly earlier warning of inflationary episodes than the AR(1), with probabilities somewhat higher during these high inflation events.

Turning to the decision-theoretic based evaluation, we examine forecast performance from the perspective of issuing preemptive inflation event warnings. Recall that the (end

²¹Amisano and Giacomini (2007) discuss the limiting distribution of related test statistics. Here we ignore the uncertainty about the model parameters and combination weights.

of evaluation) Total Economic Loss, TEL , can be expressed as:

$$TEL = Ln_{10} + C(n_{01} + n_{00}).$$

We plot TEL against R for the VAR with the AR(1) benchmark normalised to 1 in Figure 12. Relative to the benchmark, the VAR specification performs well. The gain to the VAR over the benchmark, for values of R between 0.2 and 0.4, lies in the range 15% to 45%. That is, the relative TEL lies between 0.85 and 0.55. With R greater than 0.2, the performance advantage drops (non-monotonically) as R rises.

Hence, from the perspective of our communication framework for inflation targeting, the usefulness of the VAR exhibits sensitivity to the parameter R , which reflects the relative cost of anticipated and unanticipated inflation events. Whether a candidate forecast is “good” from an economic perspective depends on the critical probability that would trigger an inflation event warning by the central bank. Unfortunately, RMSPE and average logarithmic predictive scores abstract from the communication actions (and their relative costs) available to the policymaker in our framework. As a result, the sensitivity of forecast performance from an economic perspective would be missed when examining more conventional statistical measures of forecast performance.

5 Conclusions

Inflation targeting central banks devote considerable resources to modelling and assessing the probabilities of inflation events. Using two examples, we have shown that strong forecast performance by conventional statistical metrics is not sufficient for forecast performance differentials to be economically significant. The first example considered the UK inflation forecasts published by the Bank of England operating within a strict inflation targeting regime. The second example evaluated the NZ inflation forecasts generated from

vector autoregressions using a sample drawn from the period in which the Reserve Bank operated within a flexible inflation targeting framework. Forecast performance from an economic perspective varied with the degree of asymmetry in the costs associated with anticipated and unanticipated inflation events in both examples. More conventional statistical analyses of forecast performance, in both cases, would have misled a policymaker about the sensitivity of the forecast differential from an economic perspective. Specifically, for some parameters of the loss function used in our examples, the performance gains over the benchmark models were negligible from the perspective of an inflation targeting central bank.

APPENDIX 1: Output Gaps

We utilise a univariate unobserved components state-space model which decomposes a time series into trend and cyclical components, extending the analysis of (for example) Harvey and Jaeger (1993) to incorporate Bayesian estimation. As noted in the main text, this particular off-model filtering approach is often used by central banks in practice, including the Reserve Bank of New Zealand. Univariate filters are less efficient but offer smaller risk of specification error.²²

Decompose the $T \times 1$ vector of log GDP observations y_t into trend, cycle and idiosyncratic components given by μ_t , $\psi_{n,t}$ and ϵ_t respectively:

$$y_t = \mu_t + \psi_{n,t} + \epsilon_t \quad t = 1, \dots, T \quad (1)$$

where $\epsilon_t \sim NID(0, \sigma_\epsilon^2)$. The trend is specified as:

$$\mu_t = \mu_{t-1} + \beta_{t-1} \quad \beta_t = \beta_{t-1} + \zeta_t \quad (2)$$

with $\zeta_t \sim NID(0, \sigma_\zeta^2)$.

The cycle is defined:

$$\begin{bmatrix} \psi_{1,t} \\ \psi_{1,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{1,t-1} \\ \psi_{1,t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad (3)$$

$$\begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \sim NID \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\kappa^2 & 0 \\ 0 & \sigma_\kappa^2 \end{bmatrix} \right) \quad (4)$$

²²See Harvey et al (2007) for a multivariate unobserved components state-space approach estimated by Bayesian methods.

$$\begin{bmatrix} \psi_{1,t} \\ \psi_{1,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{1,t-1} \\ \psi_{1,t-1}^* \end{bmatrix} + \begin{bmatrix} \psi_{i-1,t-1} \\ \psi_{i-1,t-1}^* \end{bmatrix}, i = 2, \dots, n \quad (5)$$

where κ_t and κ_t^* are uncorrelated noise disturbances with the same variance, σ_κ^2 . The parameter ρ is a damping factor, with $0 < \rho < 1$. A smaller damping factor implying a less amplified cycle (for a given σ_κ^2). The parameter λ controls the frequency of the cycle. For example, λ equals 0.3 would be a cycle of 21 quarters. We assume that the cycle is second order in our application, $n = 2$.

We estimate the two cyclical parameters ρ and λ , and three variance parameters σ_ϵ , σ_ζ , and σ_κ using Bayesian methods to produce posterior distributions for each parameter, and then draw from these distributions to produce estimates of the cyclical component of (log) GDP. We iterate the MCMC algorithm 50,000 times, and burn the first 48,000.

Our priors are similar to those utilised by Harvey et al (2007) in their analysis of US GDP for a second-order cycle. The posterior means for the parameters 0.64, 0.29, 11.1, 4.9 and 9.8 for ρ , λ , σ_ϵ , σ_ζ , and σ_κ , respectively. Using the posterior means, the cyclical component constructed from this process is similar to that constructed using an HP filter with lambda equal to 1600. We also experimented with a prior specification that allowed for greater flexibility in the trend. This variant produced a similar forecast performance for New Zealand inflation in bivariate vector autoregressions.

APPENDIX 2: Ensemble Forecasting with Vector Autoregressions

The forecasts for New Zealand inflation are generated from an ensemble of vector autoregressions (VAR) following Garratt et al (2011), using “real-time” data.

Since the analysis in this paper is concerned with one step ahead forecasts only, the output gap forecasts do not influence the forecast for inflation. Hence, we consider only the inflation equation of each VAR, which can be written:

$$\pi_t = \alpha_{1,p} + \sum_{p=1}^P \beta_{1,p} \pi_{t-p} + \sum_{p=1}^P \gamma_{1,p} y_{t-p} + \varepsilon_{1,p,t}, \quad (6)$$

where inflation is defined as the log difference in the price level and the output gap (cyclical component of log GDP) is denoted y_t .

We allow for the maximum number of lags, p , to vary between 1 and P , where P equals 4. With four VAR forecasts in each ensemble, we construct the predictive densities using a linear opinion pool. The ensemble densities for inflation are defined by the convex combination:

$$p(\pi_t) = \sum_{p=1}^P w_{p,t} g(\pi_t | I_{p,t}), \quad t = \underline{t}, \dots, \bar{t}, \quad (7)$$

where $g(\pi_t | I_{p,t})$ are the 1-step ahead forecast densities from model p , $p = 1, \dots, 4$ of inflation π_t , conditional on the information set I_t , and the evaluation period runs from \underline{t} to \bar{t} . That is, from (forecast targets) 2000.01 to 2013.03 in our New Zealand forecasting application. The publication delay in the production of real-time data ensures that this information set contains lagged variables, here assumed to be dated $t - 1$ and earlier. The non-negative weights, $w_{p,t}$, in this finite mixture sum to unity. Since each VAR produces a forecast density that is Student-t, the combined density defined by the linear opinion pool is a mixture.

We use the logarithmic score to measure density fit for each individual VAR specifica-

tion. The logarithmic scoring rule assigns a high score to a density forecast with a high probability for the realised value and can be interpreted as a measure of the Kullback-Leibler distance. The logarithmic score of the i^{th} density forecast, $\ln g(\pi_t^o | I_{p,t})$, is the logarithm of the probability density function $g(\cdot | I_{p,t})$, evaluated at the outturn π_t^o .

The recursively constructed weights for the VAR ensemble are then given by:

$$w_{p,t} = \frac{\exp \left[\sum_{\underline{t}}^{t-1} \ln g(\pi_t | I_{p,t}) \right]}{\sum_{p=1}^4 \exp \left[\sum_{\underline{t}}^{t-1} \ln g(\pi_t | I_{p,t}) \right]}, \quad t = \underline{t}, \dots, \bar{t}. \quad (8)$$

References

Aastveit, K.A., K. Gerdrup, A.S. Jore and L.A. Thorsrud (2014) “Nowcasting GDP in real time: A density combination approach”, *Journal of Business and Economic Statistics*, 32(1), p 48-68.

Amisano, G. and R. Giacomini (2007) “Comparing density forecasts via likelihood ratio tests”, *Journal of Business and Economic Statistics*, 25(2), p 177-190.

Adolfson, M., M.K. Andersson, J. Linde, M. Villani and A. Vredin (2007) “Modern forecasting models in action: improving macroeconomic analyses at central banks”, *International Journal of Central Banking*, 3(4), p 111-144.

Bank of England *Inflation Report*, various issues.

Bao, Y., T-H. Lee and B. Saltoglu (2007) “Comparing density forecast models”, *Journal of Forecasting*, 26, p 203-225.

Berge, T. and O. Jorda (2011) “The classification of economic activity into expansions and recessions”, *American Economic Journal: Macroeconomics*, 3(2), p 246-277.

Berrocal, V., A.E. Raftery, T. Gneiting and R. Steed (2010) “Probabilistic weather forecasting for winter road maintenance”, *Journal of the American Statistical Association*, 105, p 522-537.

Bjørnland, H., K. Gardrup, A.S. Jore, C. Smith, L.A. Thorsrud (2011) “Weights and pools for a Norwegian density combination”, *The North American Journal of Economics and Finance*, 22(1), p 61-71.

Clark, T.E. (2011) “Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility”, *Journal of Business and Economic Statistics*, 29(3), p 327-341.

Clements, M.P. (2004) “Evaluating the Bank of England density forecasts of inflation”, *The Economic Journal*, 114, p 844-866.

Coe, P. and S.P. Vahey (2015) “Predicting the onset of the Great Recession with historical experts”, unpublished manuscript.

Coenen, G. and A. Warne (2014) “Risks to price stability, the zero lower bound, and forward guidance: A real-time assessment”, *International Journal of Central Banking*, June, p 7-54.

Conflitti, C., C. De Mol and D. Giannone (2015) “Optimal combination of survey forecasts”, *International Journal of Forecasting*, forthcoming.

Del Negro, M., R.B. Hasegawa and F. Schorfheide (2014) “Dynamic prediction pools: An investigation of financial frictions and forecasting performance”, Federal Reserve Bank of New York Staff Report 695.

Diks, C., V. Panchenko and D. van Dijk (2011) “Likelihood-based scoring rules for comparing density forecasts in tails”, *Journal of Econometrics*, 163, p 215-230.

Diebold, F.X. and R.S. Mariano (1995) “Comparing predictive accuracy”, *Journal of Business and Economic Statistics*, 13, p 253-263.

Edge, R.M. and J.B. Rudd (2012) “Real-time properties of the Federal Reserve’s output gap”, Discussion Paper No 86, Federal Reserve Board, Washington, D.C.

Franta, M., J. Barunik, R. Horvath and K. Smidkova (2014) “Are Bayesian fan charts useful? The effect of the zero lower bound and evaluation of financial stability stress tests”, *International Journal of Central Banking*, March, p 7-54.

Galbraith, J.W. and S. van Norden (2011) “Kernel-based calibration diagnostics for recession and inflation probability forecasts”, *International Journal of Forecasting*, 27(4), p 1041-1057.

Galbraith, J.W. and S. van Norden (2012) “Assessing gross domestic product and inflation probability forecasts derived from Bank of England fan charts”, *Journal of the Royal Statistical Society Series A - Statistics in Society*, 175 (3), p 713-727.

Garratt, A., K. Lee, M.H. Pesaran and Y. Shin (2003) “Forecast uncertainties in macroeconomic modelling: an application to the UK economy”, *Journal of the American Statistical Association*, 98, p 829-838.

Garratt, A., G. Koop, E. Mise and S.P. Vahey (2009) “Real-time prediction with UK monetary aggregates in the presence of model uncertainty”, *Journal of Business and Economic Statistics*, 27(4), p 480-491.

Garratt, A., J. Mitchell, S.P. Vahey and E.C. Wakerly (2011) “Real-time inflation forecast densities from ensemble Phillips curves”, *The North American Journal of Economics and Finance*, 22(1), p 77-87.

Garratt, A., S.P. Vahey and E.C. Wakerly (2012) “Forecasting the probability of exceeding the US debt ceiling”, presented at the SNDE Annual Symposium.

Garratt, A., J. Mitchell and S.P. Vahey (2014) “Probabilistic forecasting for the Federal Reserve”, Warwick Business School, Working Paper No 7.

Gneiting, T. and R. Ranjan (2011) “Comparing density forecasts using threshold- and quantile-weighted proper scoring rules”, *Journal of Business and Economic Statistics*, 29(3), p 411-422.

Granger, C.W.J. and M.H. Pesaran (2000a) “Economic and statistical measures of forecast accuracy”, *Journal of Forecasting*, 19, p 537-560.

Granger, C.W.J. and M.H. Pesaran (2000b) “A decision-based approach to forecast evaluation” in W.S. Chan, W.K. Li and H. Tong (Eds.), *Statistics and Finance: An Interface*, London: Imperial College Press.

Greenspan, A. (2004) “Risk and uncertainty in monetary policy”, *American Economic Review*, Papers and Proceedings, May, p 33-40.

Harvey, A.C. and A. Jaeger (1993) “Detrending, stylised facts and the business cycle” *Journal of Applied Econometrics*, 8(3), p 231-247.

Harvey, A.C., T.M. Trimbur and H.K. van Dijk (2007) “Trends and cycles in economic time series: A Bayesian approach”, *Journal of Econometrics* 140(2), p 618-649.

Harvey, D., S. Leybourne and P Newbold (1997) “Testing the equality of prediction mean squared errors”, *International Journal of Forecasting*, 13, p 281-291.

Katz, R.W. and A.H. Murphy (1997) “Forecast value: prototype decision-making models”, in R.W. Katz and A.H. Murphy (Eds.), *Economic value of Weather and Climate Forecasts*, Cambridge University Press.

Kruger, F., T. Clark and F. Ravazzolo (2015) “Using entropic tilting to combine BVAR forecasts with external nowcasts”, Working Paper No 1439, Federal Reserve Bank of Cleveland, *Journal of Business and Economic Statistics*, forthcoming.

Kryshko, M., F. Schorfheide and K. Sill (2010) “DSGE model-based forecasting of non-modelled variables”, *International Journal of Forecasting*, 26(2), p 348-373.

Labbé F. and H. Pepper (2009) “Assessing recent external forecasts”, Reserve Bank of New Zealand Bulletin, 74(4), p 19-25.

Lahiri, K. and J.G. Wang (2007) “The value of probability forecasts as predictors of cyclical downturns”, *Applied Economics Letters*, 14(1), p 11-14.

Luciani, M. and L. Ricci (2015) “Nowcasting Norway”, *International Journal of Central Banking*, 10(4), p 215-248.

Murphy, A.H. (1977) “The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation”, *Monthly Weather Review*, 105, p 803-816.

Orphanides, A. and S. van Norden (2002) “The unreliability of output-gap estimates in real time”, *The Review of Economics and Statistics*, 84(4), p 569-583.

Orphanides, A. and S. van Norden (2005) “The reliability of inflation forecasts based on output gap estimates in real time”, *Journal of Money, Credit and Banking*, 37(3), p 583-601.

Reserve Bank of New Zealand *Monetary Policy Statement*, various issues.

Reserve Bank of New Zealand *Policy Targets Agreement*, <http://www.rbnz.govt.nz/>.

Turner, J. (2006) “An assessment of recent Reserve Bank forecasts”, *Reserve Bank Bulletin*, 69(3), p 38-43.

Wallis, K.F. (2003) “Chi-squared tests of interval and density forecasts, and the Bank

of England's fan charts", *International Journal of Forecasting*, 19, p 165-175.

Wallis, K.F. (2004) "An assessment of Bank of England and National Institute inflation forecast uncertainties", *National Institute Economic Review*, No 189, p 64-71.

Wallis, K.F. (2005) "Combining density and interval forecasts: a modest proposal", *Oxford Bulletin of Economics and Statistics*, 67, p 983-994.

Figure 1: Bank Forecasts and UK Inflation

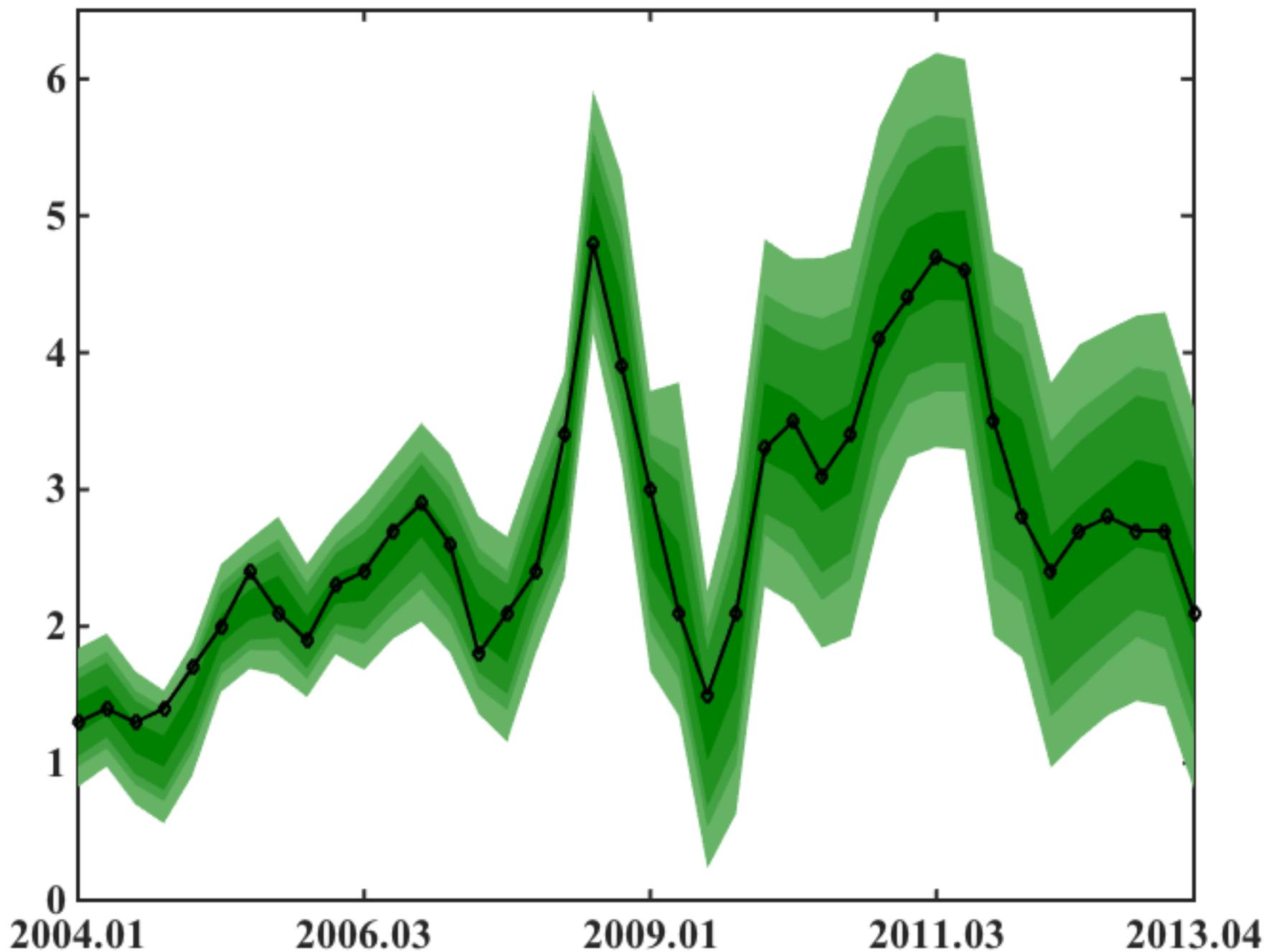


Figure 2: PITs, Bank of England

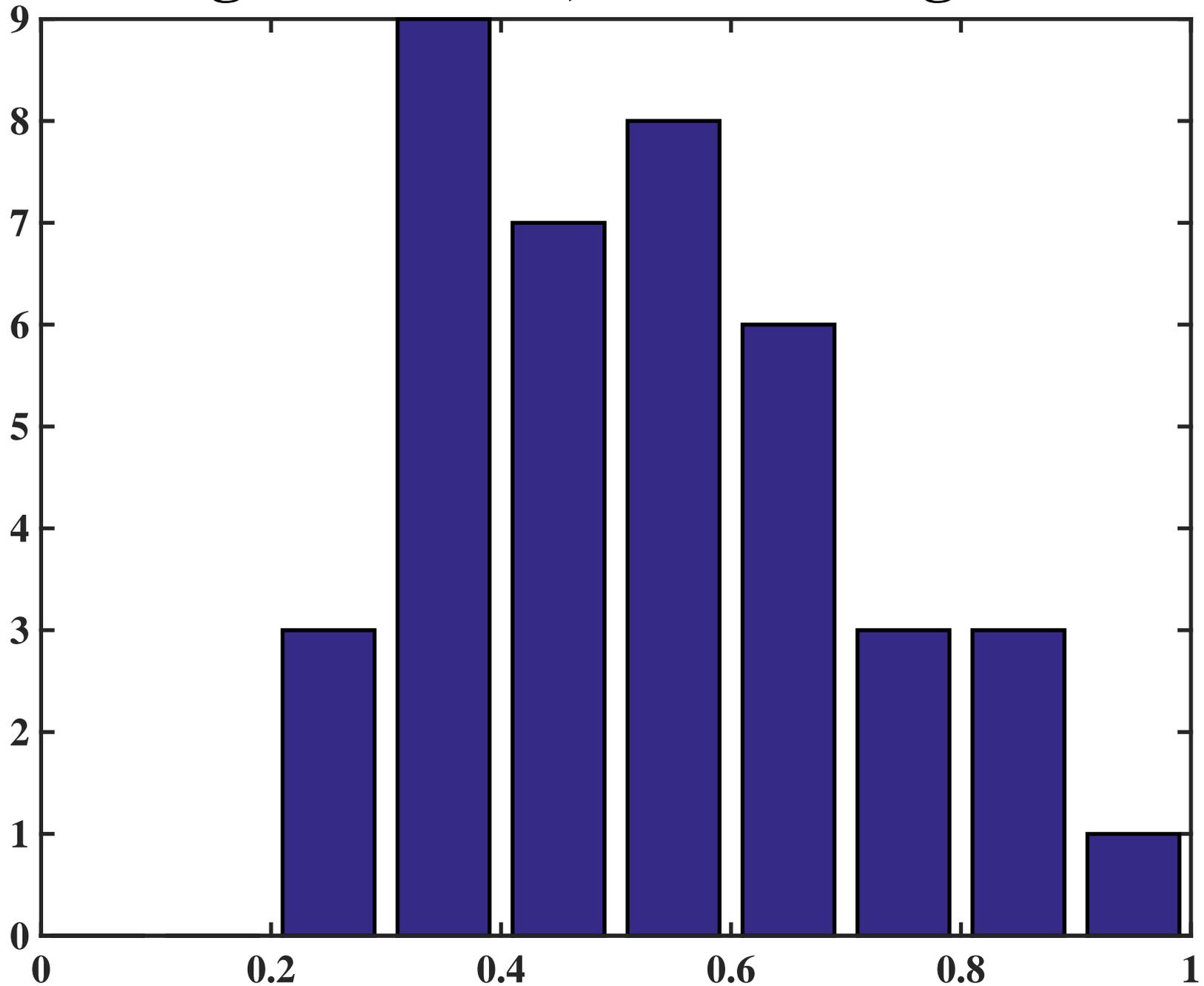


Figure 3: AR(1) Forecasts and UK Inflation

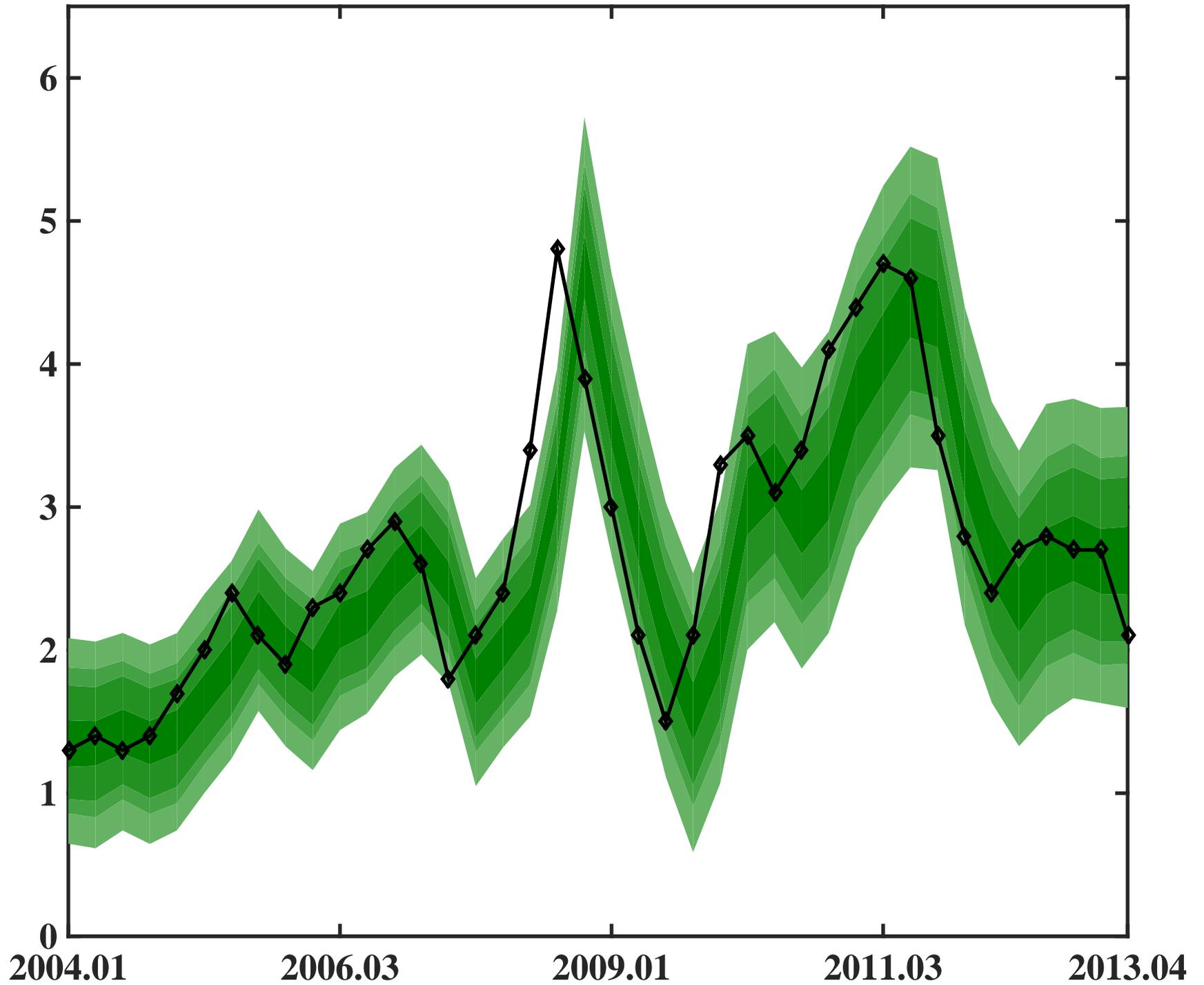


Figure 4: PITs, AR(1)

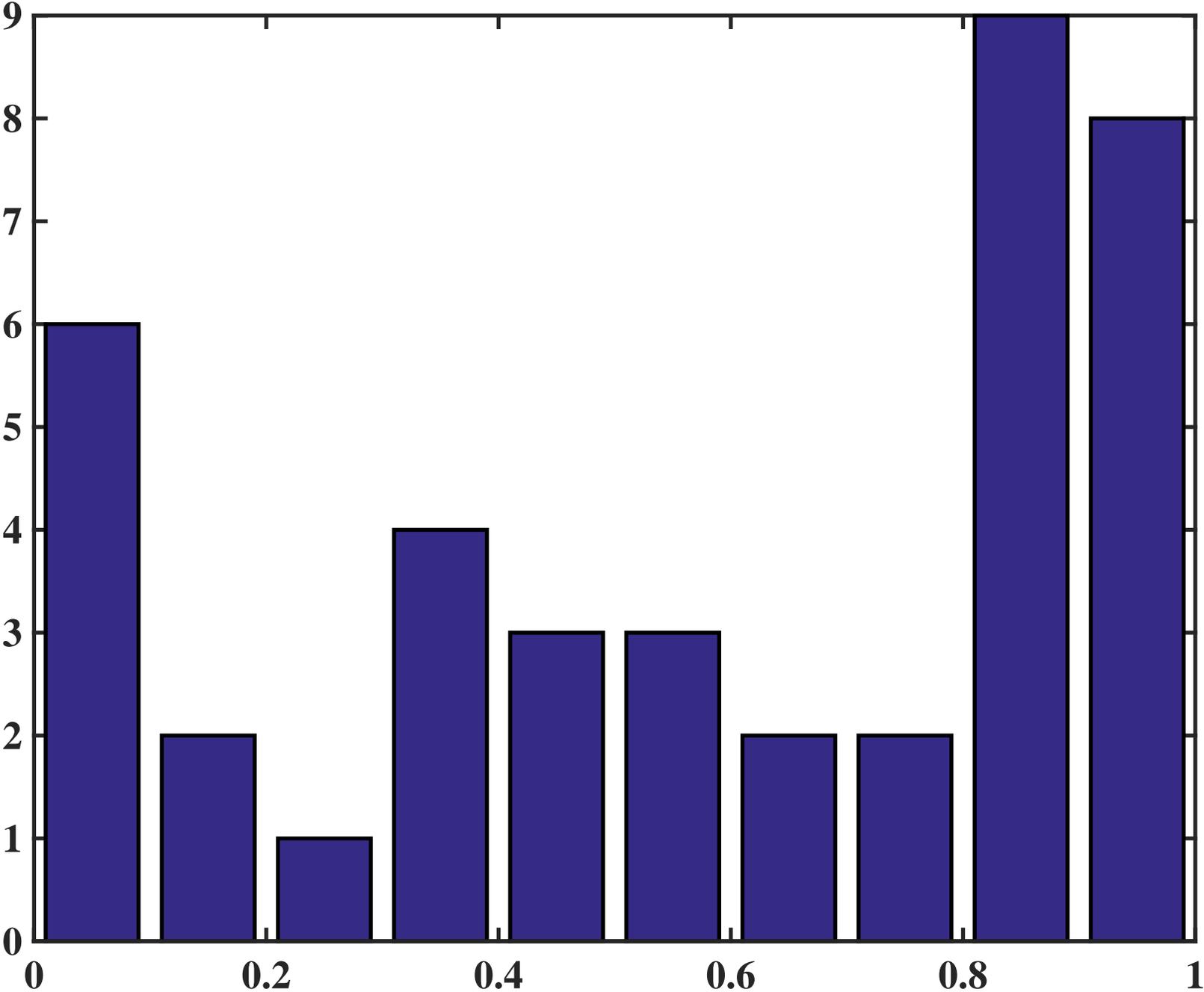


Figure 5: Total Economic Loss (TEL), UK

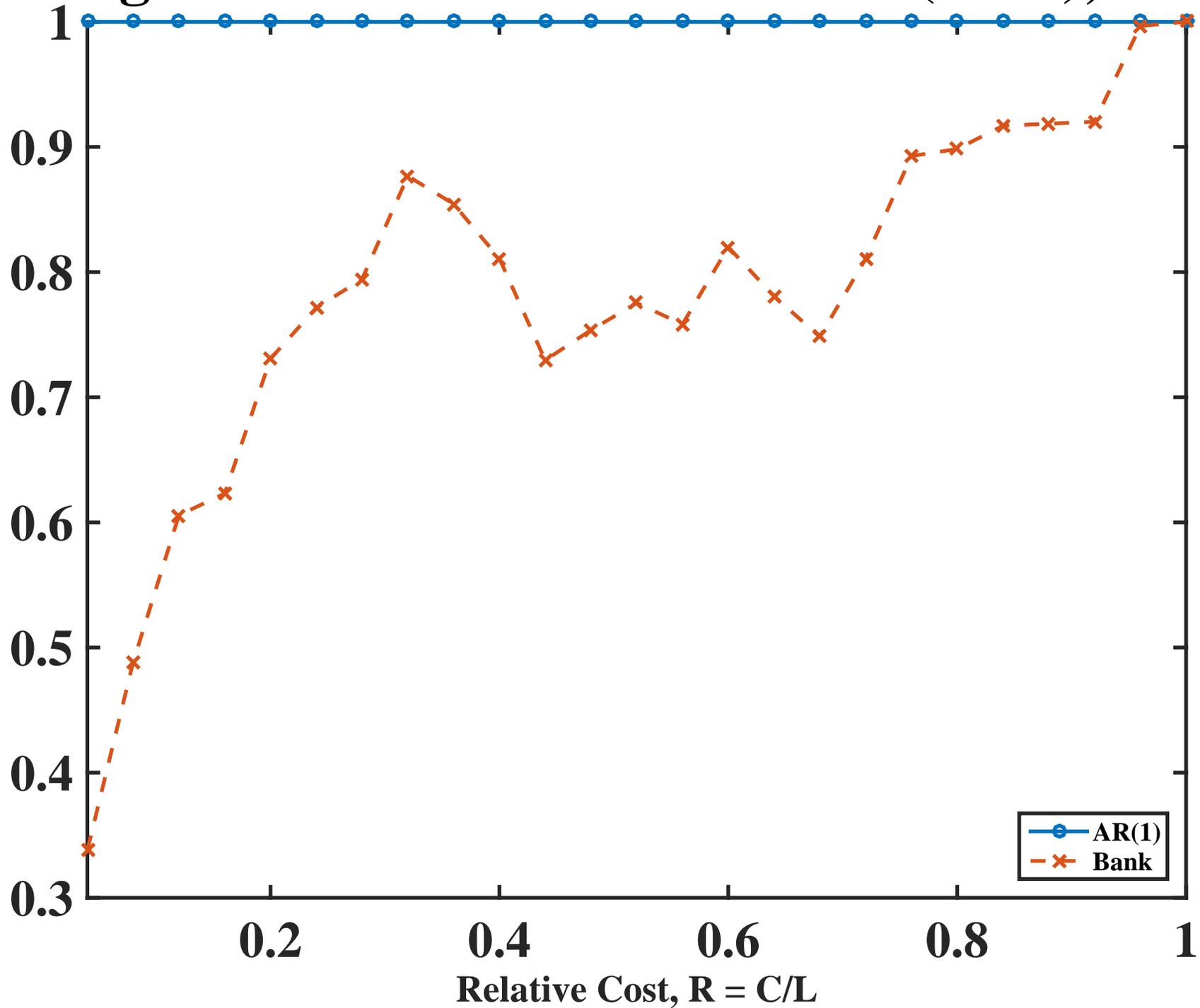


Figure 6: Probability of UK Inflation > 3%

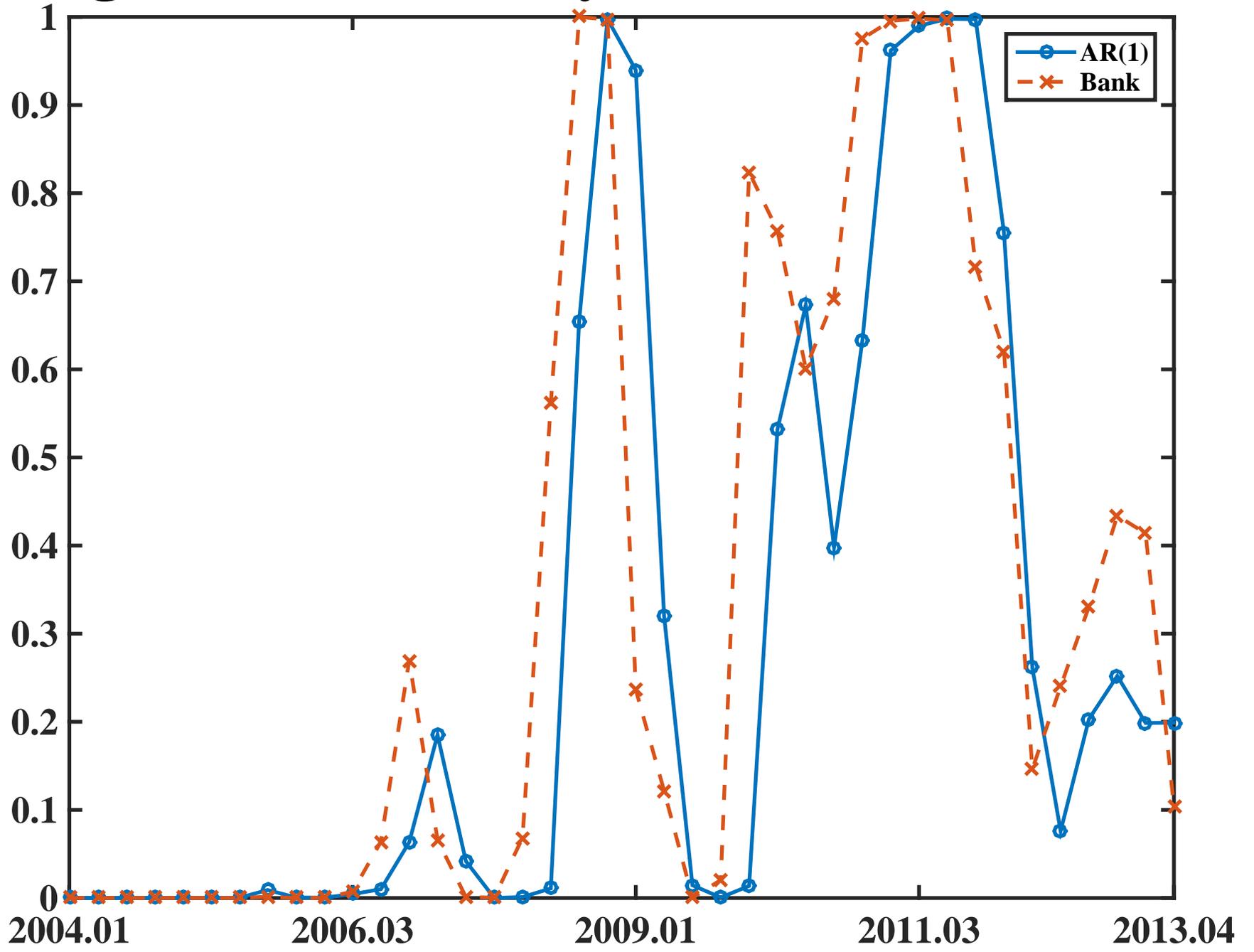


Figure 7: VAR Forecasts and NZ Inflation

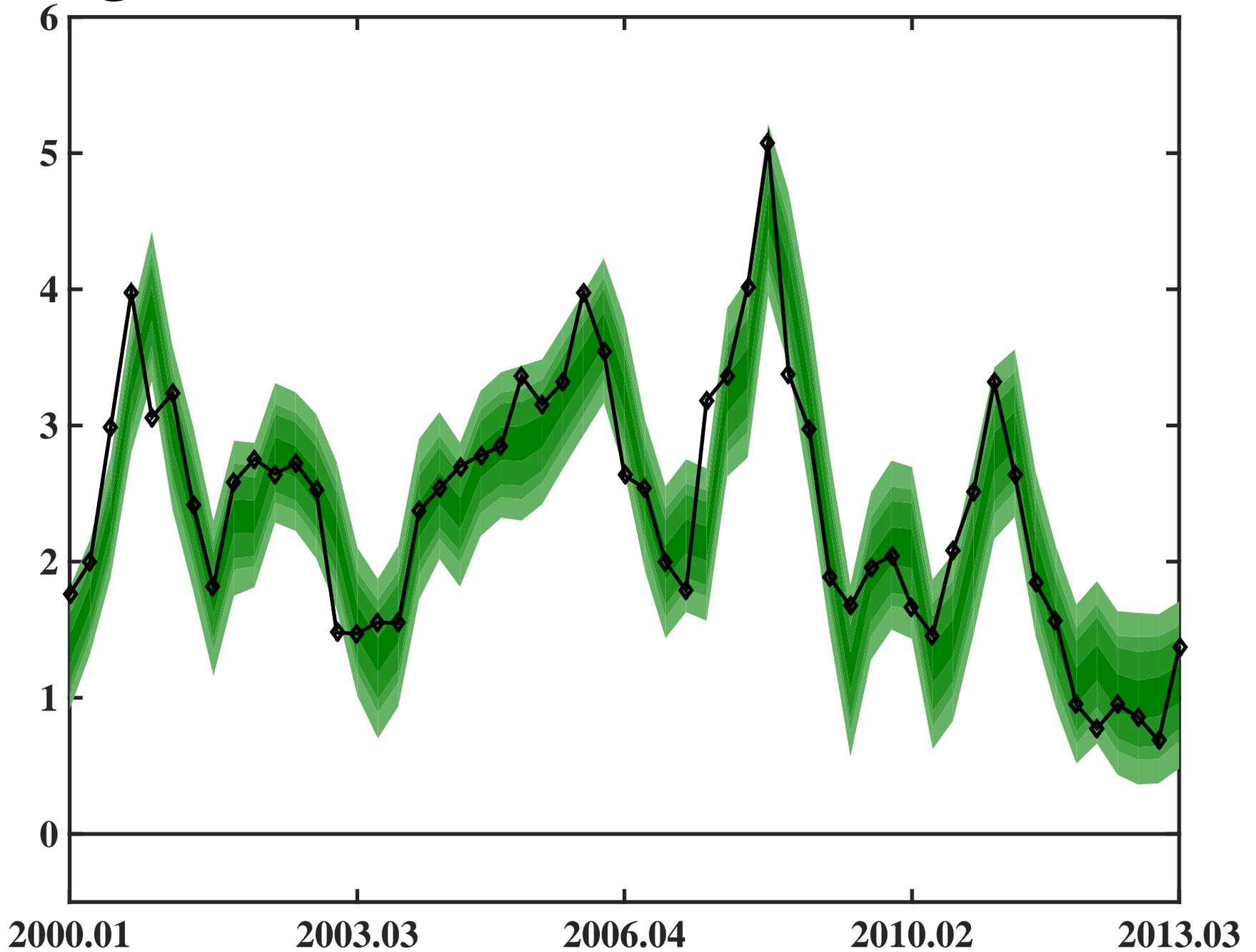


Figure 8: AR(1) Forecasts and NZ Inflation

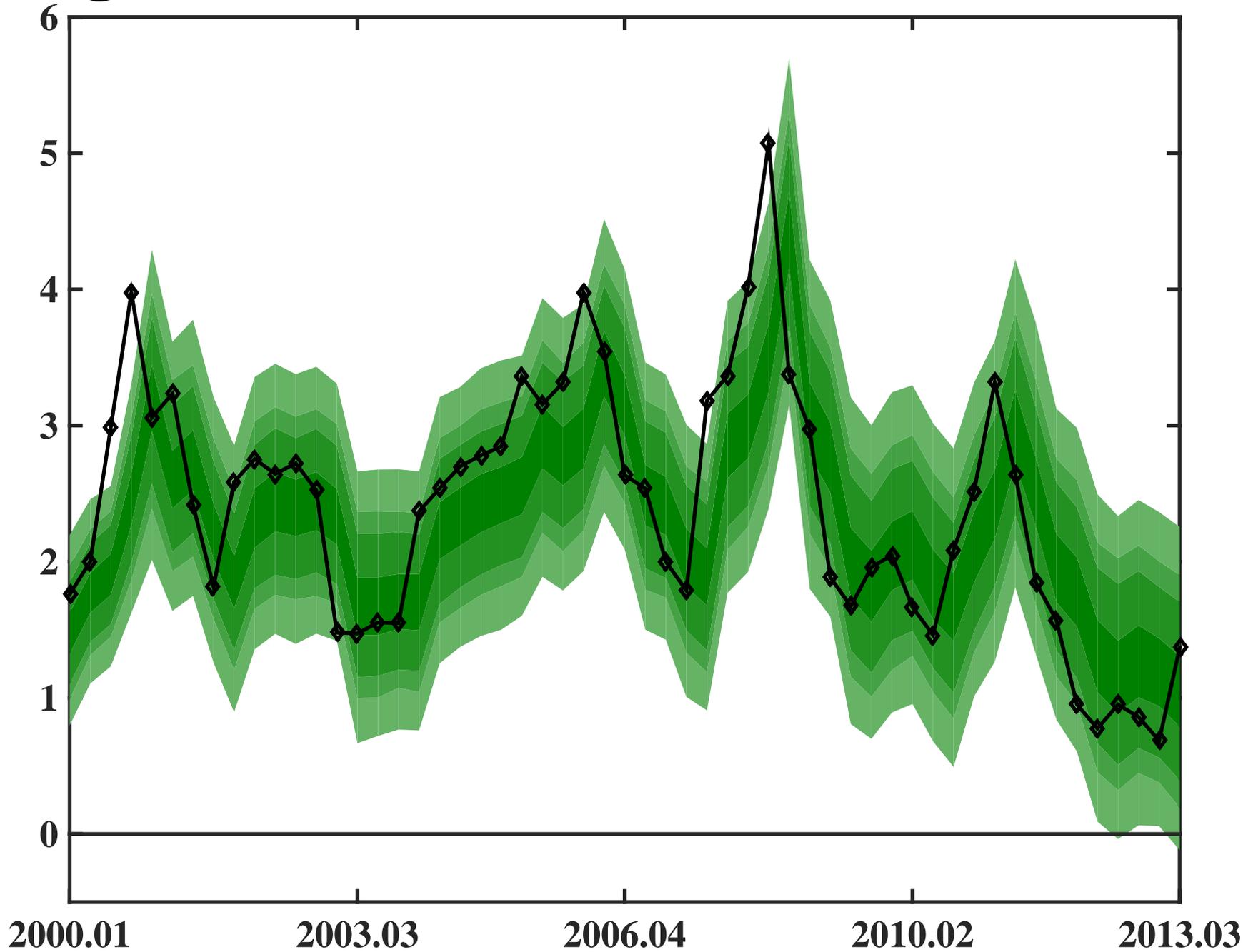


Figure 9: PITs, VAR

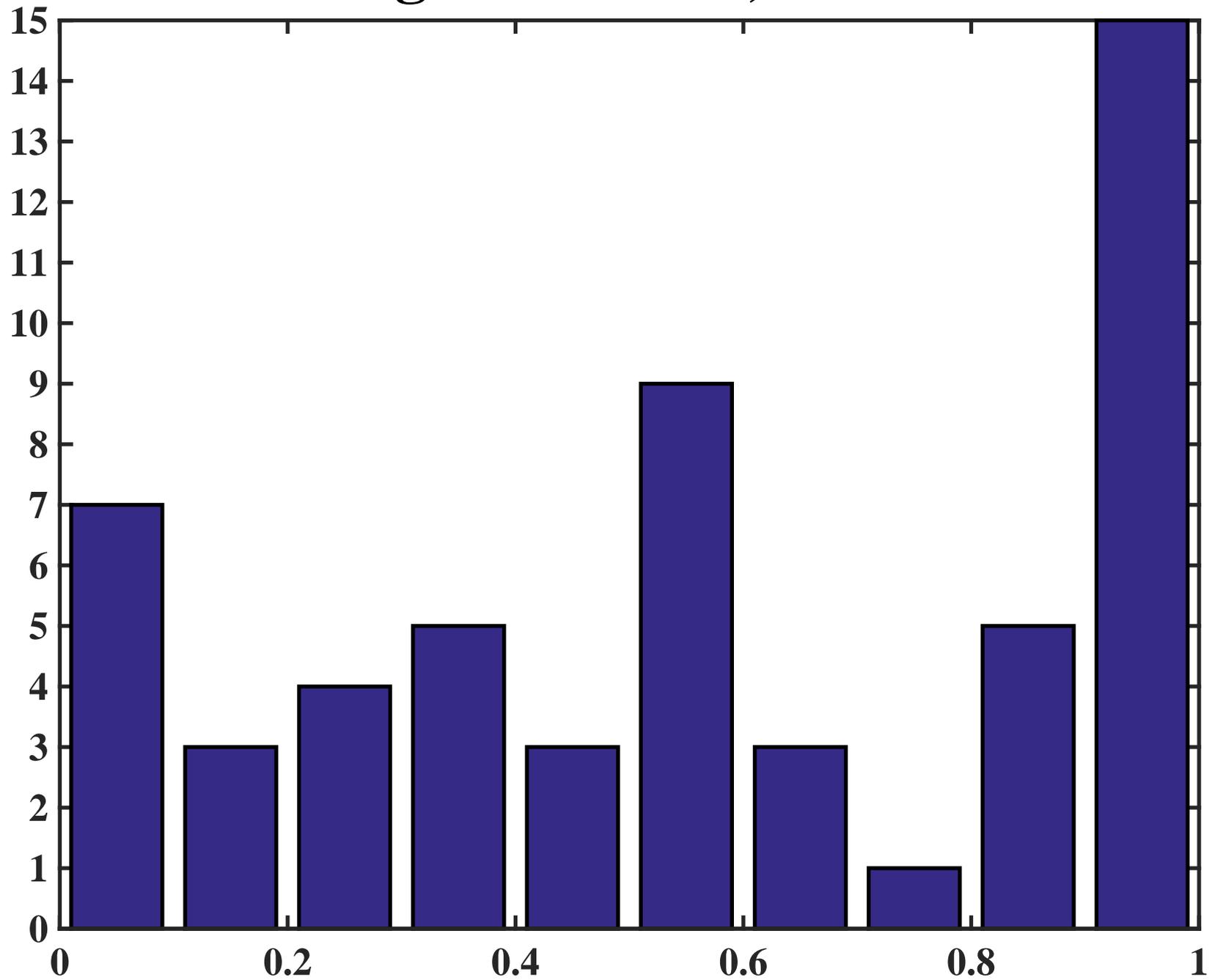


Figure 10: PITs, AR(1)

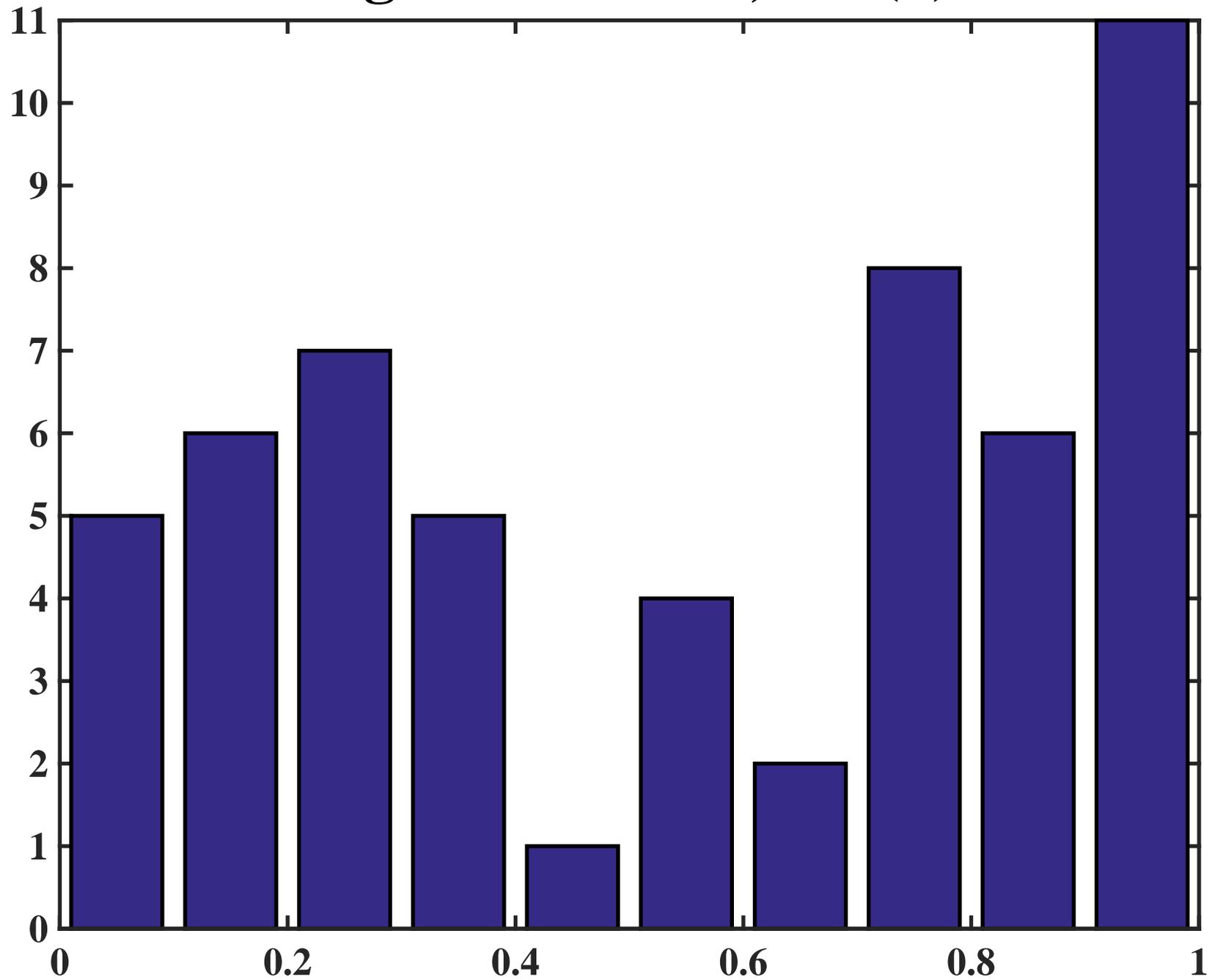


Figure 11: Probability of NZ Inflation > 3%

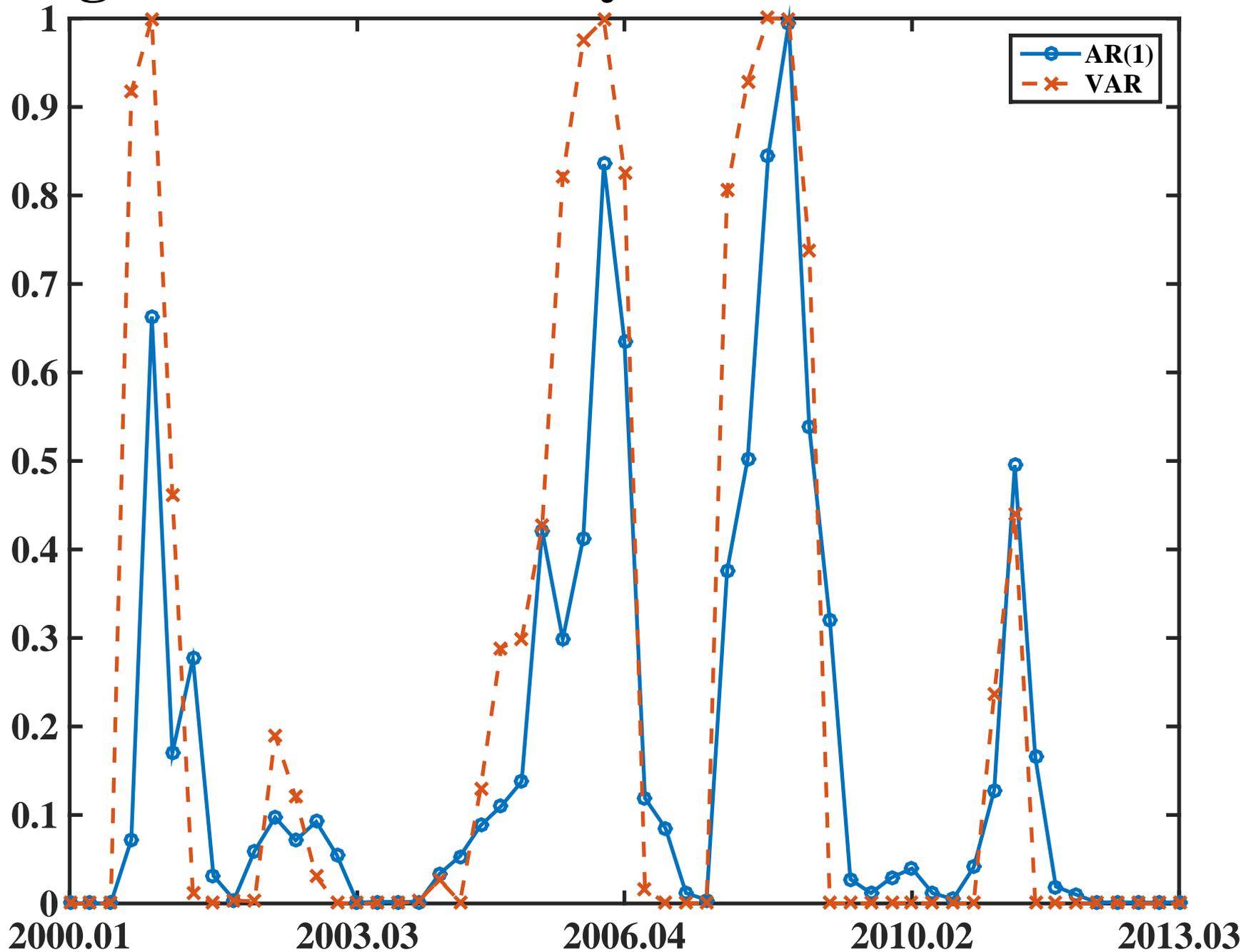


Figure 12: Total Economic Loss (TEL), NZ

