

Sveriges Riksbank's inflation interval forecasts 1999–2005*

Michael Lundholm

Department of Economics, Stockholm University

June 21, 2010

Abstract

Are Sveriges Riksbank's inflation (CPI and KPIX) interval forecasts calibrated in the sense that the intervals cover realised inflation with the stated *ex ante* coverage probabilities 50, 75 and 90 percent? In total 150 interval forecast 1999:Q2–2005:Q2 are assessed for CPI and KPIX. The main result is that the forecast uncertainty is understated, but there are substantial differences between individual forecast origins and inflation measures.

JEL: C53, E31 and E37.

Keywords: Inflation, forecast, interval forecast, forecast uncertainty.

1 Introduction

Since 1993 the objective of Sveriges Riksbank (the central bank of Sweden) has been price stability. In 1995 a self-adopted explicit inflation target of a yearly 2 percent increase in consumer prices with a tolerance band of ± 1 was implemented. In order to determine whether a change in mainpolicy rate is necessary the Bank forecasts the inflation rate measured as changes in consumer price index (CPI) and underlying inflation (KPIX; before November 12, 2007 called UND1X) with 2 year forecast horizon. Since 1999 these point forecasts are accompanied with interval forecasts for 50, 75 and 90 percent coverage with a forecast horizon of 1 – 25 months.¹ Until 2005:Q2 these

*Correspondence to Michael Lundholm, email <mlu@ne.su.se>. Address: Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden. I am grateful for comments from Jim Albrecht, Mahmood Arai, Martin Flodén, Matthew Lindquist and Peter Thoursie as well as from seminar participants at Sveriges Riksbank.

¹Since only sometimes a 26 month horizons was forecasted, these will not be used in the study.

forecasts where made four times a year and conditional on an unchanged policy rate.²

Sveriges Riksbank applies a judgemental inflation forecast procedure which is described in Blix and Sellin (1998, 1999) and briefly in Berg (2000). But whereas the corresponding point forecasts have been evaluated in several studies there is hitherto no assessment about the reliability of the interval forecast.³ The purpose of this study is therefore to throw light on the question whether these interval forecasts are appropriately calibrated. That is, are the interval forecasts for (say) 50 percent coverage actually contain actual inflation in 50 percent of the cases?

These forecasts constitute a data set consisting of 25 forecast origins with three interval forecasts at each origin (50, 75 and 90 percent coverage) for two measures of inflation (CPI and KPIX). Each forecast has a horizon of 25. This sums up to 150 individual forecasts each of which with 25 observations, or if all forecast origins are pooled 6 forecasts with 625 observations in each.

The most apparent strategy would be to employ the test procedure suggested by Christoffersen (1998). For each forecast an indicator variable is created, taking the value one when actual inflation is inside the forecasted interval and else zero. Then the following tests are performed: (i) a test for unconditional coverage, (ii) for independence (serial correlation) and (iii) for conditional coverage (testing coverage but conditioning on independence).

In this study these three tests are performed on each of the 150 interval forecast series. Since the independence and conditional coverage tests require that a series of conditional probabilities can be calculated, the corresponding conditioning events must occur at least once in the data; if not the test cannot be performed. As a consequence there will be some missing values on the series of test statistics. This will actually occur, which indicates 25 observations are too few to guarantee that all events actually are observed. Therefore, even if all events are observed, the number of observations may be too small for the independence and conditional coverage tests to be reliable; see Andrés and Tejedor (2003). This is the motivation to pool the forecast origins into 6 forecast series and only perform the test for unconditional coverage.

The main result of the present study using pooled data is that the interval forecasts of Sveriges Riksbank have a tendency to understate the forecast uncertainty. The reason for failing the test is that the forecasted intervals

²For a recent review of forecast based monetary policies see Woodward (2007).

³Recently Lundholm (2010a) showed using the maximum entropy bootstrap for inference that forecasts for horizons longer than one year tended to have a positive bias. Previous evaluations, e.g., Blix *et al.* (2002), Konjunkturinstitutet (2002), Jansson and Vredin (2003), Bergvall (2005) and Andersson *et al.* (2007) (where only Konjunkturinstitutet (2002) and Bergvall (2005) did not formally originate from Sveriges Riksbank although the staff at the National Institute of Economic Research responsible for Konjunkturinstitutet (2002) had a professional background at Sveriges Riksbank) showed that Sveriges Riksbank was no worse or better than other forecasters in terms of bias.

are too narrow; Sveriges Riksbank underestimates the forecast uncertainty.

Today there is a small but growing literature assessing different type of uncertainty estimates in the inflation forecasts produced by central banks. Using 2-year-ahead *probability forecasts* published by Sveriges Riksbank for a part of the above time period, Dowd (2004) concludes that they understate the forecast uncertainty. Wallis (2003, 2004) and Clements (2004) assess the forecasts of the Bank of England’s Monetary Policy Committee (MPC) using the published information about the two-piece normal distribution in the density forecasts. The main result is that the MPC overstates the forecast uncertainty.

This paper is organised as follows. In section 2 the data is described. Section 3 describes the framework for assessment used. In section 4 the indicator data set used in the study is defined from the original data. Section 5 contains the results and section 6 concludes the paper.

All estimations were performed using R version 2.11.1 (2010-05-31).⁴ Lundholm (2010d) is a technical documentation accompanying this paper. It contains the econometric code with comments, detailed information about versions of the econometric software and packages and a more detailed presentation of the results.

2 Data

Data (interval forecasts and corresponding outcomes) are publicly available as the R package **sifds** (Lundholm, 2010b). The package and its availability as well as the data and its sources are described in greater detail in Lundholm (2010c).

In package **sifds** data consists of yearly inflation rate point and interval forecasts from Sveriges Riksbank.⁵ Inflation rates are measured as relative 12 month change in CPI (the standard consumer price index) as well as KPIX (CPI with temporary effects excluded) for each month during the forecasted time period 1999:M05–2005:M07; see Figure 1 for the inflation rates during the forecasted period. The forecasts were published with a 3 month interval between the origins during 1999:Q2–2005:Q2.⁶ The forecast horizons were from 1 month up to 25 or 26 months.

⁴See R Development Core Team (2010).

⁵The package also contains point forecasts, but no interval forecasts, from Konjunkturinstitutet (the National Institute of Economic Research).

⁶The forecasts have continued after 2005:Q2, but the data set does not include these because Sveriges Riksbank changed (i) a basic assumption underlying the forecasts and (ii) the number of forecast origins (i.e. the frequency) from 4 to 3 each year. The basic assumption that changed in 2005 was that rather assuming that their main policy rate would remain unchanged over the forecast horizon Sveriges Riksbank started to forecast its own policy rate.

3 Framework for assessment

The log likelihood test procedure is described in detail in Christoffersen (1998) and reviewed by Wallis (2003). An interval forecast is an interval $[L_t(p), U_t(p)]$, where $L_t(p)$ and $U_t(p)$ are the lower and upper limits of the point forecast for the coverage probability $p \in \{0.5, 0.75, 0.9\}$, where $t = 1, \dots, n$, is the the date of the forecasted event given the forecast origin. In the case of pooled data t is an index identifying the forecast origin and the date of the forecasted event. Let y_t be the observed inflation at date t . A sequence of indicators $\{I_t(p)\}$ is then created such that $I_t = 1$ if $y_t \in [L_t(p), U_t(p)]$ and else $I_t = 0$.

Let the subsequence of indicators $\langle I_t, I_{t-1} \rangle$ be a subset of $\{I_t\}$ and let n_{ij} be the number of subsequences such that $\langle I_t, I_{t-1} \rangle = \langle j, i \rangle \forall i, j = 0, 1$. Since we in order to make this definition lose one observation (the first) we have $n_1 = n_{01} + n_{11} = \sum_{t=2}^n I_t$ as the number of outcomes inside the interval and $n_0 = n_{00} + n_{10} = (n - 1) - n_1$ as the number of observations outside the interval. Finally, denote the sample mean of the indicator variable $\pi = n_1 / (n_0 + n_1)$. Similarly, let $\pi_{ij} = n_{ij} / n_i = n_{ij} / (n_{i0} + n_{i1})$.

The definition of correct coverage is then that $\{I_t(p)\} \stackrel{iid}{\sim} \text{Bern}(p) \forall t$; see Christoffersen (1998, Lemma 1 and Definition 1). Three tests are performed; (i) unconditional coverage, (ii) independence and (iii) coverage conditional on independence.

In the unconditional test we test $H_0 : \pi = p$ against $H_1 : \pi \neq p$. The test statistic is then the log-likelihood ratio

$$\text{LR}_{\text{uc}} = -2 \log \left(\frac{(1-p)^{n_0} p^{n_1}}{(1-\pi)^{n_0} \pi^{n_1}} \right) \sim \chi^2(1). \quad (1)$$

If the test statistic is sufficiently large, i.e. $\text{LR}_{\text{uc}} > \chi_{0.95}^2(1) \approx 3.84$, then we reject null hypothesis and conclude that the width of the interval forecast are not appropriately calibrated. This will happen if π deviates too much from p in either direction. Note that the test is only defined if we have observations of both types (inside and outside the forecast intervals); i.e. $\pi \notin \{0, n-1\}$.

To control for (first order) autocorrelation an independence test is made where we test $H_0 : \{I_t(p)\}_{t=1}^n$ is independent with the probability of observing an inflation outcome inside the interval π and outside the interval $1 - \pi$ against $H_1 : \{I_t(p)\}_{t=1}^n$ is a Markov chain with transition probability matrix given by $\pi_{ij} \forall i, j \in \{0, 1\}$. The test statistic is

$$\text{LR}_{\text{ind}} = -2 \log \left(\frac{(1-\pi)^{n_{00}+n_{10}} \pi^{n_{01}+n_{11}}}{(1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}^{n_{11}}} \right) \sim \chi^2(1), \quad (2)$$

with critical value for the statistic $\chi_{0.95}^2(1) \approx 3.84$. Note that the test is only defined if $n_{ij} \notin \{0, n-1\} \forall i, j \in \{0, 1\}$.

Conditioning the coverage test on independence, we test $H_0 : \{I_t(p)\}_{t=1}^n$ is independent with probability of observing an inflation outcome inside the interval p and outside the interval $1 - p$ against the alternative hypothesis of the independence test above.

$$\text{LR}_{\text{cc}} = -2 \log \left(\frac{(1-p)^{n_{00}+n_{10}} p^{n_{01}+n_{11}}}{(1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}^{n_{11}}} \right) \stackrel{asy}{\sim} \chi^2(2), \quad (3)$$

with critical value for the statistic $\chi_{0.95}^2(2) \approx 5.99$. Given that we have conditioned on the first observation we have $\text{LR}_{\text{cc}} = \text{LR}_{\text{uc}} + \text{LR}_{\text{ind}}$.

4 Indicator data sets

We will construct two different types of data sets:

1. Each of the 25 forecast origins (1999:Q2–2005:Q2), for each of the two inflation measures (CPI and KPIX), for each of the three *ex ante* coverage probabilities (50, 75 and 90 percent) and with forecast horizons 1 – 25 months is treated as a separate data set. The forecast horizon 26 months ahead is not used since it was not forecasted by Sveriges Riksbank at all forecast origins and using it would imply considerably less than 25 observations. This means that there will be $25 \times 2 \times 3 = 150$ different data sets with 25 observations in each.
2. The above data sets are pooled so that all forecast origins with the same inflation measure and coverage probability are pooled into one data set. This means that there will be $2 \times 3 = 6$ data sets with $25 \times 25 = 625$ observations in each.

The test procedure requires that a number of non-zero conditional probabilities can be calculated. If some events do not occur the probabilities are undefined and the tests cannot be performed. If so, this is a clear indication that the number of observations in each data set (25) is too small. The problem also exists even if we have non-zero conditional probabilities for all events since the absolute number of observed outcomes for some events may be too small for the test to be reliable. For instance, at the 90 percent coverage level the expected number of observations outside the interval is less than three ($0.1 \times 25 = 2.5$) and recent research shows that even the standard rule of thumb with 5 observations for the two-tailed tests may be too low (Andrés and Tejedor, 2003). This motivates the use of pooled data even if only the unconditional coverage test then can be performed.

The means of the indicator variables are plotted in Figure 2. A general conclusion from these three plots (one for each coverage probability) is that the mean for the CPI indicator series almost always is below the *ex ante* coverage probability. For KPIX the outcome is more mixed. For shorter

Table 1: Mean of indicators for pooled data.

Inflation measure	90%	75%	50%
CPI	75.8	53.6	34.1
KPIX	79.7	65.3	44.5

Table 2: Test statistics for the unconditional coverage test using pooled data. Critical value $\chi_{0.95}^2(1) \approx 3.84$.

Inflation measure	90%	75%	50%
CPI	104.46	134.28	63.93
KPIX	59.03	29.56	7.42

forecast horizons the means are below but for longer horizons above the *ex ante* coverage probability. For the 75 and 90 percent coverage probabilities the difference between short and long horizon in this sense is about one year.

The means of the pooled data are found in Table 1 which confirms the tendency that the means are below the coverage probabilities. The question is if they deviate significantly from the coverage probability?

5 Results

The results for the pooled data are in in Table 2: All series fail the unconditional coverage test because they have too narrow intervals. The best result is achieved by the KPIX 50 percent coverage where the difference between observed and expected coverage is only 5 percentage points. Generally, the CPI intervals do worse than the KPIX intervals.

Results (i.e., the LR_{cc} test statistic) for the 150 individual forecast origins are reported in Figure 3–5. All numerical values are also available in Lundholm (2010d, Tables 3–5). Whenever a conditional probability is undefined the test statistic is not calculated.

The most important series are of course those with a 90 percent coverage probability. In Figure 3 we see that almost all interval forecast up to a one year forecast horizon fail the conditional coverage test. Above one year the result is mixed. For CPI there are several forecasts that pass the test but for KPIX the test statistic is undefined in many cases. For the 75 percent coverage probability (Figure 4) the results for KPIX are almost the same as but with more forecasts with longer forecast horizons passing the test. For CPI the results are in a sense opposite; forecasts passing the test tend to have short horizon whereas the longer horizons do not pass the test. The

forecasts with 50 percent coverage probability perform best (Figure 5): The KPIX forecast tend to pass the test but with mixed results for CPI. In latter case a number of longer forecast horizons do not pass the test.

The general result is that the interval forecast in very few instances do not pass the the test and the reason is that the coverage is too low.

6 Conclusions

This study has tested whether the inflation interval forecasts of Sveriges Riksbank are calibrated according to the *ex ante* stated coverage probabilities. In total 150 different forecast series with 25 observations in each were tested for conditional coverage. Due to limitations in data, observations from all forecast origins were pooled for each forecasted inflation measure (CPI and and KPIX) and each coverage probability (50, 75 and 90 percent) and these series were also tested for unconditional coverage.

The general result is that the interval forecast in very few instances do not pass the the test and the reason is that the coverage is too low. That is, Sveriges Riksbank tend to underestimate the forecast uncertainty. This is also what Dowd (2004) but it is opposite the results regarding the Bank of England (Monetary Policy Committee, MPC) where Wallis (2003, 2004) conclude that the MPC overstates the forecast uncertainty.

In Lundholm (2010a) the corresponding point forecasts from Sveriges Riksbank where evaluated in terms of bias. There the conclusion was that there was a clear difference between shorter horizons (negative bias) and longer horizons (positive bias). Such a clear difference can here only be seen in the mean of the indicator variable for the 90 percent coverage where the sample mean is closer to the *ex ante* for the longer horizon. However, with only 25 observations the test statistic for conditional coverage in most cases is undefined and no test can be performed. Pooling data over forecast horizons the tendency from the shorter horizons with too narrow interval dominates.

One possible explanation for these results are the bias in the point forecasts since they position the interval forecasts. Even if the interval forecasts are well calibrated a downward or upward position of a correct interval may lead a failure in these tests. Another explanation is that there is a true underestimation in the point forecast uncertainty due to the variability in the actual inflation during the forecasted period; see Figure 1.

References

Andersson MK, Karlsson G, Svensson J (2007). “En utvärdering av Riksbankens prognosförmåga.” *Penning- och valutapolitik*, **2007**(3), 59–74. URL http://www.riksbank.se/upload/Dokument_riksbank/

Kat_publicerat/PoV_sve/sv/andersson_karlsson_svensson2007_3svny.pdf.

- Andrés AM, Tejedor IH (2003). “On the minimum expected quantity for the validity of the chi-squared test in 2×2 tables.” *Journal of Applied Statistics*, **27**, 807–820. URL <http://www.informaworld.com/smpp/ftinterface~content=a713674807~fulltext=713240930~frm=content>.
- Berg C (2000). “Inflation forecast targeting: The Swedish experience.” In MI Blejer, A Ize, AM Leone, S Werlang (eds.), “Inflation targeting in practice: Strategic and operational issues and application to emerging market economies,” pp. 28–36. International Monetary Fund. URL www.imf.org/External/Pubs/FT/seminar/2000/Targets/strach4.pdf.
- Bergvall A (2005). “Uvärdering av Konjunkturinstitutets prognoser.” Specialstudie Nr 5, mars 2005, Konjunkturinstitutet., URL http://www.konj.se/download/18.da9067102c97d16a280003793/specialstud_5.pdf.
- Blix M, Friberg K, Åkerlind F (2002). “An evaluation of forecasts for the Swedish economy.” *Penning- och valutapolitik*, **2002**(3), 39–73. URL http://www.riksbank.se/upload/Dokument_riksbank/Kat_publicerat/Artiklar_PV/pv02_3_artikel3.pdf.
- Blix M, Sellin P (1998). “Uncertainty bands for inflation forecasts.” Working paper 65, Sveriges Riksbank, URL <http://www.riksbank.se/upload/1013/98nr65.pdf>.
- Blix M, Sellin P (1999). “Inflation forecasts with uncertainty intervals.” *Penning- och valutapolitik*, **1999**(2), 12–28. URL http://www.riksbank.se/upload/Dokument_riksbank/Kat_publicerat/Artiklar_PV/qr99_2_artikel2.pdf.
- Christoffersen PF (1998). “Evaluating interval forecasts.” *International Economic Review*, **39**, 841–862. URL <http://www.jstor.org/page/termsConfirm.jsp?redirectUri=/stable/pdfplus/2527341.pdf>.
- Clements MP (2004). “Evaluating the Bank of England density forecasts of inflation.” *Economic Journal*, **114**, 844–866. URL <http://www.jstor.org/page/termsConfirm.jsp?redirectUri=/stable/pdfplus/3590241.pdf>.
- Dowd K (2004). “The Swedish inflation fan charts: An evaluation of the Riksbank’s inflation density forecasts.” **2**, 73–82.

- Jansson P, Vredin A (2003). “Forecast-based monetary policy: The case of Sweden.” *International Finance*, **6:3**, 349–380. URL <http://dx.doi.org/10.1111/j.1367-0271.2003.00122.x>.
- Konjunkturinstitutet (2002). “Peningpolitiken 1999–2001.” Specialstudie Nr 1, mars 2002, Konjunkturinstitutet., URL http://www.konj.se/download/18.596dfd2811604b2cde9800017179/Specialstud_1.pdf.
- Lundholm M (2010a). “Are inflation forecasts from major Swedish forecasters biased?” Research paper in economics 2010:10, Department of Economics, Stockholm University, URL http://swopec.hhs.se/sunrpe/abs/sunrpe2010_0010.htm.
- Lundholm M (2010b). *sifds: Swedish inflation forecast data set*. R package version 0.9, URL <http://CRAN.R-project.org/package=sifds>.
- Lundholm M (2010c). “sifds: Swedish inflation forecast data set 1999:Q2–2005:Q2.” Research paper in economics 2010:4, Department of Economics, Stockholm University, URL http://swopec.hhs.se/sunrpe/abs/sunrpe2010_0004.htm.
- Lundholm M (2010d). “Sveriges Riksbank’s inflation interval forecasts 1999–2005: Technical documentation.” Unpublished manuscript, Department of Economics, Stockholm University, URL http://people.su.se/~lundh/projects/interval/interval_techreport.pdf.
- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Wallis KE (2003). “Chi-squared tests of interval and density forecasts, and the Bank of England’s fan charts.” *International Journal of Forecasting*, **19**, 165–175. URL [http://dx.doi.org/10.1016/S0169-2070\(02\)00009-2](http://dx.doi.org/10.1016/S0169-2070(02)00009-2).
- Wallis KE (2004). “An assessment of Bank of England and National Institute inflation forecast uncertainties.” *National Institute Economic Review*, **189**, 64–71. URL <http://ner.sagepub.com/cgi/reprint/189/1/64>.
- Woodward M (2007). “The case for forecast targeting as a monetary policy strategy.” *Journal of Economic Perspectives*, **21**(4), 3–24. URL <http://www.jstor.org/page/termsConfirm.jsp?redirectUri=/stable/pdfplus/30033749.pdf>.

Figure 1: Yearly changes in consumer prices (CPI and KPIX) 1999:M5-2007:M07.

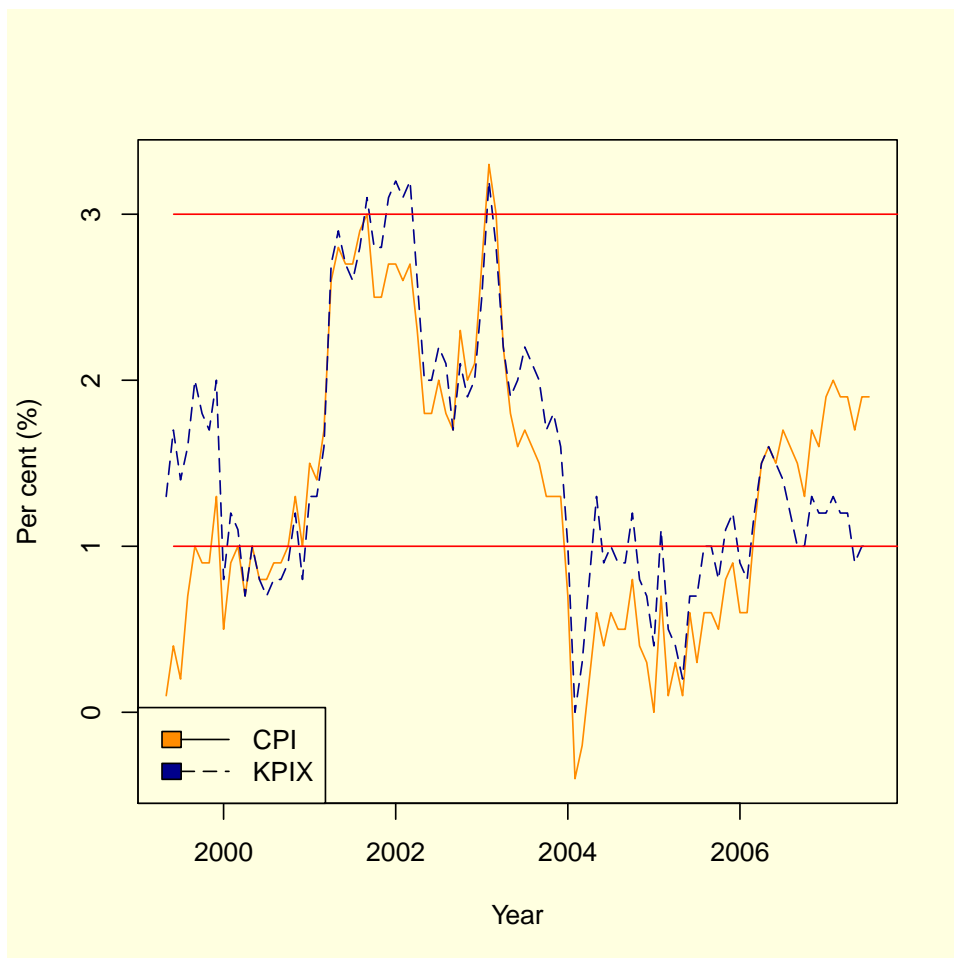


Figure 2: Average coverage $p \in \{0.50, 0.75, 0.90\}$. Horizons 1 – 25.

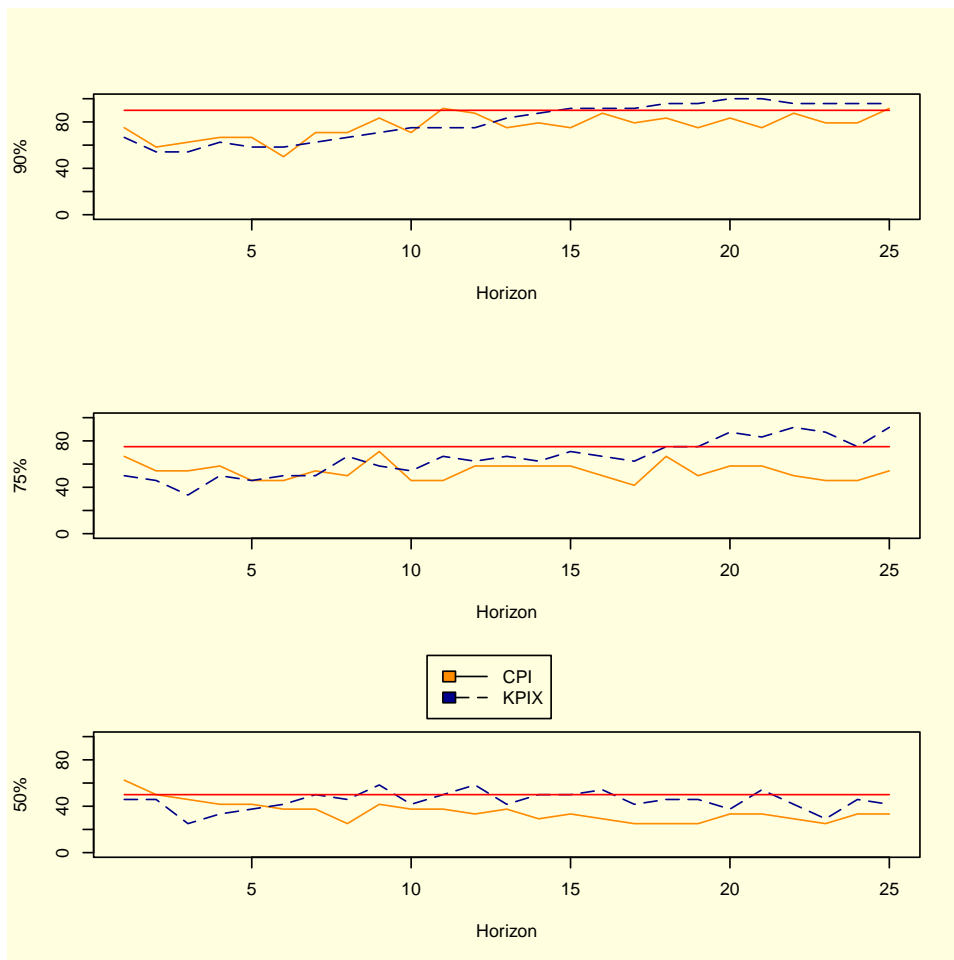


Figure 3: LR_{cc} for CPI and KPIX 90 percent coverage. Horizons 1 – 25.

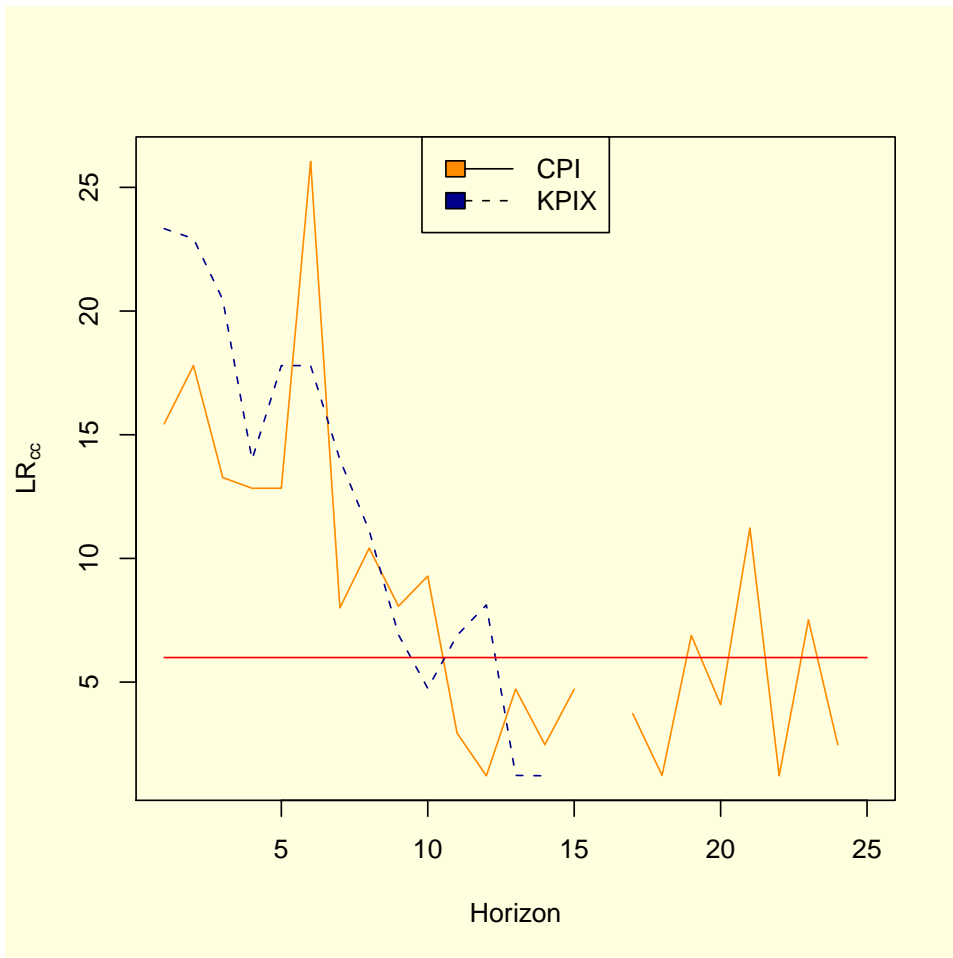


Figure 4: LR_{cc} for CPI and KPIX 75 percent coverage. Horizons 1 – 25.

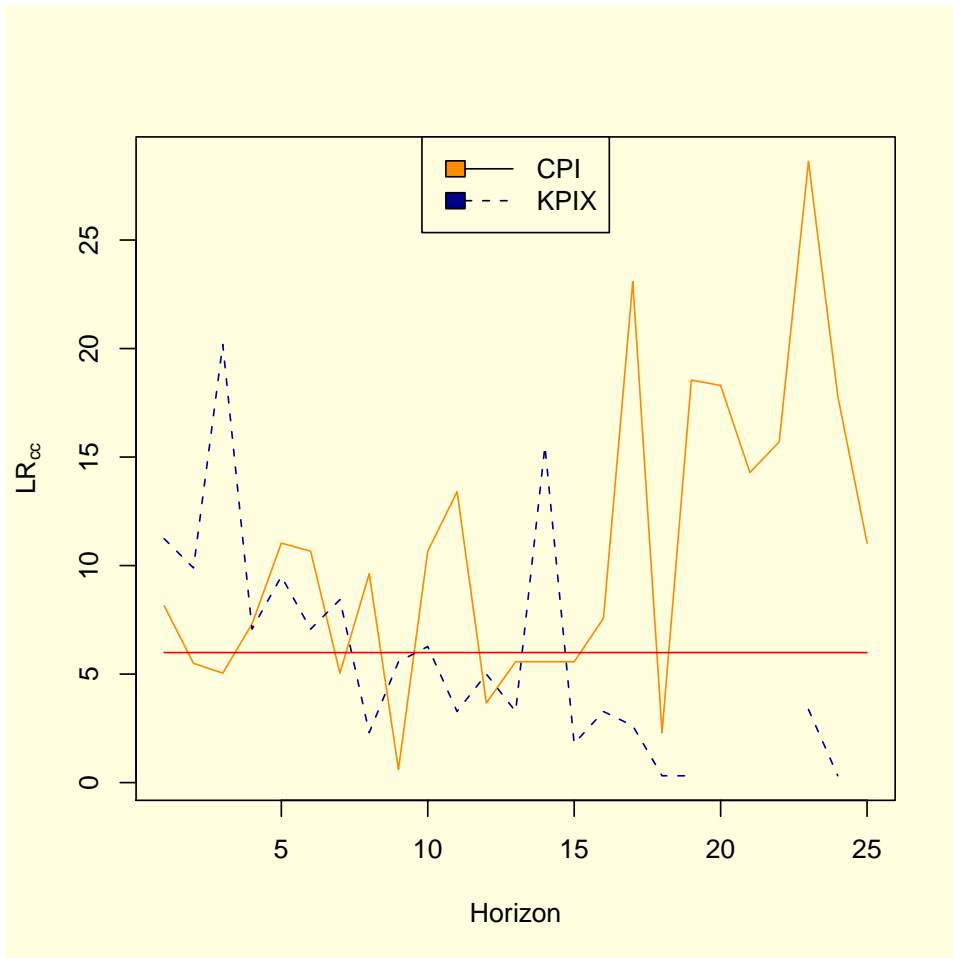


Figure 5: LR_{cc} for CPI and KPIX 50 percent coverage. Horizons 1 – 25.

